

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 6月10日現在

機関番号：62615

研究種目：若手研究（B）

研究期間：2010～2012

課題番号：22700113

研究課題名（和文） 大規模放送映像アーカイブマイニングに基づく顔情報の構造化

研究課題名（英文） Face Information Structuring Based on Data Mining from Large-Scale Broadcast Video Archive

研究代表者

武 小萌（WU XIAOMENG）

国立情報学研究所・コンテンツ科学研究系・特任研究員

研究者番号：20462179

研究成果の概要（和文）：本研究では、多様性に富んだ大規模放送映像アーカイブに放送メタデータの階層的構造化を目指し、視覚特徴と放送メタデータの統合に基づいた顔認識・構造化技術高度化を実現した。照明条件、姿勢及び表情による特徴量変動に頑強な顔画像照合法を提案した一方、放送メタデータと映像間の相関関係をもとに、顔認識をネットワークの経路を伴った因果関係の発生確率を定量的に予測する課題に帰着させ、放送メタデータ・ベイジアンネットワークに基づいた新たな顔分類法を実現した。

研究成果の概要（英文）：In this project, we achieved a higher-level facial recognition and structuration technique based on the integration of visual information and broadcasting metadata. The technique was applied to a large-scale broadcast video archive with broadcasting metadata of great variety. While proposing a facial image matching algorithm robust against the variation in illumination condition, pose, and expression, we formulated the classic facial recognition task into a quantitative prediction problem that estimates the event probability of cause-effect relationship accompanied by network path. The solution was achieved in the form of a novel facial classification approach based on a Bayesian network constructed on top of broadcasting metadata, which leverages the correlation between the broadcasting metadata and the broadcast video in a both effective and efficient manner.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2010年度	1,100,000	330,000	1,430,000
2011年度	1,000,000	300,000	1,300,000
2012年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	3,100,000	930,000	4,030,000

研究分野：総合領域

科研費の分科・細目：情報学、メディア情報学・データベース

キーワード：マルチメディア

## 1. 研究開始当初の背景

- (1) 人物認識を目標とする従来の顔照合法には、より安定した撮影条件を前提とし

たものが多い一方、放送映像中の顔画像は、撮影視点や照明条件、人物の姿勢・表情の相違等の影響を受けやすく、視覚特徴による認識が困難である。

- (2) 放送映像における内容理解に利用可能なコンテンツとしては、膨大な放送番組やCM等の情報を、検索や取得を簡単にするために、細かく記述した放送メタデータがある。放送メタデータの利用法としては、文字放送字幕を利用した話題検出・追跡、並びに、一次メタデータと映像の統合による二次メタデータ生成等が提案・実証されている。

## 2. 研究の目的

- (1) 多様性に富んだ大規模放送映像アーカイブに放送メタデータの階層的構造化及び顔画像の検出・追跡の実現
- (2) 視覚特徴に基づいた、照明条件、姿勢及び表情による特徴量変動に頑強な顔画像照合法の実現
- (3) 視覚特徴と放送メタデータの統合、放送メタデータと映像間の相関関係の抽出、並びに、それに基づいた顔認識・構造化技術高度化の実現

## 3. 研究の方法

- (1) 顔画像検出には、Haar 特徴に基づいたカスケード分類器を、顔画像追跡には、KLT (Kanade-Lucas-Tomasi) 特徴点軌跡解析による顔領域追跡法を、大規模放送映像アーカイブへの適用を行った。顔画像照合にテレビコマーシャル等の同一映像断片が影響を与える可能性が考えられ、放送映像からテレビコマーシャルを自動的に検出できる二段時間的共起ハッシュアルゴリズムを提案した。
- (2) 顔情報構造化に必要な顔画像系列の照合には、Local Binary Pattern (LBP) や Local Phase Quantization (LPQ) 等を含む六種類の画像特徴量に基づいた顔画像照合法を提案した。また、顔画像照合は、照明条件、姿勢及び表情による特徴量変動が性能に大きく影響するが、Fukunaga-Koontz Transform (FKT) 法を画像特徴量に適用し、クラス内における要素の分散をより低く抑えると同時に、クラス間の分離の判別力を高める新たな特徴選択法を提案した。
- (3) 放送メタデータの階層的構造化を実現するには、ジャンルや登場人物等のメタデ

ータを変数にし、多数の独立頂点集合により構成される有向グラフを構築した。顔分類を実現するには、メタデータ有向グラフによるネットワークをもとに、顔認識をネットワークの経路を伴った因果関係の発生確率を定量的に予測する課題に帰着させ、放送メタデータ・ページアンネットワークに基づいた新たな顔分類法を提案した。

## 4. 研究成果

- (1) Haar 特徴に基づいたカスケード分類器及び KLT (Kanade-Lucas-Tomasi) 特徴点軌跡解析による顔領域追跡法を大規模放送映像アーカイブへ適用し、95.2%の平均適合率を実現した。テレビコマーシャル検出の実験では、提案した二段時間的共起ハッシュアルゴリズムを10時間の放送映像に適用した(表1)。提案手法は、より高い適合率を実現したのみでなく、既存研究より10倍以上の処理速さも本実験で実証された。

表1 テレビコマーシャル検出実験結果

	検出適合率	処理時間
提案手法	95.1%	4 秒
Berrani	77.5%	40 秒
Dohring	83.5%	108 秒

表2 視覚特徴に基づいた顔認識実験結果

特徴	Baseline	提案手法
LBP	65.2	70.6
Gabor C2	51.3	63.6
TPLBP	64.6	71.1
FPLBP	56.9	59.5
LPQ	63.9	68.5
RI-LPQ	59.7	62.8
Total	60.3	64.9

- (2) 視覚特徴に基づいた顔画像照合の実験は、Labeled Faces in the Wild Home (LFW) データセットを対象に行った。表2は実験結果を示している。提案した特徴選択法は、クラス間の分離の判別力を最大化にするための特徴量投影軸を学習することができるため、照明条件や姿勢等の環境変数の変動における影響をより低く抑えられることが本実験で実証された。
- (3) 放送映像メタデータに基づいた顔分類を評価するために、2011年に放送された10個のテレビ番組を収集した。本データセットは、284本の放送映像を含み、映像に登場した人物の数が176人となり、顔画像系列の数が15,158個とな

る。その中から、7,441 個の顔画像系列を学習データとし、残った 7,717 個を実験データとする。本データセットを対象に行った実験の結果は表 3 に示している。放送メタデータを利用した場合の顔分類は、視覚特徴のみに基づいた顔分類法より、11.4%も高い平均適合率を実現したことがわかる。また、実験結果から、ドラマや落語等の放送番組に対して放送メタデータの顔分類に与える有用な影響が少ない一方、ニュースやスポーツ、バラエティ等の放送番組に対して放送メタデータが顔分類のパフォーマンスを遥かに向上させることができたことも本実験で実証された。

表 3 メタデータに基づいた顔分類結果

	適合数	誤認数	正解率
提案手法	6,621	1,096	85.8%
Baseline	5,743	1,974	74.4%

(4) 顔画像照合にテレビコマーシャル等の同一映像断片が影響を与える可能性が考えられ、放送映像からテレビコマーシャルを自動的に検出できる二段時間的共起ハッシュアルゴリズムを提案した。記憶装置の大容量化に伴い、大規模の放送映像を蓄積して利用できるようになりつつあり、なかでもテレビコマーシャルは、地上波民間放送において全放送時間における 20%の比率を保有している。放送映像解析や市場調査等の分野には、蓄積された膨大な放送映像アーカイブからテレビコマーシャルを効率的且つ効果的に検出・同定する技術が求められる。本研究は、超高速の二段時間的共起ハッシュアルゴリズムを提案し、テレビコマーシャルの大規模放送映像アーカイブからの検出・同定を目指す。1 か月間のアーカイブに対する処理時間がわずか 42 分であった提案手法は、98.1%の検出精度と 97.4%の位置特定精度を達成した。また、検出された大量のテレビコマーシャルにおける経時分布を可視化することで、広告市場に関する潜在的知識の発見を可能にする新たなアプリケーションを実現した。本アプリケーションは、4 か月間 5 チャンネルという莫大の量の放送映像へ提案手法を適用した結果をもとにしたもので、提案手法の実用性とその知識発見への高い応用可能性を実証した。図 1 と図 2 は、アルコールと自動車の CM における放送頻度の分布を示している。アルコール CM の場合、放送自主規制により分布が規則正しい形状になっていることがわかる。また、自動車 CM の場合は、

主にサラリーマンが食事を取る時間帯を狙い、テレビ局が CM を放送させている傾向が見える。

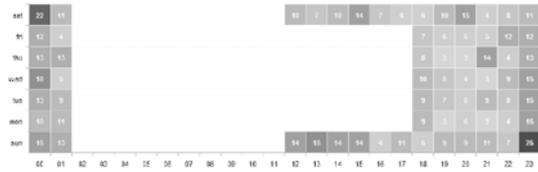


図 1 アルコール CM の放送頻度分布

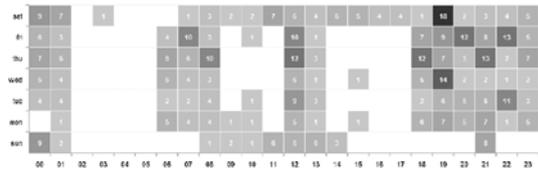


図 2 自動車 CM の放送頻度分布

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計 2 件)

- ① 武小萌、佐藤真一、Commercial mining based on temporal recurrence hashing algorithm and bag-of-fingerprints model, 18<sup>th</sup> IEEE International Conference on Image Processing, 2011.9.11, pp. 2529-2532, Brussels, Belgium, DOI:10.1109/ICIP.2011.6116177
- ② 武小萌、佐藤真一、Temporal recurrence hashing algorithm for mining commercials from multimedia streams, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 査読有、2011.5.22, pp. 2324-2327, Prague, Czech, DOI: 10.1109/ICASSP.2011.5946948

[図書] (計 1 件)

- ① 武小萌、Sebastien Poullot、佐藤真一、CRC Press、Multimedia Security: Watermarking, Steganography, and Forensics (Chapter 1. Multimedia Duplicate Mining toward Knowledge Discovery)、2012、423

6. 研究組織

(1) 研究代表者

武 小萌 (WU XIAOMENG)

国立情報学研究所・コンテンツ科学研究系・

特任研究員

研究者番号：20462179

(2) 研究分担者

なし

(3) 連携研究者

なし