

超高速テレビコマーシャル検出と知識発見への応用

Ultrahigh-Speed TV Commercial Detection, Extraction, and Matching

武小萌
Xiaomeng Wu

佐藤真一
Shin'ichi Satoh

何がわかる？

本研究は、二段階時間的共起ハッシュアルゴリズムを提案し、地上波民間放送時間の20%も占めるテレビコマーシャルを超高速且つ正確に検出するだけでなく、テレビコマーシャルの放送分布を明らかにし、既存研究で効率良く把握し難かった企業経営の実態を捉えることも目標としています。

どんな研究？

二段階時間的共起ハッシュアルゴリズムは、訓練データやクエリーを必要とせず、地域・時代によって異なりうる予備知識も依存しなく、超高速且つ正確にテレビコマーシャルの探知から、位置特定、並びにクラスタリングまで全自動的に実現する手法である。

実験・評価

10時間 → 右図参照

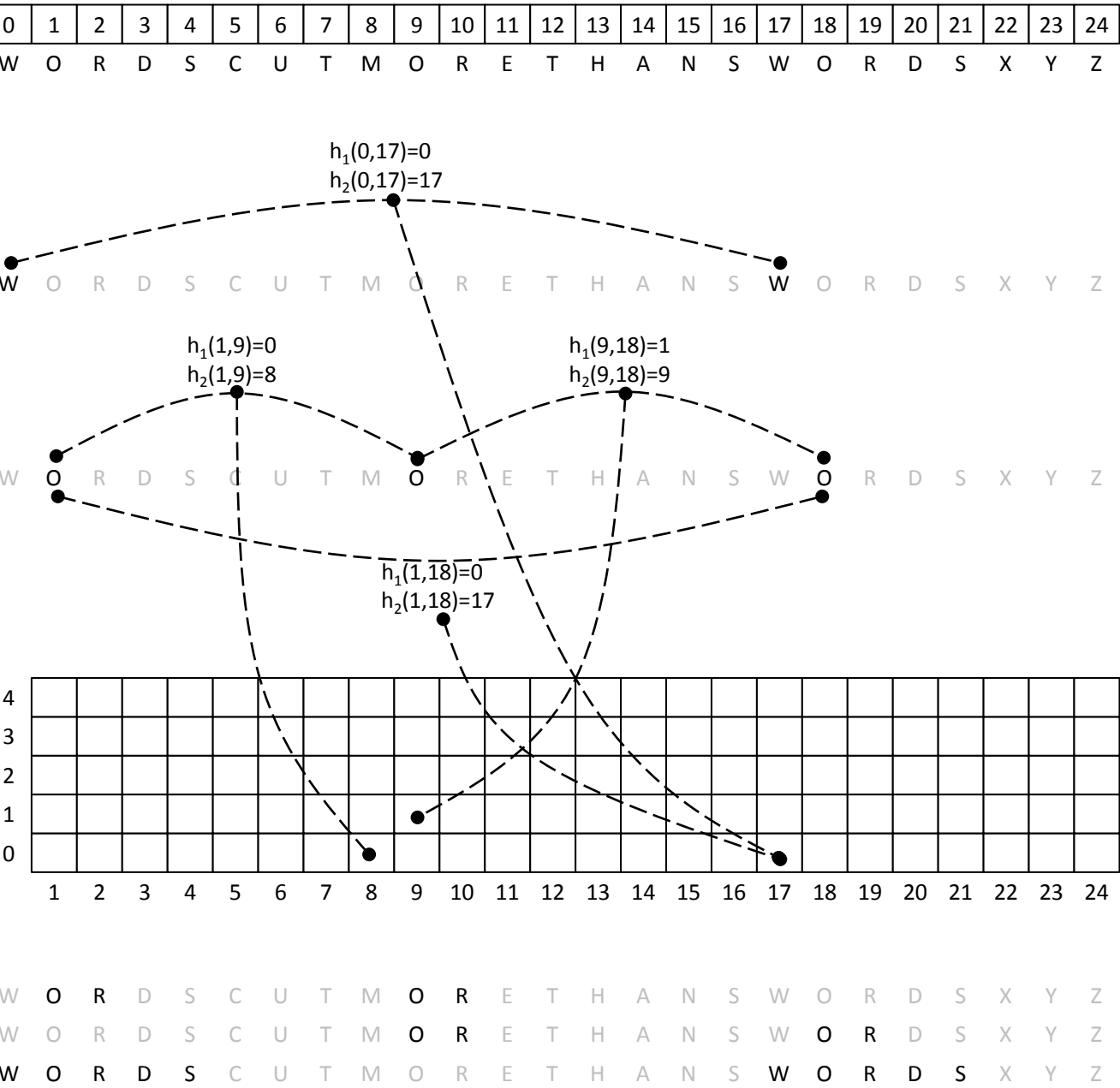
1か月間 → 50分（画像）；42分（音声）

5年間 → 21時間（4スレッド並行計算の場合）

検出精度：検出されたテレビコマーシャルクラスタを、重なり率の最も高いグラウンドトルースクラスタに割り当て、正しく割り当てられたテレビコマーシャルの割合が基準とする
位置特定精度：テレビコマーシャルが放送映像のどのフレームから始まりどのフレームまで終わるかを、どれだけ正確に検出できたかを評価するための基準

	P_s (%)	R_s (%)	F_s (%)	P_f (%)	R_f (%)	F_f (%)	t (sec.)
TRHA-V*	90.6	100	95.1	97.2	96.3	96.8	4
BERRANI	72.2	83.7	77.5	99.2	83.5	90.7	40
DOHRING	74.4	95.1	83.5	98.2	94.4	96.3	108
	P_s (%)	R_s (%)	F_s (%)	P_f (%)	R_f (%)	F_f (%)	t (sec.)
TRHA-A*	93.9	99.5	96.6	99.4	92.0	95.6	6
HAITSMA	82.2	98.0	89.4	97.9	86.7	92.0	40
	P_s (%)	R_s (%)	F_s (%)	P_f (%)	R_f (%)	F_f (%)	t (sec.)
TRHA*	96.6	99.5	98.1	97.3	97.4	97.4	10

超高速テレビコマーシャル検出法



マルチスレッドにおける検出結果の統合

コマーシャルクラスタの特徴を計算するにはまず、各コマーシャル映像を構成したFingerprintに注目する。Fingerprintが各コマーシャル映像に存在する相対位置の平均を、該当Fingerprintのテンポラルコードと定義する。クラスタから派生したFingerprintの数がNである場合、各Fingerprintのテンポラルコードを結合することにより、一本のN次元の特徴ベクトルが生成され、これをテンポラルコードベクトルと定義し、該当クラスタの特徴として用いる。

コマーシャルクラスタ間の照合を行うには、二つのクラスタが共有するFingerprintに注目する。共有されるFingerprintの数がMである場合、各クラスタから共有されるFingerprintのテンポラルコードのみを抽出し、二本のM次元のテンポラルコードベクトルが生成される。同じテレビコマーシャルのクラスタであっても、位置特定の不正確さからFingerprint相対位置間のずれが生じる場合があるため、テンポラルコードベクトル間の距離を直接に計算するのではなく、ダイナミックプログラミング照合法を用いる。クラスタ間の照合において、Fingerprintの類似性のみでなく、Fingerprintの時間的配置構造の類似性も照合基準に入れるため、復号エラー等の影響に強い性質を持つ。

	n_u	r_u (%)	n_o	r_o (%)
BOFs	3	1.31	10	4.37
HAITSMA	14	6.11	10	4.37

Large Vocabulary Quantization for Instance Search

朱才志 佐藤真一
Cai-Zhi ZHU Shin'ichi SATOH

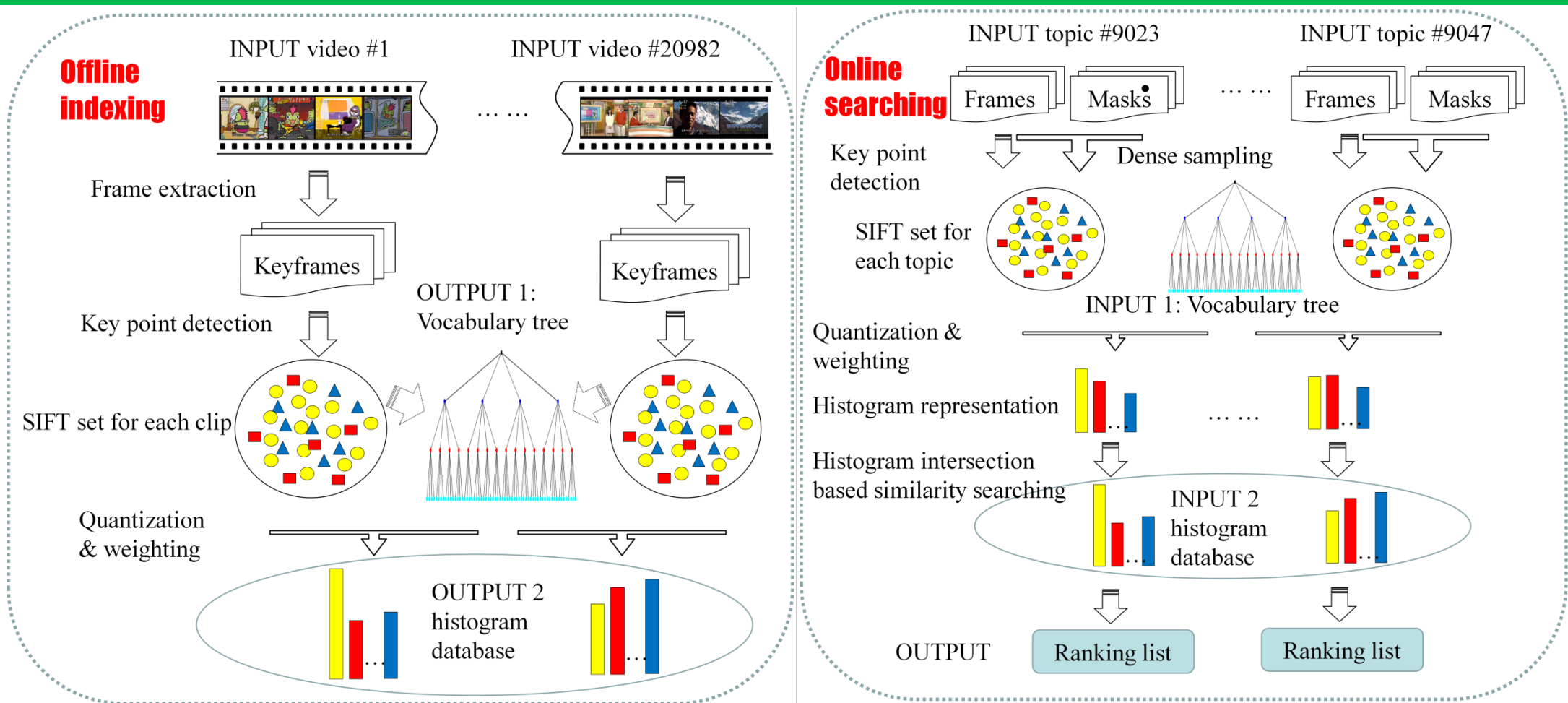
Inspirations

- ❖ The visual BOW analogy of text retrieval is very efficient for image retrieval. [J.Sivic, ICCV'03]
- ❖ Large vocabulary size improves retrieval. [D.Nister, CVPR'06]
- ❖ Query-to-Class (not Image-to-Image) distance is optimal; quantization degrades discriminability. [O.Boiman, CVPR'08]
- ❖ Hierarchical tree based pyramid intersection computes partial matching between feature sets without penalizing unmatched outliers. [K.Grauman, NIPS'06]

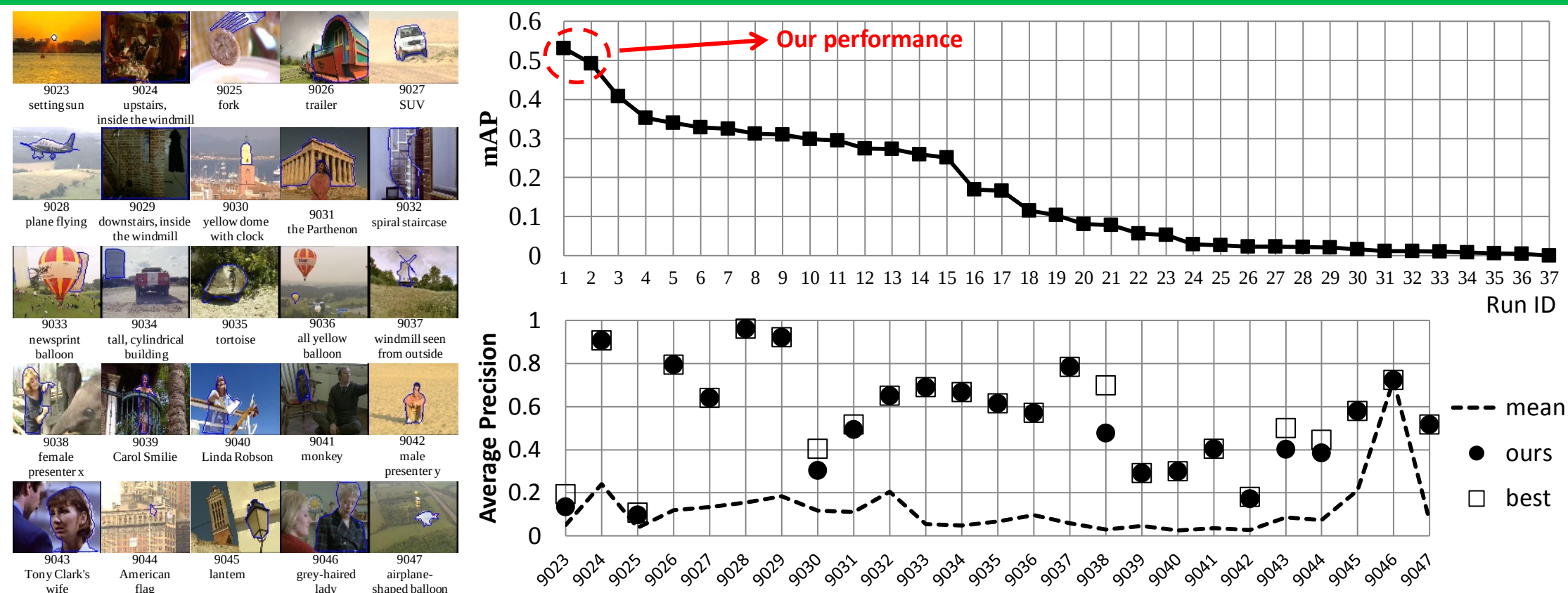
Features

- ❖ Building BOW framework upon hierarchical k-means based large vocabulary (1M) quantization.
- ❖ Matching similarity between topics and video clips.
- ❖ Balancing both context and object regions while computing similarity distance.
- ❖ Computing histogram intersection on hierarchically weighted histogram of codewords for ranking .

Pipeline of the Algorithm



Performance on TRECVID INS 2011 dataset



視覚的情報の役割に着目したニュースショット分類による 震災テレビ報道の分析

Classifying TV News Shots Based on the Visual Roles and
Its Application to the Aftermath of the Great East Japan Earthquake

片山 紀生
Norio KATAYAMA

孟 洋
Hiroshi MO

佐藤 真一
Shin'ichi SATOH

何かわかる？

テレビ報道は日常の身近な情報源のひとつであり、特に、東日本大震災以降は、国民ひとりひとりの生活に大きな影響を与える情報源となっています。震災以降のテレビ報道を長期間に渡って分析することにより、テレビ報道において視覚的情報がどのような役割を果たしているのか分析・考察することを目的としています。

どんな研究？

震災後のテレビ報道におけるニュースショットの分類と分析を行うためのプラットフォームを構築し、ニュースショットの出現傾向の分析を行いました。ショットの検出は計算機クラスタを用いた自動処理によって行い、ショットの分類は専用のメタデータ付与環境を実装し、目視による人手の作業によって行いました。

研究の目的と方法

テレビ報道は日常の身近な情報源のひとつであり、主に言語的情報と視覚的情報から構成される。テレビ報道の内容は、主として言語的情報によって伝えられるが、新聞記事など文字を主体とするメディアと比較したとき、視覚的情報はテレビ報道の大きな特徴のひとつである。そこで本研究では、ニュース番組中のショットを対象に視覚的情報の役割を考慮した分類を行うことにより、テレビ報道において視覚的情報がどのような役割を果たしているのかについて分析・考察を試みている。特に、東日本大震災以降は、テレビに限らず報道全般が国民ひとりひとりにとって生活に大きな影響を与える情報源となっており、震災以降のテレビ報道を対象とすることによって、甚大かつ深刻な災害におけるテレビ報道の特徴を明らかにするとともに、具体性の高い分析・考察結果の導出を目的としている。そこで、震災後のテレビ報道におけるニュースショットの分類と分析を行うためのプラットフォームを構築し、ニュースショットの出現傾向の分析を行った。ショットの検出は計算機クラスタを用いた自動処理によって行い、ショットの分類は専用のメタデータ付与環境を実装し、目視による人手の作業によって行った。東日本大震災後の半年間の NHK ニュース 7（延長時は先頭の30分間）に対してショット分類を行ったところ、学生アルバイト 6 人で分担してのべ約200時間の作業を必要とした。そして、得られた分類結果を用いてショットの出現傾向について分析した結果、モノログや説明用補助資料の出現頻度はニュースの内容によって大きく変動することがあり、震災後のテレビ報道の中では、特に原発事故の報道において顕著に現れていることが明らかになった。この結果は、テレビ報道において視覚的情報がどのように活用されているのかを考察する上で、具体的かつ客観的な手掛かりを与えるものである。

分類結果

ショット分類項目

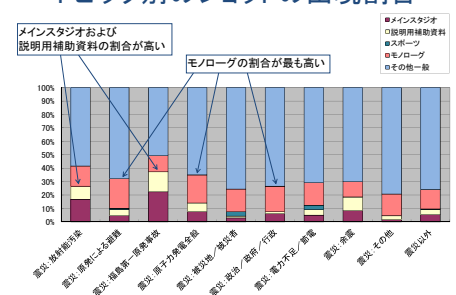
- ・オープニング
- ・メインスタジオ(アナウンサーによる番組進行等)
- ・説明用補助資料(説明用に作られた図／表／地図／リスト／模型／CG等)
- ・天気予報コーナー
- ・写真・映像
 - ・スポーツ中の場面
 - ・話している人の大写真(モノログ/会見/インタビュー等)
 - ・その他一般

ニュースの内容ではなく視覚的情報に基づく分類。
作業時間と正確さのため、複雑過ぎないよう配慮。

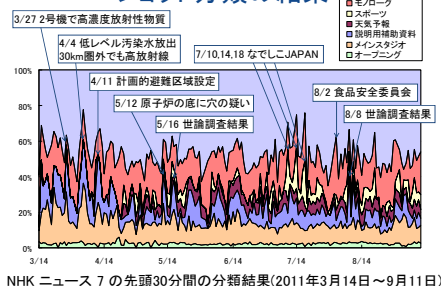
ニューストピックによる分類

- ・オープニング
- ・天気予報コーナー
- ・スポーツ
- ・震災
 - ・放射能汚染(流出した放射性物質による汚染)
 - ・原発事故による避難(避難地域、避難者、一時帰宅など)
 - ・福島第一原発事故(原子炉冷却、汚染水処理、損害賠償など)
 - ・原子力発電全般(再稼働問題など)
 - ・被災地／被災者／被災地支援(原発に限らないもの)
 - ・政治／政府／行政(原発に限らないもの、復興計画など)
 - ・電力不足／計画停電／節電
 - ・余震
 - ・その他(日本経済への影響など)
- ・震災以外

トピック別のショットの出現割合

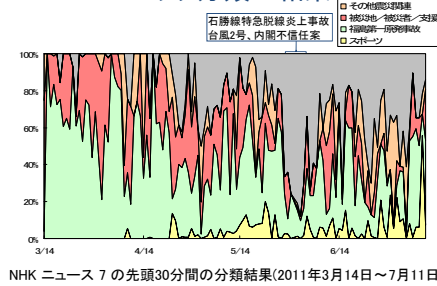


ショット分類の結果



NHK ニュース 7 の先頭30分間の分類結果(2011年3月14日～9月11日)

トピック分類の結果



NHK ニュース 7 の先頭30分間の分類結果(2011年3月14日～7月11日)

共通ショットの検出結果

震災当日に複数チャンネルで放送された映像



同一チャンネルでのみ複数回放送された映像の例



発表文献

1. 片山紀生, 孟 洋, 佐藤真一:「ニュースショットクラウド: 出現チャンネル数によって視覚的有用性を評価したニュース映像の一覧表示方式」, 電子情報通信学会技術研究報告(PRMU) 110(414) 165-170 (2011年2月).
2. 片山紀生, 孟 洋, 佐藤真一:「視覚的情報の役割に着目したニュースショット分類による震災テレビ報道の分析」, 情報処理学会研究報告 2012-IFAT-106(6) 1-8 (2012年3月).