

実走行車内単語音声データベース CENSREC-3 と 共通評価環境の構築

藤本 雅清¹ 中村 哲¹ 武田 一哉² 黒岩 眞吾³ 山田 武志⁴ 北岡 教英⁵
山本 一公⁶ 水町 光徳⁷ 西浦 敬信⁸ 佐宗 晃⁹ 宮島 千代美² 遠藤 俊樹¹

¹ATR 音声言語コミュニケーション研究所 ²名古屋大学 ³徳島大学 ⁴筑波大学
⁵豊橋技術科学大学 ⁶信州大学 ⁷九州工業大学 ⁸立命館大学 ⁹産業技術総合研究所

あらまし 本稿では、SLP 雑音下音声認識評価ワーキンググループの活動成果として、自動車内音声認識の評価用データベース CENSREC-3 と、標準評価スクリプトによるベースライン評価結果について述べる。CENSREC-3 の音声認識タスクは、実走行車内での孤立単語音声認識であり、音声データの収録は、接話マイクロホンと遠隔マイクロホンの 2 種類を用いて、3 種類の走行速度と 6 種類の車内環境を組み合わせた 16 種類の環境下で行っている。CENSREC-3 では、これら様々な環境したで収録された音声データを用いた 6 種類の評価環境を提供する。

キーワード: 雑音下音声認識, 共通評価フレームワーク, 自動車内音声データベース, CENSREC-3

CENSREC-3: Data Collection for In-Car Speech Recognition and Its Common Evaluation Framework

Masakiyo Fujimoto¹ Satoshi Nakamura¹ Kazuya Takeda² Shingo Kuroiwa³
Takeshi Yamada⁴ Norihide Kitaoka⁵ Kazumasa Yamamoto⁶ Mitsunoti Mizumachi⁷
Takanobu Nishiura⁸ Akira Sasou⁹ Chiyomi Miyajima² Toshiki Endo¹

¹ATR Spoken Language Translation Research Laboratories ²Nagoya University ³University of Tokushima
⁴University of Tsukuba ⁵Toyohashi University of Technology ⁶Shinshu University
⁷Kyushu Institute of Technology ⁸Ritsumeikan University
⁹National Institute of Advanced Industrial Science and Technology

Abstract This paper introduces a common database, an evaluation framework, and its baseline recognition results for in-car speech recognition, CENSREC-3, as an outcome of IPSJ-SIG SLP Noisy Speech Recognition Evaluation Working Group. CENSREC-3 is designed as the evaluation framework of isolated word recognition in real driving car environments. Speech data was collected using 2 microphones, a close-talking microphone and a hands-free microphone, under carefully controlled 16 different driving conditions, i.e., combinations of 3 car speeds and 5 car conditions. CENSREC-3 provides 6 evaluation environments which are designed using speech data collected in these car conditions.

Keywords: noisy speech recognition, common evaluation framework, in-car speech database, CENSREC-3

1 はじめに

近年の音声認識技術の進歩の大部分は、統計モデルの導入と大規模コーパスの収集により支えられ、また、1980 年代に開始された音声認識に関する米国国防総省の DARPA プロジェクト [1] の成功により、不特定話者大語彙連続音声認識のシステム構成基本技術が確立された。しかしながら、現在の音声認識装置を実際に利用されるような雑音のある環境で利用すると、未だ著しい性能劣化が避けられない。

このような音声認識の音響環境に対する頑健性の問題に対しては、これまでに米国 DARPA 主催の SPINE [2] と、欧州における AURORA[3]-[9] の 2 つのプロジェクトが進められた。

SPINE (SPeech recognition In Noisy Environments) は、米国 DARPA 主催で 2000 年に SPINE1, 2001 年に SPINE2 が開催された。タスクは、オペレータと雑音のある環境にいる兵士との英語の自由発話の対話である。このプロジェクトにより多々の雑音下音声認識における多くの技術が開発されたが、タスクが複雑で、要因と各種手法による改善効果の解析が困難なことが問題であった。

一方、欧州の通信技術標準化団体 ETSI [10] 傘下の AURORA グループが、分散音声認識のための雑音下音声認識の前処理の技術標準化を行っており、これに並行して、雑音下音声認識の発展、標準化のためのコーパスと HTK [11] を利用した評価環境が整備された。現在、AURORA2 [3],

AURORA3 [4]-[7], AURORA4 [8, 9] が公開されている。

このような欧米の動きに対し、筆者らは、2001年10月に情報処理学会 音声言語情報処理研究会内に、有志によるワーキンググループを作り、雑音下音声認識の評価のための議論を進めてきた。ワーキンググループの活動成果として筆者らは、2003年7月に AURORA2 の数字を日本語に翻訳し、同一の雑音データを付与した AURORA-2J [12] を作成し、配布を行った。

本稿では、AURORA-2J に続く雑音下音声認識の標準環境である、CENSREC-3 (Corpus and Environments for Noisy Speech REcognition) の概要と、標準評価スクリプトによるベースライン性能について述べる。CENSREC-3 の音声認識タスクは、実走行車内での孤立単語音声認識であり、様々な環境下で収録された音声データを用いた6種類の評価環境を提供する。また、本稿では、CENSREC-3 の ETSI ES 202 050 フロントエンド (Advanced front-end) [13] による評価結果についても述べる。

2 音声データの収録

CENSREC-3 のデータには、名古屋大学統合音響情報研究拠点 CIAIR で収録されたデータベースの一部を用いる [14, 15]。

2.1 収録語彙

CENSREC-3 では、自動車内での単語音声の認識を対象とする。評価用の単語音声データには、表1に示す50単語を用い、音響モデルの学習データには、自動車内で収録された音素バランス文を用いる。

表1: 収録50単語の一覧

デジタルロッカー	認証開始	2001年1月1日
山田太郎	検索終了	暗証番号
0123	4567	8901
2345	6789	コンテンツ
映画	羊たちの沈黙	サウンドオブミュージック
ゲーム	バックマン	音楽
JPOP	今週のトップテン	ジャンル別検索
ポップス	ロック	ビートルズ
選曲	イエスタデー	レットイットビー
配信開始	フェリー案内	時刻表
第二便を予約	ネットニュース	トピックス
音声読み上げ	天気予報	交通情報
神奈川県	横浜市	中区
東京都	世田谷区	首都高速
東北自動車道	セブンイレブン	ユニクロ
スターバックス	ホテル一覧	パシフィックホテル
予約表	サービス終了	

2.2 音声データの収録環境

自動車内音声の収録は、特別に設計された実験車両を用いて行う。実験車両には、運転席周辺に5本のマイクロホンが図1に示すような位置に設置されており、3, 4番はダッシュボード上、5, 6, 7番は天井に設置されている。また、1番は接話マイクロホンである。これらのマイクロホンの

内、CENSREC-3 では、1番の接話マイクロホンと6番の遠隔マイクロホンで収録された音声を用いる [15]。それぞれのマイクロホンには、SONY ECM77B を用いている。

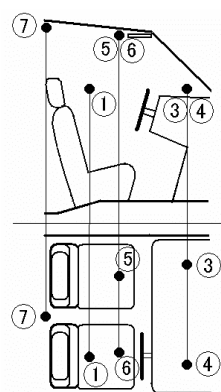


図1: マイクロホンの設置位置 (上段): 側面, (下段): 真上

評価データの収録は、表2に示す、3種類の走行速度 (アイドリング, 低速 (市街地) 走行, 高速走行) と、6種類の車内環境 (通常走行, ハザード On, エアコン (Low), エアコン (High), オーディオ On, 窓開) を組み合わせた16種類の環境で行う [14]。評価データの発話者数は18名 (男性8名, 女性10名) であり、収録音声の総数は、マイクロホン一本あたり14,216発話である。

表2: 評価データの収録環境

走行速度	車内環境
アイドリング	通常走行, ハザード On, エアコン (Low), エアコン (High), オーディオ On, 窓開
低速走行	通常走行, エアコン (Low), エアコン (High), オーディオ On, 窓開
高速走行	通常走行, エアコン (Low), エアコン (High), オーディオ On, 窓開

一方、学習データの収録は、アイドリング, 低速走行の条件で行い、車内環境は通常走行のみである [15]。学習データの発話者は、293名 (男性202名, 女性91) であり、収録データの総数は、マイクロホン一本あたり14,050発話である。ここで、学習データである音素バランス文の収録は、アイドリング時には発話内容が記された原稿を読み上げるにより行う。しかし、走行時では原稿を見て読み上げを行うことは不可能であるので、ヘッドセットマイクロホンのイヤホンから発話内容を指示することにより、発話を行う。またこの際、発話者が指示された1文を記憶して発話することは困難であるので、1文をいくつかの文節に区切って指示を行う。よって、走行時の音声は、1文の内容が文節単位に分割された複数のデータに収録される。

以上のような音声の収録において、収録条件は評価データ, 学習データともに、標準化周波数16kHz, 語長16bitであり、バイトオーダーはリトルエンディアンである。

表 3: 各評価環境で用いられる学習データ

評価環境	Condition 1		Condition 2		Condition 3		Condition 4		Condition 5		Condition 6	
	接話	遠隔	接話	遠隔	接話	遠隔	接話	遠隔	接話	遠隔	接話	遠隔
マイクロホン												
アイドリング	○	○	○	—	—	○	—	○	○	—	○	—
低速走行	○	○	○	—	—	○	—	—	○	—	—	—

表 4: 各評価環境で用いられる評価データ

評価環境	Condition 1		Condition 2		Condition 3		Condition 4		Condition 5		Condition 6	
	接話	遠隔	接話	遠隔	接話	遠隔	接話	遠隔	接話	遠隔	接話	遠隔
マイクロホン												
アイドリング	○	○	○	—	—	○	—	—	—	—	—	—
低速走行	○	○	○	—	—	○	—	○	—	○	—	○
高速走行	○	○	○	—	—	○	—	○	—	○	—	○

表 5: 各評価環境における学習データ数

走行速度	マイクロホン	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
アイドリング	接話	3608	3608	—	—	3608	3608
	遠隔	3608	—	3608	3608	—	—
	合計	7216	3608	3608	3608	3608	3608
低速走行	接話	10442	10442	—	—	10442	—
	遠隔	10442	—	10442	—	—	—
	合計	20884	10442	10442	—	10442	—
合計		28100	14050	14050	3608	14050	3608

3 評価環境の設計

CENSREC-3 では、前章にて述べた、様々な環境で収録された音声データを用いて、6種類の音声認識評価環境 (Condition 1~6) を構成する*。表3, 4の○印は、各々の評価環境で用いるデータを示しており、これら6種類の評価環境の内、Condition 1, 2, 3は、学習データと評価データのマイクロホン種別及び、収録環境が一致している。これらの環境は、AURORA3のWell-matched conditionに相当する。Condition 4は、学習データと評価データでマイクロホン種別が一致しているが、収録環境が異なっており、この評価環境は、AURORA3のModerate-mismatched conditionに相当する。また、Condition 5, 6は、マイクロホン種別、収録環境(の一部)が共に異なっており、これらはAURORA3のHigh-mismatched conditionに相当する[4]-[7]。以上の6種類の評価環境におけるデータ量は、表5, 6に示す通りである。

4 ベースライン評価

4.1 評価用スクリプト

評価用ベースラインスクリプトは、HTK [11] を用いてHMMの学習、評価実験が容易に行えるように作成し、以下のようなベースライン評価の仕様を設計した。

- スクリプトは全てPerlで作成し、Perl ver.5以降に対応。
- 評価実験の時間を短縮するため、複数の計算機での並列処理機能を提供する。並列処理は、特徴抽出、学

習、認識の全てに対応しており、スクリプトの設定ファイルにリモートホスト名を追加するだけで容易に設定可能である。

- 音声認識は音素HMMにより行い、標準の単語辞書と、図2に示すEBNF記法の認識文法を提供する。
- 音素HMMは、Triphone HMM (2,000状態)であり、各HMMの状態数は5(出力分布を持つ状態は3)、状態あたりの混合分布数は32である。
- 単語辞書において、連続母音の長母音化が生じる単語の場合、連続母音と長母音の2種類の発声規則を登録する(「認証」の場合「 $n i N sh o u$ 」と「 $n i N sh o:$ 」の2種類を登録)。
- 特徴量は、HTKのHCopyにより抽出された、12次MFCC + log-power + Δ MFCC + Δ log-power + $\Delta\Delta$ MFCC + $\Delta\Delta$ log-powerの39次元とする。分析条件は、 $1-0.97z^{-1}$ のプリエンファシス、ハミング窓、24次元のメルフィルタバンク、20msの分析フレーム長、10msのフレームシフトとする。また、Cepstral Mean Subtractionは行わない。
- 自動車雑音特有の低周波成分に対処するため、メルフィルタバンク分析時に250Hz以下の低周波成分を取り除く。

```
$words = デジタルロッカー | 認証開始 | ...
        ... | サービス終了;
([silB] $words [silE])
```

図 2: EBNF 記法による認識用文法

*学習データの内、遠隔マイクロホンで収録されたデータは有償による配布とする。CENSREC-3は1枚のDVDに無償、有償全てのデータを収録して配布する。有償(遠隔マイクロホン)データを利用するには、ライセンス料の支払いが必要となる。

表 6: 各評価環境における評価データ数

走行速度	マイクロホン	車内環境	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
アイドリング	接話	通常走行	898	898	—	—	—	—
		ハザード On	900	900	—	—	—	—
		エアコン (Low)	887	887	—	—	—	—
		エアコン (High)	900	900	—	—	—	—
		オーディオ On	896	896	—	—	—	—
		窓開	899	899	—	—	—	—
		合計	5380	5380	—	—	—	—
	遠隔	通常走行	898	—	898	—	—	—
		ハザード On	900	—	900	—	—	—
		エアコン (Low)	887	—	887	—	—	—
		エアコン (High)	900	—	900	—	—	—
		オーディオ On	896	—	896	—	—	—
		窓開	899	—	899	—	—	—
		合計	5380	—	5380	—	—	—
合計		10760	5380	5380	—	—	—	
低速走行	接話	通常走行	848	848	—	—	—	—
		エアコン (Low)	850	850	—	—	—	—
		エアコン (High)	895	895	—	—	—	—
		オーディオ On	849	849	—	—	—	—
		窓開	897	897	—	—	—	—
		合計	4339	4339	—	—	—	—
		遠隔	通常走行	848	—	848	848	848
	エアコン (Low)		850	—	850	850	850	850
	エアコン (High)		895	—	895	895	895	895
	オーディオ On		849	—	849	849	849	849
	窓開		897	—	897	897	897	897
	合計		4339	—	4339	4339	4339	4339
	合計			8678	4339	4339	4339	4339
	高速走行	接話	通常走行	900	900	—	—	—
エアコン (Low)			900	900	—	—	—	—
エアコン (High)			900	900	—	—	—	—
オーディオ On			899	899	—	—	—	—
窓開			898	898	—	—	—	—
合計			4497	4497	—	—	—	—
遠隔			通常走行	900	—	900	900	900
		エアコン (Low)	900	—	900	900	900	900
		エアコン (High)	900	—	900	900	900	900
		オーディオ On	899	—	899	899	899	899
		窓開	898	—	898	898	898	898
		合計	4497	—	4497	4497	4497	4497
		合計		8994	4497	4497	4497	4497
合計			28432	14216	14216	8836	8836	8836

4.2 ベースライン認識結果

表 7 に、評価環境 Condition 1~6 の車内環境毎の詳細なベースライン認識性能を示す[†]。表 7 より、接話マイクロホンで収録された音声の認識率は全体的に高いことがわかる。一方、遠隔マイクロホンで収録された音声の認識率は全体的に低い。特に低速、高速時の窓開環境では認識率の低下が著しく、改善の余地が十分に認められる。

また、研究機関毎の認識性能比較を容易にするため、表 8 のような Microsoft Excel にて作成されたスプレッドシートを配布する。表 8 の上段は、各評価環境のベースライン

[†]HTK の仕様により、複数の計算機を用いて学習を行うと、計算機の台数によっては HMM パラメータの推定結果が異なったものになり、表 7 のベースライン認識率が得られない場合がある（評価環境によっては ±1% 程度の違いが現れる可能性がある）。表 7 の結果は 4 台の計算機（OS: Red Hat Linux release 7.2）を用いて得られた結果であり、CENSREC-3 ではこの結果をベースラインとする。尚、この現象は再現性があり、4 台の計算機を用いると、表 7 と同じ結果が得られる。

性能とその平均を示しており、中段に自身の手法による認識結果（単語正解精度）を入力する。中段に認識結果を入力すると、下段にベースライン性能との相対的な改善性能（誤り改善率）が自動で出力される。

表 8 に示した中段、下段の値は、ETSI ES 202 050 フロントエンド (Advanced front-end) [13] による評価結果であり、これらの結果より ETSI ES 202 050 フロントエンドの性能の高さがうかがえる。

4.3 評価カテゴリー

CENSREC-3 においても、AURORA-2J と同様に、バックエンドの変更（HMM の学習方法、トポロジーの変更、特徴量の変更など）に対して、その度合に応じたカテゴリーを設定する。バックエンドを変更した結果を発表する場合、以下に示すカテゴリーから一つを選び、発表でそれを示す必要がある。バックエンドを変更しない場合は、カテゴリー

表 7: CENSREC-3 ベースライン認識性能の詳細 (%)

走行速度	マイクロホン	車内環境	Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6
アイドリング	接話	通常走行	99.89	100.00	—	—	—	—
		ハザード On	99.33	99.89	—	—	—	—
		エアコン (Low)	99.55	100.00	—	—	—	—
		エアコン (High)	97.78	99.44	—	—	—	—
		オーディオ On	98.77	99.67	—	—	—	—
		窓開	99.11	99.33	—	—	—	—
	遠隔	通常走行	99.44	—	99.78	—	—	—
		ハザード On	98.78	—	98.89	—	—	—
		エアコン (Low)	90.19	—	94.02	—	—	—
		エアコン (High)	53.56	—	53.44	—	—	—
		オーディオ On	81.47	—	81.36	—	—	—
		窓開	89.66	—	89.88	—	—	—
	全環境	99.07	99.72	—	—	—	—	
	低速走行	接話	通常走行	100.00	100.00	—	—	—
エアコン (Low)			100.00	100.00	—	—	—	—
エアコン (High)			97.99	98.77	—	—	—	—
オーディオ On			98.82	99.41	—	—	—	—
窓開			99.11	98.55	—	—	—	—
全環境			99.17	99.33	—	—	—	—
遠隔		通常走行	98.00	—	99.17	88.21	56.60	45.99
		エアコン (Low)	90.82	—	94.12	77.41	54.35	35.18
		エアコン (High)	62.57	—	60.11	41.79	43.46	28.83
		オーディオ On	79.27	—	78.56	65.02	47.47	37.57
		窓開	64.66	—	65.33	45.60	23.97	15.27
		全環境	78.73	—	79.10	63.17	44.92	32.33
全環境		88.95	99.33	79.10	63.17	44.92	32.33	
高速走行		接話	通常走行	99.89	99.89	—	—	—
	エアコン (Low)		99.67	99.89	—	—	—	—
	エアコン (High)		97.67	99.22	—	—	—	—
	オーディオ On		99.78	99.78	—	—	—	—
	窓開		96.66	95.21	—	—	—	—
	全環境		98.53	98.80	—	—	—	—
	遠隔	通常走行	92.33	—	95.56	64.78	29.67	21.78
		エアコン (Low)	85.11	—	89.44	48.22	30.67	19.89
		エアコン (High)	59.67	—	55.22	37.33	40.78	22.44
		オーディオ On	78.31	—	79.20	49.72	30.03	23.92
		窓開	24.83	—	21.83	15.37	7.80	6.46
		全環境	68.07	—	68.27	43.10	27.80	18.90
	全環境	83.30	98.80	68.27	43.10	27.80	18.90	
	全環境	88.43	99.31	78.36	52.95	36.20	25.50	

0. となる。カテゴリ内で性能比較を行なうことで、各手法の性能比較をより適切に行なうことができる。尚、下記のカテゴリは、AURORA-2J のカテゴリ設定に一部変更を加えたものとなっている。

カテゴリ 0. ベースラインスクリプトを全く変更しない場合。

カテゴリ 1. 標準 HMM と同じトポロジーの HMM だが、識別学習等、学習方法を変更している場合。このカテゴリの認識時のコストは、ベースラインと全く同じである。その他の実験条件はベースラインと同じ条件に従う。

カテゴリ 2. ベースラインスクリプトと同じトポロジーの HMM で、認識時の適応技術を導入している場合。話者適応、環境適応、1 状態 1 混合の雑音 HMM を用いた PMC 等がここに含まれ、認識時に適応を行

なうことによる認識コストの増加がある。その他の実験条件はベースラインと同じ条件に従う。

カテゴリ 3. 混合数や状態数等の HMM トポロジーを変更している場合。ただし、モデル単位はベースラインと同じ (CENSREC-3 では triphone) であることを条件とし、2 状態以上の雑音モデルを用いた PMC 等がこれに相当。その他の実験条件はベースラインと同じ条件に従う。

カテゴリ 4. 認識デコーダがベースラインスクリプトと同じ (CENSREC-3 では hvite) であることを条件にどのような処理も許される場合。モデル単位の変更や、文法・辞書の書き換え等がこれに相当。

カテゴリ 5. 規定無し。提供されるデータベース内であれば、どんな処理でも許される。認識デコーダの変更も許容。

表 8: CENSREC-3 スプレッドシートと ETSI ES 202 050 フロントエンドによる評価結果

CENSREC-3 Evaluation Results						
CENSREC-3 Baseline Results (%)						
Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6	Average
88.43	99.31	78.36	52.95	36.20	25.50	63.46
CENSREC-3 Word Accuracy (%)						
Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6	Average
95.48	99.62	91.95	86.63	83.70	73.85	88.54
CENSREC-3 Relative Improvement						
Condition 1	Condition 2	Condition 3	Condition 4	Condition 5	Condition 6	Average
60.93%	44.93%	62.80%	71.58%	74.45%	64.90%	68.63%

カテゴリ B. 提供されるデータ以外のデータを使用する場合．評価データは提供されているものを用いる．

5 まとめ

本稿では，自動車内音声認識の評価用データベースである CENSREC-3 と，標準評価スクリプトによるベースライン評価結果，ETSI ES 202 050 フロントエンドによる評価結果について報告した．

筆者らの今後の活動展開として，ロンバード効果を含む音声を対象とした CENSREC-1.5 (AURORA-2.5J)，自動車内での孤立単語認識を対象とした CENSREC-3 に対して，自動車内での連続数字認識を対象とした CENSREC-2 (AURORA-3J) を順次公開，配布する予定にしている．

CENSREC データベースが対象とする認識タスク，評価環境は，現在作成中の CENSREC-1.5 等を含めて，以下のような変遷を辿っている．

CENSREC-1 人工データ，連続数字認識，加法性雑音
CENSREC-1.5 人工データ，連続数字認識，加法性雑音，ロンバード効果

CENSREC-2 実走行車内音声データ，連続数字認識
CENSREC-3 実走行車内音声データ，孤立単語認識

今後，非定常雑音下及び，残響下での認識，大語彙連続音声認識など，雑音環境と認識タスクを徐々に難しくした評価環境を設計し，データベースとして公開する予定である．また，標準的な雑音データベースの設計，単語正解精度以外の評価指標の検討，雑音対策用の標準ツールの開発，配布などについても検討を行う予定である．尚，AURORA-J/CENSREC に関する最新の情報は以下の URL を参照されたい．

AURORA-J/CENSREC Web site:
<http://sp.shinshu-u.ac.jp/CENSREC/>

謝辞 本研究は，情報通信研究機構の研究委託により実施したものである．

参考文献

- [1] DARPA project Web site, <http://www.nist.gov/speech/publications/>
- [2] SPINE Web site, <http://elazar.itd.nrl.navy.mil/spine/>
- [3] H.G.Hirsch and D.Pearce, "The AURORA Experimental Framework for the Performance Evaluations of Speech Recognition Systems under Noisy Condition," Proc. ISCA ITRW ASR2000, pp. 18-20, Paris, France, Sep. 2000.
- [4] AU/378/01, "Danish SpeechDat-Car Digits Database for ETSI STQ-Aurora Advanced DSR," Aalborg University, Jan. 2001.
- [5] AU/225/00, "Baseline Results for subset of SpeechDat-Car Finnish Database for ETSI STQ WI008 Advanced Front-end Evaluation," Nokia, Jan. 2000
- [6] AU/273/00, "Description and Baseline Results for the Subset of the Speechdat-Car German Database used for ETSI STQ Aurora WI008 Advanced DSR Front-end Evaluation," Texas Instruments, Dec. 2001.
- [7] AU/271/00, "Spanish SDC-Aurora Database for ETSI STQ Aurora WI008 Advanced DSR Front-End Evaluation: Description and Baseline Results," UPC, Nov. 2000.
- [8] AU/337/01, "Experimental Framework for the Performance Evaluation of Speech Recognition Front-Ends on a Large Vocabulary Task: Version 1.0," Ericsson, June 2001.
- [9] AU/345/01, "Large Vocabulary Evaluation of Front-ends: Baseline Recognition System Description, Final Report," Mississippi State University, Jan. 2002.
- [10] ETSI Web site, <http://www.etsi.org/>
- [11] HTK Web site, <http://htk.eng.cam.ac.uk/>
- [12] S.Nakamura et al., "AURORA-2J: An Evaluation Framework for Japanese Noisy Speech Recognition," IEICE Transactions on Information and Systems, Vol.E88-D, No.3, Mar. 2005. (to appear)
- [13] ETSI ES 202 050 V1.1.1, "Distributed speech recognition; advanced front-end feature extraction algorithm; compression algorithms," 2002.
- [14] 武田 一哉 他 "走行状況別車内音声データベースとその予備評価," 音講論集, 3-P-10, pp. 185-186, Mar. 2002.
- [15] K.Takeda et al., "Construction and Evaluation a Large In-Car Speech Corpus," IEICE Transactions on Information and Systems, Vol.E88-D, No.3, Mar. 2005. (to appear)