

雰囲気メタファによる街の偏在情報の集約・提示システム

宮森 恒[†] 水口 充[†] 河合 由起子[†] 是津 耕司^{††} 木俵 豊^{††}

[†] 京都産業大学 コンピュータ理工学部 ^{††} 情報通信研究機構 知識処理グループ

1 はじめに

将来のコピキタス情報社会において、人々は大量の情報を利用するだけでなく、実世界においてさまざまな情報を発信することが予想されている。これまでも、レストランやホテルなどについて人々が感じたことなどが評判情報として発信され、口コミ検索や評判検索の形で利用されてきた [1]。しかし、これまでの評判情報は、特定の店や住所などに紐付けられた特定の対象物に対する評判であり、実世界における任意の領域に対応した情報ではない。本稿では、携帯端末等により街のさまざまな場所で人々が感じたことをつぶやきのような短い発話として位置・時刻情報とともに発信される状況を想定し、街の任意の時空間的な範囲における特徴的なイベントや状況を雰囲気メタファとして集約・提示するシステムを提案する。提案システムによって、道の凍結など日常生活における重要情報を時空間的な領域形状に依存せずに効率よく取得できることが期待される。

2 提案システムの概要

まず、雰囲気メタファ A を以下の通り定義する。

$$A = \{R, F\}$$

ここで、 R は時空間範囲、 F はその範囲内の特徴量の集約を表す。雰囲気メタファは、任意の時空間範囲に紐付けられた顕著な/共通した特徴量を表現するものであり、これを地図上に射影することで、各範囲に対応する街角や時代がもつ雰囲気 (賑わいや佇まいを含む) を直感的に把握できるようにすることを意図している。

特に本稿では、携帯端末等でさまざまな場所から発信される位置、時刻、発話を利用することを想定し、雰囲気メタファ A を以下通り構成する。

$$A = \{R, F\} \text{ where } R = \{t, c, r\}, F = \{l\}$$

ここで、 t, c, r はそれぞれ時区間、時区間に含まれるデータ群の空間的な中心位置、時区間に含まれるデータ群で c から最も離れたデータまでの距離を表し、 l は

発話を集約したラベル (文字列) を表す。

以下、システムの概要について説明する。

まず、入力データとして、GPS による位置・時刻付きの音声および画像データを記録可能な端末を利用し、20 代の被験者 14 名が京都市内を 8 月と 12 月の一定期間散策してサンプルデータを収集した。収集されたデータは、音声認識と軽度の手作業修正により、図 1 に示すようなデータ列 (以後、1 データ列を 1 記事と呼ぶ) として蓄積される。

Data ID	Data Type	Time	Latitude	Longitude	User ID	Message
1828	VOICE	2008/08/17 14:25:16	34.9818998	135.7669117	0001	
1829	PHOTO	2008/08/17 15:01:07	34.9838137	135.7669117	0001	
1830	VOICE	2008/08/17 15:01:07	34.9818911	135.7669117	0001	京都駅南口、でっかい電車が来りました。
1831	VOICE	2008/08/17 15:01:07	34.9818911	135.7669117	0001	西宮線南口、工事中です。はっは。

図 1: 入力データ例 (左から、データ ID, データタイプ (音声/写真), 時刻, 緯度・経度, ユーザ ID, 発話を表す)

次に、雰囲気メタファを以下の手順で自動生成する。

1. 位置情報に基づくクラスタリング

ある時区間 t (週, 月単位) に含まれるデータの位置情報に基づき、Ward 法による階層クラスタリングを行い、各クラスタについて中心位置 c 、距離 r を得る。生成されたクラスタ階層から、全体のクラスタ数が $2 \sim N$ (N は予め与えられたクラスタ数) となるように統合する。

2. ラベル生成

各クラスタに含まれる発話テキストを形態素解析し、名詞、形容詞、感動詞に該当する語を抽出する (「えー」等のフィラーとみなせる感動詞は除去する)。各品詞で出現頻度をとり、出現頻度の最も高い形容詞あるいは感動詞、および、名詞から、「形容詞/ 感動詞+名詞」の形でラベル l を生成する。

3. 雰囲気メタファの提示

異なる時空間詳細度で雰囲気メタファ A を生成し、各 A を中心 c 、半径 r の円、および、ラベル l として地図上に表示する。

図 2, 3 にシステム画面例を示す。画面左上のカレンダーでは、該当時区間を 1 週間、1 ヶ月単位で選択でき、画面右上では、表示クラスタ数や記事表示モードを選択できる。記事表示モードとは、各記事の位置や発話・写真を表示するかどうかを選択するものであり、図 2 では各記事の発信された位置が、図 3 においては、位置だけでなく発話や写真が表示されている。

Study on Aggregation and Browsing of Ubiquitous City Information using Atmospheric Metaphor

[†] Hisashi MIYAMORI (miya@cse.kyoto-su.ac.jp)

[†] Mitsuru MINAKUCHI (mmina@cse.kyoto-su.ac.jp)

[†] Yukiko KAWAI (yukiko@cc.kyoto-su.ac.jp)

^{††} Kouji ZETTSU (zetsu@nict.go.jp)

^{††} Yutaka KIDAWARA (kidarawa@nict.go.jp)

Faculty of Computer Science and Engineering, Kyoto Sangyo University ([†])

Knowledge Clustered Group, National Institute of Communications and Technology (^{††})

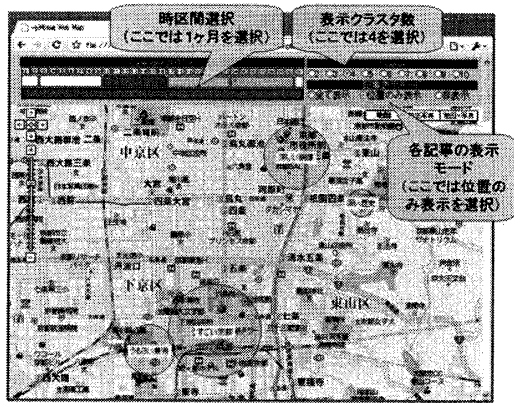


図 2: 雰囲気メタファによる集約・提示の画面例 (1)

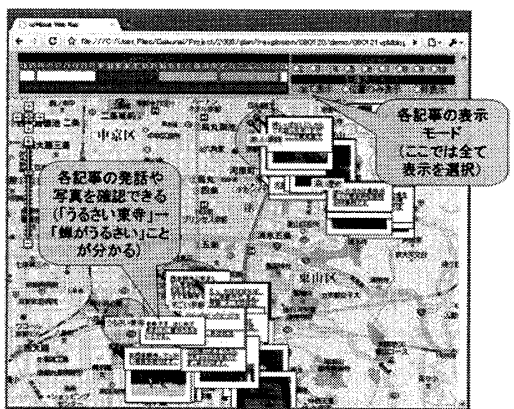


図 3: 雰囲気メタファによる集約・提示の画面例 (2)

例えば、図 2 の地図左下の雰囲気メタファのラベルは「うるさい東寺」となっているが、図 3 で各記事の発話を参照すると、「東寺です」「蝉がうるさいとです」等の発話を確認でき、その場所で起きている事象の特徴を把握できる。他にも「寒い凍結」などのラベルから、道の凍結状況を把握することができ、日常生活上重要な情報を効率よく取得できることを確認した。

3 考察

提案システムについて考察と課題を述べる。

今回生成した雰囲気メタファは、同一時区間における発話の位置情報のクラスタリングに基づく簡便なものであった。しかし、クラスタリングは位置情報だけでなく、その領域で起きているイベント名や状況を表現した語、各人が発するセンチメント等、発話を構成するさまざまな要素を利用することが考えられる。今後、位置情報だけでなく、発話内容に基づくクラスタリングを組み合わせた雰囲気メタファ生成手法について改良を進めていく必要がある。

また、今回の実装では、発話時刻は、ある 1 点の時刻を想定してきたが、展示会やイベント紹介などある

一定期間にわたって継続的に行われる内容について掲載したサイト等を参照する場合、時刻ではなく時間範囲として扱われるべき発話も存在する。今後、時間の開始点・終了点も考慮した実装を検討する必要がある。

さらに、特定の状況の増加・減少傾向を示す時間変化タイプを検出・分類したり、利用者クエリにより対象データを適宜絞り込むことで、雰囲気メタファの機能のバリエーションを増強していく必要がある。

一方、雰囲気メタファの提示としては、各クラスタに含まれる全データを内包する凸包で表現する方法 [2] も考えられるが、領域の大きさによっては地図上で認識しづらくなる場合があり、今回は認識がより容易な円によって表現することとした。また、ユビキタスコンテンツを実世界内で利活用するためには、その場、その時に、雰囲気メタファを動的に提示・閲覧する手法が必要となる。特に、携帯端末などの表示領域が限定された画面においては、表示可能な雰囲気メタファの数も限られており、表示範囲外にある雰囲気メタファについても利用者に効率よく気づかせる手法が必要になると考えられる。表示すべき雰囲気メタファの選別方法や、表示範囲外にある雰囲気メタファの可視化手法、ラベル、領域への色や動きといった効果の付与手法等についてさらに検討する必要がある。

なお、携帯端末から発信されたつぶやきに相当する短いテキストを位置情報とともに発信し、順次地図上に表示するシステムとして携帯 twitter [3] があるが、あくまで各発話を地図と対応付けて表示しているのみであり、情報を集約しているわけではない点で本システムとは異なるものである。

4 まとめ

街のさまざまな場所や時刻における特徴的なイベントや状況を、雰囲気メタファとして集約・提示するシステムを提案した。提案システムは、その時空間における顕著な、あるいは、共通的な特徴を文字と領域の大きさとで表現し、凍結している道の的確な把握など、特定の対象物によらない任意形状の領域の状況把握を可能とする点で有効であることを確認した。

謝辞

本研究の一部は科研費特定領域 19024073、および、科研費基盤 (C)20500120 の助成を受けたものである。

参考文献

- [1] 立石 健二, 石黒 義英, 福島 俊一: インターネットからの評判情報検索, 情報処理学会研究報告 2001-NL-144-11, Vol.2001, No.69, pp. 75-82, 2001.
- [2] O.Masutani, H.Iwasaki: BEIRA: A Geosemantic Clustering Method for Area Summary, LNCS4831, WISE2007, pp.111-122, 2007.
- [3] 携帯 twitter. <http://onosendai.jp/twitter/>