

# 目的のものが映っているシーンを探す

## ～静止画クエリによる映像検索技術～

### どんな研究

特定の人物/物体/場所などの静止画を検索キー(クエリ)として、目的のものが映っているシーンを大量の映像の中から検索する技術の研究です。このような映像検索は**インスタンスサーチ**と呼ばれ、映像検索分野の国際的なコミュニティTRECVIDの重点課題の一つになっています。

### どこが凄い

新たに確率的なコンテンツベースの映像検索手法を考案しました。本技術では検索クエリであるインスタンスと各映像間の適合性を確率的に評価し、大量の映像をランキングします。これによりTRECVIDのデータを使用した実験で、**現在世界最高レベルの検索精度**を記録しています。

### めざす未来

雑音があったり、目的のものの見た目が異なっても情報検索を可能にするスマートなメディア探索技術の実現をめざします。この技術は、身のまわりの音・画像・映像を用いて、大量の音や映像の中から必要な情報を瞬時に探し出せるメディア社会の実現に寄与します。

### ・インスタンスサーチ

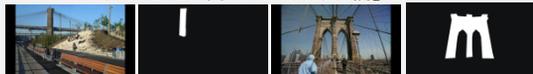
インスタンス例「スティーブン コルベア」



インスタンス例「ペプシコーラの丸いロゴ」



インスタンス例「ブルックリン橋」



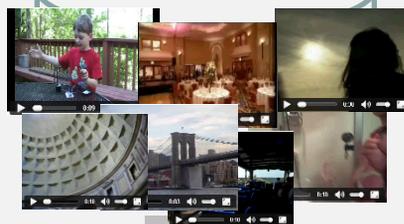
インスタンス例「パントン神殿」



インスタンス = 特定の人物・物体・場所

「システムはインスタンスが映っている映像をデータベースから検索し、適合度順でランキングして検索結果をユーザに提示する」

大規模映像データベース



「ブルックリン橋」

「ペプシコーラの丸いロゴ」

「スティーブン コルベア」

1. 映像No. 65213

2. 映像No. 55238

3. 映像No. 20169

4. 映像No. 1963

...

1000. 映像No. 62469

検索結果(映像ランキング)

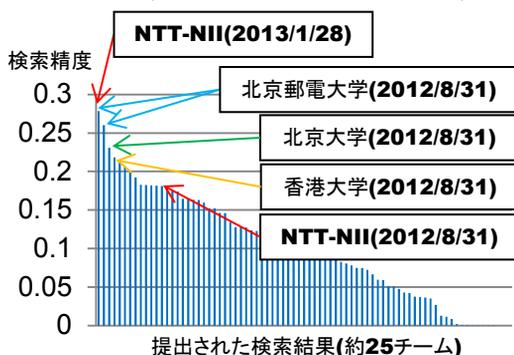
### ・提案手法(BM25M: BM25 with masked query)



● : 強調特徴点(技術のポイント)  
インスタンスに特有か否か、及び各映像に対する識別性の観点からクエリ画像群の特徴点を重み付けし、それらを多く含む映像を検索する

インスタンスと各映像間の適合性を確率的に評価

### ・検索精度(TRECVID 2012 dataset)



・BM25Mの映像検索タスクに対する有効性を確認  
・世界最高レベルのインスタンス検索精度を記録

### 関連文献

[1] M. Murata, T. Izumitani, H. Nagano, R. Mukai, K. Kashino, S. Satoh, "NTT Communication Science Laboratories and National Institute of Informatics at TRECVID2012: Instance Search and Multimedia Event Detection Tasks," in *Proc. TRECVID2012 Workshop*, 2012.

[2] 村田真哉, 永野秀尚, 向井良, 柏野邦夫, 佐藤真一, "画像をクエリとしたインスタンス映像検索," 信学技報 PRMU2012-175, pp. 215-219, 2013.

### 連絡先

村田 真哉 (Masaya Murata) メディア情報研究部 メディア認識研究グループ  
E-mail: murata.masaya[at]lab.ntt.co.jp ({at}の部分をもに置き換えてください)