

ブースティングを用いた野球実況中継に対するメタデータの作成

佐古 淳[†] 滝口 哲也^{††} 有木 康雄^{††}

[†] 神戸大学大学院自然科学研究科 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

^{††} 神戸大学工学部 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

E-mail: [†]sakoats@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

あらまし 大量のマルチメディア・コンテンツの中から欲しい情報を抽出するためには、コンテンツに対して適切なメタデータを付与することが必要である。本稿では、スポーツ実況中継、特に野球実況中継について、ラジオ音声の音声認識結果を用いてメタデータを作成する手法について述べる。単に、音声認識結果中のキーワードによってメタデータを作成した場合、わき出しが問題となる。また、ネットワーク文法を用いる場合、柔軟な表現を受理することが困難になるうえ、構築に多くのコストが必要となる。そこで、本研究では、ブースティングを用いることにより音声認識結果からメタデータを推定するための素性を学習し、素性の重み付き投票によって付与すべきメタタグ情報を推定する手法を提案する。推定すべきメタタグ情報は多クラスであることから、Multi-class AdaBoost を用いて推定を行う。実験により、単語正解精度 65.0%の音声認識結果から、メタタグ正解率 87%を得ることができた。また、未知語が発話された場合でも、メタタグ情報の推定精度はほとんど低下しないことから、未知語に対して頑健な手法であると言える。

キーワード ブースティング, メタデータ, 音声認識, SVM

Creating Meta-data to Baseball Live Games Based on Boosting

Atsushi SAKO[†], Tetsuya TAKIGUCHI^{††}, and Yasuo ARIKI^{††}

[†] Graduate School of Science and Technology, Kobe University Rokkodaicho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan

^{††} Faculty of Engineering, Kobe University Rokkodaicho 1-1, Nada-ku, Kobe, Hyogo, 657-8501 Japan

E-mail: [†]sakoats@me.cs.scitec.kobe-u.ac.jp, ^{††}{takigu,ariki}@kobe-u.ac.jp

Abstract It is necessary to provide appropriate meta-data to multimedia contents to retrieve them. In this paper, we propose the method to create meta-data using speech recognition results. As the multimedia contents, we focus on sports live speeches, especially baseball commentary speeches. To produce meta-data, we employ the boosting method instead of using simple keywords or a network grammar since keywords are weak with misrecognition and a network grammar cannot retrieve flexible expression. Experimental results showed the 87% of the meta-tag accuracy on the speech recognition results with 65.0% of word accuracy. Moreover another experiment indicated that the proposed method is robust for unknown words.

Key words Boosting, Meta-data, Speech recognition, SVM

1. はじめに

近年、デジタルテレビや WWW などの発展により、映像や音声など、多くのマルチメディア・コンテンツを所

有することが可能となってきた。このような大量のコンテンツに対しては、ユーザーが欲しい情報を検索できる必要がある。また、すべてのコンテンツを視聴するには時間がかかりすぎるため、要点だけ、または好みのシー

ンだけを抜き出して視聴したいという要求も存在する。

このような要求を満たすためには、映像や音声などのコンテンツに対してメタデータを作成しておく必要がある。コンテンツが大量に存在するため、人手でメタデータを作成することは現実的ではない。計算機が自動的にメタデータを付与できることが必要である。

本研究では、マルチメディア・コンテンツのひとつとして、スポーツ実況中継、特に野球実況中継を対象としてメタデータの作成を行った。メタデータとして、イベントの種別、実況か解説者との会話かなどの種別（以後メタタグ情報と記載）を付与する。すなわち、本研究の目的は、実況音声からメタタグ情報を識別し、メタデータとして付与することである。メタデータの詳細については2.章で述べる。

野球中継に対してメタデータを作成する研究としては、カメラワークを抽出して映像を構造化する手法 [1] や、映像中のテロップを解析する手法 [2]、クローズドキャプションを用いる手法 [3] などが提案されている。しかし、野球の実況中継に対して映像認識を行うことは困難である。そこで、アナウンサーの実況中継音声認識し、この言語情報を元にメタデータを作成するものとする。特に、テレビではなくラジオの実況中継音声を用いる。これは、映像がないために、ラジオ音声の方が情報量が多いためである。ラジオ実況中継音声を用いてメタデータを作成する研究として、従来、単語系列とともに状況を推定する音声認識 [4] について研究を行ってきた。これは、状況の推定、すなわちメタタグ情報の推定と音声認識を統合できるという点に利点があったものの、キーワード、もしくはキーワードの共起によりメタタグ情報の推定を行っていたため、推定能力自体は高精度なものとは言えなかった。

本研究では、これに対し、ブースティングを用いて高精度にメタタグ情報を推定する手法を提案する。提案手法では、メタタグ情報であることを特徴づける素性・ないことを特徴づける素性をそれぞれ学習し、重み付き投票によって識別を行う。また、メタタグ情報を識別するための素性を自動的に学習させることにより、人間がキーワードと考えないような単語をも識別のための特徴として用いることが可能となっている。

以下、次章で野球中継におけるメタデータ・メタタグ情報について述べる。3.章で提案手法であるブースティングによるメタタグ情報の推定について述べ、4.章で実験について述べる。最後に5.章でまとめる。

2. 野球中継に対するメタデータ

2.1 メタデータの種別

まず、本研究で用いたメタデータの種別について述べる。野球中継に対するメタデータとして、表1のメタ

表1 メタデータの種別

Table 1 List of meta-data.

中継に起因するメタデータ
解説者との会話 投球 実況イベント 実況一般 ボールカウントの解説 アウトカウント・ランナーの解説
ルールに基づくメタデータ
ストライク ボール アウト 三振 ファール フォアボール 牽制球 盗塁 ヒット ホームイン ツーベース

タグ情報を用いた。中継に起因するメタデータのうち、“解説者との会話”は、アナウンサーが実況ではなく、解説者を行う会話に対して付与した。これらの発話は、実況ではなく、試合進行とは無関係であるため、要約を行う際に有効な指標となると考えられる。“投球”は、ピッチャーが投球を行ったことを示す発話に対して付与した。投球は、野球において、試合進行の起点となる重要な動作であり、映像のリプレイなどでも投球を起点に再生される場合が多い重要な情報である。“実況イベント”は、バッターがボールを打った後の実況全体に対して付与した。ボールがグラウンド上にあり、行方の分からない状況を実況した発話であるため、アナウンサーの興奮度が大きくなっている。“実況一般”は、その他、細かく分類することの出来ない発話に付与した。

ルールに基づくメタデータは、カウントに関するもの、及びイベントとなっている。“ストライク”は、投球の結果、ストライクとなったことを示す発話に対して付与した。“ボール”、“アウト”等も同様である。学習データ中に現れなかったイベントについては省略した（インフィールドフライ等）。

2.2 学習コーパスの仕様

次に、本研究で用いたコーパスの仕様について述べる。実況中継音声は、ラジオの音声を用いた。録音は、アナウンサーにヘッドセットを装着した状態で行った。球場の歓声などはノイズとして重畳はしているものの、音声認識に大きく影響をあたえるほどではない。解説者との会話も存在するが、解説者の声は含まれていない。発話速度が速い、言い間違いが多い、発音がなまけているなどの特徴があり、音声認識にとっては困難なタスクとなっている。

人手による書き起こしテキストにおいて句点で区切られた単位を発話の単位とした。音声認識、及びメタタグ情報識別はともに、一発話毎に行った。メタタグ情報識別のための学習データは、発話毎にメタタグ情報ラベルを人手で付与し、作成した。発話は2233発話であり、総発話時間は約1時間半であった。

3. 提案手法

提案手法の概要を図1に示す。入力音声に対して、音声認識を行い、認識結果に対して、Decision Stumpsに基づ

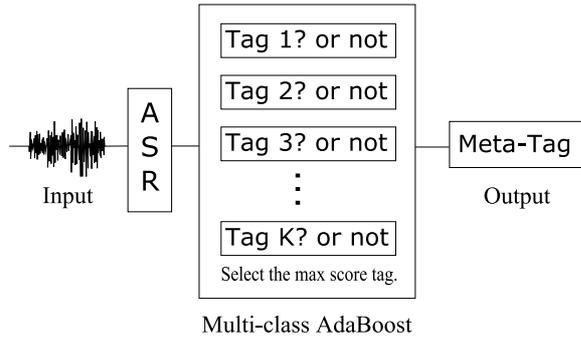


図 1 提案手法の概要

Fig. 1 System flow of proposed method.

く AdaBoost により識別を行う。ただし, AdaBoost は 2 値の識別器であるため, 多クラスであるメタタグ情報の識別を行うには拡張が必要である。本研究では, one-vs-rest 法による Multi-class AdaBoost によって多クラスのメタタグ情報の識別を行った。以下, 次節で AdaBoost と, その拡張である one-vs-rest の Multi-class AdaBoost について述べる。

3.1 AdaBoost

AdaBoost は, いくつもの識別器を組み合わせてひとつの高度な識別器を構成する *ensemble learning method* のひとつである。Schapire ら [5] が提案している学習のアルゴリズムを図 2 に示す。図中, I は, $I(\text{true})$ ならば 1, $I(\text{false})$ ならば -1 となる。 ϵ_t が 0.5 未満の弱学習器を見つけ続けることができれば, 学習誤差 0 の最終学習仮説を生成できる。また, 未知のサンプルに対する汎化誤差も小さくできることが実験的に報告されている [6], [7]。一方, 雑音を有するサンプルの場合, 過学習を起こすことが報告されている。これに対しては, AdaBoost の学習過程をマージン最大化ととらえ, SVM における Soft Margins の概念を導入した手法も提案されている [8], [9]。本研究では, 認識結果を扱うため, サンプルには多くの雑音に乗っているものと考えられる。このことから, 通常の AdaBoost ではなく, Soft Margins 付きの AdaBoost を用いることとした。

AdaBoost を用いたテキスト分類手法としては, 文献 [5], [10] などが提案されている。これらの文献では, テキスト分類のための弱学習器として, Decision Stumps が用いられている。Decision Stumps とは, ある素性の有無に基づいて分類を行う単純な手法である。素性には, 単語や単語 bi-gram, ラベル付き順序木などが用いられる。学習時には, 学習サンプルを最もうまく分類するような“素性”を選択し, その際の重みを得る。すなわち, 弱識別器 $h_t(\mathbf{x})$ は, 文書 \mathbf{x} 中に素性 t があるか, ないかによって $+1$, または -1 を出力する関数となる。識別は, 学習した T 個の弱識別器による重み付き投票によって行う。

Input: n examples $Z = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$

Initialize: $w_1(\mathbf{z}_i) = 1/n$ for all $i = 1 \dots n$

Do for $t = 1, \dots, T$,

(1) Train a base learner with respect to weighted example distribution w_t and obtain hypothesis $h_t : \mathbf{x} \mapsto \{-1, 1\}$

(2) Calculate the training error ϵ_t of h_t :

$$\epsilon_t = \sum_{i=1}^n w_t(\mathbf{z}_i) \frac{I(h_t(\mathbf{x}_i) \neq y_i) + 1}{2}.$$

(3) Set

$$\alpha_t = \log \frac{1 - \epsilon_t}{\epsilon_t}$$

(4) Update example distribution w_t :

$$w_{t+1}(\mathbf{z}_i) = \frac{w_t(\mathbf{z}_i) \exp\{\alpha_t I(h_t(\mathbf{x}_i) \neq y_i)\}}{\sum_{j=1}^n w_t(\mathbf{z}_j) \exp\{\alpha_t I(h_t(\mathbf{x}_j) \neq y_j)\}}.$$

Output: final hypothesis:

$$f(\mathbf{x}) = \frac{1}{\|\alpha\|_1} \sum_{t=1}^T \alpha_t h_t(\mathbf{x}).$$

図 2 AdaBoost のアルゴリズム

Fig. 2 AdaBoost algorithm.

AdaBoost による識別は, 2 値の識別しか行うことができない。そこで, 多クラスの識別を行うため拡張する必要がある。本研究では, one-vs-rest 法により多クラスへの拡張を行う。

まず, 特定のクラス k に着目し, 学習データをクラス k とそれ以外に分類する。クラス k の学習データには $+1$, それ以外には -1 のラベルを付与し, クラス k に対する強識別器 $y_k(\mathbf{x})$ を学習する。これを繰り返し, クラスの数を K 個とすると, 強識別器 $y_1(\mathbf{x})$ から $y_K(\mathbf{x})$ まで学習を行う。識別は, 以下の式により行う。

$$\hat{k} = \underset{k}{\operatorname{argmax}} f_k(\mathbf{x}). \quad (1)$$

3.2 キーワード・フィルタリング

野球実況中継においては, 解説者との会話中や選手名など, システムにとって未知の単語がしばしば出現する。このような未知語が別の単語に誤認識され, メタタグ情報の推定に悪影響を及ぼす恐れがある。

通常, このような問題に対してはキーワード・スポッティングが用いられる。キーワード・スポッティングでは, キーワードを認識・受理するネットワーク文法と, キーワード以外を認識するガベージモデルが用いられる。特に, ガベージモデルを構築する手法として, キーワードとは異なるドメインの発話集合から抽出した頻出音節系列を用いる手法が効果的であると報告されている [11]。

しかし, キーワード・スポッティングを用いる場合, ネットワーク文法を構築する必要があることに加え, どのような学習データに基づいてガベージモデルを構築するか決定することが困難である。本研究では, キーワー

ド、ガページとともに野球に関連する単語になることから、効果的なガページモデルの学習は困難である。

これに対し、提案手法では、以下の機構により自動的にガページモデルに相当するモデルが構築されるものと考えられる。まず、ブースティングによる学習の結果、単語は以下の3種類に分類される。

- メタタグ情報 k であると投票する単語
- メタタグ情報 k ではないと投票する単語
- 投票に関与しない単語

このうち、投票に関与しない単語が大半を占める。未知語が発話された場合、“投票に関与しない単語”として認識されれば、メタタグ情報の推定には悪影響を及ぼさないことになる。すなわち、“投票に関与しない単語”はガページモデルとして働くものと考えられる。実験の結果、未知語は高い割合で“投票に関与しない単語”に認識されていることがわかった。実験の詳細は4.2.3節で述べる。

この機構は、発話中のキーワードを発見するためのものではない。そのためキーワード・スポッティングとは別の概念であると考えられる。本稿では、投票に関与する単語のみを拾い出すという点からキーワード・フィルタリングと呼ぶものとする。

4. 実験

4.1 クリーンな書き起こしテキストを用いたメタデータの付与

システムの上限值を調査するため、人手によって書き起こされたクリーンな（認識誤りの含まれていない）テキストに対するメタデータ付与実験を行った。比較手法として、SVMを用いた。かな漢字混じりの書き起こしテキストを Mecab [12] によって形態素に分割し、AdaBoost, SVM でそれぞれ学習を行った。なお、SVM においては単語の出現頻度をベクトル化し、特徴とした。AdaBoost においては、素性として、uni-gram、及び bi-gram を用いた。判別実験は、10 fold のクロスバリデーション法により行った。メタタグ情報の正解率、及び、メタタグ情報の再現率、適合率を求めた。実験結果を表2に示す。

ブースティングを用いた場合と SVM を用いた場合では、同等の性能を示している。ほとんど全ての条件において、約9割の精度で識別ができています。識別に失敗しているものは、出現頻度の低い表現を用いた発話や、人手でもラベルの付与に迷うような発話などであった。

4.2 音声認識結果を用いたメタデータの付与

次に、音声認識結果を用いて、メタデータを作成する実験を行った。まず、音声認識の条件と結果について述べる。

4.2.1 音声認識条件と結果

ベースラインの音響モデルは、日本語話し言葉コーパス (CSJ: Corpus of Spontaneous Japanese) モニター

表2 クリーンな書き起こしテキストによるメタタグ情報識別結果

Table 2 Results of meta-tag discrimination using clean transcription.

	AdaBoost(uni)	AdaBoost(bi)	SVM
正解率	0.93	0.94	0.91
投球 (F 値)	0.98	0.99	0.99
会話	0.97	0.97	0.96
実況イベント	0.86	0.90	0.86
カウント	0.95	0.94	0.95
ランナー	0.90	0.92	0.86

表3 音響分析条件と HMM の仕様

Table 3 Condition of acoustic analysis and HMM specification.

サンプリング周波数	16kHz
特徴パラメータ	MFCC(25 次元)
フレーム長	20ms
フレーム周期	10ms
窓タイプ	ハミング窓
タイプ	244 音節
H 混合数	32 混合
M 母音 (V)	5 状態 3 ループ
M 子音+母音 (CV)	7 状態 5 ループ

版 [13] のうち、男性話者 200 名の講演音声を用いて作成した。音響分析条件と HMM の仕様を表3に示す。これらの条件で音響モデルを作成し、さらに、MLLR+MAP [14] により音響モデル適応を行った。音響モデル適応は、オープン、クローズドそれぞれの条件で行った。ただし、オープンにおいても、テストセットと同一話者の適応データを用いた。適応データの分量は、約1時間半であった。

言語モデルは、野球実況中継音声の書き起こしテキストから tri-gram モデルを作成した。書き起こしテキストは、話し言葉を書き起こしたものである。学習コーパス中にテストセットを含めたものをクローズド、含めなかったものをオープンとした。異なり単語数は約3,000、コーパスサイズは約8万形態素であった。

音声認識結果の単語正解精度、及びキーワードの F 値を表4に示す。ここで、キーワードは AdaBoost によって学習された素性を用いた（次節参照）。音響モデルについては、オープンであっても話者は同一であるため、言語モデルをクローズにする方が効果が大きくなっている。キーワードの F 値は、ほぼ単語正解精度に比例する結果が得られた。以後、これらの認識結果を用いて実験を行った。

4.2.2 音声認識結果に対するメタデータの付与

前節によって得られた認識結果に対し、提案手法によるメタタグ情報の推定を行った。学習とテストはクリーンな書き起こしテキストの場合と同様に 10 fold のクロス

表 4 音声認識結果の単語正解精度.

Table 4 Word accuracy of the speech recognition results.

音響モデル	言語モデル	単語正解精度	キーワード F 値
オープン	オープン	65.0%	0.80
クローズド	オープン	70.4%	0.84
オープン	クローズド	73.4%	0.85
クローズド	クローズド	78.3%	0.88

バリデーション法により行った．素性は，AdaBoost では uni-gram，及び bi-gram を用いた．SVM ではクリーンな書き起こしテキストの場合と同様に，単語の出現頻度をベクトル化したものを特徴として用いた．音響モデル・言語モデルともにオープンの条件における音声認識結果を用いた場合のメタタグ情報識別結果を表 5 に示す．また，このとき Adaboost (bi-gram) によって学習された素性語を表 6 に示す．

識別性能については，AdaBoost (uni-gram, bi-gram)，SVM とともに大きな差はみられない．クリーンテキストの場合に比べて精度は低下するものの，高精度な識別ができています．ただし，実況イベントの識別については大幅な性能の低下が見られた．これは，ボールを打った後の発話であるため，アナウンサーが興奮した状態で発話を行い，そのために音声認識精度が極端に低くなるのが原因であると考えられる．同様に，ヒットやアウトのような場合にも識別性能の低下がみられた．

次に，音声認識率とメタタグ情報識別率との関係を図 3 に示す．横軸は音声認識精度であり，音響モデル・言語モデルのオープン・クローズドの条件を変化させることによって変化させたものである．“認識結果で学習”は，音声認識結果を用いてブースティングを行い，識別素性を学習したものである．“クリーンテキストで学習”は，クリーンテキストを用いて識別素性を学習し，これを用いて認識結果の識別を行った．図より，認識結果を用いて識別素性の学習を行った方が高精度にメタタグを識別できることがわかる．これは，認識誤りの傾向まで含めて学習が行われたためと考えられる．また，クリーンテキストを用いて学習を行った場合，音声認識精度の低下とともに，メタタグ識別性能も低下する傾向がみられる．これに対し，認識結果を用いて学習を行った場合，低下の傾向が緩やかであった．音声認識結果を用いて学習を行うことにより，認識性能に対して頑健なメタタグ識別ができるものと考えられる．

4.2.3 キーワード・フィルタリングの効果

本手法において未知語に対する頑健性を検証するために実験を行った．未知語は，音声認識辞書から“投票に関係しない単語”をランダムに削除することによって生成した．未知語を含む辞書を用いて再度音声認識を行い，その結果に対して，メタタグ情報の識別を行った．識別

表 5 音声認識結果によるメタタグ情報識別結果

Table 5 Results of meta-tag discrimination using recognized transcription.

	AdaBoost(uni)	AdaBoost(bi)	SVM
正解率	0.84	0.87	0.84
投球 (F 値)	0.94	0.96	0.98
会話	0.90	0.93	0.89
実況イベント	0.61	0.71	0.67
カウント	0.93	0.93	0.92
ランナー	0.86	0.90	0.86

表 6 AdaBoost によって選択された素性語

Table 6 Selected features by AdaBoost.

メタタグ情報識別素性例
あたり あんまり きのう ほんと よく アウト インサイド ストライク スリー ツーアウト バッター ボール ワン 一塁 回っ 外れ 甘い 監督 詰まっ 球 空振り 牽制 三振 始まり 送球 打ち 直球 変化球 方向

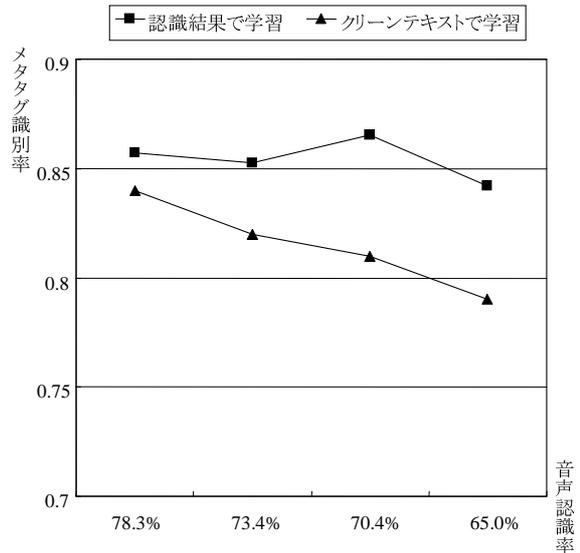


図 3 音声認識性能とメタタグ識別性能の関係

Fig. 3 Relations between ASR performance vs meta-tag discriminating performance.

器の学習はやり直さず，未知語なしの条件で学習した識別器をそのまま用いた．AdaBoost に用いた素性は単語 uni-gram である．未知語率を 10% から 50% まで変化させて実験を行った．結果を図 4 に示す．

未知語が投票に関係する単語として認識されてしまう割合は 12% ~ 15% 程度であることがわかる．また，わき出している単語は，比較的投票重みの小さいものが多くみられた．未知語なしの場合のメタタグ正解率が 0.84 であるのに対し，未知語率を 10% ~ 50% と上げた場合でも 0.83 ~ 0.82 と，メタタグ正解率の変化は微少である．このことから，未知語に対して頑健にメタタグ情報の識

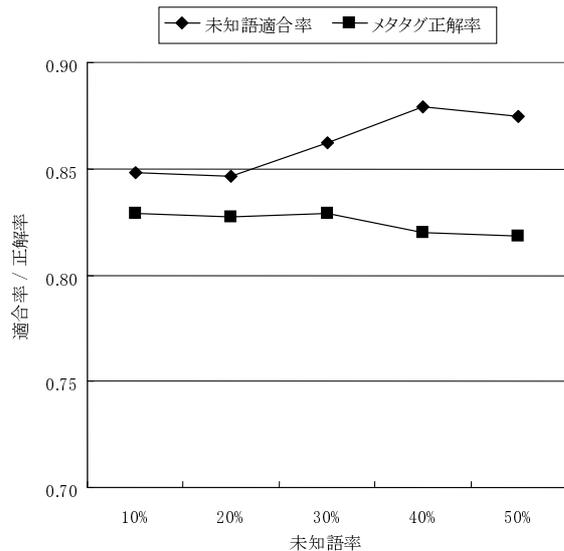


図 4 未知語率とメタタグ識別性能の関係

Fig. 4 Relations between out-of-vocabulary rate vs meta-tag discriminating performance.

別ができるものと考えられる。

5. まとめ

本稿では、ブースティングを用いて野球実況中継に対してメタデータを作成する手法について述べた。ブースティングにより得られた素性に基づいて重み付き投票を行うことにより、87%の割合で正しくメタタグ情報を識別することが可能であった。特に、メタデータとして重要な投球タグについては、F 値 0.96 と高い精度での識別が可能であった。

AdaBoost の学習は、音声認識結果を用いて行うことにより、認識率が低下した場合でもメタタグ情報の識別性能を維持することが可能であった。また、未知語率を 50%まで変化させた場合でも、未知語なしの場合とほとんど変わらない精度でメタタグ情報を識別することが可能であった。認識誤りや未知語に対して、頑健にメタタグ情報を識別できるものと考えられる。

一方で、実況イベントについては F 値 0.71 と精度の低下が見られた。これは、音声認識が、イベント時の実況のような興奮した音声に対応できなかったためと考えられる。今後の課題として、音響情報との統合によりメタデータを付与すること、推定したメタタグ情報を音声認識にフィードバックすること、興奮した場合の音声に対応することがあげられる。

文 献

- [1] 山本拓, 佐藤宏介, 千原國宏, “野球中継映像における各種プレイシーンの自動検索/編集システム,” 2000 信学総大, 情報・システム 2, D12-77, p.247, 2000.
- [2] 館山公一, 川嶋稔夫, 青木由直, “野球中継におけるシーン検索,” 第 3 回知能情報メディアシンポジウム論文集,

- pp.195-202, 1997.
- [3] 新田直子, 馬場口登, 北橋忠広, “言語の画像の情報統合によるスポーツ映像からの人物・アクション・イベント抽出,” 信学技報, PRMU99-256, 2000.
- [4] 佐古 淳, 有木 康雄, “知識を用いた音声認識による野球実況中継の構造化,” 第 6 回音声言語シンポジウム, SP2004-136, pp.85-90, 2004-12.
- [5] R.Schapire, Y.Freund, P.Bartlett, and W.Lee, “Boosting the margin: A new explanation for the effectiveness of voting methods,” Annals of Statistics, vol.26, no.5, pp.1651-1686, Oct. 1998.
- [6] Y.Freund and R.Schapire, “Experiments with a new Boosting algorithm,” Proc. 13th International Conference on Machine Learning Bari, Italy Morgan Kaufmann, pp.148-146, July 1996.
- [7] H.Schwenk and Y.Bengio, “Adaboosting neural networks,” Proc. ICANN'97, vol.1327 of LNCS Berlin Springer, pp.967-972, Oct. 1997.
- [8] G.Ratsch, T.Onoda, and K.-R. Muller, “Soft Margin for AdaBoost,” Machine Learning, vol.42, no.3, pp.287-320, March 2001.
- [9] 小野田崇, “Boosting の過学習とその回避,” 電子情報通信学会論文誌, Vol.J85-D2, No.5, pp. 776-784, 2002 年 5 月.
- [10] 工藤拓 / 松本裕治, “半構造化テキストの分類のためのブースティングアルゴリズム,” 情報処理学会論文誌, Vol.45, NO.9, 2004 年 9 月.
- [11] 河原達也, 石塚健太郎, 堂下修司, “発話検証に基づく音声操作プロジェクトとそれによる講演の自動ハイパーテキスト化,” 情報処理学会論文誌, Vol.40, No.4, pp.1491-1498, 1999.
- [12] “MeCab: Yet Another Part-of-Speech and Morphological Analyzer,” <http://chasen.org/taku/software/mecab/>
- [13] 古井貞熙, 前川喜久雄, 伊佐原均, “『話し言葉工学』プロジェクトのこれまでの成果と展望,” 第 2 回話し言葉の科学と工学ワークショップ, pp.1-6, 2002.
- [14] 緒方淳, 有木康雄, “音素事後確率に基づく信頼度を用いた音響モデルの教師なし適応,” 信学技報, SP2001-105, 2001.