

# 動向情報の根拠探索のための テレビ番組からの図表画像検出手法の検討

梅澤 啓史<sup>1,a)</sup> 宮森 恒<sup>1,b)</sup>

**概要:** 本稿では、テレビ番組の映像データから図表画像を検出する手法を検討する。著者らは、ソーシャルネットワーク上の言説、特に、ある対象の時間的変化を記述した動向情報を対象に、その信頼性判断を支援する根拠の一つとして、テレビ番組で用いられた図表画像を提示する手法を提案している。先行研究では、クエリとして入力された動向情報の根拠となりうる図表画像がテレビ番組中に存在すれば、一定の良好な精度で関連図表を検出できるという結果が得られた。しかし、一部の図表画像の検出精度やカバレッジの向上が必要という課題が残った。本稿では、図表画像の統計的な性質やエッジに基づく特徴などを考察し、様々な検出手法を比較検討する。

## 1. はじめに

現在、ネット上では、膨大な情報が発信・蓄積され続けている。利用可能な情報量の増大は、一見すると利便性が向上しているようにも見えるが、実際は、それら情報は内容の検証や整理をされないままネット上に放置されているようなものであり、情報が増えるほど、内容の精査や正しい情報の選別はますます困難になるといわざるを得ない。ソーシャルメディア等の発達により、情報を発信する手段は普及する一方、信頼度の高い情報を選別する手段は十分に整備されているとは言い難い。

Web コンテンツの信頼性判断を支援する方法としては、これまでに、情報の発信者や社会的意見、情報外観(参考文献や連絡先の明記等)といった多角的な観点からの情報をユーザに提示する手法 [2] や、評価対象の Web コンテンツとそれに関連する Web コンテンツをデータ対で表現し、その support 関係の強さで対象コンテンツの信憑性を評価する手法 [3] などが提案されている。

筆者らは、これまで、ソーシャルネットワーク上の言説、特に、ある対象の時間的変化を記述した動向情報を対象に、その信頼性判断を支援する根拠の一つとして、テレビ番組で用いられた図表画像を提示する手法を提案している [1]。先行研究では、クエリとして入力された動向情報の根拠となりうる図表画像がテレビ番組中に存在すれば、一定の良

好な精度で関連図表を検出できるという結果が得られた。しかし、一部の図表画像の検出精度やカバレッジの向上が必要という課題が残った。

そこで、本稿では、図表画像の統計的な性質やエッジに基づく特徴などを考察し、図表画像の検出手法について比較検討する。

## 2. 動向情報の根拠探索

ここでは、先行研究である、テレビ画像中の図表画像を利用した動向情報の根拠探索について概要を説明する [1]。

まず、動向情報とは、一般に、商品の売り上げ推移や内閣支持率の変化等、ある事柄や数値についての時間的変化を表現するデータあるいは言語情報を指す。ここでは特に後者の言語情報で表現されたものを対象とする。

また、図表画像とは、その中に折れ線グラフや円グラフ等の図表が主として出現しているテレビ番組中の画像フレームのことである。テレビ番組の図表は、予め番組スタッフが政府統計等の複雑な統計データ表から、最適なデータを選び出し、視聴者の誰が見ても理解しやすいように噛み砕いた形に視覚化されているため、一般ユーザが一目で内容を確認するという作業には非常に適している。また、テレビ番組は、プロの番組制作者によって作られており、公共性を保つ必要があるため、一定の信頼度が担保された情報源と考えることができる。

全体の処理手順は以下の通りである。

まず、インデックス生成部は、図表画像、キーワードを抽出し、メタファイルを生成する。以下の手順で処理を行う。

(1) テレビ番組を録画して動画、字幕情報(クローズド

<sup>1</sup> 京都産業大学  
Kyoto Sangyo University, Motoyama, Kamigamo, Kita-ku,  
Kyoto, 603-8555, Japan

a) g1044129@cse.kyoto-su.ac.jp

b) miya@cse.kyoto-su.ac.jp

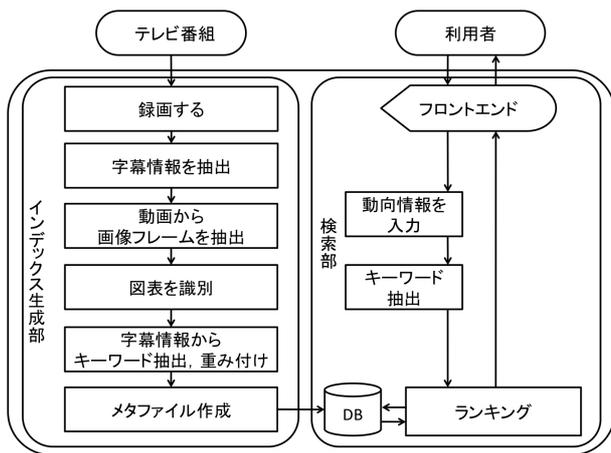


図 1 システム構成図 [1]

表 1 正例, 負例の件数

グラフの種類	正例数	負例数
棒グラフ	155 件	1300 件
円グラフ	371 件	1300 件
折れ線グラフ	194 件	1300 件

キャプション) を抽出する

- (2) 動画から一定間隔おきに画像フレームを抽出する
- (3) 図表画像識別器を用いて, 抽出した画像フレームに図表が含まれるかどうか判定する (図表が含まれると判定されたフレームを図表画像と呼ぶ)
- (4) 字幕情報から, 図表画像の前後 N 秒に出現するキーワードを抽出し, 重み付けする
- (5) 抽出した図表画像, 重み付けしたキーワードをインデックス化し, DB に格納する

検索部では, 利用者が与えたクエリを受け取り, DB 内の図表画像をランキングし, 結果を出力する.

- (1) 利用者が入力した動向情報クエリからキーワードを抽出する
- (2) クエリから抽出したキーワードと, 字幕情報から抽出したキーワードの合致度によって, 図表画像をランキングする
- (3) ランキング結果を出力する

### 3. 図表画像の検出

#### 3.1 正例, 負例データの収集

録画したテレビ番組, および, Web から人手で図表画像を収集し, 正例とする [1]. 図表が使用されている番組のほとんどがニュースであったため, 今回は, ニュース番組を中心に作業を進めた. 図表の種類は, 棒グラフ, 円グラフ, 折れ線グラフを対象とした. 同様に, 負例もテレビ番組から人手で収集した. また, ここでは, 128x128 のサイズに正規化した画像を学習データとした. 以上により, 表 1 のような正例, 負例を収集した.

#### 3.2 図表画像の検出に有効な特徴量と学習

図表画像を検出するにあたり, 一般の写真と図表の分類について検討した文献 [4] を参考にした. 図表を含む画像は, 一般に, 平坦なテキストチャで表されることが多く, 図表以外の画像は, 複雑なテキストチャが含まれ, 多くの色が混在しているという統計的な性質があることが指摘されている. そこで, 写真と図表のヒストグラムを利用した特徴量を考えることにより, テレビ番組を対象とした場合でも, 図表検出に有用である可能性があると考えた.

ヒストグラムに基づく特徴量 (Pct と呼ぶ) としては, 以下の 3 種類を用いることとする.

**Pct2Pk** 画像の輝度ヒストグラムにおける 2 つの最大の山が占めるヒストグラム範囲の割合. 山とは, 全画素数の 0.5% より大きい画素数を含み, 2 つの最小値 (ヒストグラム範囲の 10% を超えない変曲点) の間にある連続したビンとして定義される.

**Pct0.5** 画素数の 0.5% より大きいヒストグラムビンの割合.

**Absdiff** 全ての連続したビンの頻度差の絶対値和. ヒストグラム頻度の変動割合を表す.

また, 先行研究 [1] で用いたエッジに基づく特徴量 (HOG 特徴量) [5] を用いることとする. HOG 特徴量 [5] は, 一つの局所領域におけるエッジ情報に着目した特徴量であり, 輝度勾配の方向についてヒストグラムをとったものである. 人物の検出等に広く利用されているが, 高次元ベクトルとなる傾向があり, 計算時間がかかるという性質をもつ.

これら特徴量を特徴ベクトルとして抽出し, 機械学習により識別器を構築する. ここでは, 樹木モデルの集団学習により, 高精度な分類を行う Random Forest [7] を用いて学習することとする.

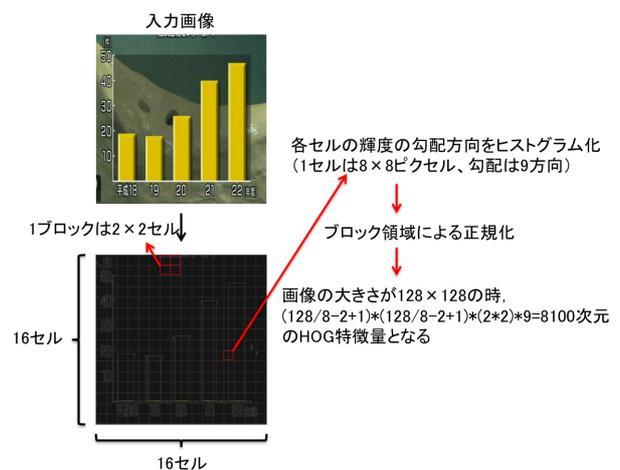


図 2 棒グラフの正例, 負例から HOG 特徴量抽出 \*1

\*1 画像は 2011 年 10 月 31 日 NHK 総合 NHK ニュース 7 から得られたもの

表 2 適合率の比較

手法	棒グラフ	円グラフ	折線グラフ
Pct	0.393 (44/112)	0.593 (220/371)	0.277 (33/119)
Hog	0.913 (116/127)	0.975 (347/356)	0.907 (127/140)
Pct+SMOTE	0.844 (1140/1351)	0.848 (1373/1619)	0.840 (1476/1757)
Hog+SMOTE	0.990 (1237/1250)	0.987 (1472/1492)	0.877 (1546/1565)

表 3 再現率の比較

手法	棒グラフ	円グラフ	折線グラフ
Pct	0.284 (44/155)	0.593 (220/371)	0.170 (33/194)
Hog	0.748 (116/155)	0.935 (347/371)	0.655 (127/194)
Pct+SMOTE	0.919 (1140/1240)	0.925 (1373/1484)	0.951 (1476/1552)
Hog+SMOTE	0.998 (1237/1240)	0.992 (1472/1484)	0.996 (1546/1552)

表 4 F 尺度の比較

手法	棒グラフ	円グラフ	折線グラフ
Pct	0.330	0.593	0.211
Hog	0.823	0.955	0.760
Pct+SMOTE	0.880	0.885	0.892
Hog+SMOTE	0.994	0.989	0.992

## 4. 実験と考察

ここでは、テレビ番組の動画から図表画像をどの程度正確に抽出できるか比較する。

### 4.1 実験データ

録画したテレビ番組、Web から人手で収集した正例、負例をもとに、ヒストグラムに基づく特徴量 (Pct と呼ぶ)、および、エッジに基づく特徴量 (Hog と呼ぶ) を抽出する。また、収集した正例、負例のデータ数には偏りがあり、十分な分類精度が得られない恐れがあるため、SMOTE[6] を用いて人工的に正例データを生成し、両者が均衡したデータ数となるように配慮した。

### 4.2 評価方法

棒グラフ、円グラフ、折れ線グラフそれぞれの場合について、各特徴量で学習した識別器を用いて、10-fold 交差確認を行い、精度を適合率 (Precision)、再現率 (Recall)、F 尺度によって評価する。適合率、再現率は、以下の式で算出する。

$$Precision = \frac{tp}{tp + fp} \quad (1)$$

$$Recall = \frac{tp}{tp + fn} \quad (2)$$

ここで、 $tp$  は正しく抽出できた正解画像数、 $fp$  は間違っ  
て抽出した不正解画像数、 $fn$  は抽出できなかった正解画像数である。

### 4.3 結果・考察

比較結果を表 2~表 4 に示す。

Pct と Hog を比較すると、正例負例が不均衡であるにもかかわらず、どのグラフについても Hog で高い識別結果が得られている。特に、適合率では比較的良好な結果が得られているものの、再現率は低くなる傾向があることがわかった。エラー分析については、まだ十分に作業が進んでいないため、発表当日に報告する予定である。

また、SMOTE により、データ数を均衡させた場合の Pct と Hog の結果を比較すると、この場合もいずれのグラフについても Hog で高い識別結果が得られている。これは、Hog のようなエッジに基づく特徴の方がグラフの特徴をよく捉えており、識別に有効であることを示している。ただし、Pct が 3 次元の特徴量であるのに対し、本稿における Hog は 8100 次元の特徴量であることを考慮すると、Pct のような単純な特徴量を用いても、正例が十分に収集できれば、良好な精度を達成しうる能力をもっていることがわかる。一般に、Hog 特徴量は高次元で計算時間がかかるため、より低次元の特徴量で高い精度を達成することは有用である。

今後、Pct や Hog を改善した特徴量や他の特徴量との組み合わせについて、さらに比較実験を進めていく予定である。

## 5. 関連研究

Web コンテンツの信頼性判断支援の研究は活発に行われている。

その一つに、Web コンテンツの信頼性を判断する基準として、「情報内容、情報発信者、情報外観、社会的評価」といった 4 つの観点から関連情報をユーザに提示することで、Web コンテンツの信頼性判断を支援する研究 [2] がある。情報内容については、Web ページの本文に書かれている内容に着目しており、情報発信者については、発信者の身元に着目した所属の分類やその分野での専門性の有無に着目している。情報外観については、情報ソースやデザイン、連絡先の明記などの Web ページの外観に着目しており、社会的評価については、他の利用者がその情報についてどのような見方をしているかに着目している。

また、データ対間の関係分析に着目した Web コンテンツの信憑性評価の研究が挙げられる [3]。ここでは、評価対象の Web コンテンツとそれに関連する Web コンテンツのデータ対で表現し、その support 関係の強さで対象コンテ

ソツの信憑性を評価するというモデルを導入している。

図表画像の検出については、文書画像を対象に、写真と描画の識別をするために、ヒストグラムに基づく特徴を用いた研究がある [4]。ここでは、ヒストグラムの特徴を表す 3 つの尺度を検討し、写真と描画を区別する尺度の順位は、2 つの最大の山が占めるヒストグラム範囲の割合 (Pct2Pk) が最も高いことを示している。

また、テレビ番組を対象に、HOG 特徴量と AdaBoost を組み合わせることで図表画像を識別する研究がある [1]。ここでは、棒グラフ、円グラフ、折線グラフについて識別器を構築しており、棒グラフ、円グラフについては、比較的良好な検出ができるものの、折線グラフについては十分な精度が得られず、改良が必要であるという課題が残った。

## 6. まとめ

本稿では、テレビ番組の映像データから図表画像を検出する手法を検討した。図表画像のヒストグラムに基づく特徴、および、エッジに基づく特徴についてどのような検出性能の差があるかを比較した。その結果、次元数の高いエッジに基づく特徴の方が良好な結果が得られるものの、ヒストグラムを利用した単純な特徴でも、正例が十分に収集できれば、良好な精度を達成しうることがわかった。

今後は、特徴量の組み合わせや改良について検討し、別の学習方法の利用も考慮しながら、さらに比較実験を進めていく予定である。

## 参考文献

- [1] 佐伯 隆太, 宮森 恒: テレビ番組に基づく Web コンテンツの信頼性判断支援システムの提案, 第 4 回データ工学と情報マネジメントに関するフォーラム (第 10 回日本データベース学会年次大会) (DEIM2012), B3-5, 2012
- [2] Miyamori Hisashi, Akamine Susumu, Kato Yoshikiyo, Kaneiwa Ken, Sumi Kaoru, Inui Kentaro, Kurohashi Sadao, "Evaluation data and prototype system WISDOM for information credibility analysis," Internet Research, Vol.18, No.2, pp.155-164, 2008.
- [3] 山本祐輔, 田中克己: データ対間のサポート関係分析に基づく Web 情報の信憑性評価, 情報処理学会論文誌 Vol.3 No.2, pp.61-79, 2010
- [4] Simske, S.J., "Low-resolution photo/drawing classification: metrics, method and archiving optimization," IEEE International Conference on Image Processing 2005(ICIP 2005), Vol.2, II,534-7, pp.11-14, 2005
- [5] Navneet Dalal, Bill Triggs, "Histograms of Oriented Gradients for Human Detection", International Conference on Computer Vision & Pattern Recognition, Vol. 2, pp.886-893, 2005
- [6] Chawla, Nitesh V. and Bowyer, Kevin W. and Hall, Lawrence O. and Kegelmeyer, W. Philip, "SMOTE: synthetic minority over-sampling technique", Journal of Artificial Intelligence Research, Vol.16, No.1, pp.321-357, 2002
- [7] Leo Breiman, "Random Forests", Machine Learning, pp.5-32, 2001