A Delicious Recipe Analysis Framework for Exploring Multi-Modal Recipes with Various Attributes

Weiqing Min[1], Shuqiang Jiang[1,2], Shuhui Wang[1], Jitao Sang[3], Shuhuan Mei[1]

[1] Key Lab of Intelligent Information Processing, Institute of Computing Technology, CAS, China

[2]University of Chinese Academy of Sciences, China

[3]National Lab of Pattern Recognition, Institute of Automation, CAS, China

Contact: minweiqing@ict.ac.cn, sqjiang@ict.ac.cn

Motivation

- Diverse food attributes: the cuisine and course attributes are discrete or categorical while the value of each flavor attribute is continuous.
- Rich modality information: recipe images, the ingredient line, and so on.
- We propose to utilize various types of attributes and multi-modal information for multi-dimensional food analysis and understanding.

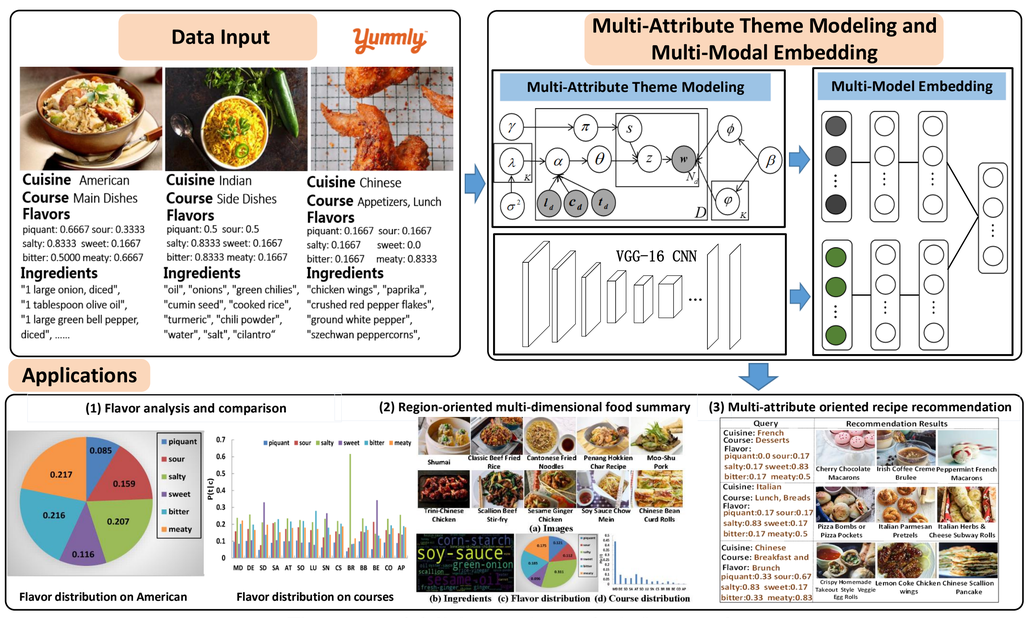
Some recipe examples from Yummly.

Main contributions

- Propose a delicious recipe analysis framework, which utilizes various types of attributes and multi-modal information to enable multi-dimensional food analysis and applications.
- Present a wide variety of applications, including 1) flavor analysis and comparison, 2) region-oriented multi-dimensional food summary, and 3) multi-attribute oriented recipe recommendation.

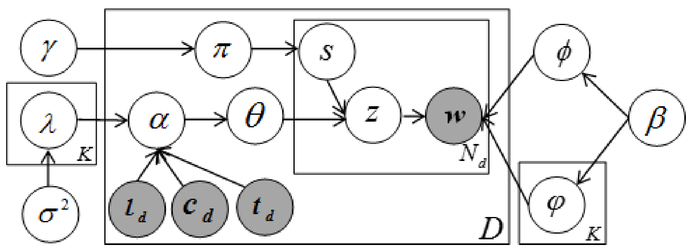
Framework

- **Input**: multi-modal recipes from Yummly with three types of attribute features, including flavor, cuisine and course attributes.
- **Method**: first develop a Multiple Attribute Theme Modeling (**MATM**) method to model the correlation between the ingredients and these food attributes, and then utilize the Multi-Modal Embedding (**MME**) method to create joint representation based on learned ingredient theme features and deep visual features.
- **Applications**: (1) flavor analysis and comparison, (2) region-oriented multi-dimensional food summary and (3) multi-attribute oriented recipe recommendation.



The proposed delicious recipe analysis framework.

1. Multiple Attribute Theme Modeling (MATM)



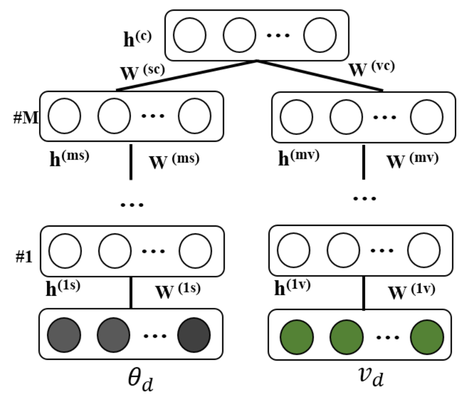
The multi-attribute theme model.

- θ_d of each recipe is no longer drawn from a Dirichlet prior with fixed hyper-parameters α.
- The elements of $\boldsymbol{\alpha}_d$, $\alpha_{dk} = \exp(\mathbf{f}_d^T \boldsymbol{\lambda}_k)$, where $\boldsymbol{\lambda}$ is the feature matrix with $K \times (L + C + T + 1)$. $\mathbf{f}_d$ is the concatenation of different attribute vectors. $\mathbf{f}_d = [\mathbf{l}_d; \mathbf{c}_d; \mathbf{t}_d; 1]$.

After the model inference, we can obtain the theme-attribute feature matrix $\{\hat{\boldsymbol{\lambda}}_k\}_{k=1}^K$ and the following parameters:

$$\hat{\pi}_{d,s} = \frac{n_{d,s} + \gamma}{\sum_{s'=0}^{1} n_{d,s'} + 2\gamma} \quad \hat{\varphi}_{k,w} = \frac{n_{k,w} + \beta}{\sum_{w'=1}^{W} n_{k,w'} + W\beta}$$
$$\hat{\phi}_w^{bg} = \frac{n_w^{bg} + \beta}{\sum_{w'=1}^{W} n_{w'}^{bg} + W\beta} \quad \hat{\theta}_{d,k} = \frac{n_{d,k} + \exp(\mathbf{f}_d^T \hat{\boldsymbol{\lambda}}_k)}{\sum_{k=1}^{K}(n_{d,k} + \exp(\mathbf{f}_d^T \hat{\boldsymbol{\lambda}}_k))} \tag{1}$$

2. Multi-Modal Embedding (MME)



The multi-modal embedding.

We utilize multi-modal DBM to correlate recipe images and ingredients.

3. Applications

- Flavor Analysis and Comparison $p(t|l) = \sum_k \frac{v_{t,k} p(t)}{\sum_{t' \in T} v_{t',k} p(t')} \psi_{l,k}$
- Region-Oriented Multi-dimensional Food Summary.

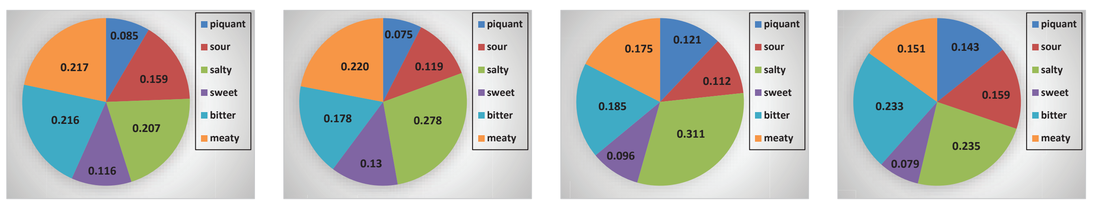
$$Rep_{l_q}(w) = n(l_q, w) \frac{\sum_k \hat{\varphi}_{kw} \psi_{l_q,k}}{\sum_k \hat{\varphi}_{kw} \psi_{l_q,k} + \hat{\phi}_w^{bg}}, p(w|l_q) = \frac{Rep_{l_q}(w)}{\sum_{w'} Rep_{l_q}(w')}$$
$$sim(l_q, I_d) = \tau \frac{\boldsymbol{\psi}_{l_q}^T \hat{\boldsymbol{\theta}}_d}{\|\boldsymbol{\psi}_{l_q}\| \|\hat{\boldsymbol{\theta}}_d\|} + (1 - \tau) \frac{\mathbf{v}_{l_q}^T \mathbf{v}_d}{\|\mathbf{v}_{l_q}\| \|\mathbf{v}_d\|} \tag{2}$$

- Multi-Attribute Oriented Recipe Recommendation $JsSim(\boldsymbol{\vartheta}_\mathbf{f}, \boldsymbol{\theta}_d) = \exp\{-D_{js}(\boldsymbol{\vartheta}_\mathbf{f} || \boldsymbol{\theta}_d)\}$.

Experimental Evaluation

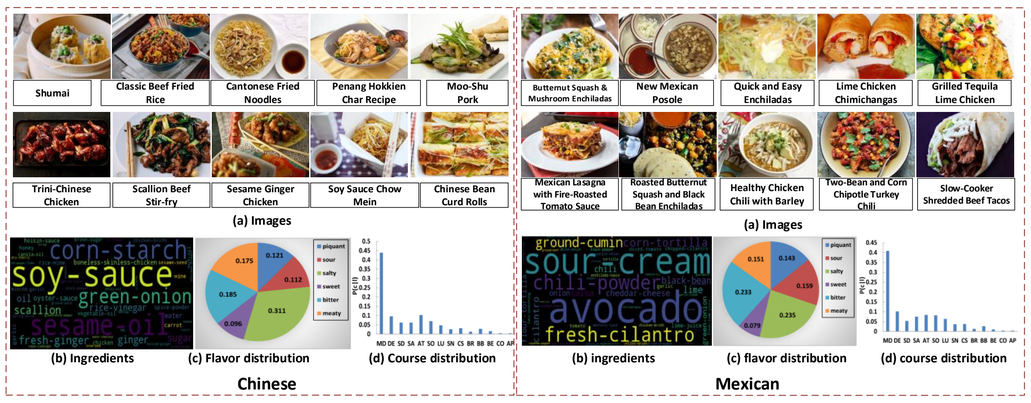
Dataset: 44,204 recipe items, 10 cuisines, 14 courses1 and 6 flavors.

Experimental Results

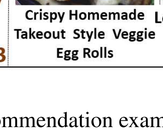
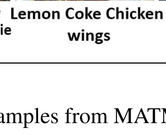
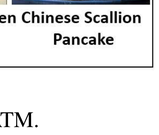


(a) American (b) French (c) Chinese (d) Mexican

Flavor distributions in different regions



Multi-dimensional summary for Chinese and Mexican food.

Query	Recommendation Results		
Cuisine: French Course: Desserts Flavor: piquant:0.0 sour:0.17 salty:0.17 sweet:0.83 bitter:0.17 meaty:0.5	 Cherry Chocolate Macarons	 Irish Coffee Creme Brulee	 Peppermint French Macarons
Cuisine: Italian Course: Lunch, Breads Flavor: piquant:0.17 sour:0.17 salty:0.83 sweet:0.17 bitter:0.17 meaty:0.5	 Pizza Bombs or Pizza Pockets	 Italian Parmesan Pretzels	 Italian Herbs & Cheese Subway Rolls
Cuisine: Chinese Course: Breakfast and Flavor: Brunch piquant:0.33 sour:0.67 salty:0.83 sweet:0.17 bitter:0.33 meaty:0.83	 Crispy Homemade Takeout Style Veggie Egg Rolls	 Lemon Coke Chicken wings	 Chinese Scallion Pancake

Some recommendation examples from MATM.

Conclusions

- Present a recipe analysis framework to incorporate multi-modal information, various types of attribute features for multi-dimensional food analysis.
- The derived attribute-theme representation and multi-modal correlation has demonstrated its effectiveness via three applications in flavor analysis, food summary and recipe recommendation.