

## 音声ドキュメント処理ワーキンググループ 活動報告

相川清明<sup>1</sup>, 秋葉友良<sup>2</sup>, 伊藤慶明<sup>3</sup>, 河原達也<sup>4</sup>,  
中川聖一<sup>2</sup>, 南條浩輝<sup>5</sup>, 西崎博光<sup>6</sup>, 胡新輝<sup>7</sup>,  
松井知子<sup>8</sup>, 山下洋一<sup>9</sup>

情報処理学会音声言語情報処理研究会音声ドキュメント処理ワーキンググループの活動のうち、2009年度から2011年度に亘って活動した第2期の活動を中心に報告する。第2期ではInterspeech2010スペシャルセッション、NTCIR-9、情報処理学会論文誌特集号を企画した。本報告では、それらの企画プロセスを中心に述べる。

### Progress Report of Spoken Document Processing Working Group

Kiyoaki Aikawa<sup>1</sup>, Tomoyoshi Akiba<sup>2</sup>, Yoshiaki Ito<sup>3</sup>,  
Tatsuya Kawahara<sup>4</sup>, Seiichi Nakagawa<sup>2</sup>, Hiroaki Nanjo<sup>5</sup>,  
Hiromitsu Nishizaki<sup>6</sup>, Xinhui Hu<sup>7</sup>, Tomoko Matsui<sup>8</sup>,  
Yoichi Yamashita<sup>9</sup>

This report describes the activities of SLP Spoken Document Processing Working Group (SDPWG). The SDPWG was organized in 2006. The working group was reorganized in 2009. This report mainly describes the activities of the second period of the SDPWG. The SDPWG organized Interspeech2010 special session, NTCIR-9 workshop, and the special issue on spoken document processing of IPSJ journal. This report focuses the planning process of these activities.

#### 1. はじめに

音声ドキュメント処理ワーキンググループは音声言語情報処理研究の活性化を図るため2006年4月に結成され、3年ずつ2期に亘って活動を行った。2011年12月の現時点は2期目の最終年度である。また、第2期ワーキンググループのメンバーは本報告の著者の10名である。

音声言語情報処理研究の1つの有望な課題として音声ドキュメント検索があった。音声ドキュメント検索とは、テキストによる書き起こしが無く、ラベルやタグ情報では十分な検索ができない音声データを音声認識によってテキスト化して検索するものである。音声認識は精度が急速に向上してきたものの、まだ十分とは言えない。一方、パソコン内部の個人のファイルや外部の様々なメディアには音声を含む各種講演や対話、講義録や動画データが爆発的に増加し、その検索が望まれていた。

情報検索研究の活性化のため、米国ではTREC (Text REtrieval Conference)、国内ではNTCIR (NII Test Collection for IR Systems) において様々な評価セットが構築されてきた。音声ドキュメント検索はTRECにおいて既に実施されたことがあるが、日本語版音声ドキュメント検索テストコレクションの構築が望まれていた。

音声認識誤りの存在下で目的のドキュメントを検索する課題は、実用的であり、かつ、音声言語情報処理研究に新たな研究課題を提供するものと考えられ、本ワーキンググループが組織された。本報告は主として2009年4月に開始された2期目の活動について報告する。以後音声ドキュメント処理ワーキンググループのことを(SDPWG: Spoken Document Processing Working Group) と表記する。

なお、本報告において音声ドキュメント処理関係論文あるいは講演論文の投稿件数分析に計上したものは、明示的に音声ドキュメント処理の課題を対象としたものに限っており、音声入力による情報検索、特に検索を目的としない講演音声認識、対話システム、QAシステムは除外した。

<sup>1</sup> 東京工科大学 Tokyo University of Technology

<sup>2</sup> 豊橋技科大学 Toyohashi University of Technology

<sup>3</sup> 岩手県立大学 Iwate Prefectural University

<sup>4</sup> 京都大学 Kyoto University

<sup>5</sup> 龍谷大学 Ryukoku University

<sup>6</sup> 山梨大学 University of Yamanashi

<sup>7</sup> 情報通信研究機構 National Institute of Information and Communication Technology

<sup>8</sup> 統計数理研究所 The Institute of Statistical Mathematics

<sup>9</sup> 立命館大学 Ritsumeikan University

## 2. SDPWG 活動記録

### SDPWG 設立まで

2005年5月27日 第1回 SLP 研究運営委員会  
新しいWGの立ち上げについて議論が行われた。音声情報検索関係のWGの必要性が議論され、立ち上げる方向で進めることになった。

2006年2月3日 第3回 SLP 研究運営委員会  
音声情報検索に関する新WG創設

### 第1期 SDPWG

2006年5月11日 第1回会合  
データベース検討

2006年7月20日 講義データ（音声・ビデオ）収録アンケート依頼配布

2006年8月4日 第2回会合  
講義データ（音声・ビデオ）収録アンケート結果→利用難いのでCSJ利用  
CSJ認識結果分析&クエリ試作

2006年9月7日 検索クエリ収集と適合性判定作業開始

2006年10月20日 第3回会合  
テストコレクション設計方針決定  
音声ドキュメント処理ワークショップ映像収録決定

2006年11月17日 第4回会合  
検索クエリ案の問題分析

2006年12月21-2日 音声言語シンポジウムにてテストコレクション発表 1)

2007年2月26-27日 第1回音声ドキュメント処理ワークショップ 2)

2007年5月25日 第5回会合

テストコレクション作成に向けて収集されたクエリと評価尺度の検討

2007年9月19-21日 日本音響学会秋季研究発表会音声Aスペシャルセッション〔音声ドキュメント処理〕にて中川が招待講演 3)

2007年10月19日 第6回会合  
SDR タスク検討: 検索単位、評価尺度  
STD タスク設計

2007年8月20日 情報処理学会に論文誌特集号提案

2008年2月29日-3月1日 第2回音声ドキュメント処理ワークショップ

2008年5月28-30日 LREC2008で発表 4)

2008年12月9-10日 音声言語シンポジウムにてSDPWG活動報告 5)

2009年2月 情報処理学会「音声ドキュメント処理」特集号発行 6)

2009年2月27-28日 第3回音声ドキュメント処理ワークショップ

### 第2期 SDPWG

2009年2月7日 第1回会合  
2期WG設立準備

2009年3月13日 SDPWGホームページ開設 7)

2009年3月17日 第2回会合  
CSJテストコレクション配布方法  
音声ドキュメント処理ワークショップ収録データ書き起こし  
Interspeech2010スペシャルセッション提案の検討開始

2009年5月21日 第3回会合  
SLP研究運営委員会における第2期SDPWG提案について  
STD検討経過、4モーラ以上、1時間に1回くらいの出現頻度、評価指標

音声ドキュメント処理ワークショップ講演の書き起こし具体化  
CSJ テストコレクション配布準備  
NTCIR 提案検討開始

2009年5月22日 SLP 研究運営委員会にて第2期 SDPWG 提案

2009年9月15日 第4回会合  
テストコレクション配布方法整理  
NTCIR タスク計画の確認  
STD 検索語プロトタイプ設定

2009年10月 APSIPA で発表 8)

2009年10月23日 SLP 研究会発表 STD タスク 9)

2009年12月21日 第5回会合  
Interspeech2010 スペシャルセッション提案準備  
**SDR テストコレクション公開開始**

2010年1月25日 Interspeech2010 スペシャルセッション提案

2010年2月22日 Interspeech2010 スペシャルセッション採択

2010年2月26-27日 第4回音声ドキュメント処理ワークショップ

2010年3月8日 第6回会合  
Interspeech2010 スペシャルセッション採択とスケジュール確認  
テストコレクション配布用ページ設立

2010年5月8日 SLP 研究会発表 STD テストコレクション 10)

2010年7月16日 NTCIR-9 に SDR/STD タスク”SpokenDoc”提案

2010年7月31日 **STD テストコレクション公開開始**  
全講演用既知語セット 100 語、コア講演用既知語・未知語セット各 50 語

2010年7月22-24日 SLP 研究会 秋葉「音声ドキュメント検索」解説 11)

2010年8月12日 NTCIR-9 に提案の”SpokenDoc”採択

2010年9月14日 第7回会合  
NTCIR-9 スケジュール確認

2010年9月28日 **Interspeech2010 スペシャルセッション**  
発表 テストコレクションについて 12)13)

2010年12月20-21日 音声言語シンポジウム発表 NTCIR について 14)15)

2010年12月21日 第8回会合  
NTCIR スケジュール確認  
論文誌特集号の企画

2011年3月7日 第5回音声ドキュメント処理ワークショップ

2011年8月19日 情報処理学会に論文誌特集号提案

2011年9月21日 第9回会合  
NTCIR 進捗  
情報処理学会論文誌特集号について

2011年10月3日 **情報処理学会論文誌特集号募集開始**

2011年12月8-9日 **第9回 NTCIR ワークショップ**

2011年12月19-20日 音声言語シンポジウム発表 SDPWG 活動報告

### 3. テストコレクション

#### 3.1 企画

テストコレクションの作成は第1期に開始された。音声ドキュメント検索(SDR: Spoken Document Retrieval)セットは第1期終了時にほぼ完成していた。第2期では音

声検索語検出(STD: Spoken Term Detection)のテストコレクションを完成させた。

当初、データベースとして各大学で収録されている講義録を用いる案も提出された。利用できるコンテンツの調査も行ったが、講義資料に含まれる著作権のある画像の問題、教室内の学生が映像に映っていることなど肖像権の問題があることがわかった。

上記の検討の結果、検索対象として日本語話し言葉コーパス(CSJ)を用いることとした。このうち講演を検索対象とし、全講演 2702 (学会講演 987、模擬講演 1715)、および、そのサブセットでアノテーションデータのある一部すなわち、コア講演 177 (学会講演 70、模擬講演 107) をテストコレクションに用いることにした。

### 3.2 SDR (Spoken Document Retrieval: 音声ドキュメント検索)

短文による検索クエリにより音声ドキュメントである講演音声やその一部分であるパッセージを検索するものである。検索単位としては1講演、15発話、30発話単位、1文などが検討された。クエリ案として多数を収集し、その中からテストコレクションに適切なものを選択した。

### 3.3 STD (Spoken Term Detection: 音声検索語検出)

1つまたは複数の単語を与えて音声ドキュメント中から、その単語の存在する部分を検索するものである。この課題では、テストコレクションを作成するに当たり、検索対象として単語の長さ、出現頻度について検討を要した。

## 4. 音声ドキュメント処理ワークショップ

2007年2月に豊橋技術科学大学主催・情報処理学会音声言語情報処理研究会共催で、「第1回音声ドキュメント処理ワークショップ」が開催された。これは、第1期SDPWGの成果である。音声ドキュメント処理ワークショップは豊橋技術科学大学の主催で以後、毎年開催され、2011年3月には、言語処理学会の協賛を得て第5回ワークショップが開催された。毎年20件前後の発表がある。発表者に収録の承諾を頂いた講演はEZプレゼンターを用いて収録し書き起こしを行っている。表1に現在までの音声ドキュメント処理ワークショップでの発表件数を示す。

表1 音声ドキュメント処理ワークショップ開催歴

回	開催年月日	発表件数
1	2007年2月26-27日	24
2	2008年2月29日-3月1日	23
3	2009年2月27-28日	22
4	2010年2月26-27日	22
5	2011年3月7日	16

## 5. 第1回情報処理学会論文誌特集号

2007年度に「音声ドキュメント処理」と題する情報処理学会論文誌特集号を企画し、2009年2月に発行した。投稿論文数は29で採録論文数は13であった。表2にそれらの論文のカテゴリの内訳を示す。この時期には音声認識のための言語モデルに関する論文が多く、STDとSDRの区別はあまり意識されていない。

表2 第1回情報処理学会「音声ドキュメント処理」特集号論文カテゴリ

カテゴリ	論文数
音声合成	1
情報抽出、構造化	4
検索	3
言語モデル、トピック推定	3
コーパス分析	2

## 6. 第2回情報処理学会論文誌特集号

2011年8月に2回目の「音声ドキュメント処理」特集号提案書を情報処理学会に提出し、9月の編集委員会において承認された。現時点で論文募集公告が学会ページに掲載されており、投稿締め切りを2012年5月31日、特集号掲載予定を2013年2月号としている。

この特集号では、講演・講義・会議・放送・対話・翻訳などの音声ドキュメント・音声コーパスに関するデータベース・分析、音声認識・理解、音響モデル・言語モデル、

ドキュメント検索、ターム検索、要約、質問応答・音声対話、メタデータ化・コンテンツ化、大規模音声データベースの処理・応用・利用システム、音声ドキュメント処理関連技術に関する論文を募集している。

## 7. Interspeech2010 スペシャルセッション

Interspeech2010 は 2010 年 9 月に幕張で開催された。ここに SDPWG から "Open Vocabulary Spoken Document Retrieval" と称するスペシャルセッションを提案して採択された。このセッションには 15 件の投稿があり、8 件が採択された。通常 Interspeech のセッションは 2 時間が単位で 20 分の発表 6 件が標準であるが、優れた論文が多く、15 分×8 件の構成となった。このスペシャルセッションに投稿された論文の多くはタイトルに STD/SDR の区別が明記されていた。特に STD 関係の論文が多かった。発表論文のカテゴリ分類を表 3 に示す。

表 3 Interspeech2010 オープンボキャブラリ音声ドキュメント検索論文カテゴリ

STD/SDR	カテゴリ	論文数
STD&SDR	テストコレクション	1
STD	言語情報	2
STD	尺度・検出手段	4
SDR	コーパス分類	1

## 8. NTCIR-9 ワークショップ

日本においても TREC と同様に評価型ワークショップが国立情報学研究所(NII: National Institute of Informatics)で企画されている。SDPWG では、このワークショップにおいて認められることが、音声ドキュメント処理研究の更なる活性化につながると考えた。2010 年度に国立情報学研究所(NII)の主催する評価型ワークショップ NTCIR の NTCIR-9 にタスクを提案し、2010 年 8 月に採択となった。その後、2011 年 5 月まで試運転であるドライラン、7 月まで本実験であるフォーマルランによる評価実験が行われた。音声ドキュメントとしては日本語話し言葉コーパス(CSJ)のうち、学術講演及び模擬講演 2702 を対象とした、SDR と STD の評価セットを用意した。SDR では 5 チーム、STD では 7 チームが参加した。2011 年 12 月に成果発表会があり、オーラルで 4 件、ポスターで 10 件の発表がある。

## 9. 音声ドキュメント処理研究の活性化

### 9.1 情報処理学会

現在までに情報処理学会に投稿された SDR および STD に関係した論文等はジャーナル 13 件、自然言語処理 (NL 研究会) 4 件、音声言語情報処理 (SLP 研究会)15 件となっている。なお、NL と SLP 研究会報告は共催時の発表のため、以下の表 4 の報告数とは異なる。このうちジャーナル論文は特集号の欄で紹介したので、ここでは研究会に報告されたものを分類する。最も古いもので参考文献 1) があるが、2009 年 2 月の特集号発行を契機として急速に発表が増加している。

表 4 情報処理学会研究会で報告された SDR・STD 関係報告

SDR/STD	カテゴリ	報告数
SDR	言語処理	6
SDR	方式・システム	4
SDR	尺度	1
SDR	解説・報告	4
STD	解説	1

### 9.2 電子情報通信学会

現在までに電子情報通信学会の情報システム論文誌和文及び英文に投稿された、SDR もしくは STD 関係の論文は少ない。表 5 にそれらの論文のカテゴリを示す。

表 5 電子情報通信学会論文誌に投稿された SDR/STD 関係論文

カテゴリ	論文数
SDR	3
QA	2
データベース・分類	2

### 9.3 音響学会

日本音響学会研究発表会における音声ドキュメント処理関係の発表件数を表 6 に示す。2009 年度以降は音声ドキュメント処理もしくは音声情報検索のセッションが毎回設置されている。

表6 日本音響学会研究発表会における音声ドキュメント処理関係発表  
CSJ-SS 日本語話し言葉コーパススペシャルセッション  
SDP-SS 音声ドキュメント処理スペシャルセッション  
S 音声ドキュメント処理または音声情報検索セッション設置

開催年	音響処理	言語処理	検索方式	要約・データ	計	備考
2005 秋	2	-	1	1	4	
2006 春	5	3	-	3	11	CSJ-SS
2006 秋	3	-	1	-	4	
2007 春	-	2	1	3	6	
2007 秋	3	2	4	2	11	SDP-SS
2008 春	1	1	-	-	2	
2008 秋	1	1	-	1	3	
2009 春	3	1	4	1	9	S
2009 秋	-	2	1	2	5	S
2010 春	6	2	6	1	15	S
2010 秋	4	2	2	1	9	S
2011 春	3	4	2	-	9	S
2011 秋	6	-	2	-	8	S

## 10. おわりに

2006年度から3年ずつ2期に亘って継続した音声ドキュメント処理ワーキンググループの活動について報告した。本報告では、特に、様々な企画のプロセスについて詳しく記述した。この報告がこれからのワーキンググループの企画や遂行に役立つことを期待している。

**謝辞** CSJ データベースの利用についてご検討いただいた国立国語研究所の前川喜久雄博士に感謝する。

## 参考文献

- 1) 伊藤 克亘, 相川 清明, 秋葉 友良 伊藤 慶明, 河原 達也, 南條 浩輝, 西崎 博光, 安田 宜仁, 山下 洋一, "音声ドキュメント検索評価のためのテストコレクションの試作", 情報処理学会研究報告, 2006-SLP-64, pp.137-142, (2006-12).
- 2) 秋葉 友良, 相川 清明, 伊藤 慶明, 河原 達也, 南條 浩輝, 西崎 博光, 安田 宜仁, 山下 洋一 伊藤 克亘, "音声ドキュメント検索評価テストコレクションの試作と基本検索性能評価", 第1回音声ドキュメント処理ワークショップ講演論文集, p.73-80, (2007-02).
- 3) 中川 聖一, "(招待講演) 音声ディクテーションから音声ドキュメント処理へ", 日本音響学会講演論文集, 1-3-1, pp.1-4, (2007-09).
- 4) Tomoyosi Akiba, Kiyooki Aikawa, Yoshiaki Itoh, Tatsuya Kawahara, Hiroaki Nanjo, Hiromitsu Nishizaki, Norihito Yasuda, Yoichi Yamashita, Katunobu Itou, "Test Collections for Spoken Document Retrieval from Lecture Audio Data", LREC2008, Session P22, (2008-05).
- 5) 秋葉友良, 相川清明, 伊藤慶明, 河原達也, 南條浩輝, 西崎博光, 安田宜仁, 山下洋一, 松井知子, 胡新輝, 中川聖一, 伊藤克亘, "SLP 音声ドキュメント処理ワーキンググループ活動報告", 情報処理学会研究報告, Vol.2008 No.123, 2008-SLP-074, pp.115-120, (2008-12).
- 6) Tomoyosi Akiba, Kiyooki Aikawa, Yoshiaki Itoh, Tatsuya Kawahara, Hiroaki Nanjo, Hiromitsu Nishizaki, Norihito Yasuda, Yoichi Yamashita, Katunobu Itou, "Construction of a Test Collection for Spoken Document Retrieval from Lecture Audio Data", IPSJ Journal Vol. 50 No. 2 501-513 (2009-02).
- 7) <http://www.cl.ics.tut.ac.jp/~sdpwg/index.php?SDPWG%20Wiki>
- 8) Tomoyosi Akiba, Kiyooki Aikawa, Yoshiaki Itoh, Tatsuya Kawahara, Hiroaki Nanjo, Hiromitsu Nishizaki, Norihito Yasuda, Yoichi Yamashita, Katunobu Itou, "Developing an SDR test collection from Japanese lecture audio data", APSIPA ASC 2009, TA-SS1-2 0196, (2009-10).
- 9) 伊藤慶明, 西崎博光, 胡新輝, 南條浩輝, 秋葉友良, 相川清明, 河原達也, 中川聖一, 松井知子, 山下洋一, "音声中の検索語検出のためのテストコレクション構築 -中間報告-", 情報処理学会研究報告, Vol.2009-SLP-78, pp.1-8, (2009-10).
- 10) 西崎博光, 胡新輝, 南條浩輝, 伊藤慶明, 秋葉友良, 河原達也, 中川聖一, 松井知子, 山下洋一, 相川清明, "Spoken Term Detection のためのテストコレクション構築とベースライン評価", 情報処理学会研究報告, Vol.2010-SLP-81 No.13, pp.1-8, (2010-05).
- 11) 秋葉友良, "音声ドキュメント検索の現状と課題", 情報処理学会研究報告, Vol.2009-SLP-82, No.10, pp.1-8, (2010-07).
- 12) [http://www.interspeech2010.org/program/session\\_4910.html](http://www.interspeech2010.org/program/session_4910.html)
- 13) Yoshiaki Itoh, Hiromitsu Nishizaki, Xinhui Hu, Hiroaki Nanjo, Tomoyosi Akiba, Tatsuya Kawahara, Seiichi Nakagawa, Tomoko Matsui, Yoichi Yamashita, Kiyooki Aikawa, "Constructing Japanese test collections for spoken term detection", INTERSPEECH2010, pp.677-680, (2010-09).
- 14) <http://research.nii.ac.jp/ntcir/index-ja.html>
- 15) 秋葉友良, 西崎博光, 相川清明, 河原達也, 松井知子, 伊藤慶明, 胡新輝, 中川聖一, 南條浩輝, 山下洋一, "NTCIR-9 SpokenDoc: 音声検索語検出と音声ドキュメント検索の評価枠組の設計", 情報処理学会研究報告, Vol.2010-SLP-84 No.18, pp.1-5, (2010-12).