

論文 / 著書情報
Article / Book Information

論題(和文)	講演シーン検索における検索語出現状況に基づくレーザーポインタ情報のフィルタリング
Title(English)	Filtering the Laser Pointer Information based on Keyword Occurrence in Presentation Scene Retrieval
著者(和文)	仲野 亘, 小林 隆志, 勝山 裕, 直井 聡, 横田 治夫
Authors(English)	Wataru Nakano, Takashi KOBAYASHI, Yutaka KATSUYAMA, Satoshi NAOI, Haruo YOKOTA
出典(和文)	DBSJ Letters, Vol. 5, No. 2, pp. 1-4
Citation(English)	DBSJ Letters, Vol. 5, No. 2, pp. 1-4
発行日 / Pub. date	2006, 9
権利情報 / Copyright	本著作物の著作権は日本データベース学会に帰属します。 Copyright (c) 2006 Database Society of Japan, DBSJ.

講演シーン検索における検索語出現状況に基づくレーザーポインタ情報のフィルタリング

Filtering the Laser Pointer Information based on Keyword Occurrence in Presentation Scene Retrieval

仲野 亘[♡] 小林 隆志[◇] 勝山 裕[▲]
直井 聡[▲] 横田 治夫[◇]

Wataru NAKANO Takashi KOBAYASHI
Yutaka KATSUYAMA Satoshi NAOI
Haruo YOKOTA

我々はこれまで、講義・講演における資料とその撮影動画をメタデータによる統合コンテンツとして蓄積し、その特性を利用して高度なシーン検索を提供する UPRISE を提案してきた。本稿では、講師が用いるレーザーポインタの認識結果から、スライド中の文字列を強調する目的以外で行われたレーザーポインタ照射の情報をフィルタリングすることで、それらが検索に与える影響を軽減させる手法を提案する。スライド中、音声中、およびその両方の検索語出現状況を考慮する 3 種のフィルタリング手法を提案し、実際の講義コンテンツを用いた評価実験により提案手法の有用性を示す。

We have proposed a search mechanism named UPRISE (Unified Presentation Slide Retrieval by Impression Search Engine) for unifying presentation contents which consist of lecture videos and presentation slides. In this paper, we propose a technique for reducing the impacts on irrelevant pointing in automatic recognized pointer information by filtering out information not for emphasizing words in slides. We propose three methods to filter the laser pointer information based on keyword occurrence in slides, in speech and both. We evaluate our approach by applying it to actual presentation contents.

1. はじめに

近年、動画や文書、音声ストリームなどの複数のメディアによるコンテンツを統合し、それらを蓄積、検索するシステムが数多く研究、および提案されており [1, 2, 3, 4], e-Learning など、様々な用途に用いられている。特に e-Learning 用のコンテンツに対しては、利用者が必要とするコンテンツを検索できるだけでなく、コンテンツのどの箇所から視聴すべきかを効果的に発見できることが重要である。

そのような検索を実現するために、我々は教育コンテンツの統合、蓄積、および統合コンテンツに対する高度な検索機能を実現するシステムである UPRISE(Unified Presentation Slide Retrieval by Impression Search Engine) を提案してきた [5]。

UPRISE では、動画ストリームを資料スライドの切り替えタイミングによってシーンという単位に分割し、各シーンとそこで使用された資料スライドを対応付けることでそれらを統合す

る。また、各シーンに対して、対応する資料スライドの文字・構造情報、シーンの時間長の情報などから検索用インデックスを作成することで、高度な検索を可能としている。従来用いられてきたスライド中の文書検索ではなく、スライドの切り替えタイミングによってシーンを分割し、それらを検索の単位とすることで、動画中で講師がバックトラックをしたり、巻き戻りがあったりすることなどにより複数のシーンで同一のスライドが使用されている場合でも、それらを異なるシーンとして区別することができるという利点がある。

我々はこの UPRISE の検索手法において、講師が用いるレーザーポインタなどのポインティング情報に着目し、ポインティング情報を統合することで検索精度を向上させる検索手法を提案してきた [6]。しかし、レーザーポインタの照射には講師の手のぶれや、複数概念間の関連を軌跡で示すもの、指示対象を明確に意図しないあいまいなものなど、スライド中の文字列を強調する目的で行われるもの以外の照射も多く見られるが、従来の手法では、これらの検索に考慮するべきではない照射も様にレーザーポインタ情報として扱い、シーン検索に利用してきた。

そこで本稿では、そのような検索に考慮するべきではないレーザーポインタ情報を他の情報を用いて除去し、スライド中の文章を強調する目的で使用されたレーザーポインタだけを選び取ることで、検索の精度向上を実現する手法を提案する。提案手法では、複数キーワードのスライド中の出現の有無や、講師の発話した音声中の出現の有無を考慮することで、各レーザーポインタ照射のフィルタリングを行う。また、それらのフィルタリングしたレーザーポインタ情報を従来の検索手法に統合した手法を提案する。

2. UPRISE の概要

UPRISE では、メタデータを用いてコンテンツを管理、統合する。メタデータには、動画のどの時刻にスライドの切り替えが起こったかというシーン情報と、その際にどのスライドを用いていたかという同期情報、スライドに含まれる文字列情報に対するインデックスを含める。これらの情報を保持するメタデータによってコンテンツを緩く結合することにより、個々のコンテンツが持つ情報に修正を加えることなくコンテンツの同期表示を実現し、柔軟な統合を可能にしている。UPRISE のシステムの詳細については [7] を参照されたい。

UPRISE では、動画中に同じスライドが複数回出現する場合にそれらを異なるシーンとして区別し、個別に適合度を算出する。これにより、それぞれのプレゼンテーションは対応する動画のシーンの集合として抽象化され、プレゼンテーション中の任意のシーンが検索可能になる。

以下では、検索に用いる基本的な適合度と、音声情報を統合した適合度について説明する。

2.1 基本となる適合度 I_c

UPRISE の検索機能は、検索キーワードに対する適合度をシーンごとに算出し、上位のシーンから表示する。この適合度のうち、最も基本的なものは適合度 I_c である。 I_c はスライドの文書構造、シーンの時間の長さ、前後シーンの文脈の 3 種類の情報を元に、以下の式により算出される。

$$I_c(s, k, \theta, \delta, \varepsilon_1, \varepsilon_2) = \sum_{\gamma=s-\delta}^{s+\delta} E(\gamma - s, \varepsilon_1, \varepsilon_2) \cdot T(\gamma)^\theta \cdot I_p(\gamma, k)$$

ここで、 δ は考慮する前後シーンの範囲を定めるパラメタであり、 $E(x, \varepsilon_1, \varepsilon_2)$ は前後関係の強弱を定める関数である。 $E(x, \varepsilon_1, \varepsilon_2)$ は以下のように定義される。

$$E(x, \varepsilon_1, \varepsilon_2) = \begin{cases} \exp(\varepsilon_1 x) & (x < 0) \\ \exp(-\varepsilon_2 x) & (x \geq 0) \end{cases}$$

適合度 I_c では、シーンの適合度はその前後 δ の範囲から影響を受け、 ε が小さいほど影響を受けやすくなる。なお、この I_c

♡ 学生会員 東京工業大学 大学院 情報理工学専攻 計算工学専攻
wnakano@de.cs.titech.ac.jp

◇ 正会員 東京工業大学 学術国際情報センター
{tkobaya,yokota}@cs.titech.ac.jp

▲ 株式会社 富士通研究所
{katsuyama,naoi.satoshi}@jp.fujitsu.com

のパラメータ群 $(s, k, \theta, \delta, \varepsilon_1, \varepsilon_2)$ は UPRISE における適合度関数の基本パラメータ群であるため、本稿では以降これを Φ として簡略化表現する。

2.2 音声情報を考慮した適合度

講義における講師の発話内容は、スライド中の文字列情報と同様に直接的にそのシーンの内容を表している。さらに、同一スライドを用いたシーンにおいても、発話内容によってそのシーンの差別化を行うことができる。そこで我々は音声情報を利用した適合度の提案を行ってきた [8]。

[8] では、まず音声認識によって講義中の音声情報を抽出する。そして、あるシーン s 中でキーワード k が発話された回数を $skc(s, k)$ (Spoken Keyword Count) とした。この $skc(s, k)$ を適合度 I_c と統合することにより、音声情報を利用した適合度を提案した。

3. レーザーポインタ情報を考慮した適合度

講師は、主に資料スライド中のある部分を強調する目的でレーザーポインタを用いる。そこで、我々は検索キーワードに対してレーザーポインタが多く当たっているシーンはそのキーワードに対してより適切であると考え、それらのレーザーポインタ情報を考慮した適合度を提案してきた [6]。本節では、レーザーポインタ情報の抽出手法と、レーザーポインタ情報を統合した適合度について説明する。

3.1 レーザーポインタ情報の抽出

撮影動画より画像認識によって抽出したポインタの光点座標 [9] に対して、スライド上で最も近い行の文字列を取得する。これは、レーザーポインタの照射が必ずしも目標の単語が最も近くなるような座標に対して行われるわけではなく、単語単位で取得するのではうまくキーワードと関連付けられないと考えたためである。

次に、同じ行の文字列を取得した連続の光点を 1 回のレーザーポインタ照射と定義し、一つのレーザーポインタ情報として統合する。ここで、レーザーポインタはある 1 行に対して正確に当て続けることが容易でないため、レーザーポインタの光点は対象行から外れてしまうことが多い。そこで、一回のレーザーポインタに対し、近傍の数行を次候補として取得しておく。最も近い行とその付近のいくつかの行を組にしてレーザーポインタ情報とすることで、講師の意図と光点とのずれをある程度解消することができる。

また、この 1 回のレーザーポインタに相当する部分をサブシーンと定義する。したがって、各シーンはより詳細な内容ごとにまとまったサブシーンを複数持つことになり、各サブシーンはある特定のシーンに属することになる。

3.2 レーザーポインタ情報と I_c の統合

3.1 節において抽出したレーザーポインタ情報を、適合度 I_c に統合する。まず、レーザーポインタは確実に講師の意図どおりに当たるわけではなく、抽出した情報も誤差を含む。この問題を解消するため、対象行以外に候補となる行を持つということはすでに述べた。そこで、レーザーポインタの照射回数はその情報の信頼度を考慮し、キーワードが全候補行に含まれていたときに 1 とするような、回数の期待値 $H(l, q)$ として数値化する。 $H(l, q)$ はサブシーン q のレーザーポインタが、行 l に照射された回数の期待値であり、全ての行の $H(l, q)$ を合計すると 1 となる。

こうして得られたレーザーポインタごとの照射回数期待値に対し、各レーザーポインタが当たっていた時間を掛け合わせることで、レーザーポインタの照射時間の期待値が得られる。この照射時間の期待値をシーンごとに合計したものを、 $phd(s, k)$ (Pointer Hit Duration) とする。ただし、 s はシーン、 k はキーワードである。シーン s 、キーワード k における $phd(s, k)$ の式

を以下のように定義する。

$$phd(s, k) = \sum_{q_i \in s} \sum_{l=1}^{L(s)} H(l, q_i) \cdot T(q_i)$$

ここで、 $T(q_i)$ はサブシーン q_i の時間を表す。すなわち、サブシーン q_i に対応するレーザーポインタの照射時間を表す。

この phd を適合度 I_c と統合することにより、レーザーポインタ情報を利用した適合度を提案してきた。これまでの報告、実験では、シーンごとの時間情報である $T(s)$ に対し、レーザーポインタの時間の期待値である $phd(s, k)$ を足し合わせ、レーザーポインタが当たっていたときにそのシーンの時間に加算するという統合手法が最も有効であるということがわかっていく。この適合度を $I_{c[d+phd]}$ とし、以下の式で定義する。

$$I_{c[d+phd]}(\Phi, \omega_d) = \sum_{\gamma=s-\delta}^{s+\delta} E(\gamma - s, \varepsilon_1, \varepsilon_2) \cdot \{T(\gamma) + \omega_d \cdot phd(\gamma, k)\}^\theta \cdot I_p(\gamma, k)$$

ここで、 ω_d は phd の適合度計算における影響度合いを調節するパラメータである。例えば、 $\omega_d = 10$ の場合、レーザーポインタが 1 秒当たるとはそのシーンが 10 秒伸びることに相当する。

4. レーザーポインタ情報のフィルタリング

3. 節で述べたように、従来の手法では、レーザーポインタ情報は抽出した情報全てを同じ条件で利用していた。そのため、スライド中文字列の直接的な強調という目的以外で行われた照射も適合度の計算に用いていた。そこで、抽出したレーザーポインタ照射の中から、そのような強調目的以外の照射を排除するために、 phd を別の情報を用いてフィルタリングする。以下では、複数キーワードのスライド中の出現条件によるフィルタリングと、講義の音声情報を利用したフィルタリングの 2 種類の手法を提案する。また、それら 2 種類の条件を同時に考慮したフィルタリングについても提案を行う。

4.1 複数キーワードのスライド中の出現条件によるフィルタリング

UPRISE では、検索語を形態素解析して形態素ごとに適合度を算出し、その和を求めることで検索語の適合度を求める。この形態素解析によって複数に分割される検索語を用いた検索や、複数のキーワードによる AND 検索など、複数の単語を用いた検索を行う機会が多い。このとき、従来の適合度算出手法では、分割された単語それぞれへのレーザーポインタの影響を考慮していたため、スライド中に複数の単語全てが出現していなくても、レーザーポインタにより適合度が加算されていた。

しかし、キーワード全てを含まない行に向けてのレーザーポインタ照射は、その検索キーワードに対しての強調の意味を持たないと考えられる。そこで、複数の単語による検索の際に、その単語全てがスライド上になければ、つまり、検索キーワード中の単語 k のうち、1 つでも $I_p(s, k) = 0$ となるものがあつた場合、そのシーンにおけるキーワードへのレーザーポインタ照射は無視する、という手法を提案する。この条件でフィルタリングを行った $phd(s, k)$ を、 $phd|_p(s, k)$ とする。

また、この $phd|_p(s, k)$ を、3.2 節で述べた適合度 $I_{c[d+phd]}$ に統合する。 $I_{c[d+phd]}$ の $phd(s, k)$ を $phd|_p(s, k)$ に置き換えたものを $I_{c[d+phd|_p]}$ と定義する。

4.2 音声情報を用いたフィルタリング

検索キーワードに対してレーザーポインタが当たっているにもかかわらず、講師がそのシーンの中でキーワードを発話していない場合、そのレーザーポインタ照射は検索キーワードを説明する補助として用いられたのではないと考える。そこで、あるシーン中でキーワードが発話されていない場合、つまりそのシーンの音声中出现回数である $skc(s, k) = 0$ の場合には、その

シーンにおけるキーワードへのレーザーポインタ照射は無視するという手法を提案する．この条件でフィルタリングを行った $phd(s, k)$ を, $phd_{js}(s, k)$ として定義する．

この phd_{js} を適合度 $I_{c[d+phd]}$ に統合する．さらに, [8] で行ったように, skc による適合度の加算分も考慮する．この適合度を $I_{c[d+phd_{js}, p+skc_{jp}]}$ とし, 以下の式で表す．

$$\begin{aligned}
 & I_{c[d+phd_{js}, p+skc_{jp}]}(\Phi, \omega_d, \psi) \\
 &= \sum_{\gamma=s-\delta}^{s+\delta} E(\gamma - s, \varepsilon_1, \varepsilon_2) \cdot I_{d[d+phd_{js}, p+skc_{jp}]}(\gamma, k, \theta, \omega_d, \psi) \\
 & I_{d[d+phd_{js}, p+skc_{jp}]}(s, k, \theta, \omega_d, \psi) \\
 &= \{T(s) + \omega_d \cdot phd_{js}(s, k)\}^\theta \cdot \{I_p(s, k) + \psi \cdot skc_{jp}(s, k)\}
 \end{aligned}$$

ここで, ψ は skc が適合度に与える影響度合いを調節するパラメタである．また, $skc_{jp}(s, k)$ はシーン s で用いられるスライドの文字列中にキーワード k が出現する場合のみ $skc(s, k)$ を計算するという関数である．

4.3 2条件を同時に考慮したフィルタリング

4.1, 4.2 節で用いた条件はそれぞれが独立である．そこで, 両者の条件を満たす場合にのみ, つまり全ての単語 k で $I_p(s, k) \neq 0$ であり, かつ $skc(s, k) \neq 0$ のときのみレーザーポインタ情報を考慮することを考える．この条件でフィルタリングを行った $phd(s, k)$ を, $phd_{jps}(s, k)$ として定義する．

$I_{c[d+phd_{jps}, p+skc_{jp}]}$ における $phd_{js}(s, k)$ を $phd_{jps}(s, k)$ に置き換えたものを, 適合度 $I_{c[d+phd_{jps}, p+skc_{jp}]}$ とする．

5. 実験

実際の講義のコンテンツを UPRISE に登録し, 登録したコンテンツに対して各適合度ごとの検索実験を行った．以下ではその実験に関して説明し, 実験結果に対して考察を行う．

5.1 実験に用いたデータ

実験では, データベースについての講義(全 11 回), 計算機アーキテクチャについての講義(全 12 回)をコンテンツ化し, 検索対象とした．

講義の音声情報の抽出には, 連続音声認識ソフトウェア Julius¹ を用い, 言語モデルと音響モデルとして, 山崎が日本語話し言葉コーパス (CSJ)[10] の学会講演と模擬講演データから, 作成した言語モデルと音響モデル [11] を用いた．言語モデル作成時の形態素解析には, 茶筌, 形態素解析用の辞書として, ipadic を用いている．また, 認識用の辞書として, 学習データ(約 300 万語)での出現頻度が高い順から選んだ 22,860 語を登録し, 資料スライド中から辞書に含まれていない単語を追加したものを講義ごとに作成し, 音声認識に利用した．

5.2 実験

提案手法の評価を行うため, 5.1 節で登録したコンテンツに対し, キーワードについて説明しているシーンを実際に検索する実験を以下の条件の下で行った．

- 基本となるパラメタ Φ は $\theta = 0.4, \delta = 4, \varepsilon_1 = 5.0, \varepsilon_2 = 0.5$ とし, skc の影響の強弱を表すパラメタ ψ は 1 に固定した．
- レーザーポインタの各光点に対し 5 つの候補行を取得し, 照射回数期待値 $H(l, q)$ を第 1 候補から順に 0.4, 0.3, 0.15, 0.10, 0.05 という値に設定した．
- 各適合度ごとに 124 種類のキーワードを検索した．
- 各適合度に対して, phd の影響の強弱を表すパラメタ ω_d を 1 から 30 まで 5 刻みに変更し, 計 7 回の計測を行った．
- キーワードに対して最もよく解説していると判断したシーンをそのキーワードの正解シーンとした．

評価に際しては, 今回の実験では正解シーンを各キーワードに対して 1 つとしていることから, 平均逆数順位 (Mean reciprocal

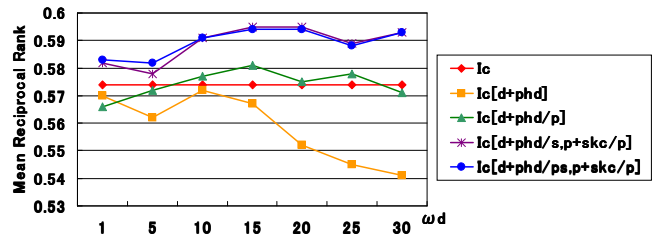


図 1: ω_d を変化させたときの適合度ごとの MRR
Fig.1 MRR of the two Lectures

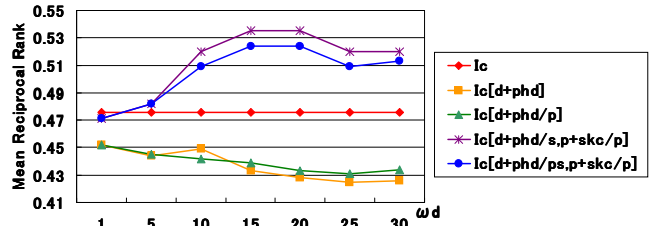


図 2: データベースに関する講義における MRR
Fig.2 MRR of the Lecture about Database

rank:MRR) を用いた [12] . MRR は質問ごとに最初に出現した正解の順位の逆数を求め, それらを全質問にわたって平均することで定義される．

5.3 実験結果と考察

提案手法である $I_{c[d+phd_{jp}]}$, $I_{c[d+phd_{js}, p+skc_{jp}]}$, $I_{c[d+phd_{jps}, p+skc_{jp}]}$ の 3 種類の適合度と, 従来のフィルタリングを行わない適合度である I_c , $I_{c[d+phd]}$ の比較を行った．

図 1 は, 2 講義の検索キーワードを用いた検索による, 各適合度の MRR の変化を示したグラフである．グラフより, 提案手法である phd をフィルタリングして用いている適合度は全体的に検索精度が向上していることがわかる．特に, 従来の $I_{c[d+phd]}$ ではパラメタ ω_d を大きくした場合に検索精度が下がるという傾向があったが, フィルタリングした phd を用いた適合度では, 少なくとも今回の ω_d の範囲では十分な精度を保っている．このことは, これまで ω_d を大きくすると不要なレーザーポインタ照射による適合度計算への影響が顕著になり, 精度を下げていたが, それらをフィルタリングしたために, ω_d を大きくしても精度を維持できている, ということであると考える．

図 2, 3 はそれぞれ, 計算機アーキテクチャに関する講義とデータベースに関する講義のそれぞれで検索を行い, 各適合度ごとに MRR を算出したグラフである．図 2 が示すように, 従来の手法では, データベースに関する講義はレーザーポインタ情報を利用することで検索精度が下がってしまっていた．これは, 基本適合度である I_c の値が計算機アーキテクチャに関する講義と比較して低めであることから考えて, この講義が, スライド文字列情報に基づいたランク付けにはあまり適していな

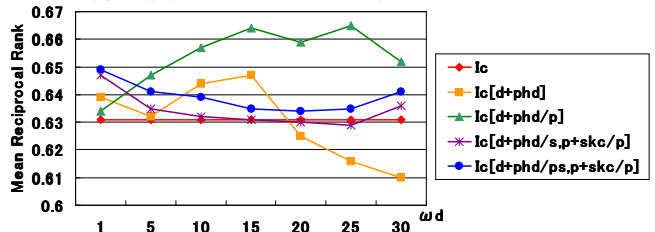


図 3: 計算機アーキテクチャに関する講義における MRR
Fig.3 MRR of the Lecture about Computer Architecture

¹<http://julius.sourceforge.jp/>

いスライド構造, および講義のトピック構造をしていたためである。しかし, レーザーポインタ情報を音声情報でフィルタリングすることにより, レーザーポインタ情報の影響度を高めると, 検索精度が大幅に向上するという結果を得た。このことから, この講義では特に, 検索キーワードの強調以外の目的で行われたレーザーポインタ照射を, 音声情報によって除去出来ているということがわかる。

また, 計算機アーキテクチャに関する講義では, 図3が示す通り, 従来の手法でもレーザーポインタ情報は検索に良い影響を与えており, それに複数キーワードのスライド上での出現の有無によるフィルタリングを行うことで, 全適合度の中で最も良い精度を達成している。このことは, この講義がデータベースに関する講義とは特性が異なり, スライド構造を基本として適合度を算出する手法に適した講義だということを示している。

このように, 2つの講義だけを取り上げてみても, その講義ごとに特性は様々である。これらは講義形態も若干異なっており, 例えば, データベースに関する講義はしばしば演習を行う形式であるが, 計算機アーキテクチャに関する講義では演習を全く行っていない。このような講義形態による特性の分析や, それに基づいた適合度の選択やパラメタの調節などを行うことが必要であり, それらは今後の課題である。

6. まとめと今後の課題

本稿ではUPRISEの検索精度を向上させるために, レーザーポインタ情報を複数キーワードのスライド中における出現の有無と, キーワードの音声情報中における出現の有無の2種類の情報を用いてフィルタリングし, 有用なレーザーポインタ情報だけを検索に利用する手法を提案した。さらに実際の講義をコンテンツ化してUPRISEに登録し, 検索実験を行った。

実験の結果, フィルタリングした *phd* を用いた適合度は従来の適合度に比べて精度が向上していることを確認した。また, 講義ごとにMRRの推移を算出し, 講義ごとの特性の違いを考察した。講義ごとの特性をより詳細に調べるためには, 今後さらに異なる講義や話者のコンテンツを蓄積し, 様々な講義に対して実験を行うことが必要であることがわかった。

本研究のその他の今後の課題としては, まず, 今回の提案では音声情報をレーザーポインタ情報のフィルタリングに利用する形で統合したが, レーザーポインタ情報と音声情報をより関連付け, 一つの適合度計算手法としてまとめることでより高度な検索が可能になると考えている。

次に, 今回の実験では音声情報として, 音声認識によって抽出した情報を用いた。音声認識の際の誤りや, 音声認識に用いた辞書とスライド上の単語間での表記ゆれ, 辞書中の英単語の読み付けができていないことなどを解決することで, より精度の高い音声情報をフィルタリングに用いることが出来ると考える。

さらに, 提案した適合度では音声情報やスライド文字列中でのキーワードの特定性を考慮していない。これらの特定性をシーン検索において考慮することは有効であり [8, 13], 今回の提案手法においてもこれらの特定性を考慮することでさらに精度が向上すると考える。

【謝辞】

本研究で用いたJuliusと音響, 言語モデルの使用にあたりご協力頂いた, 東京工業大学大学院情報理工学専攻計算工学専攻の古井貞熙教授, 岩野公司助手, 山崎裕紀氏に感謝致します。なお, 本研究の一部は, 文部科学省科学研究費補助金特定領域研究(15017233,16016232,18049026), 独立行政法人科学技術振興機構CREST, および東京工業大学21世紀COEプログラム「大規模知識資源の体系化と活用基盤構築」の助成により行われた。

【文献】

- [1] R. Müller and T. Ottmann. The "Authoring on the Fly" system for automated recording and replay of

(tele)presentations. *Multimedia Syst.*, Vol. 8, No. 3, pp. 158–176, 2000.

- [2] Carnegie Mellon University The Informedia Project. Informedia ii digital video library. <http://www.informedia.cs.cmu.edu/>.
- [3] G. D. Abowd. Classroom 2000: an experiment with the instrumentation of a living educational environment. *IBM Syst. J.*, Vol. 38, No. 4, pp. 508–530, 1999.
- [4] A. Fujii, K. Itou, and T. Ishikawa. Lodem: A system for on-demand video lectures. *Speech Communication*, Vol. 48, No. 5, pp. 516–531, 2006.
- [5] H. Yokota, T. Kobayashi, T. Muraki, and S. Naoi. UPRISE: Unified Presentation Slide Retrieval by Impression Search Engine. *IEICE Trans. on Info. and Syst.*, Vol. E87-D, No. 2, pp. 397–406, 2004.
- [6] W. Nakano, Y. Ochi, T. Kobayashi, Y. Katsuyama, S. Naoi, and H. Yokota. Unified Presentation Contents Retrieval Using Laser Pointer Information. In *Proc. of SWOD*, pp. 170–173, 4 2005.
- [7] 小林隆志, 村木太一, 直井聡, 横田治夫. 統合プレゼンテーションコンテンツ蓄積検索システムの試作. 電子情報通信学会論文誌, Vol. J88-D-I, No. 3, pp. 715–726, 3 2005.
- [8] 岡本拓明, 仲野亘, 小林隆志, 直井聡, 横田治夫, 岩野公司, 古井貞熙. プレゼンテーション蓄積検索システムにおける講義・講演音声情報を利用した適合度の改善. データ工学ワークショップ, pp. DEWS2006-6C-o1, Mar 2006.
- [9] Y. Katsuyama, N. Ozawa, J. Sun, H. Takebe, T. Kobayashi, H. Yokota, and S. Naoi. A New Solution for Extracting Laser Pointer Information from Lecture Videos. In *Proc. of E-learn2004*, pp. 2713–2718, 10 2004.
- [10] 国立国語研究所. 日本語話し言葉コーパス. <http://www2.kokken.go.jp/~csj/public/>.
- [11] 山崎裕紀. 講義音声認識の高精度化に関する研究. 東京工業大学工学部卒業論文, 2 2005.
- [12] 酒井哲也. よりよい検索システム実現のために. 情報処理, Vol. 47, No. 2, pp. 147–158, 2006.
- [13] H. Yokota, T. Kobayashi, H. Okamoto, and W. Nakano. Unified contents retrieval from an academic repository. In *Proc. of International Symposium on Large-scale Knowledge Resources LKR2006*, pp. 41–46, 2006.

仲野 亘 Wataru NAKANO

平 17 東工大・工・情報工卒. 同大学院・情報理工・計算工・修士課程在学中. 日本データベース学会学生会員.

小林 隆志 Takashi KOBAYASHI

平 9 東工大・工・情報工学卒. 平 11 同大学院・情報理工・計算工学・修士課程了. 平 16 同大学院・同専攻・博士課程了. 平 14 同大学術国際情報センター・助手. 工博. 日本データベース学会, 日本ソフトウェア科学会, 情報処理学会, ACM 各会員.

勝山 裕 Yutaka KATSUYAMA

昭 59 東北大・工・通信工卒. 昭 61 同大学院・工・情報工・修士課程了. 同年富士通(株)入社. 電子情報通信学会員.

直井 聡 Satoshi NAOI

昭 58 慶大・工・電工卒. 昭 60 同大学院・工・電気工学・修士課程了. 同年(株)富士通研究所入社. 平 13 東工大・学術国際情報センター・客員助教授. 平 17 同センター・客員教授. 工博. 電子情報通信学会, 情報処理学会 各会員.

横田 治夫 Haruo YOKOTA

昭 55 東工大・工・電物卒. 昭 57 同大学院・情報・修士課程了. 同年富士通(株). 同年6月(財)新世代コンピュータ技術開発機構研究所. 昭 61 (株)富士通研究所. 平 4 北陸先端大・情報・助教授. 平 10 東工大・情報理工・助教授. 平 13 東工大・学術国際情報センター・教授. 工博. 日本データベース学会理事, 電子情報通信学会, 情報処理学会, 人工知能学会, IEEE, ACM 各会員.