

P18：局所的類似性に着目した特徴的調理手順の抽出

二宮あかり，尾崎知伸（日本大学）

背景

食材の代わりを知りたい

似ているレシピを探したい

簡単なレシピを知りたい

献立が思いつかない

レシピデータを効果的に利用したい

各レシピが持つ特徴を抽出し、
レシピデータを分類・構造化することが重要

レシピには多様な側面があり、
総合的に考慮して類似性を測ることが大切である

多様な側面からの類似性を適切に
統合した距離定義が必要

表現学習による分散表現の獲得

各対象（単語，文，図など）に，その対象の特徴を表すn次元の
実ベクトルを割り当てる技術

word2vec

doc2vec

単語や文章の
分散表現を獲得

検索
クラスタリング
例外発見
など

潜在的な距離を獲得し利用することができる

これまでの取り組み

●フローグラフを用いたレシピ分散表現

A. Ninomiya and T. Ozaki: Learning Distributed Representation of Recipe Flow Graphs via Frequent Subgraphs, The 11th Workshop on Multimedia for Cooking and Eating Activities (CEA'19), pp.25-28, 2019.

二宮あかり，尾崎知伸：部分グラフに基づくレシピフローグラフ分散表現の比較評価，人工知能学会第117回知識ベースシステム研究会，SIG-KBS-B901-02, pp.7-13, 2019.

●調理テキストと調理画像系列を用いたベクトル系列表現

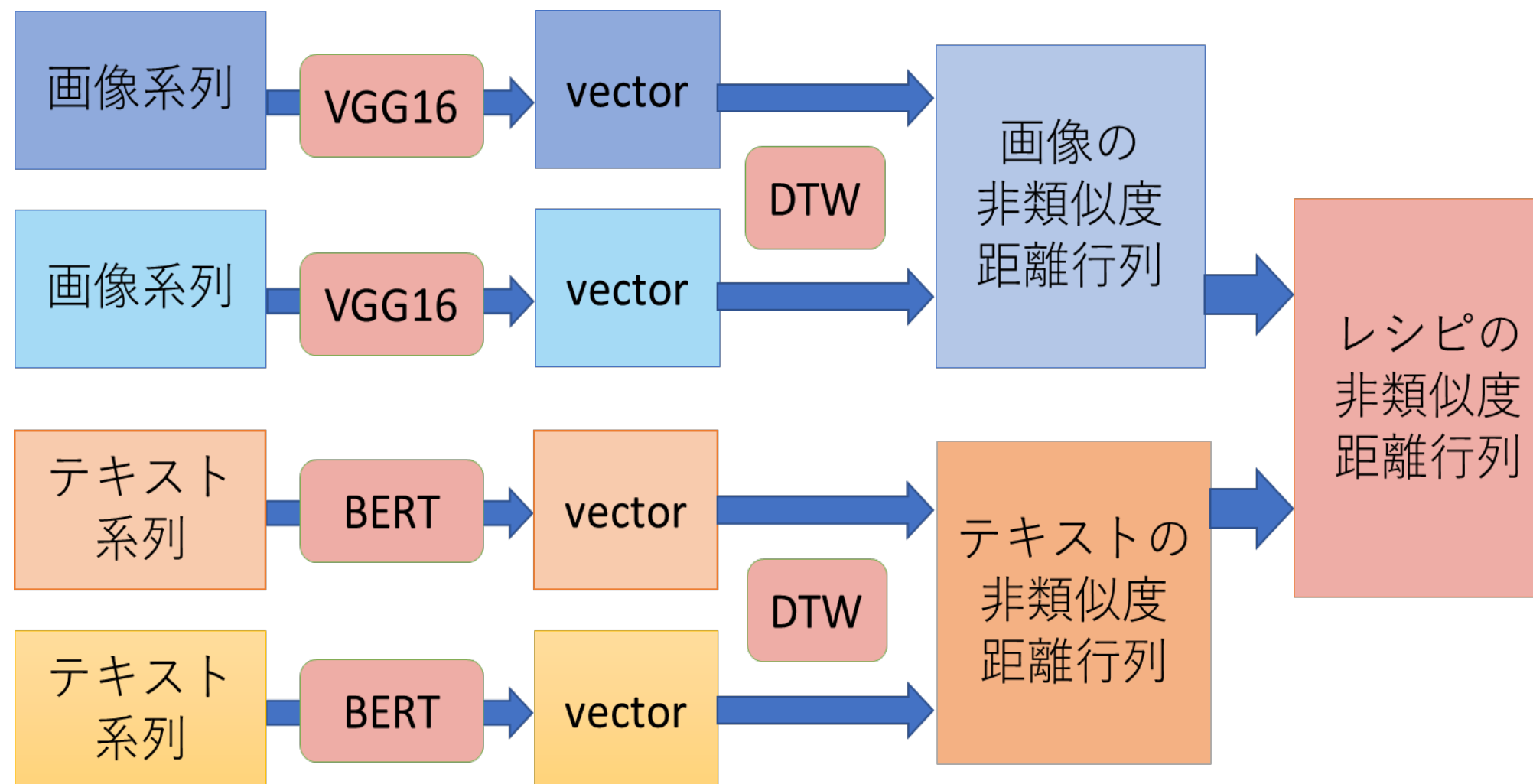
二宮あかり，尾崎知伸：レシピテキストと調理画像系列の埋め込みを用いたケーキレシピの分析，人工知能学会第119回知識ベースシステム研究会，SIG-KBS-B903-01, pp.1-5, 2020.

A. Ninomiya and T. Ozaki: Cooking recipe analysis based on sequences of distributed representation on procedure texts and associated images, (submitted to the 11th Workshop on Multimedia for Cooking and Eating Activities)

調理テキストと調理画像系列を用いたベクトル系列表現

調理手順

1. ○○を切る
2. ▲▲を切る
3. ▲▲と○○を混ぜる
4. 油を敷いて強火で炒める



実験・結果

●レシピのタイトルが「オムライス」，「コロッケ」，「ハンバーグ」，
「チャーハン」で終わるレシピ：1000件ずつ(合わせて4000件)

▶提案手法：テキスト系列+画像系列(w=0.5)

▶ベースライン1：テキスト系列のみ

▶ベースライン2：画像系列のみ

▶ベースライン3：完成品画像のみ

内的妥当性・外的妥当性

| 提案手法 | CPCC | Rand Index(#of clusters) | | | | |
|------|-------|--------------------------|-------|-------|-------|-------|
| | | 4 | 5 | 6 | 7 | 8 |
| 提案手法 | 0.177 | 0.605 | 0.612 | 0.628 | 0.663 | 0.667 |
| ベース① | 0.263 | 0.589 | 0.632 | 0.635 | 0.650 | 0.667 |
| ベース② | 0.218 | 0.594 | 0.647 | 0.650 | 0.659 | 0.665 |
| ベース③ | 0.111 | 0.665 | 0.663 | 0.672 | 0.676 | 0.693 |

単純な分類においては外観に着目することが有効である
テキスト系列や画像系列は料理に関わらず過程で類似する点が影響する

今後の取り組み

レシピを構造化した上で，レシピ全体の距離を測定
⇒中身がどこからどこまで似ているのかを見ることはできない

中身の繋がりをみるために，局所的な類似性を測りたい
⇒分岐点の着目を考える

共通部分は，
使いまわしできる典型的な手順
と考えられる

共通した手順

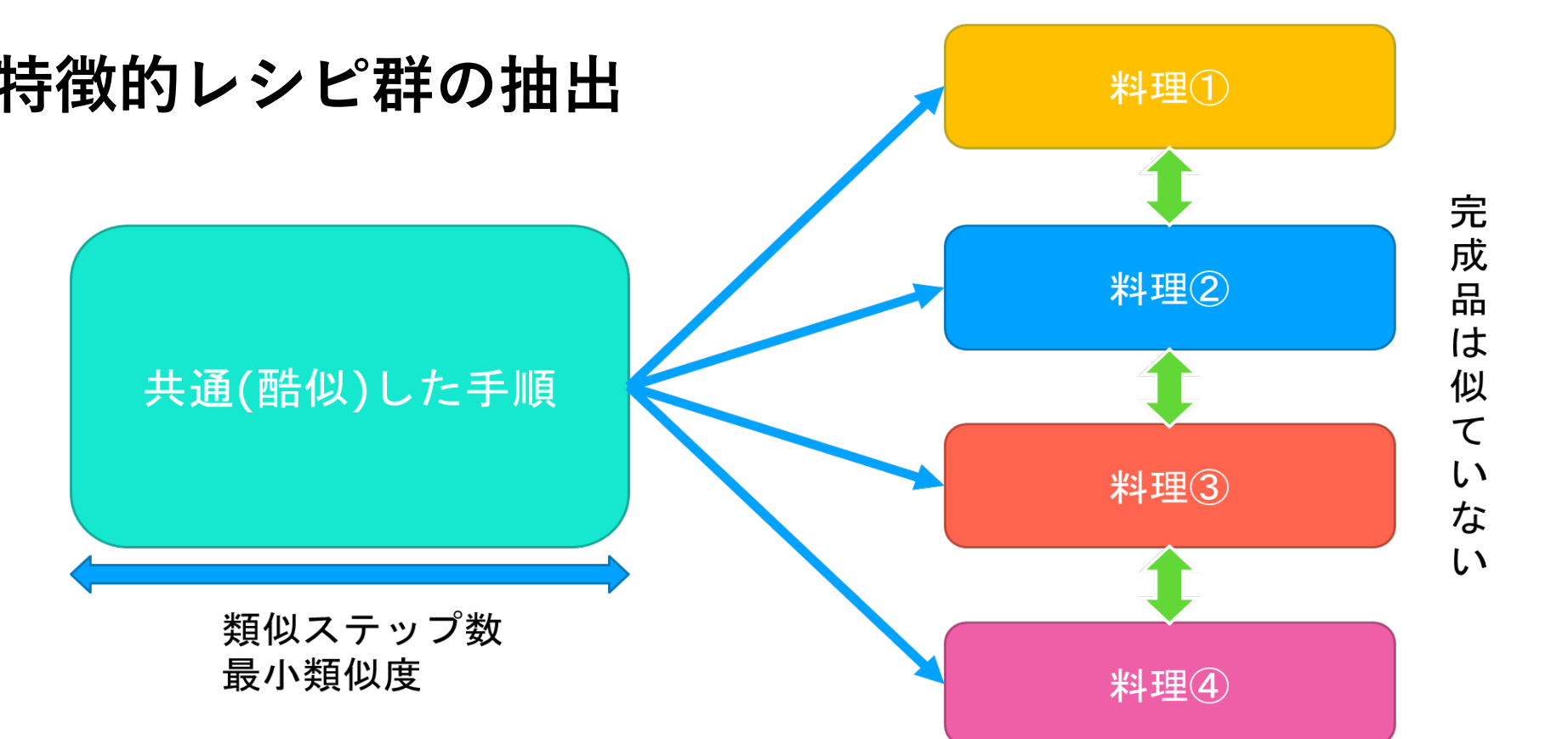
料理の見た目、
完成品を変えたい時に
どのステップで変えるか？

レシピ全体の計量ではなく、**部分的**な計量を行う

◆手順のどこまで似ているのか，似ていないのかを全体的に整理
⇒ **系統樹の作成**

◆レシピ間に見られる特徴的なレシピ集合の発見
⇒ **特徴的調理手順の発見**

特徴的レシピ群の抽出



テキスト系列・グラフ系列・画像系列の相互比較

