

論文 / 著書情報
Article / Book Information

論題	GMM-Supervectorを用いた映像の高速セマンティック検索システム
著者	井上中順, 篠田浩一
出典	第18回画像センシングシンポジウム講演論文集, Vol. , No. , DS2-08
発行日 / Issue date	2012, 6
Note	第18回画像センシングシンポジウム講演論文集より転載

GMM-supervector を用いた映像の高速セマンティック検索システム Fast Semantic Video Search System Using GMM-supervectors

井上 中順† 篠田 浩一†

Nakamasa Inoue†, Koichi Shinoda†

†東京工業大学大学院情報理工学研究科計算工学専攻

†Department of Computer Science, Tokyo Institute of Technology

E-mail: {inoue, shinoda}@ks.cs.titech.ac.jp

1 目的・背景

近年，デジタル録画機器やネットワークの発展にと
もない，膨大な映像データが利用可能となった．必
要な映像をデータベースの中から効率よく発見する
ためには，映像の意味に基づいた検索技術が必要
である．本研究では，「人が楽器を演奏している」，
「飛行機が飛んでいる」などの意味的タグを映像に自
動付与し，関連映像を検索するシステムを提案する．

2 提案手法

本手法では，混合ガウス分布スーパーベクトル
(GMM-supervector)[1]を用いて映像に意味的タグ
を付与する．GMM-supervector は一般物体認識に
おける Bag-of-visual-words (BoW)を確率的枠組み
に発展させた手法である．本手法のアルゴリズムは
以下の通りである．

1. フレーム画像から局所特徴量 HOG を抽出する．
2. HOG から GMM パラメータを推定する．
3. GMM パラメータを結合して，GMM-supervector
を作成する．
4. 線形 SVM を用いて識別を行う．

本システムでは，[1]で提案されている GMM パラメ
ータ推定の高速化手法に加え，識別器に線形 SVM
を用いることで更なる高速化を行った．システムの外
観を図1に示す．入力された映像の2秒に1枚のフレ
ーム画像に対して GMM-supervector を抽出し，映
像の意味的タグを付与するとともに，関連映像ショット
を表示している．



図1. システムの外観

3 評価実験

映像解析に関するワークショップ TRECVID の
Semantic Indexing タスクのデータセットを用いて，
30 種類の意味的タグの自動付与実験を行った．映
像データはインターネットアーカイブから集められ
たものである．TRECVID 2011 の評価尺度に基づい
てシステムの評価を行った結果，検出精度の Mean
Average Precision は 0.107 となった．また，1枚の
入力フレーム画像に意味的タグを付与する際の計算
時間は，平均 0.39 秒(1.86GHz Intel Core 2 Duo
のノート PC で測定)であった．

参考文献

- [1] N. Inoue, and K. Shinoda. “A Fast MAP Adaptation Technique for GMM-supervector based Video Semantic Indexing”, In Proc. of *ACM Multimedia*, 2011.