

サポートベクターマシンの内部パラメータに基づく 事例選択手法の提案と映像境界検出問題への応用

Proposal of selection of training data based on internal parameter of SVM
and its Application to Shot Boundary Detection

高取 大地*¹ 橋本 和夫*¹ 寺邊 正大*¹ 松本 一則*²
Daichi TAKATORI Kazuo HASHIMOTO Masahiro TERABE Kazunori MATSUMOTO

NGUYEN, Dung Duc*²
Dung Duc NGUYEN

*¹東北大学大学院情報科学研究科
Graduate School of Information Sciences, Tohoku University

*²株式会社 KDDI 研究所
KDDI R&D Laboratories, Inc

Support Vector Machine (SVM) is known as a general learning technique that derives classifier with high classification performance. Machine learning is more difficult when unsuitable examples to learn are included in the training data. By removing unsuitable examples to learn in the training data appropriately, this paper proposes the method to improve the generalization of the classifier induced by SVM. Our method improves generalization without any significant loss of accuracy by relearning SVM. The effectiveness of the proposed method was proven through the experiment on shot boundary detection.

1. まえがき

近年、動画共有サイトの発展や、HDD レコーダの普及にと
もない、マルチメディアデータベース、映像アーカイブが急増
している。これらに含まれている映像データを有効に管理、再
利用を行うためには、映像データを場面ごとに整理して管理す
ることが必要となる。そのためには、映像データから場面の境
界、すなわちショットの境界を検出する必要がある。

しかし、手作業によるショット境界の検出は多大なコスト
を要する。このため、ショット境界を自動的に検出する技術が
求められている。このショット境界の自動検出は映像境界検出
(Shot Boundary Detection) と呼ばれている。

映像データを構成する最小単位はフレームと呼ばれ、1 枚の
静止画像として示される。そして、1 つのカメラにより撮影さ
れた連続したフレームによりショットが構成される。そして、
映像データは複数のショットより構成される。ショットが切り替
わった最初のフレームをショット境界として定義する。すなわ
ち、映像境界検出問題とは、ショットが切り替わった最初のフ
レームを検出することに相当する。フレーム、ショット、ショッ
ト境界の例を図 1 に示す。

映像境界検出問題は TRECVID[1] でも取り上げられたこと
などから、多くの研究者に取り組みられるテーマとなっている。
著者は、これまでに映像境界検出問題を、映像データの各
フレームを事例とし、ショット境界とそれ以外に分類する 2 値
分類問題に帰着させ、サポートベクターマシン (以下、SVM)
学習を適用することにより、優れた性能を持つ分類モデルが得
られることを報告している [2]。

SVM 学習は 2 値分類問題に対して有効な分類モデル導出手
法であることが知られており、連続値データの扱いが容易であ
り、非線形分類問題に対して性能の良い分類モデルを導出する
ことから、特徴量化した映像データに対する映像境界検出問題
に適した学習手法であるといえる。SVM 学習対象となるデー
タには、分類境界を決定づけるような学習事例が十分に含まれ
ている必要がある。一方、分類境界を決定づけるような学習事

例にミスラベル事例や、クラスの判断があいまいな事例が含ま
れていた場合、導出される分類モデルの性能は低下する。

映像境界検出問題の学習データは、手作業でショット境界か
否かのラベル付けを行うことにより生じるミスラベル事例や、
人間でもショット境界と判断してよいか迷うあいまいな特徴を
持つ事例が含まれる。これらの事例を学習データに含めると導
出される分類モデルの性能が悪化する。よって、学習データに
対して事例選択を適用し、学習対象として不適切なこれらの事
例を除去する必要がある。

ミスラベル事例や、ラベルの判断があいまいな事例が存在
した場合、わずかしが存在しないこれらの事例にも対応した分
類モデルの導出を試みることから、分類モデルに対して大き
な影響を与えるため、SVM 分類モデルにおいて、サポートベ
クター (以下、SV) となり、かつ、SV に付与される内部パラ
メータ (以下、内部変数) であるラグランジュ定数が大きな値
となる。筆者らは、この特徴に注目し、SV に付与される内部
変数が大きい事例を、ミスラベル事例やラベルの判断があい
まいな事例など学習に適さない事例として選択、除去すること
を考えた。

本論文では、学習データに対して導出した SVM 分類モデル
の内部変数に基づいて事例を選択し、学習に不適切な事例を除
去した学習データを用いて、再度 SVM 分類モデルを導出する
ことにより、最終的に性能の良い分類モデルを導出する手法を
提案する。また、実験を通じて、提案手法により最終的に導出
される分類モデルの汎化の程度と、分類性能について評価を
行った結果を報告する。

2. 研究の背景

2.1 映像境界検出問題

映像境界検出問題とは、映像中に存在する各ショットの切り
替え点であるショット境界を自動的に検出する問題である。本
研究では、ショットが切り替わった最初のフレームをショット境
界として定義する。ショットとショット境界の例を図 1 に示す。

映像境界検出問題は映像データをフレームごとに分割し、各
フレームがショット境界であるか否かを判定する問題であるこ
とから、2 値分類学習問題に帰着することが可能である。

連絡先: 高取 大地, 東北大学大学院情報科学研究科,
takatori@shino.ecei.tohoku.ac.jp

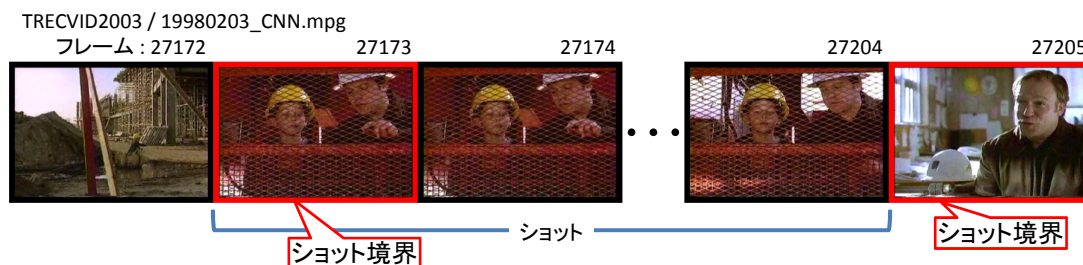


図 1: ショット境界

映像境界検出問題では特徴量選択の工夫により性能を改善することが可能であるが、本研究では特徴量選択の部分については所与とし、分類モデル導出手法を改良することにより性能を改善する。

2.2 サポートベクターマシン

2 値分類問題に対する有効な分類モデルの導出手法として SVM が提案されている。SVM は線形分類モデル、高次元空間への写像、最適化を組み合わせることにより、線形分類不可能な問題に対して効率よく分類モデルを導出することができる。また、凸 2 次計画法により問題を解くため局所解が存在しないという利点がある。さらに、通常、高次元空間における機械学習は低次元空間における機械学習と比べて計算量が大きくなるが、SVM では比較的高速に分類モデルを導出することが可能である。

SVM 学習は学習事例 x_i 、クラス $y_i = \{-1, +1\}$ の学習事例集合 $S = ((x_1, y_1), \dots, (x_l, y_l))$ に対して、式 1 に示す分類モデルとして示される。

$$f(x) = \sum_{i=1}^l \alpha_i y_i K(x_i, x) + b. \quad (1)$$

式 1 は、式 2, 3, 4 からなる最適化問題を解くことにより得られる。ここで、 α_i は SVM 分類モデルの内部変数であり、 $0 < \alpha_i$ に対応する学習事例 x_i が SV となる。また、 $K(x, z)$ はカーネル関数であり、高次元空間への写像を陰に定義している。

$$\text{目的関数} \quad \max_{\alpha, b} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j), \quad (2)$$

$$\text{制約条件} \quad \sum_{i=1}^l \alpha_i y_i = 0, \quad (3)$$

$$0 \leq \alpha_i \leq C, \quad (4)$$

$$i = 1, \dots, l.$$

3. 提案手法

3.1 提案手法の概要

本研究では、SVM 学習により導出された分類モデルの内部変数に基づいて、学習データに事例選択を行い学習に不適切な事例を除去し、更に、事例選択後の学習データを用いて再度 SVM 学習を行うことにより、より性能の良い分類モデルを導出する手法を提案する。

3.2 SVM 分類モデルの内部変数に基づく事例選択

3.2.1 反復解法による SVM 分類モデル導出

SVM 分類モデルを得る最適化問題の目的関数式 2 は 2 次凸関数であるため、反復解法により最適解を導出し、SVM 分類

モデルを導出することが可能である。反復解法の基本的手法である最急勾配上り法を用いて双対形式の最適解を求めるアルゴリズムを表 1 に示す。なお、 $\eta_i > 0$ は学習率を示す。

repeat ループの停止条件は双対目的関数の増加量が予め定めた閾値より小さくなった場合、Karush-Kuhn-Tucker 条件を満たした場合などが考えられる。

この手法はバイアス b を予め固定して計算する手法である。しかし、 b はカーネル関数の一部として置き換えることが可能であるため、最適化問題に制約条件を追加することとはならない。

表 1: 最急勾配上り法を用いた SVM 学習アルゴリズム

```

 $\alpha \leftarrow 0$ 
repeat : 停止基準を満たすまで
  for  $i = 1 ; i \leq l ; i++$ 
     $\alpha_i \leftarrow \alpha_i + \eta_i \left( 1 - y_i \sum_{j=1}^l \alpha_j y_j K(x_i, x_j) \right)$ 
     $\alpha_i \leftarrow \min(0, \alpha_i)$ 
     $\alpha_i \leftarrow \max(\alpha_i, C)$ 
  end for
end repeat
return  $\alpha$ 

```

3.2.2 内部変数 α_i に基づく事例選択

SVM 分類モデルの導出過程において内部変数 α_i の更新は、表 1 中の式 5 により行われる。

$$\alpha_i \leftarrow \alpha_i + \eta_i \left(1 - y_i \sum_{j=1}^l \alpha_j y_j K(x_i, x_j) \right). \quad (5)$$

ここで $1 - y_i \sum_{j=1}^l \alpha_j y_j K(x_i, x_j)$ は、学習事例 x_i が分類モデルからみて一般的ではなく学習が困難な事例であれば正、一方、学習が容易な事例であれば負の値をとる。すなわち、 α_i の更新では学習事例 x_i が学習困難であれば増加し、学習容易であれば減少する。よって、repeat ループ終了時点において、 α_i の大きさにしたがって学習の困難さをランク付けすることが可能である。

ミスラベル事例やクラスの判断があいまいな事例は、一般的に学習困難な事例であるため、 α_i が大きくなる。よって、内部変数 α_i に基づいて、この値が大きい、すなわち、学習困難である SV を、学習に不適切な事例として検出することができる。

筆者らは、 α_i と事例の学習困難さとの関係に注目し、除去する事例の選択を行い、選択した事例を除去した後の学習データに対して再学習を行うことにより、最終的に性能の良い分類モデルを導出する手法を提案する。具体的には、提案手法では SVM 学習により導出された分類モデルにおいて、 α_i が大

きい学習が困難な SV を除去すべき事例として検出する．そして、選択された事例を除去した後の学習データに対して再度 SVM 学習を行うことにより、より性能の良い分類モデルを導出する．ここで、性能の良い分類モデルとは、より単純で、高い分類精度を有する分類モデルをさす．

3.3 提案手法の流れ

映像境界検出問題に対応した分類モデルの性能向上を目的として、次に示す手順により事例選択を行い、SVM 分類モデルの再学習を行い改良する．フローチャートを図 2 に示す．

1. SVM 学習 学習データ (D1) に対して SVM 学習を行い、分類モデル (M1) を導出する．
2. 事例選択 分類モデル (M1) の内部変数 α_i に基づき、事例選択を行い、選択された学習が困難な事例を学習データ (D1) より除去する．事例選択後の学習データを学習データ (D2) とする．
3. SVM 再学習 事例選択後の学習データ (D2) に対して再度 SVM 学習を行い、分類モデル (M2) を導出する．

なお、事例選択を行う基準により、選択、除去される学習事例数は異なるため、結果として再学習により導出される分類モデルの性能は異なる．

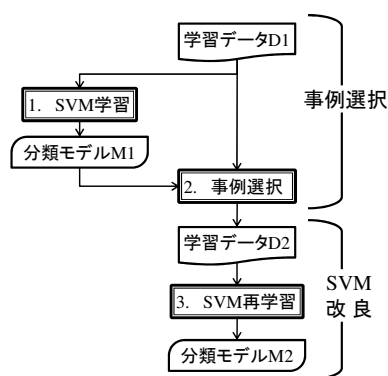


図 2: 提案手法

4. 実験

4.1 目的

提案手法による分類モデルの性能向上について評価するために実験を行った．事例選択前の学習データ、すなわち全ての学習データを用いて導出された分類モデルと、提案手法により事例を選択、除去した後の学習データを用いて再学習を行い導出された分類モデルの性能比較を行い、提案手法による事例選択の有効性について評価した．また、検出基準と導出される分類モデルの性能との関係についてもあわせて評価した．SVM 分類モデルの性能評価は、SV 数によるモデルの複雑さの評価、映像境界検出における再現率、適合率、 F 値という分類精度の評価により行う．

4.2 実験ツール

提案手法の検証を行うための SVM ツールとして LibSVM[3] を著者らが映像境界検出問題用に改良したものをを用いる．

SVM 学習における内部変数 α_i の上限を示すパラメータは $C = 20$ に固定する．また、カーネル関数にはガウス (Gaussian) カーネル $K(x, z) = \exp(-\gamma \|x - z\|^2)$ を用いる．カー

ネルパラメータは $\gamma = 0.2$ に固定する．なお、これらのパラメータは著者らの事前実験により得られた結果より最適値を選択している．

4.3 実験データ

4.3.1 TRECVID データについて

実験データとして、TRECVID 2003, 2004, 2005, 2006 にて使用された映像データを用いる．この映像データは ABC, CNN など海外の主要ニュース番組を録画したものであり映像境界検出問題の評価用データとしては一般的なものである．なお、本来配布されている映像データには映像境界検出問題用のラベルがあらかじめ割り当てられているが、ミスラベル事例が多く含まれている．実験で用いるデータは著者らが明らかなミスラベル事例のラベルを手作業で修正したものをを用いる．ただし、いくつかのミスラベル事例や、クラスの判断があいまいな事例は残っていると推測される．

4.3.2 学習データとテストデータ

TRECVID 2003 のデータセットを用いて、事例選択前の学習データを用いて導出された SVM 分類モデルと、提案手法による事例選択後の学習データを用いて導出された分類モデルについて TRECVID 2004, 2005, 2006 の映像データを用いて分類テストを行う．学習データ、テストデータの概要を表 2 に示す．各データセットに含まれる各映像データのフレーム総数を事例数、ショット境界総数を境界数とする．学習データに用いる TRECVID 2003 については番組や記録された日時の異なる 8 種類の映像データが提供されている．実験では、これら 8 種類の映像データのうち 2 種類を 1 組とした 4 組のデータを作成した．そして、このうち 3 組 6 種類の映像データからなる 4 とおりの学習データセットに対して、それぞれ学習を行い、4 つの分類モデルを導出した．次節に示す実験結果は、これら 4 つの分類モデルについてテストを行った結果の平均値である．

表 2: 学習データ, テストデータ

	データ	事例数	境界数
学習	データセット 1	308199	1910
	データセット 2	311508	1833
	データセット 3	314052	1920
	データセット 4	314601	1807
テスト	テストデータ	1960048	7517

4.4 事例選択の判定基準

提案手法は事例選択を行う際の最適な判定基準について、理論的に求めることはできない．そこで、提案手法における事例選択による SV 除去の判定基準について、導出された分類モデルにより誤分類される SV, 分類モデルにおいて $\alpha = 20$ となる SV, $\alpha \geq 15$ となる SV, $\alpha \geq 10$ となる SV, $\alpha \geq 5$ となる SV の 5 とおりを設定し、提案手法中の判定基準と導出される分類モデルの性能について確認すべく、実験を行った．

5. 実験結果

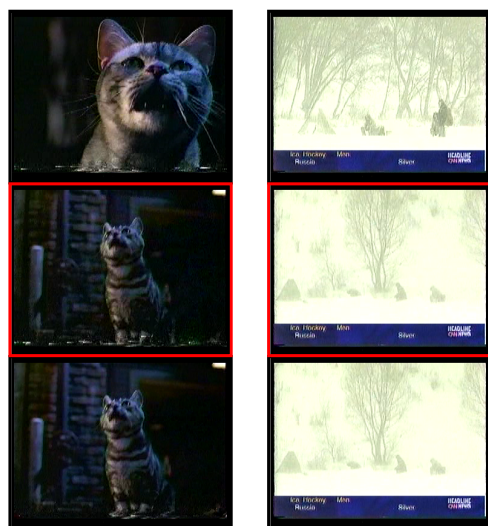
5.1 事例選択により除去される事例

判定基準 $\alpha = 20$ を用いた場合の事例選択において、除去対象となる例を図 3 に示す．図 3(a) は確認の結果、ミスラベル事例であった例である．このフレームは本来ショット境界であるが、ショット境界を示すラベルは付与されていなかった．

また、図 3(b) は、クラスの判断があいまいな事例として事例選択により除去すべき事例である．このシーンは 1 フレ

ムでシーンが変化しており、シーン境界に対応するラベルが付与されているが、シーン変化の前後ともひどい吹雪のシーンであり、特徴が極めて似ているおり、クラスの判断があいまいな事例である。

このように提案手法の事例選択により、ミスラベルである、あるいはクラスの判断があいまいである除去すべき事例が除去されている。



(a) ミスラベル事例 (b) あいまいな事例

図 3: 事例選択により除去対象となる映像

5.2 再学習による SVM の性能の変化

事例選択前の学習データを用いて導出された分類モデルと提案手法による事例選択後の学習データを用いて導出された分類モデルの SV 数, 再現率, 適合率, F 値を表 3 に示す。

事例選択により, より多くの SV を除去することで, 分類境界がより疎となり, より少ない SV 数で分類モデルを表現されるため, 事例選択後に再導出した分類モデルは, より単純なものになる。

事例選択後に再導出した分類モデルの再現率は向上している。特に, 判定基準を緩和し, より小さい内部変数 α に対応する SV を除去することにより, 再現率が向上している。

一方, 適合率は再現率とトレードオフの関係にあることが知られており, 再現率の上昇にともない, 下降する。また, 再現率と適合率の調和平均である F 値はほとんど変化していない。

しかし, 映像境界検出問題ではショット境界の検出に主眼を置くため, 再現率の向上を目的としている。提案手法を適用することにより単純で, 高い再現率を持つ, 性能の良い分類モデルを導出することができることから, 提案手法は映像境界検出問題に適した手法であることが示された。

また, 今回用いた判定基準では, 最も高い再現率を示した $\alpha \geq 10$ を採用すべきである。

5.3 多数クラスを中心とした事例選択

映像境界検出問題は極めてクラスの偏りが大きい問題である。よって, 多数クラスと少数クラスに異なる判定基準を用いた事例選択を行うことで, より性能の良い分類モデルの導出が可能であると考えられる。

ここでは, 少数クラスの事例選択の判定基準を最も緩やかな選択基準である「誤分類される SV」に固定し, 多数クラスの判定基準を変化させた場合の分類モデルの性能について実験により評価した。

表 3: 実験結果

	判定条件	SV 数	再現率	適合率	F 値
事例選択前	—	1276.25	0.8643	0.9150	0.8890
事例選択後	誤分類	1267.50	0.8727	0.9066	0.8893
	$\alpha = 20$	1270.00	0.8780	0.8976	0.8877
	$\alpha \geq 15$	1267.25	0.8794	0.8962	0.8877
	$\alpha \geq 10$	1222.25	0.8872	0.8923	0.8897
	$\alpha \geq 5$	1180.50	0.8854	0.8884	0.8869

表 4: 実験結果 (多数クラス中心の事例選択)

	判定条件	SV 数	再現率	適合率	F 値
事例選択前	—	1276.25	0.8643	0.9150	0.8890
事例選択後	誤分類	1267.50	0.8727	0.9066	0.8893
	$\alpha = 20$	1231.25	0.8905	0.8855	0.8880
	$\alpha \geq 15$	1228.75	0.8913	0.8878	0.8895
	$\alpha \geq 10$	1197.25	0.9007	0.8780	0.8892
	$\alpha \geq 5$	1168.50	0.9104	0.8633	0.8862

事例選択前の学習データを用いて導出された分類モデルと, 事例選択後の学習データを用いて導出された分類モデルの SV 数, 再現率, 適合率, F 値を表 4 に示す。

前節の同じ判定基準の場合の結果と比較すると, SV 数が減少しており, かつ, 再現率が向上している。このように, 多数クラスを中心とした事例選択を行うことにより, クラスの偏りを意識しない事例選択と比較して, より単純, かつ, 高い再現率を有する分類モデルを導出することができる。この特徴は, 映像境界検出問題に適したものである。

また, 多数クラスを中心とした事例選択における判定基準では, 最も高い再現率を示した $\alpha \geq 5$ を採用すべきである。

6. まとめ

本論文では SVM 分類モデルにおける内部変数 α に基づき学習データの事例選択を行い, 事例選択後の学習データに対して SVM 学習を行うことにより, より性能の良い SVM 分類モデルを導出する手法を提案した。

さらに, 提案手法を映像境界検出問題に適用し, TRECVID のデータを用いた実験を行った。ラベルを手作業で修正したデータに対して提案手法を適用し, 事例選択を行うことにより, 学習に不適切な事例を的確に除去できる。また, 再導出後の SVM 分類モデルはより単純な分類モデルとなり, かつ, 映像境界検出の再現率が向上させることがわかった。さらに, 提案手法による事例選択において, 多数クラスの学習事例を除去することが, 分類モデルの性能改善に有効であることを確認した。

参考文献

- [1] TREC Video Retrieval Evaluation.
<http://www-nlpir.nist.gov/projects/trecvid/>
- [2] Kazunori Matsumoto, Masaki Naito, Keiichiro Hoashi, Fumiaki Sugaya. SVM-Based Shot Boundary Detection with a Novel Feature. ICME 2006. pp.1837–1840. 2006.
- [3] Chih-Chung Chang, Chih-Jen Lin. LIBSVM – A Library for Support Vector Machines.
<http://www.csie.ntu.edu.tw/~cjlin/libsvm/>