

AURORA-2J/AURORA-3J データベースとその評価ベースライン

山本一公¹, 中村 哲², 武田一哉³, 黒岩眞吾⁴, 北岡教英⁵
山田武志⁶, 水町光徳², 西浦敬信⁷, 藤本雅清⁸

¹信州大学, ²ATR 音声言語コミュニケーション研究所, ³名古屋大学,
⁴徳島大学, ⁵豊橋技術科学大学, ⁶筑波大学, ⁷和歌山大学, ⁸龍谷大学

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, 雑音下音声認識評価用共通データベース AURORA-2J と, その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. AURORA-2J は, AURORA プロジェクトの AURORA-2 データベースの日本語版として設計され, 標準評価スクリプトも AURORA-2 で配布されているスクリプトをベースとして開発されている. この共通評価フレームワークにより, 各機関における雑音環境下音声認識手法の性能を容易に比較することが可能となり, 雑音環境下音声認識手法の発展を促すことができると考えられる. また, 自動車内における数字 / コマンド発声データベースである AURORA-3J の開発進捗状況についても述べる.

キーワード 雑音下音声認識, AURORA プロジェクト, 共通評価フレームワーク

AURORA-2J/AURORA-3J Corpus and Evaluation Baseline

Kazumasa Yamamoto¹, Satoshi Nakamura², Kazuya Takeda³,
Shingo Kuroiwa⁴, Norihide Kitaoka⁵, Takeshi Yamada⁶,
Mitsunori Mizumachi², Takanobu Nishiura⁷, Masakiyo Fujimoto⁸

¹Shinshu University, ²ATR Spoken Language Translation Research Labs.,
³Nagoya University, ⁴University of Tokushima, ⁵Toyohashi University of Technology,
⁶University of Tsukuba, ⁷Wakayama University, ⁸Ryukoku University

Abstract This paper introduces a common database, an evaluation framework, and its baseline recognition result for noisy speech recognition, AURORA-2J, as an outcome of IPSJ-SIG SLP Noisy Speech Recognition Evaluation Working Group. AURORA-2J is designed as Japanized version of the AURORA-2 database and the evaluation framework is based on the AURORA-2 baseline scripts. This common evaluation framework enables to compare various noisy speech recognition techniques on a common ground. We hope more development of noisy speech recognition techniques using this evaluation framework. We also describe about AURORA-3J, digits and speech command database in car environments.

key words noisy speech recognition, AURORA project, common evaluation framework

1 はじめに

近年の音声認識技術の進歩は, 統計モデルの導入と大規模コーパスの収集によりもたらされたといえる. しかし, 実際には 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 プロジェクトで開発された様々な技術に負うところが大きい. このプロジェクトでは, 参加した研究チームに同一の学習データとテストデータを用いて同一

のタスクに対する性能改善を競わせる. この手法により, 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しようとする, 未だ著しい性能劣化が避けられないことも事実であり, 性能改善の余地があると言わざるを得ない.

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も, これまでに米国 DARPA 主催

の SPINE プロジェクト [1] と、欧州における AURORA プロジェクト [2] の 2 つの研究プロジェクトが進められた。

まず、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、タスクが特殊すぎることと、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であることが問題であった。

一方、欧州の通信技術標準化団体 ETSI 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に関するスペシャルセッションを Eurospeech で開催することとなった [3]。このグループは ETSI のもと、標準化に向けて技術標準化を行っていたが、これに並行して、さらに雑音下音声認識の発展のため、標準化のためのコーパス (TI digit+Noise) とそれを認識するための HTK を利用した標準スクリプト、標準スクリプトで得られるベースライン性能からの性能改善率を求める Microsoft Excel Spread Sheet を研究者に配布し、これに関するスペシャルセッションを Eurospeech, ICSLP で企画した。これまでに、TI digit に雑音を付与した連続数字音声認識タスクである AURORA-2, 自動車内の連続数字 / コマンドタスクである AURORA-3 をそのスクリプトと共に配布している [4, 5]。この AURORA のメリットは、タスクが連続数字と比較的小さく、1) 大語彙連続音声認識に比べて簡単であること、2) ベースライン性能が配布される HTK スクリプトにより容易に得られること、があげられる。さらに、Wall Street Journal タスクをベースとした雑音下大語彙連続音声認識タスクである AURORA-4 も配布が開始されている [6]。

筆者らは、2001 年の 10 月に情報処理学会音言語情報処理研究会の中に、有志によるワーキンググループを作り、雑音下日本語音声認識の評価のための議論を進めてきた [7, 8]。このワーキンググループの目的は、雑音下音声認識の要素技術のアクセスメントのための計画、標準コーパスの構築、共通評価手法の開発、標準パッケージの配布である。まず初めにタスクが簡単で評価が明解であることを考慮し、AURORA グループの活動にならうことにし、AURORA 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 を作成することとした。また、AURORA-3 に対しては、独自のデータセットで自動車内日本語連続数字とコマンド語のデータを収録した。

本稿では、第 2 章において、AURORA-2J コーパスの収録とその標準スクリプト、ベースライン性

能に関して述べる。第 3 章においては、AURORA-3J 自動車内音声コーパスとその評価方法の検討について述べる。第 4 章においては、雑音下音声認識技術が公平に評価されるための技術カテゴリーについて延べ、最後に第 5 章で今後の方向について述べる。

2 AURORA-2J

AURORA-2J は、雑音環境下連続英語数字音声認識タスクの共通評価フレームワークである AURORA-2 の日本語版である。本章では、AURORA-2J コーパスの収録と、その共通評価環境およびベースライン性能について述べる。

2.1 収録する数字列

AURORA-2J の発声リストは AURORA-2 と同一のものを用いた。また、話者数、男女比も同一で話者毎の発声リストも同一となっている。ただし、発声者は日本人、数字の読みは日本語でこの 2 点が AURORA-2 と異なる。

表 1 に各数字の読みを示す。AURORA-2 では、“0” に対し /zero/、/oh/ の 2 種類の発音が定義されており、発声リストおよびファイル名は “Z” および “O” と明確に区別されている。日本語の場合、“0” は「ぜろ」「れい」「まる」等と発声されるが、電話番号やクレジットカード番号を電話でオペレータ等に伝える場合、「ぜろ」と「まる」の比率が高い。そこで、AURORA-2J では “Z” を「ぜろ」、 “O” を「まる」と発声させた。また同様の理由で、「し」「しち」等の読みは採用しなかった。一方、“2” や “5” の長母音化に関しては発声者の自由とした。

2.2 音声データ収録

音声データの収録には、ヘッドセットマイクロフォン (Sennheier MHD25) と、USB オーディオインターフェイス (Edirol UA-5) を装着した Windows PC を用いた。収録は防音室で、PC に接続された CRT モニターに映し出される数字列を話者が読み上げる形で行なった。この時点での音声データは、サンプリング周波数 16kHz の Microsoft WAV フォーマットとして保存された。

2.3 ダウンサンプリングおよび雑音重畳

AURORA-2J では AURORA-2 のデータ作成方法と同一の手法により音声に雑音を重畳した。そのために必要となる雑音信号 (Subway, Babble, Car, Exhibition, Restaurant, Street, Air-

表 1: 数字の読み

Digit	AURORA-2	AURORA-2J
1	one	/ichi/
2	two	/ni/
3	three	/saN/
4	four	/yoN/
5	five	/go/
6	six	/roku/
7	seven	/nana/
8	eight	/hachi/
9	nine	/kyuH/
Z	zero	/zero/
O	oh	/maru/

port, Station の 8 種類), 各種フィルタ (G.712, MIRS の 2 種類), およびソースプログラムとスクリプトファイルは, すべて AURORA プロジェクトから提供を受けた. この提供より AURORA-2J では, AURORA-2 と全く同じ条件によるデータの作成を実現した. AURORA-2J では, 収録した 16kHz の日本語音声信号を 8kHz にダウンサンプリングした後に, AURORA-2 プロジェクトから提供を受けた雑音信号と各種フィルタを用いて, 雑音重畳作業およびフィルタリング作業を行った. 注意点として, AURORA-2 プロジェクトから提供を受けた雑音重畳プログラムには, オーバーフローを防止するため, 雑音重畳後の信号が 16bit を越えた場合には, その信号の最大値による正規化処理を行っていた. よって AURORA-2J でも AURORA-2 と同様に, 雑音重畳後の信号が 16bit を越えた場合には, その信号の最大値による正規化処理を行っている.

2.4 学習 / テストデータの構成

学習およびテストデータの構成は, AURORA-2 のものをそのまま採用している.

学習データは, clean training (クリーン音声によるモデル学習), multicondition training (雑音重畳音声による学習) 共に 110 名, 8,440 発話 (男女 55 名, 4,220 発話ずつ) である. clean training の場合はこのデータに雑音を重畳しないで学習を行ない, multicondition training の場合は 4 種類の雑音 (Subway, Babble, Car, Exhibition) を 5 種類の SNR レベル (clean, 20dB, 15dB, 10dB, 5dB) で重畳した音声 (各雑音・SNR で 422 発話ずつの学習データとなる) を用いて学習を行なう. チャ

ネルフィルタとして G.712 の中で規定されているフィルタを用いている.

テストセットは大別して,

テストセット A 雑音は Subway, Babble, Car, Exhibition. チャネルフィルタは G.712. clean training ではチャネル条件がクローズ, multicondition training では雑音条件およびチャネル条件がクローズ.

テストセット B 雑音は Restaurant, Street, Airport, Station. チャネルフィルタは G.712. clean training, multicondition training 共にチャネル条件のみクローズ.

テストセット C 雑音は Subway, Street. チャネルフィルタは MIRS. clean training, multicondition training 共にチャネル条件がオープン.

の 3 種類となっている. 基本となるテストデータは 104 名, 4,004 発話 (男女 52 名, 2002 発話ずつ) で, テストセット A / B ではこれを 4 分割し各種雑音を 7 種類の SNR レベル (clean, 20dB, 15dB, 10dB, 5dB, 0dB, -5dB) で重畳, テストセット C では半分の 2,002 発話をさらに 2 分割して各雑音を重畳している (各雑音・フィルタ条件に対して 1,001 発話). 同じ雑音・フィルタ条件ならば, SNR が違って発話内容は同じである.

2.5 評価用スクリプト

評価用ベースラインスクリプトは, AURORA-2 と同様に HTK[9] を用いて HMM の学習および認識実験を行なうよう, AURORA-2 で配布されているスクリプトをベースとして作成されている. HMM トポロジー, 特徴量など複数の条件で実験を行ない, 様々な議論を重ねた結果, AURORA-2 を踏襲する形でベースラインスクリプトの仕様を以下のように定めた.

- スクリプトは sh(bsh) スクリプトであり, 一部 (初期モデル生成プログラムなど) は perl スクリプトで書かれている.
- HMM は先に述べた 10 数字 (11 モデル) と, 長さの異なる 2 種類の無音 (sil, sp) の計 13 モデルである.
- 数字 HMM は 18 状態 (出力分布を持つ状態は 16), 長い無音モデル (sil) は 5 状態 (同じく 3 状態), 短い無音モデル (sp) は 3 状態 (同じく 1 状態) のモデルである. sp の出力分布は sil の真中の状態と共有される.
- 各状態のガウス混合分布は 20 混合 (無音モデルは 36 混合) である.

- ベースラインの特徴パラメータは、HTK の HCopy により特徴抽出された MFCC (12 次元) + Δ MFCC (12 次元) + $\Delta\Delta$ MFCC (12 次元) + log power (1 次元) + Δ power (1 次元) + $\Delta\Delta$ power (1 次元) の計 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、25ms の分析フレーム長、10ms のフレームシフトとする (ただし、64Hz 未満は使用しない: LOFREQ=64 と設定している)。

2.6 ベースライン性能と認識性能比較

表 2 にベースラインとなる認識性能を示す。表の上半分は clean training, 下半分は multicondition training の結果である。各雑音毎の Average は、SNR 20dB ~ 0dB の平均値であることに注意されたい。

この結果は、認識結果を集計するため、共通の Microsoft Excel Spread Sheet として配布する。この Spread Sheet に各機関で得られる認識率を入力すれば、認識実験結果をベースライン性能からの相対性能として自動的に集計することができる。これによって機関毎の認識性能比較を容易に行なうことができる。

3 AURORA-3J

AURORA-3J は、AURORA-3 に対応する実環境 (自動車内) 日本語連続数字とコマンド語のデータベースである。本章では、AURORA-3J コーパスの収録と検討中の評価方法について述べる。

3.1 データ収録

AURORA-3J は、名古屋大学 CIAIR (Center for Integrated Acoustic Information Research) で収録されている自動車内音声コーパス [10] の一部 (日本語連続数字とコマンド語) を用いて構築される。100 話者の約 80,000 単語発話が収録されている。

3.2 評価方法

ETSI の AURORA-3 データベースでは、アイドリング、低速 (市街地) 走行、高速走行の 3 つの環境で音声データの収録が行われており、それぞれの環境において接話マイク (Closed-talk microphone: CT), ハンズフリーマイク (Hands-free microphone: HF) の 2 種類のマイクを用いて収録が行われている。AURORA-3 では、以上の環境で収録された音声データを用いて、Well-matched condition (WM), Medium-mismatched condition (MM), High-mismatched

表 3: AURORA-3 の各評価セットにおける学習データ

評価セット	WM		MM		HM	
	CT	HF	CT	HF	CT	HF
使用マイク						
アイドリング						
低速走行						
高速走行						

表 4: AURORA-3 の各評価セットにおける評価データ

評価セット	WM		MM		HM	
	CT	HF	CT	HF	CT	HF
使用マイク						
アイドリング						
低速走行						
高速走行						

condition (HM) の 3 つの評価セットが定義されており、それぞれ評価データ、学習データには表 3, 4 に示すようなデータが用いられている (表中の印が使用されているデータである)。

AURORA-3J の評価方法としては、AURORA-3 と全く同じ方法をとる予定にしている。AURORA-3J にて収録された音声データは AURORA-3 に含まれる収録環境を全て網羅しているため、表 3, 4 に示した AURORA-3 と全く同じ評価環境を作成することが可能である。また、AURORA-3 に収録された音声データは数字音声のみであるが、AURORA-3J には数字音声に加えて単語音声が含まれている。このため、AURORA-3J では、数字音声のみの評価と単語音声のみの評価の 2 つの評価を行うこととする。

数字音声の評価では、収録された 100 名分の数字音声を学習データと評価データに分割し、表 3, 4 に示したような評価環境を構築する。評価の際の音響モデルの構造は AURORA-3 と同様とし、以下のような構造を持つものとする。

- 数字 HMM: 16 状態 3 混合分布 (Whole word HMM)
- 無音 HMM: 3 状態 6 混合分布
- ショートポーズ HMM: 1 状態 6 混合分布 (無音 HMM の第 2 状態と共有)

単語音声の評価は、音素 HMM を用いた単語音声認識により評価を行うが、AURORA-3J に含まれる音声データのみでは、学習データが不足するので、外部の音声データを利用する。現在、名古屋大学にて収録された車内音声データベースを利用

表 2: AURORA-2J ベースライン認識性能

Clean Training (%Acc)														
	A					B					C			Overall
	Subway	Babble	Car	Exhibition	Average	Restaurant	Street	Airport	Station	Average	Subway M	Street M	Average	
Clean	99.72	99.58	99.82	99.60	99.68	99.72	99.58	99.82	99.60	99.68	99.82	99.67	99.75	99.69
20 dB	96.90	80.80	89.59	95.90	90.80	84.86	88.51	82.17	82.29	84.46	91.50	92.26	91.88	88.48
15 dB	76.27	56.83	58.16	75.41	66.67	61.10	65.39	57.80	55.01	59.83	70.80	75.39	73.10	65.22
10 dB	47.16	38.63	38.86	41.65	41.58	40.50	42.59	41.93	37.98	40.75	43.51	47.28	45.40	42.01
5 dB	25.27	23.16	20.79	21.97	22.80	21.06	23.79	26.16	22.25	23.32	25.91	25.03	25.47	23.54
0 dB	12.28	8.16	10.38	11.97	10.70	9.89	13.75	12.68	9.84	11.54	13.72	13.60	13.66	11.63
-5 dB	7.43	4.35	7.25	7.90	6.73	1.90	8.56	4.77	5.46	5.17	8.81	8.74	8.78	6.52
Average	51.58	41.52	43.56	49.38	46.51	43.48	46.81	44.15	41.47	43.98	49.09	50.71	49.90	46.17

Multicondition Training (%Acc)														
	A					B					C			Overall
	Subway	Babble	Car	Exhibition	Average	Restaurant	Street	Airport	Station	Average	Subway M	Street M	Average	
Clean	99.79	99.64	99.67	99.75	99.71	99.79	99.64	99.67	99.75	99.71	99.69	99.55	99.62	99.69
20 dB	99.63	99.67	99.70	99.57	99.64	98.62	99.46	98.90	97.99	98.74	99.51	99.40	99.46	99.25
15 dB	99.26	99.40	99.37	98.83	99.22	96.90	97.58	96.45	94.11	96.26	99.17	98.37	98.77	97.94
10 dB	98.25	97.43	97.94	97.38	97.75	86.83	89.57	91.29	84.94	88.16	96.90	93.71	95.31	93.42
5 dB	93.89	89.78	92.16	92.32	92.04	68.56	71.28	77.72	76.18	73.44	87.47	80.86	84.17	83.02
0 dB	74.85	62.48	64.96	73.68	68.99	31.87	48.22	49.36	51.90	45.34	52.32	50.57	51.45	56.02
-5 dB	30.46	25.12	23.17	29.56	27.08	-3.78	18.65	16.70	16.69	12.07	21.31	14.96	18.14	19.28
Average	93.18	89.75	90.83	92.36	91.53	76.56	81.22	82.74	81.02	80.39	87.07	84.58	85.83	85.93

する予定にしておき、このデータベースは連続音声認識コンソーシアムの 2002 年度版ソフトウェアに含まれている車内環境 HMM の学習の際に用いられたデータベースである。音素 HMM の構造等については現在検討を進めている段階である。また、単語音声の評価においても、AURORA-3 と同様の評価環境を構築する。しかし、学習データである車内音声データベースでは、アイドリングと低速走行の音声のみが収録されており、高速走行の音声は収録されていない。このため、単語認識における WM condition では、学習データ、評価データそれぞれアイドリングと低速走行の音声のみを用いて評価を行う。また、HM condition においても高速走行の学習データが必要となるが、これに関しては現在検討中である。

4 評価カテゴリー

これまでに AURORA-2/3 を使用した研究結果が数多く公表されており、Eurospeech, ISCLP ではスペシャルセッションが催されて各機関での性能改善手法が競われている。しかし、AURORA は元来音声認識フロントエンドの改善・評価を目的としたプロジェクトであるにもかかわらず、これらの発表の中にはバックエンド（提供されている学習/テスト用スクリプト）の変更を伴うものが多く存在し（例えば、HMM の混合数を増加する、モデル単位を変更してコンテキスト依存モデルを導入する、AURORA-2 データベースに含まれていない雑音データを使用するなど）、これらを同一に比較評価することが難しくなっている。また、

手法が複合的な場合、各手法の認識性能改善への貢献度の切り分けが難しいという問題もある。

AURORA-2J/3J では、バックエンドの変更に對して、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 0. となる。カテゴリー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と考えられる。

カテゴリー 0. バックエンドスクリプトを全く変更しない場合。

カテゴリー 1. ベースラインスクリプトと同じトポロジーの HMM で、学習プロセスを変更している場合。識別学習の導入等がここに含まれる。このカテゴリーの認識時のコストは、ベースラインと全く同じである。

カテゴリー 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 等がここに含まれる。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。

カテゴリー 3. 混合数や状態数等の HMM のトポロジーを変更している場合。ただし、モデル単位は数字であることを条件とする。状態数が 1 状態より多い雑音モデルを用いた PMC 等がこれに相当する。

カテゴリー 4. 認識時の計算時間 (CPU 時間) がベースライン環境で使用される時間以下であることを条件にどのような処理も許される場合. 複雑なモデルを用いる代わりに特徴パラメータの次元数を減らす等した場合がこれに相当する.

カテゴリー 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.

カテゴリー B. 提供されるデータ以外のデータを使用する場合. もちろん, 評価データは提供されているものを用いるが, このカテゴリーはカテゴリー 1.~5. とは本質的に異なるものである.

5 まとめと今後の計画

本稿では, 日本語連続数字認識をタスクとした雑音下音声認識評価用データベースである AURORA-2J, およびベースライン評価スクリプトによる評価結果について報告した. また, 自動車雑音下での数字およびコマンド発声データベース (AURORA-3J) の開発進捗についても述べた.

今後は, 雑音下音声データベースと並行して, 実際に雑音を聞きながら発声された連続数字音声データベース AURORA-2.5J の開発を進める. AURORA-2.5J は, ヘッドフォンから AURORA-2(J) と同じ雑音を聞きながら AURORA-2 と同様の連続数字を発声するもので, 加算性雑音はない状態で収録される. 従って, ロンバード効果と加算性雑音の影響を分離して分析できる. AURORA-2 と同じタスクであるため, 雑音加算による雑音下音声認識シミュレーションとの直接比較も可能である.

さらに, 大語彙連続音声認識タスクでのデータベース (AURORA-4J) も検討中である. 大語彙連続音声認識タスクの導入については, 現在の簡単なタスクでも十分な性能が得られないのにタスクを難しくすることに対する批判がある一方で, 音声認識システムのカタログスペックとしての性能評価として賛成意見もあり, 議論が割れているところである. 今後, どのような雑音下音声コーパスを構築し, どのような評価手法を開発すれば良いのか, 様々な立場からの意見を吸い上げ, ワーキンググループとしてまとめていく必要があると考えている.

また, 我々は AURORA-2 および AURORA-2J データベース上で種々の雑音抑圧手法の比較および統合実験を行っている [11, 12]. これらの手法はツールとして容易に使用・組み合わせができるように実装している. これらについても一般に公開する予定である.

参考文献

- [1] <http://elazar.itd.nrl.navy.mil/spine/>
- [2] <http://eurospeech2001.org/ese/NoiseRobust/>
- [3] ETSI standard document, "Speech processing, transmission and quality aspects (STQ); Distributed speech recognition; Front-end feature extraction algorithm; Compression algorithm," ETSI ES 201 108 v1.1.2 (2000-04), 2000
- [4] Hirsh, H. G., Pearce, D., "The AURORA experimental framework for the performance evaluations of speech recognition systems under noisy conditions," ISCA ITRW ASR2000, 2000
- [5] Pearce, D., "Developing the ETSI AURORA advanced distributed speech recognition front-end & What next," Proc. Eurospeech 2001, 2001
- [6] <http://www.elda.fr/>
- [7] 中村, 武田, 黒岩, 山田, 北岡, 山本, 西浦, 藤本, 水町, "SLP 雑音下音声認識評価ワーキンググループ活動報告," 情報処理学会研究報告, 2002-SLP-42-11, 2002.
- [8] 中村, 武田, 黒岩, 山田, 北岡, 山本, 西浦, 藤本, 水町, "SLP 雑音下音声認識評価のための WG: 評価データ収集について," 情報処理学会研究報告, 2002-SLP-45-9, 2003.
- [9] <http://htk.eng.cam.ac.uk/>
- [10] Kawaguchi, N., Takeda, K., et al., "Construction of Speech Corpus in Moving Car Environment," Proc. ICSLP 2000, pp.1281-1284, 2000.
- [11] Yamada, T., Okada, J., Takeda, K., Kitaoka, N., Fujimoto, M., Kuroiwa, S., Yamamoto, K., Nishiura, T., Mizumachi, M., Nakamura, S., "Integration of noise reduction algorithms for Aurora 2 database," Eurospeech 2003, (to appear) 2003.
- [12] 山田, 岡田, 武田, 北岡, 藤本, 黒岩, 山本, 西浦, 水町, 中村, "雑音下音声認識のための複数の前処理手法の統合とその AURORA-2J による評価," 情報処理学会研究報告, 2003-SLP-47-18, 2003.