

動向情報の根拠探索のための図表画像検出手法の改善

北森 新[†] 宮森 恒[†]

[†] 京都産業大学コンピュータ理工学部 〒603-8555 京都府京都市北区上賀茂本山

E-mail: †{g1144434,miya}@cse.kyoto-su.ac.jp

あらまし テレビ番組の映像から図表を含む画像を検出する手法を検討する。我々の研究室では、Web 上の動向情報(ある対象の増減や時間的変化を記述した情報)の真偽をユーザが効率よく判断する手段の一つとして、テレビ番組中の図表画像(円グラフや棒グラフを含む画像)を提示する手法を提案している。先行研究では、エッジに基づく特徴を利用することで一定の良好な精度で図表画像を検出できるという結果が得られた。しかし、エッジが類似したものの誤検出や、図表のエッジ上に文字が重なる場合の検出漏れが問題として残った。本稿では、画像の分割と学習を組み合わせることで、図表画像の検出精度を改善する手法を提案する。

キーワード Web コンテンツ, 情報信頼性, 動向情報, テレビ番組, 図表画像, 機械学習

1. はじめに

現在、Twitter や 2ちゃんねる等の SNS をはじめ、Web 上では膨大な情報が発信・蓄積され続けている。利用可能な情報が増え、一見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、ユーザがこれらの中から、その信頼性を見極めた上で適切に情報を取捨選択することは、ますます困難になっているといわざるを得ない。

我々の研究室では、これまで、Web 上の動向情報(ある対象の増減や時間的変化を記述した情報)の真偽をユーザが効率よく判断する手段の一つとして、テレビ番組中の図表画像(円グラフや棒グラフを含む画像)を提示する手法を提案している。これにより、もし、動向情報が、該当する正しい図表画像の内容と矛盾していた場合、その情報は誤りであると判断することができ、情報のより適切な取捨選択につながることを期待できる。

先行研究では、エッジに基づく特徴を利用することで一定の良好な精度で図表画像を検出できるという結果が得られた。しかし、エッジが類似したものの誤検出や、図表のエッジ上に文字が重なる場合の検出漏れが問題として残った。

本稿では、画像の分割と学習を組み合わせることで、図表画像の検出精度を改善する手法を提案する。

2. 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 [3] 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、他の利用者がある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 コンテンツの信頼性評価の研究が挙げられる [4]。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ンツの信頼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 [5]。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 [6] と AdaBoost を組み合わせることで図表画像を識別する研究 [1] や、図表画像の識別においてヒストグラム特徴量 [5] や k-means による局所特徴量 [7] などを比較した研究がある [2]。

これら従来研究では、図表画像の識別には、エッジに基づく特徴量を用いた際に、より良好な結果が得られることが明らかとなったが、エッジが類似したものと誤検出や、図表のエッジ上に文字が重なる場合の検出漏れが問題として残った。

本稿では、画像の分割と学習を組み合わせることで、図表画像の検出精度を改善する手法を提案する。

3. 動向情報の根拠探索

ここでは、先行研究である、テレビ番組内の図表画像を用いた動向情報の根拠探索 [1] について概要を説明する。

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。

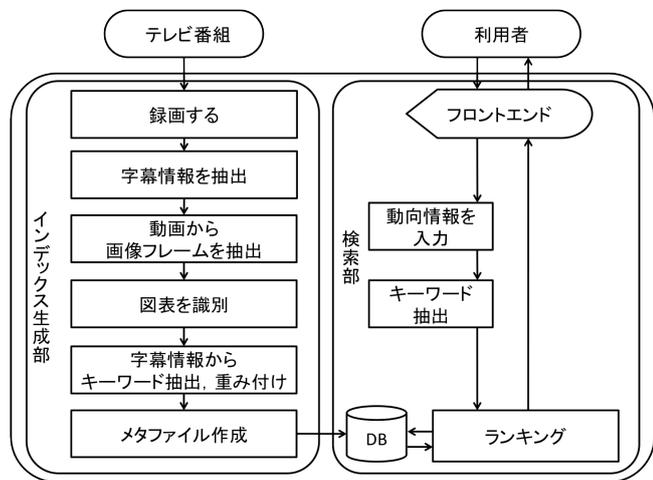


図1 システム構成図

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、メタファイルを作成する。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。大まかな処理手順は以下の通り。

- (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

4. 画像分割を用いた図表画像の検出

ここでは、提案手法である、画像の分割と学習を組み合わせた図表画像の検出手法について説明する。

従来手法の検出漏れの原因の一つに、図表画像に重畳する文字の存在が挙げられる。重畳により図表が本来もつエッジが隠

され、エッジ特徴から図表画像を十分に識別できなくなることが原因と考えられる。

一方、図表画像のエッジ全てに文字が重畳することは比較的まれで、エッジの一部は本来の図表画像の特徴を保持していることが少なくない。

そこで、予め図表画像をいくつかに分割し、分割された図表画像を用いて学習と検出を行う。そして、分割された図表画像が複数検出され、それらが分割前の相対的位置関係を保持していると判断される場合に、図表画像が検出されたと判断することとする。

なお、本稿で提案する内容は、図1のインデックス生成部にある「図表を識別」に相当する部分である。

4.1 正例、負例データの収集

表1 正例、負例の件数

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

録画されたテレビ番組、および、Webから人手で収集された円グラフ、棒グラフ、折れ線グラフの画像を正例として用いることとする[1][2]。同様の方法で収集した図表画像以外の画像については負例とする。収集した画像は、グレースケール化され、128x128ピクセルに正規化されている。ここでは、さらに、正例および負例画像を、図2のように、縦横半分に4分割した画像を学習データとして用意した。

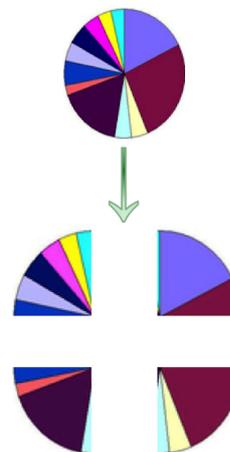


図2 図表画像の分割

4.2 検出に用いる特徴量

ここでは、先行研究で用いられた3種類の特徴量のうち精度の良かった2種類の特徴量を用いることとする。

(1) エッジに基づく特徴量

先行研究[1]で用いたエッジに基づく特徴量(HOGと呼ぶ)[6]を用いることとする。HOG特徴量[6]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。ここでは、先行研究 [1] にならい、図 3 のように特徴量の計算を行った。

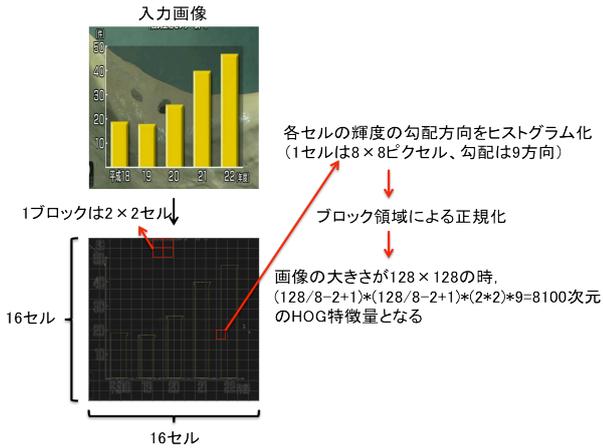


図 3 HOG 特徴量の抽出手順 (注1)

(2) K-MEANS による局所特徴量

画像を表現する特徴量を教師なし機械学習で設計する手法 [7], [9] が提案されている。この手法は、学習画像から一定の大きさのパッチをランダムにサンプリングし、パッチの k-means クラスタリングから得られる特徴量 (K-MEANS と呼ぶ) を用いることで、自然画像を高い精度で分類できることが報告されている。

特徴量学習の流れを示す。先行研究では [9] に習い、パッチサイズを 6×6 とされている。

1. 学習画像をグレースケール化し、128×128 画素にリサイズする。
2. 学習画像からパッチをランダムに 100 カ所取得する。
3. 得られたパッチに対して正規化および白色化を施す。
4. k-means クラスタリングによって特徴量を学習する。

ここで、頻繁に出現する一定の色領域をフィルタリングで除外するために、ランダムに取得するパッチは、分散が最大画素値の 10% より小さいときは除外し、パッチを再取得している。また、正規化とは、各パッチに対して画素値の平均を引き、標準偏差で割る操作であり、これにより、平均 0、分散 1 となり、コントラストが正規化される。また、白色化とは、パッチの共分散行列が単位行列となるような変換処理である。

k-means による学習では、クラスタリングの結果得られたクラスタ重心群をコードブックと呼び、コードブックとの類似度を特徴としている。

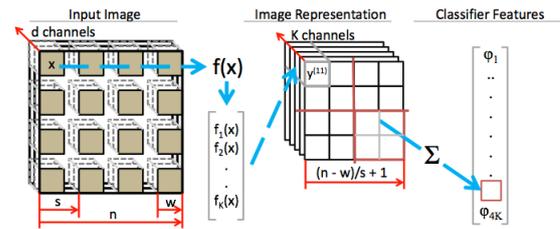
図表画像の識別

1. 入力画像からパッチを 1 画素ずつ移動させながら取得する。
2. 学習時と同様に各パッチの正規化を行う。
3. 各パッチを式 (1) により符号化する。
4. 入力画像を 2×2 の領域に分割し、各領域において 3. の総和をとる。
5. 4. で得られた特徴ベクトルを入力画像の特徴ベクトルとする。

(注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

$$f(x) = [f_1(x), f_2(x), \dots, f_K(x)] \quad (1)$$

$$f_k(x) = \max(0, \mu(z) - z_k) \quad (2)$$



ここで $z_k = \|x - c^{(k)}\|_2$ (k 番目のクラスタ中心 $c^{(k)}$ とのユークリッド距離)、 $\mu(z)$ は全クラスタ中心との距離の平均値を表す。

本稿では [9] と同様の手順を行い、コードブックサイズを 200 とし、特徴ベクトルの要素数を 800 とした。

4.3 学習および識別

学習には、樹木モデルの集団学習を行う Random Forest を用いた。

また、分割ありの提案手法では、4 分割した切片の 2 片以上で図表画像と識別された場合、全体としても図表画像と識別することとした。

5. 評価実験と考察

5.1 実験：分割の有無による図表画像の検出精度の比較

5.1.1 目的

画像を分割することにより、図表画像の検出精度が向上するかどうかを明らかにする。

5.1.2 方法

表 2 正例、負例の件数

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

表 2 で示した学習データに対して、4.2 節で示した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 (F-measure) によって評価する。適合率、再現率は、以下の式で算出する。

$$Precision = \frac{tp}{tp + fp} \quad (3)$$

$$Recall = \frac{tp}{tp + fn} \quad (4)$$

$$F\text{-measure} = \frac{2 \times precision \times recall}{precision + recall} \quad (5)$$

表 3 画像分割の有無による適合率の比較

手法	円グラフ	棒グラフ	折れ線グラフ
HOG(画像分割なし)	0.975 (347/356)	0.913 (116/127)	0.907 (127/140)
HOG(画像分割あり)	0.946 (365/386)	0.841 (138/164)	0.708 (138/195)
K-MEANS(画像分割なし)	0.842 (256/304)	0.898 (97/108)	0.811 (77/95)
K-MEANS(画像分割あり)	0.894 (356/398)	0.833 (140/168)	0.815 (172/211)

表 4 画像分割の有無による再現率の比較

手法	円グラフ	棒グラフ	折れ線グラフ
HOG(画像分割なし)	0.935 (347/371)	0.784 (116/155)	0.655 (127/194)
HOG(画像分割あり)	0.984 (365/371)	0.896 (138/154)	0.715 (138/193)
K-MEANS(画像分割なし)	0.842 (256/371)	0.626 (97/154)	0.397 (77/194)
K-MEANS(画像分割あり)	0.960 (356/371)	0.909 (140/154)	0.891 (172/193)

表 5 画像分割の有無による F 尺度の比較

手法	円グラフ	棒グラフ	折れ線グラフ
HOG(画像分割なし)	0.955	0.823	0.760
HOG(画像分割あり)	0.964	0.868	0.711
K-MEANS(画像分割なし)	0.842	0.738	0.533
K-MEANS(画像分割あり)	0.926	0.870	0.851

5.1.3 結果

画像分割ありの場合、なしの場合のそれぞれについて、各特微量で学習した識別器による検出精度の結果を、表 3, 4, 5 に示す。

実験結果より、HOG 特微量を用いた場合、円グラフ、棒グラフに対して提案手法による検出精度の向上が見られた。また、K-MEANS 特微量を用いた場合は、円グラフ、棒グラフ、折れ線グラフ全ての場合で提案手法による検出精度の向上が確認できた。

5.1.4 考察

ここでは、分割の有無による図表画像の検出結果について考察する。

分割なしの手法で検出に失敗し、分割ありの提案手法で検出に成功した例を図 5 に示す。また、分割ありの提案手法による検出漏れ、誤検出の例を、それぞれ図 6, 図 7 に示す。

a) 検出に成功した例

先行研究では文字の重畳等によるノイズや、複雑な背景に対する検出漏れが問題であったが、図表画像を分割し、ノイズの少ない領域を検出に利用する事により検出漏れを軽減できた。

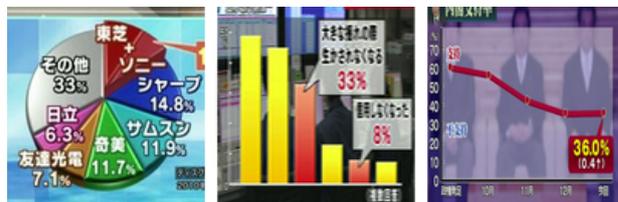


図 5 検出に成功した例

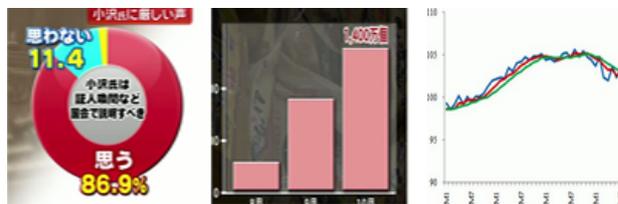


図 6 検出漏れ例



図 7 誤検出例

b) 検出漏れ・誤検出

i) 円グラフ

複数の片について大きいノイズがあるものについては検出できなかった。誤検出については、円の形に近いイラストのような画像を誤検出してしまふ傾向がみられた。

ii) 棒グラフ

複数ある棒グラフのうちの一つが大きく、分割した際に棒グラフのエッジが反映されないものは検出できない結果が得られた。また、棒グラフのように細い棒状のものが複数存在する画像については誤検出しやすい傾向がみられた。

iii) 折れ線グラフ

画像領域中の折れ線グラフのエッジのある部分が小さく、分割した際に空白部分が出てしまうものが検出漏れしやすい結果が得られた。中心に存在する物体のエッジが画像領域中の横を区切るものが誤検出される傾向がみられた。

ただし、ここで確認された誤検出の例は、図表画像を分割するだけではそもそも克服することが難しいと予想される事例であり、誤りの改善にはさらなる工夫が必要になると考えられる。

6. まとめ

本稿では、テレビ番組の映像から図表画像の検出精度を改善する手法を提案した。画像を予め分割し、各切片の識別結果に応じて最終的な検出結果を出力することにより、検出精度を改善する。実験の結果、エッジに文字が重なっている図表画像の場合等で検出漏れを改善できることがわかった。

今後は、実際のテレビ番組に対する検出精度や検出時間を調べる実験を進める予定である。

(注3) : 文献 [7] より引用

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] 梅澤 啓史, 宮森 恒: 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
- [3] 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.
- [4] 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 情報の信頼性評価, *情報処理学会論文誌* Vol.3 No.2, pp.61–79, 2010
- [5] 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
- [6] Navneet Dalal, Bill Triggs, "Histograms of Oriented Gradients for Human Detection", *International Conference on Computer Vision & Pattern Recognition*, Vol. 2, pp.886–893, 2005
- [7] A. Coates, H. Lee, and A.Y. Ng, "An analysis of singlelayer networks in unsupervised feature learning," *International Conference on Artificial Intelligence and Statistics*, pp.215–223, 2011.
- [8] P. Viola, M.J.Jones, "Rapid object detection using a boosted cascade of simple features," *CVPR*, vol.1 ,pp.511-518,(2001)
- [9] Manolis Savva, Nicholas Kong, Arti Chhajta, Li Fei-Fei, Maneesh Agrawala, Jeffrey Heer "ReVision: Automated Classification, Analysis and Redesign of Chart Images"
- [10] 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