メディア分析のための時系列画像群の 3 次元可視化

伊藤 正彦　　豊田 正史　　喜連川 優*

概要.　本稿ではソーシャルメディアおよび放送メディアを含む複数メディアから抽出された時系列画像群を可視化することで，メディア分析を可能にするシステムを提案する．画像は説明無しに多くの事実や印象を伝えることが可能であり，画像群を俯瞰することで，話題の変遷を観測できる．提案するシステムにより，メディア上で流れる画像群を時間を追って観測し，人々の考え，経験，興味の変遷，メディア間の差異等を知ることが可能になる．

1　はじめに

放送映像とウェブ情報は相互に影響し合っており，放送メディアの影響によりブログの記事になる場合もあれば，ウェブ上で話題になることにで放送メディアが取り上げる場合もある．そのため，その話題がどのメディアで広まったのか，メディア間の違いは何か等を分析することは，社会学，マーケティングなど多様な社会分析の観点から重要である．

画像を用いることにより，文章のみでは伝えることが困難な事実や印象を伝えることが可能である．それゆえ，各メディア上で流れる画像群を時間を追って観測することで，人々の考え，経験，および興味の変遷の概要を知ることが可能になると考える．

本稿では，ソーシャルメディアおよび放送映像を含む複数メディアから抽出された時系列画像群を可視化することで，メディア分析を可能にするシステムを提案する．提案するシステムでは，特定のトピックに関して，様々な観点からなる時系列画像群を Web アーカイブおよび放送映像アーカイブから抽出し，それらを観点毎に 3 次元空間に可視化する．これにより，トピックにおける流行変化を様々な観点から比較・探索できる．また，他のメディアにおける出現頻度変化を重畳表示することでメディア間差異を探索可能にする．さらに，テキスト解析による話題探索システムと連動することでトピックの詳細探索を可能にする．

2　時系列画像の 3 次元可視化

図 1 に提案システムの全体像を示す．まず，提案システムでは，トピックに関する時系列画像群の流行変化を観測するために，観点毎に時間軸上に画像群をヒストグラムのように堆積し可視化する．画像を集約する時間ウィンドウは1ヶ月，1 週間，1 日など任意に設定可能である．これにより，ある観点から見た際に，トピックがいつ頃話題になり始めたのか，いつ頃バーストしたのか，流行の変化，流行期間等が俯瞰できる．

次に，複数観点の比較を容易にするために，複数の画像ヒストグラムを 3 次元空間に並べてて配置する．このさい，3 次元空間の奥行き方向を観点軸として用いる．これにより，観点間における流行のタイミングの違い，流行の発生順，もしくは同じタイミングにおける流行の内容の差異等を観測できる．

メディア間の差異を比較するために，他のメディアにおける出現頻度変化をラインチャートで重畳表示する．図 1 の例では，放送映像における特定人物に関する言及数を画像ヒストグラムで表示し，ブログ中での言及数をラインチャートで表示している．複数のメディアを統合的に分析することで，1 メディアを分析しただけでは分からない，重要なイベントを相補的に抽出することが可能になる．

さらに，画像の詳細もしくは関連情報を探索するために，テキスト解析による話題探索システム [1] と連携し，選択した画像に関する観点および選択時間における詳細探索を実現している．

3　応用例

3.1　放送とウェブにおける反応比較

図 1 では，手前から菅直人，清水正孝（東電社長），佐藤雄平（福島県知事），谷垣禎一の各人物に関する東日本大震災後の放送映像における言及数およびブログ言及数を一日毎に集約し重畳可視化している．これにより大震災の後，菅総理を始めとする各人物のとった行動に対する，放送およびブログ上での反応の分析が可能になる．放送映像は国立情報学研究所において収集された放送映像アーカイブを用いている．また，ブログデータとしては，著者らの研究室で構築している，日本語ブログアーカイブを用いている．

図中（A）の時期では，菅首相の被災地訪問という，重要なイベントがあったにも関わらずブログではあまり反応が出ていない．また，図中（B）の時期における浜岡原発停止要請のような賛否両論の議

Copyright is held by the author(s).

* Masahiko ITOH, Masashi TOYODA, and Masaru KITSUREGAWA 東京大学 生産技術研究所

図 1. 放送映像・ブログ情報を統合した相補的なイベント可視化システム

論を呼ぶ話題では，放送に比較してブログでより大きな反応が出ていることが伺える．(C) の時期には清水社長が姿を見せないことに関してブログでのみ非常に大きな反応が出ており，メディアや話題の種類に依存した反応の違いを見ることができる．

図中 (A) (D) (E) の時期のように，菅首相の佐藤知事訪問，および清水社長の佐藤知事訪問といった複数人物が関わる話題では，同時期に複数の人物に大きな反応が出たり，複数人物間で同じ放送映像が現れることがある．このように，時間軸上で複数人物を比較することで，複数の人物が共通に関わる重要なイベントを抽出することが可能になる．

3.2 ブログにおける商品トレンドの可視化

図 2 の例では，ブログにおける話題推移の一例として商品トレンドの可視化例を示す．ブログ上の話題の推移を把握可能にするため，画像の周辺テキスト，局所特徴量，および時間情報を用いてクラスタリングを行い [2]，各クラスタを観点軸に配置し可視化した．例では,「キットカット」に関するフレーバの流行具合が確認出来る．特に「きっと勝つ」や「桜咲く」という語呂合わせ的な意味合いからか大学受験シーズンに「サクラ味」「大学芋味」などが流行していることが確認出来る．

図 2. ブログにおける商品トレンドの可視化

謝辞

本研究の一部は，文科省「Web 社会分析基盤ソフトウェアの研究開発委託事業」及び国立情報学研究所と東京大学の間の共同研究に基づく．佐藤真一教授からは貴重な放送映像アーカイブのデータをご提供頂き，ここに記して感謝の意を表します．

参考文献

[1] 伊藤正彦, 吉永直樹, 豊田正史, 喜連川優．係り受け解析を用いたブログユーザの行動・興味に関する時系列推移 3 次元可視化システム．電子情報通信学会論文誌, J96-D(7):1454–1466, 2012.

[2] 上條哲也, 豊田正史．社会分析を目的とした CGM 画像クラスタリング手法に関する一検討．第 3 回データ工学と情報マネジメントに関するフォーラム (DEIM2011), 2011.