Detecting Sensory Textures with Rheological Characteristics from Recipe Sharing Sites

Hiroshi Uehara
Faculty of Data Science
Rissho University
Saitama, Japan
uehara@ris.ac.jp

Daichi Mochihashi
Department of Statistical Inference and Mathematics
The Institute of Statistical Mathematics
Tokyo, Japan
daichi@ism.ac.jp

***Abstract*—This study tries to estimate texture of gel related dishes based on both the data from recipe sharing sites, and research results of food science. Since most of recipes are not accompanied by sufficient information about what kind of textures they realize, we propose a method to estimate characteristics of textures for each recipe by applying a joint topic model to bridge sensory texture terms in a recipe sharing site with corresponding quantitative textures resulting from food science research. The result shows that the estimated texture terms for dishes are consistent with rheology, the quantitative textures provided by related food science research.**

***Index Terms*—computational cooking, topic model, food science, rheology, food texture**

I. Introduction

Most of recipes posted on recipe sharing sites (hereafter, referred to as posted recipes) are not accompanied by sufficient information of their cooked characteristics such as taste, aroma, and texture. Therefore, users have to select their favorites with insufficient knowledge of these cooked results. Especially among these cooked characteristics, texture is the most important attribute for solid foods: it occupies top 32.1% of all the food attributes associated by subjects over a wide variety of foods with words of attributes [1]. With this respect, texture has been addressed by a wide variety of studies ranging from international standardization of textural attributes by ISO [2] to empirical texture evaluations in food science research [3]–[5]. Nevertheless, the findings of these studies are not linked to specifying texture characteristics of posted recipes.

This study tries to detect texture information of gel related dishes from collective posted recipes. Gelling agents are indispensable ingredients for controlling and preserving food texture of a wide variety of foods including sweets, sauce, dressing, and frozen foods. Our proposal enables the texture information to represent not only texture terms, but also the relationship between the texture terms and concentrations of ingredients. Then, these concentrations enable linkages between the texture terms and quantitative texture resulted from food science research. As such, this study aims to provide home cooking users with reliable information of texture, thereby enabling to find their favorite recipes in more suitable manner.

Related work is surveyed in Section II. Then, our proposal is described in Section III followed by introduction of data for our experiments in Section IV. Subsequently, the results and evaluations are described in Section V. Section VI concludes our study.

II. Related Work

A wide variety of texture terms exist to express the same textural perceptions, further, the variations are different depending on language by each country [7], [8]. These variations have been an inherent issue for objective evaluation in empirical textural studies. With this background, many of studies have adopted instrumental evaluation instead of sensory evaluations(i.e., perceptive questionnaire to subjects) to represent texture as universal quantitative attributes.

However, these attributes are not intuitively interpretable in terms of users' perceived texture. Very few quantitative research in food science have tried to interpret the quantitative results to perceived texture [13], [14] Nevertheless, proper interpretations of a wide variety of texture terms are still ambiguous, because the results of these studies are acquired from questionnaires to limited number of subjects.

Our study addresses a wide variety of texture terms from a large amount of posted recipes and enabled them to be linked to corresponding quantitative results of food science research. The result of the linkage is validated by [10], comprehensive allocations for Japanese texture terms.

Thus, our proposal bridges research on natural language processing and quantitative research in food science. Although research on computational cooking have analyzed a wide variety of data including recipe texts, and images etc., data concerning physical characteristics of food have not been addressed [12].

III. Method

In order to detect classified texture information from posted recipes, we adopt joint topic model [11], a kind of latent Dirichlet allocation algorithms (LDA). LDA identifies topics, each of which represents classified pattern of distribution. While conventional LDA represents topics by a single type of data (e.g., distribution of texture terms), joint topic model represents topics by multiple types of data. In our proposal, each joint topic comprises two types of distribution, the one represents a pattern of texture terms and the other represents concentrations of ingredients.

Once the distribution of the concentrations is acquired, empirical texture studies in food science can be evaluated in terms of the topics, because the setting in these studies are also represented as the distribution of concentration of ingredients. This method allows topics to be texture are allocated to empirical studies so that patterns of texture terms are linked to quantitative texture. The validity of these linkages can be evaluated by the Texture Profile [10]. Followings elaborate the proposal.

A. Constructing dataset

Our dataset comprises a sequence of texture terms and a distribution of concentration of ingredients on each posted recipe. In order to identify texture terms from posted recipes, we make use of Comprehensive Japanese Texture Terms[1] as a dictionary, which is constructed based on studies on Japanese texture terms including [10]. Here, each texture term is annotated by categories representing quantitative characteristics such as hardness and adhesiveness. We construct the dictionary by extracting all the texture terms belonging to the categories of hardness, cohesiveness, and adhesiveness in Comprehensive Japanese Texture Terms, in order to compare texture terms with the quantitative attributes described later. As the result, the dictionary includes 288 texture terms. Then, all the texture terms appeared in the descriptions of posted recipes are extracted by referring to the dictionary.

Meanwhile, we focus on characteristics of gel related texture as a trial so that unrelated texture to gel is excluded by word2vec. All the descriptions of retrieved posted recipes are trained by word2vec. Then, if similar words to the extracted texture terms include ingredient terms unrelated to gel, the texture terms are excluded. For example, a recipe of mousse with topping of nuts might create texture terms representing crispy. In such case, nuts appear in similar words to the texture terms of crispy.

On the other hand, the quantity of each ingredient is described in different manner depending on each recipe, such as "spoonful of", "two cups of", "g", "cc". We convert all the variations to weight (grams). In case the unit of quantity is described as a volume (i.e., "spoonful of", "cc" etc.), a specific weight against water is taken into account to convert to grams. Measuring spoons for cooking are standardized by nations. For example, the capacity of a small measuring spoon is 5 mL in Japan. We make use of these standard for the conversion. Then, concentrations of all the ingredients in each recipe are derived as the ratios of each weight against total weight in each recipe. Then, these feature x of concentrations are converted to information quantity $-\log(x)$ because x represents a ratio whose small difference will affect considerable difference of textures.

B. Topic modeling

Joint topic model is applied for acquiring topics comprised of a pattern of texture terms accompanied by corresponding

[1]Constructed by The National Agriculture and Food Research Organization in Japan https://www.naro.affrc.go.jp/org/nfri/yakudachi/terms/texture.html

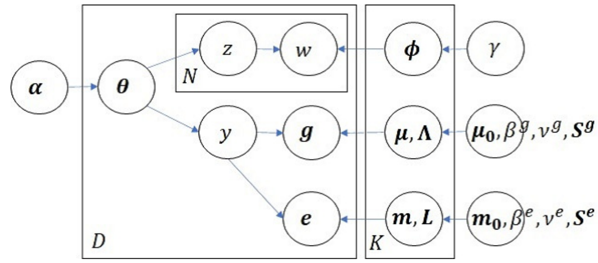


Fig. 1. A graphical model for topic of gel related texture

ingredients' concentrations. In each topic, texture terms form a categorical distribution of its own, whereas corresponding ingredients' concentrations form a multi-normal distribution. Here, an expectation of each random variable of the multi-normal distribution represents an average concentration value of each ingredient. According to the studies of gel related dishes in food science, small amount of gel related ingredients (i.e., gelatin, agar etc.) principally affect the resulting texture with subordinate effects by large mixing amount of emulsions[2] such as milk and cream [19].

Reflecting these matters, the graphical model for the proposed joint topic model is designed as in Fig.1. In Fig.1, ϕ represents the topic of texture terms, while μ and $\mathbf{\Lambda}$ represent the topic of gel ingredient concentrations. Similarly, $\mathbf{m}$ and $\mathbf{L}$ represent the topic of emulsion ingredient concentrations. All these parameters are generated through the generative procedure of the latent topics, $\mathbf{z}$ and $\mathbf{y}$, where $\mathbf{z}$ is a latent topic for a texture word $\mathbf{w}$ in each recipe, and $\mathbf{y}$ is a latent topic for $\mathbf{g}$, a vector of gel ingredient concentrations in each recipe. Both of $\mathbf{z}$ and $\mathbf{y}$ are subject to latent topic distribution,θ so that both the latent topic of texture terms and gel concentrations are linked with each other. On the other hand, the topic of emulsions is subordinate to the latent topic of gel concentrations, $\mathbf{y}$. The equation (1) corresponds to the generative model given in **Figure** 1.

$$
\begin{aligned}
&p\left(\mathbf{W}, \mathbf{Z}, \theta, \phi, \mathbf{g}, \mathbf{y}, \mu, \mathbf{\Lambda}, \mathbf{e}, \mathbf{L}\right) = \\
&\mathrm{Dir}\left(\theta|\alpha\right) \mathrm{Dir}\left(\phi|\gamma\right) \prod_{n} \mathrm{Mult}\ \left(z_{d_n}|\theta_d\right) \mathrm{Mult}\left(w_{d_n}|\emptyset_{\mathbf{k}}, z_{d_n}\right) \\
&\times \mathcal{NW}\left(\mu_{\mathbf{k}}, \mathbf{\Lambda}_{\mathbf{k}}|\mu_{\mathbf{0}}, \beta^g, \nu^g, \mathbf{S}^{\mathbf{g}}\right) \prod_{\mathbf{d}} \mathrm{Mult}\left(y_d\,|\theta_d\right) \mathcal{N}(\mathbf{g}_{\mathbf{d}}|\mu_{\mathbf{k}}, \mathbf{\Lambda}_{\mathbf{k}}, y_d) \\
&\times \mathcal{NW}\left(\mathbf{m}_{\mathbf{k}}, \mathbf{L}_{\mathbf{k}}|\mathbf{m}_{\mathbf{0}}, \beta^e, \nu^e, \mathbf{S}^{\mathbf{e}}\right) \prod_{\mathbf{d}} \mathcal{N}(\mathbf{e}_{\mathbf{d}}|\mathbf{m}_{\mathbf{k}}, \mathbf{L}_{\mathbf{k}}, y_d) \qquad (1)
\end{aligned}
$$

Then, the generative process for the model is as follows.

1) For $k \in 1, ..., K$:
   a) $\phi_{\mathbf{k}} \sim \mathrm{Dir}\left(\gamma\right)$
   b) $\mu_{\mathbf{k}}, \mathbf{\Lambda}_{\mathbf{k}} \sim \mathcal{NW}(\mu_{\mathbf{0}}^{\mathbf{g}}, \beta^g, \nu^g, \mathbf{S}^{\mathbf{g}})$
   c) $\mathbf{m}_{\mathbf{k}}, \mathbf{L}_{\mathbf{k}} \sim \mathcal{NW}(\mu_{\mathbf{0}}^{\mathbf{e}}, \beta^e, \nu^e, \mathbf{S}^{\mathbf{e}})$
2) For $d \in 1, ..., D$:
   a) Draw $\theta_d \sim \mathrm{Dir}\left(\alpha\right)$
   b) For $n \in 1, ..., N_d$:
      i) Draw $z_{dn} \sim \mathrm{Mult}\left(\theta_d\right)$
      ii) Draw $w_{dn} \sim \mathrm{Mult}\left(\phi_{\mathbf{z_{dn}}}\right)$

[2]A mixture of two or more liquids that are normally unmixable such as oil in water.

c) Draw $y_d \sim \mathrm{Mult}(\theta_d)$
d) Draw $g_d \sim \mathcal{N}(\mu_{y_d}, \Lambda_{y_d})$
e) Draw $e_d \sim \mathcal{N}(m_{y_d}, L_{y_d})$

Here,

$\phi_{\mathbf{k}}$: Parameters of word distribution in each topic, k, with hyper parameters, γ
$\mu_{\mathbf{y_d}}, \mathbf{\Lambda_{y_d}}$: Parameters of gel distribution in k, with hyper parameters, $\mu_{\mathbf{0}}^{\mathbf{g}}, \beta^g, \nu^g, \mathbf{S^g}$
$\mathbf{m_{y_d}}, \mathbf{L_{y_d}}$: Parameters of emulsion distribution in k
θ_d : Parameters of topic distribution in d(each recipe)
α : Hyper parameters of θ_d
z_{dn} : A latent topic for a texture word, w_{dn}, appeared in n^{th} place in d
y_d : A latent topic for g_d, gel concentrations in d
e_d : Emulsion concentrations in d

C. Inference of topics

Given the graphical model above, the latent topics are inferred by Gibbs sampling. And the topics concerning texture terms accompanied by concentration vectors of gels are acquired through the sampling. (2)(3)(4) are the sampling formulas.

1) Sampling z_{dn} :

$$p\left(z_{dn} = k \middle| \mathbf{W}, \mathbf{Z^{-dn}}, \alpha, \gamma, \mathbf{Y}, \mathbf{X}, \mu, \mathbf{\Lambda}, \mu_{\mathbf{0}}^{\mathbf{g}}, \beta^g, \nu^g, \mathbf{S^g}\right)$$
$$\propto (N_{dk}^{-dn} + M_{dk} + \alpha) \frac{N_{k,w_{dn}}^{-dn} + \gamma}{N_k^{-dn} + \gamma \mathbf{V}} \quad (2)$$

Here,
N_{dk} : The number of texture words belonging to k in d
N_k: The number of texture words belonging to k
N_d : The number of texture words in each recipe
N_{kw} : The number of each texture term,w, belonging to k
M_d : The number of gel ingredient vectors in d
M_{dk}: The number of gel ingredient vectors belonging to k in d
N_{dk}^{-dn} : The number of texture words belonging to k in d except dn. (Hereafter, "-" of upper suffix represents similar exceptions)

2) Sampling y_d :

$$p(y_d = k | \mathbf{W}, \mathbf{Z}, \alpha, \gamma, \mathbf{Y^{-d}}, \mathbf{G}, \mu, \mathbf{\Lambda}, \mu_{\mathbf{0}}^{\mathbf{g}}, \beta^g, \nu^g, \mathbf{S^g})$$
$$\propto \frac{N_{dk} + M_{dk}^{-d} + \alpha_k}{N_d + M_d - 1 + \sum_k \alpha_k} \times \mathcal{N}(\mathbf{e_d} | \mu_{\mathbf{y_d}}, \mathbf{\Lambda_{y_d}}) \quad (3)$$

3) Inferring topic parameters :
The topic parameters, μ_k and $\mathbf{\Lambda_k}$ for gel ingredient concentrations, are sampled along with the sampling of the latent topics above. The formula is as follows. m_k and $\mathbf{L_k}$ are subject to the same formula.

$$p\left(\mu_{\mathbf{k}}, \ \mathbf{\Lambda_k} \middle| \mathbf{Y}, \mathbf{G}, \mu^{-\mathbf{k}}, \mathbf{\Lambda^{-k}}, \mu_{\mathbf{0}}, \beta, \nu^{\mathbf{g}}, \mathbf{S^g}\right)$$
$$\propto \mathcal{N}\left(\mu_{\mathbf{k}} \middle| \mu_{\mathbf{c}}, (\beta^{\mathbf{g}} \mathbf{\Lambda_c})^{-\mathbf{1}}\right) \mathcal{W}(\mathbf{\Lambda_k} | \nu_c^g, \mathbf{S_c^g}) \quad (4)$$

Here,
$\mathcal{W}(\mathbf{\Lambda_k} | \nu_c^g, \mathbf{S_c^g})$: Wishart distribution with hyperparameters ν_c^g and $\mathbf{S_c^g}$

$(\mathbf{S_c^g})^{-1} = (\mathbf{S^g})^{-1} + \sum_{\mathbf{g_d} \in \mathbf{topic\ k}} (\mathbf{g_d} - \bar{\mathbf{g}})(\mathbf{x_d^g} - \bar{\mathbf{g}})^{\mathbf{T}} + \frac{N_k \beta g}{N_k + \beta^g} (\bar{\mathbf{g}} - \mu_{\mathbf{0}})(\bar{\mathbf{g}} - \mu_0)^{\mathbf{T}}$

$\bar{\mathbf{g}} = \frac{1}{N_k^g} \sum_{\mathbf{g_d} \in topic\mathbf{k}} \mathbf{g_k}, \ \ \mu_{\mathbf{c}} = \frac{N_k^g \bar{\mathbf{g}} + \beta^g \mu_0}{N_k^g + \beta^g}$

$\mathbf{\Lambda_c} = (N_k^g + \beta^g) \mathbf{\Lambda_k}$, $\nu_c^g = \nu^g + N_k^g$

Meanwhile, ϕ_k, the topic of texture terms is derived by the following formula after the convergence of Gibbs Sampling, together with θ_d, the topic distribution in each recipe.

$$\phi_{kv} = \frac{N_{kv} + \gamma}{N_k + \gamma V}, \theta_{dk} = \frac{N_{dk} + M_{dk}}{N_d + M_d + \sum \alpha} \quad (5)$$

4) Making linkage between topic and quantitative texture : Once the topics are acquired, patterns of texture terms are accompanied by corresponding vectors of ingredient concentrations. By comparing the vectors and the empirical settings of gel concerning research in food science [3]–[5], [15]–[17], the quantitative results of texture are accompanied by the relevant topics. Kullback–Leibler divergence is applied for deriving most similar topic to the settings of the research. Then, the quantitative texture is linked to corresponding texture terms, qualitative ones in the topics. And the linkages are validated by referring to the dictionary introduced in Section II, A. Constructing dataset, where each texture term is annotated by the category representing quantitative attributes. As described before, the gel ingredient concentrations primarily determine the resulting texture so that only the gel ingredient concentrations are used for the comparison.

TABLE I
EMPIRICAL DATA FROM RESEARCH IN FOOD SCIENCE

	Gels			**Quantitative texture (rheological unit)**		
Data	Gelain	Kanten	Agar	Hardness	Cohesive-ness	Adhesive-ness
1	0.018	0	0	0.20	0.6	0.1
2	0.02	0	0	0.3	0.59	0.04
3	0.025	0	0	0.72	0.17	0.57
4	0.03	0	0	2.78	0.31	0.42
5	0.03	0	0.03	3.01	0.35	12.6
6	0	0.008	0	2.2	0.12	0
7	0	0.01	0	3.5	0.1	0
8	0	0.012	0	5.0	0.8	0
8	0	0.02	0	5.67	0.03	0
10	0	0	0.008	1.0	0.48	0
11	0	0	0.01	1.5	0.33	0.01
12	0	0	0.012	2.7	0.28	0.02
13	0	0	0.03	2.21	0.20	1.95

IV. DATA

Gel related posted recipes are collected from Cookpad[3], the largest recipe sharing site Japan. And empirical research on gel related texture are collected from the journals of food science. Followings are the details of the data.

[3]https://cookpad.com/

A. Collecting recipes from a recipe sharing site

Among a variety of gels, gelatin, kanten (Japanese agar), and agar are widely used for home cooked dishes in Japan so that the number of the recipes in Cookpad is 45,000, 15,000 and 3,000, respectively. Although the total recipes amount to 63,000, much smaller amount of them, approx.10,000 recipes are accompanied by the descriptions including texture terms in the dictionary. All these 10,000 recipes include gels, either of gelatin, kanten, or agar, or the mixture of them with a variety of the concentrations. Then, each recipe is converted to three kinds of features, a sequence of texture terms, a vector of gel ingredient concentrations, and a vector of emulsion ingredient concentrations. The features are constructed as below.

Sequence of texture terms
: Term frequency by each texture term extracted by Section II, *A. Constructing dataset*.

Vector of gel ingredient concentrations
: Information quantity $-\log(x)$ of concentration x by each gel (i.e., gelatin, kanten, and agar). Concentrations are derived based on Section II, *A. Constructing dataset*.

Vector of emulsion ingredient concentrations
: Information quantity $-\log(x)$ of concentration x by each emulsion. Here, emulsions consist of six types: sugar, egg albumen, egg yolk, raw cream, milk, and yogurt.

Out of these recipes, the ones including large portions of ingredients unrelated to gels and emulsions are excluded. For example, if fruits occupy a large portion, the recipes might reflect texture unrelated to gels. To focus on the gel related texture in our trial, the recipes occupied by more than 10 percent of unrelated ingredients are excluded. As the result, approx.3,000 recipes become the dataset, which include 41 texture terms out of 288 terms in the dictionary.

B. Collecting empirical data from research in food science

Table I shows the empirical settings and the quantitative results collected from six research [3]–[5], [15]–[17] concerning gel related ingredients. The settings are the concentration vectors of the three types of gels, the same form of data as the gel ingredient vector in the topics. Corresponding quantitative results comprise three attributes, hardness, cohesiveness, and adhesiveness measured by rheometers, the instruments for food texture proposed by [9].

Figure 2 illustrates how rheometers work. Rheometers imitate chewing actions in a mouth by descending/ascending the disc-like probe onto a food sample. Positive forces for compressing the samples are measured during the descending actions, and negative ones are measured during the ascending actions, which represent the stickiness. F1 in **Figure** 2 indicates a peak force during a compression. If compressing force becomes larger than the resilience of a food sample, the food shape begins to collapse, and the force for the compression becomes smaller than F1. The area of negative force in b appears during ascending. Based on these principles, the quantitative attributes are defined as follows.

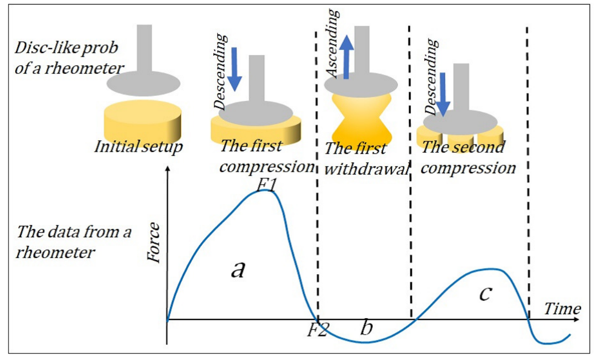


Fig. 2. Quantitative measurement by a rheometer

Hardness
: The peak force of the first compression (F1), which represents perceived hardness at the first bite.

Cohesiveness
: The ratio of the force for the second compression to the one for the first compression, c/a. Small cohesiveness represents perceived easiness of swallowing, because the small cohesiveness indicates collapsed piece of food becomes smaller by fewer bites.

Adhesiveness
: Cumulative forces during the first ascending action(b), which represents perceived stickiness.

The unit of measurements for these attributes are different depending on the research, because the unit is not necessarily standardized among the products of rheometers. So, we converted all the values of the measurement to the unit of RU (rheological unit), which is the most popular one adopted by related research.

V. Results and Evaluations

A. Linkage between texture terms and the quantitative attributes

Table II(a) shows topics derived by Section II, *C* in our method. Each topic comprises gel concentrations (the column "gels:concentrations") and corresponding texture terms (the column "Texture terms"). The texture terms are sorted by the probability of each term (the value in the parentheses). The column "# Recipes" represents the number of recipes in Cookpad belonging to each topic. A topic for each recipe is specified by the maximum probability in θ_d, which represents topic distributions for each recipe as in the formula 5.

The last column of **Table** II(a) indicates similarities to empirical results in **Table** I derived by Section II, *D* in our method. For example, topic 8 is similar to the research results 1 and 2 in terms of gels concentrations. As such, the same topic could be corresponding to multiple research results, because the concentrations of topics are not completely classified in accordance with the ones of the research results. These similarities approximate the coincidence between the quantitative texture in **Table** I and the texture terms in the topics. For example, both the gelatins, 0.018 and 0.02 in **Table** I are characterized by "furufuru"(softness) in the topic

TABLE II
TOPIC ASSIGNMENT TO EMPIRICAL DATA OF FOOD SCIENCE RESEARCH

(a) Acquired topics by the joint topic model and the assignment to the empirical data in **Table** I

Topic	Gels:concentration	Texture terms(probability)[Japanese texture terms] ordered by probability	# Recipes	**Table** I
7	gelain:0.005	Soft, tender(0.877)[yawarakai] Light, soft and fluffy(0.123)[howat]	73	
4	gelain:0.007	Light, short(0.93)[karui] Soft swollen,and fluffy(0.07)[fukkura]	74	
0	gelain:0.012	Sticky, elastic and chewy(1.0)[mochimochi]	152	
8	gelain:0.014	Soft and slightly wobbly, easy to break(1.0) [furufuru]	300	1,2
3	gelain:0.054	Hard, firm, stiff, tough, rigid(0.307)[katai] Resilient, firm and slightly sticky(0.245)[muchimuchi] Mushy; having lost its original shape(0.129)[gucha] Thick, resistant to flow(0.089)[potteri] Elastic and slightly wobbly(0.062)[burunburun] Dry, crumbly and not compact(0.06)[bosoboso] Thick and heavy, resistant to flow(0.055)[botet] Crisp; material is cut off or shear off easily(0.029)[shakusyaku] Elastic and slightly wobbly(0.022)[buruburu]	38	3,4
5	agar(0.009) gelain(0.009)	Soft elastic and slightly sticky, slightly wobbly(1.0)[purupuru]	1046	5
2	agar(0.016)	Sticky, viscous and thick(0.445)[nettori] Crispy, sound emitted by biting slightly hard foods(0.255)[purit] Thick and viscous, resistant to flow(0.21)[mottari] Crumbly and soft(0.08)[horohoro] Very sticky and viscous(0.01)[necchiri]	62	10,11,12,13
6	gelain:0.003 kanten:0.002	Soft and fluffy(1.0)[fuwafuwa]	1200	
1	kanten:0.004	Thin, loose, easy to deform(0.487)[yuruyuru] Sticky, viscous and watery(0.432) [bechat] Soft, swollen and somewhat elastic(0.027)[fukahuka] Firm and resilient(0.027)[burit]	60	
9	kanten:0.021	Heavy, dense(0.27)[dossiri] Slippery, smooth and wet surface(0.165)[churuchuru] Soft elastic and slightly sticky(0.1)[punipuni]: Soft, not taut(0.074)[kutat] Firm and resilient(0.069)[burinburin] Crunchy(0.064)[korit] Thick, heavy(0.057)[daradara] Dry and crispy(0.055)[karat] Cracking open, fizzy(0.055)[hajikeru] Heavy(0.054)[omoi]	31	6,7,8,9

(b) Empirical research results of Bavarois and Milk jelly, and the topic assignment

Dish	**Quantitative texture**			**Gel concentrations**			**Emulsion concentrations**						Assigned topic
	Hardness	Cohesive-ness	Adhesive-ness	Gelain	Kanten	Agar	Sugar	Egg albumen	Egg yolk	Raw cream	Milk	Yogurt	
Bavaois	3.860	0.809	0.095	**0.025**	0	0	0	0	0.08	0.2	0.4	0	3
Milk jelly	1.83	0.27	0.44	**0.025**	0	0	0.032	0	0	0	0.787	0	3
Data 3 in **Table** I	0.72	0.17	0.57	**0.025**	0	0	0	0	0	0	0	0	3

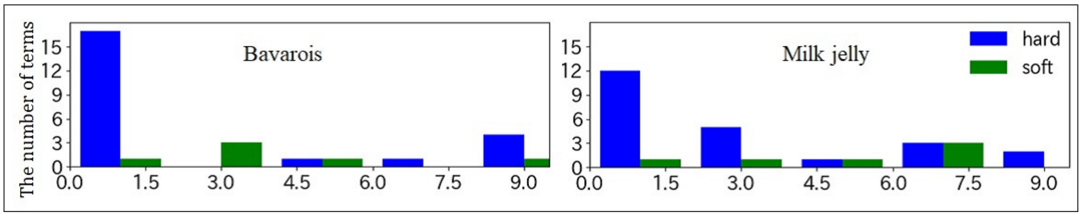
8. On the contrary, gelatins more than 0.025 in **Table** I are characterized by "katai", "muchimuchi"(hardness) in the topic 3. Four variations of kanten in **Table** I are accompanied by the same topic, because only two topics are acquired concerning kanten, and the one is more similar to all the variations than the other. The result inclines to texture terms of hardness such as "dossiri", "burinburin", "korit", coincident with the hardness of kanten in **Table** I. Thus, the texture of gel related recipes (i.e., the ones of empirical research in **Table** I) could be interpreted to corresponding texture terms by estimating the most similar topic.

B. Texture terms and quantitative attributes of Bavarois and milk jelly

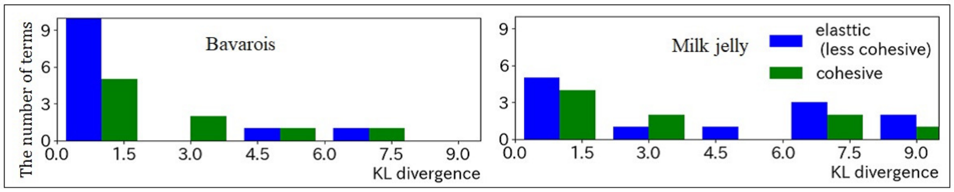
As a trial, we try to interpret quantitative texture attributes of gels mixed with emulsions to texture terms. **Table** II(b) shows the attributes of the dishes, Bavarois and Milk Jally both of which are research results of food science [20], [21]. The concentration of gels of both the research are very similar, approx. 2.5 % of gelatin, which is the same as data id 3 in **Table** I. Despite of the same concentrations, the quantitative attributes are different with each other, which might be the subordinate effects of different concentrations of emulsions.

The most similar topic for both the dishes is 3, derived by Section II, *D* in our method. While overall texture terms of topic 3 are represented in **Table** II(a), individual recipe belonging to the topic has its own texture terms. Among these recipes, we acquired similar ones to each dish by computing KL divergence of emulsion concentrations between the recipes and the dish. **Figure** 3 shows the results. The histograms show the number of recipes within topic 3 ordered by KL divergence. For example, bins of hardness represent the number of texture terms belonging to the category of hardness in Comprehensive Japanese Texture Terms as described in Section II, *A* in our method. Similarly the bins of softness represent the ones concerning softness.

Apparently the smaller the KL is, the more frequent the bins of hardness become (**Figure** 3(a)), which implies both the dishes are likely to be harder recipes among the recipes in topic 3. The result is coincident with hardness of quantitative texture in **Table** II(b), where both the dishes are much harder than pure gelatin (the third row). Meanwhile, the smaller the KL is, the more frequent the bins of elastic in case of Bavarois, but not in the case of milk jelly (**Figure** 3(b)). This contrast coincides with the gaps of quantitative cohesiveness in **Table** II(b), because strong elasticity leads to large value of cohesiveness.



(a) The number of recipes classified into hard or soft by order of KL divergence of emulsion concentrations



(b) The number of recipes classified into hard or soft by elastic or cohesive by order of KL divergence of emulsion concentrations

Fig. 3. Distribution of Cookpad recipes by order of similarities of emulsion concentration to Bavarois and Milk jelly in **Table** II(b)

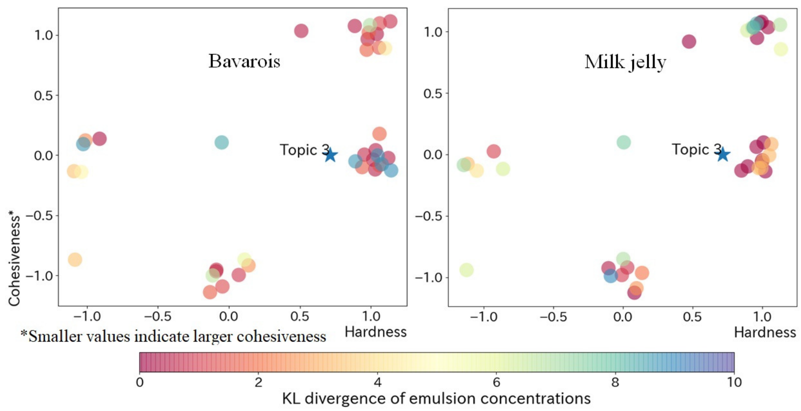


Fig. 4. Distribution of Cookpad recipes on hardness/cohesiveness axis classified by similarities of emulsion concentrations to Bavarois and Milk jelly

Here, all the recipes in the histogram can be represented by two axis, Hardness, and Cohesiveness, because softness is negative hardness, and also elasticity is negative cohesiveness. **Figure** 4 is the scatter plot of the recipes by the consolidated two axis. Colors of the plots represent KL divergence of emulsion concentration, the same as **Figure** 3. The star mark represents the result of the similar classification of texture terms for topic 3.

Red colored plots concentrate in the right area in both of Bavarois, and Milk jelly, indicating similar recipes to these dishes tend to be harder than topic 3. In addition, Red colored plots of Bavarois concentrate in the upper right area while Milk jelly concentrate in the middle right. These plots coincide with the quantitative texture in **Table** II(b) in that both the dishes are much more harder than pure gelatin dish (The third row in the table), and Bavarois indicates larger cohesiveness than Milk jelly.

Thus, the results imply the coincidence of texture terms to the quantitative characteristics in case of the gel mixed recipes, though the samples are limited.

VI. Conclusion

This study tried to acquire characteristics of texture concerning gel related recipes, as a form of texture terms combined by rheology, a quantitative texture. According to food science research, principal determinant of gel related texture is concentrations of gel ingredients with subordinate effects of emulsions. We designed a joint topic model to reflect these findings, and compared acquired topics with the findings of food science research. The results indicated coincidence with the findings as follows.

- Topics represent texture terms in accordance with types of gels and their concentrations.
- Quantitative texture of simple gel dishes was properly interpreted to corresponding texture terms based on the topics.
- Quantitative texture of emulsion-gel mixture dishes was found to be similar to texture terms reflecting the subordinate effects.

These results support the valid interpretation of gel related recipes to texture. We will apply our method to a wider variety of dishes to acquire more general validity in future. Then, we will detect rules bridging between recipe information including ingredient concentrations, cooking steps etc., and sensory textures of consumers. These rules will contribute to development of new food in food industry.

REFERENCES

[1] A.S.Szczesniak, and D.H.Kleyn, "Consumer Awareness of Texture and Other Food Attributes," Food Technology, Vol.17, pp.74-77.1963. .

[2] ISO 11036:2020, Sensory analysis — Methodology — Texture profile.

[3] H. Suzuno, S.Sawayama, and A. Kawabata, "The Texture Profiles of Jelly Fig Gel," Journal of Home Economics of Japan, Vol.43,pp. 525-531,1992.

[4] F.Kawamura, and S.Takayanagi, "Properties of Gelatin Gel and Sol Mixed with Whipped Cream," Science of cookery, Vol.13, pp. 284-288,1980.

[5] K.Okuma, H.Akabane, and N.Nakahama, "Breaking Properties of Agar Gels," Journal of Home Economics of Japan , Vol. 29, Pages 22-27,1978.

[6] A.S.Szczesniak, "Classification of Textural Characteristicsa," Journal of Food Science, Vol.28, pp.385-389, 1963.

[7] K.Nishinari, F.Hayakawa, C.F.Xia, L.Huang, J.F.Meullenet, and J.M.Sie ermann, "Comparative study of texture terms : English, French, Japanese and Chinese," .Journal of Texture Studies, Vol.20, pp.1-27,1989.

[8] B.Drake, "Sensory textural/rheological properties a polyglot list," Journal of Texture Studies, Vol.39, pp.530-568,1989.

[9] H.H. Friedman, J.E.Whitney, and A.S.Szczesniak, "The Texturometer - A New Instrument For Objective Texture Measurement," Journal of Food Science, Vol. 28, pp 390-396,1963.

[10] F.Hayakawa, Y.Kazami,,K.Nishinari, K.Ioku, S.Akuzawa, Y.Yamano,Y. Baba, and K.Kohyama, "Classification of Japanese texture terms," Journal of Texture Studies, Vol.44, pp.140-159,2013.

[11] D.Mimno, H.M.Wallach, J.Naradowsky, D.A.Smith, and A.McCallum, "Polylingual topic models," EMNLP '09, Vol.2, pp.880–889,2009.

[12] W Min, S Jiang, L Liu, Y Rui, and R Jain, "A Survey on Food Computing," ACM Computing Surveys, Vol.52, pp 1–36,2020.

[13] J.C.Meullenet, B.G.Lyon, J.A. Carpenter, and C.E.Lyon, "Relation between sensory and instrumental texture profile attributes," Journal of Sensory Studies, Vol.13, pp.77-93.

[14] A.M.Paula, and A.C.Conti-Silva, "Texture profile and correlation between sensory and instrumental analyses on extruded snacks," Journal of Food Engineering, Vol. 121, pp.9-14,2014.

[15] K.Kurimoto, H.Moritaka, K.Fujii, H.Ogoshi, N.Nakahama,"Thermal and Physical Properties of Mixed -Carrageenan-Gelatin Gels," Journal of Home Economics of Japan, Vol. 48, pp.885-892,1997

[16] F. KAWAMURA, S. NAKAJIMA, and R. KOUNO, "Properties of Gelatin Sol and Gel Mixed with Emulsion," Journal of Home Economics of Japan, Vol. 29, pp.356-361,1978.

[17] A.Murayama, "Carrageenan Jerry," Science of Cookery, Vol. 25, pp.353-359,1992.

[18] T.Farjami, and A.Madadlou, "An overview on preparation of emulsion-filled gels and emulsion particulate gels," Trends in Food Science & Technology, Vol.86, pp.85-94, 2019.

[19] T.Farjamia, and A.Madadloub, "An overview on preparation of emulsion-filled gels and emulsion particulate gels," Trends in Food Science & Technology, Vol. 86, pp. 85-94,2019.

[20] A.Kawabata, S.Sawayama, "Texture of Low-Methoxyl Pectin Bavaroisl," Japan Society of Nutrition and Food Science, Vol.27, pp.55-63,1974.

[21] M.Motegi, "Texture of jelly foods," Science of cookery, Vol.1, pp.28-35,1975.