

TR-I-0299

音声情報処理研究室
デモソフトウェア 解説書 (V1.0)
Demonstration Manual
in ATR ITL/Speech Processing Department

杉山 雅英

1993 年 2 月 24 日

内容梗概

本報告は自動翻訳電話研究所音声情報処理研究室で開発されてきたデモソフトウェアに関する解説書である。HMM, Neural Network, 音声合成に関する各種のデモに関して、そのデモ項目名、make, run の仕方、関連する研究発表文献について述べる。

© ATR Interpreting Telephony Research Labs.

© ATR 自動翻訳電話研究所

音声情報処理研究室
デモソフトウェア 解説書 (V1.0)

田川 博章
SET (株)

杉山 雅英
ATR 自動翻訳電話研究所
音声情報処理研究室

1993年2月24日

目次

1	まえがき	1
1.1	デモの起動の仕方 (demon)	3
2	Speaker-Adapt. Cont. Speech Recognition (CBM)	5
2.1	作成	5
2.2	機能説明	5
2.3	デモソフトウェアの所在	5
2.4	Directory 構造	5
2.5	Make について	8
2.5.1	Hmmlr/	8
2.5.2	ADAPT-TREE/	8
2.5.3	LR2/	8
2.5.4	VQ-CDBOOK/IOLib/	8
2.5.5	VQ-CDBOOK/LPClib/	8
2.5.6	VQ-CDBOOK/	8
2.5.7	Ctrl/	9
2.5.8	Tool/	9
2.5.9	Lib/LPC/	9
2.5.10	Lib/XKRNL11/	9
2.5.11	オプション	9
2.6	実行の仕方	9
2.7	開発メモ	10
2.7.1	original	10
2.7.2	modified	10
2.7.3	Workstation Install Memo	10
2.8	関連研究論文など	10
3	VFS Speaker Adaptation CMD-HMM-LR	13
3.1	作成	13
3.2	機能説明	13
3.3	デモソフトウェアの所在	13
3.4	Directory 構造	13
3.5	Make について	16
3.5.1	Hmmlr	16
3.5.2	Hmmlr1200	16
3.5.3	Xmenu	16
3.5.4	color, color-new	17
3.5.5	devoc	17
3.5.6	mklist	17
3.5.7	sed2	17
3.5.8	nkf	17
3.5.9	tdcmhmm1DEMO	17

3.5.10	mklogt	17
3.5.11	mktdata	17
3.5.12	WavePara34	17
3.6	実行の仕方	17
3.7	シェルスクリプトの構造	17
3.8	開発メモ	18
3.8.1	original	18
3.8.2	modified	18
3.8.3	Workstation Install Memo	18
3.9	関連研究論文など	18
4	Speaker-Adapt. Cont. Speech Recognition	21
4.1	作成	21
4.2	機能説明	21
4.3	デモソフトウェアの所在	21
4.4	Directory 構造	21
4.5	Make について	22
4.5.1	Display	22
4.5.2	libLPC.a	22
4.5.3	cut buff	22
4.5.4	Hmmlrfa-Ogv	22
4.5.5	pcm-fzy-hmmlr	23
4.5.6	map-3s	23
4.5.7	IOlib.a	23
4.5.8	WaveCorr	23
4.6	実行の仕方	23
4.7	開発メモ	23
4.7.1	original coding	23
4.7.2	modified coding	23
4.7.3	Workstation Install Memo	23
4.8	関連研究論文など	23
5	Speaker-Indep.(male) Cont. Speech Recognition	25
5.1	作成	25
5.2	機能説明	25
5.3	デモソフトウェアの所在	25
5.4	Directory 構造	25
5.5	Make について	26
5.5.1	Display	26
5.5.2	libLPC.a	26
5.5.3	cut buff	26
5.5.4	Hmmlrfa-Ogv	26
5.5.5	pcm-fzy-hmmlr	27
5.5.6	map-3s	27
5.5.7	IOlib.a	27
5.5.8	WaveCorr	27
5.6	実行の仕方	27
5.7	開発メモ	27
5.7.1	original coding	27
5.7.2	modified coding	27
5.7.3	Workstation Install Memo	27
5.8	関連研究論文など	27

6	AdaptCMixHMMLR	29
6.1	作成	29
6.2	機能説明	29
6.3	デモソフトウェアの所在	29
6.4	Directory 構造	29
6.5	Make について	36
6.5.1	Hmmlr	36
6.5.2	tdcmhmm1	36
6.5.3	tdcmhmm1DEMO	36
6.5.4	tdcmhmm1N	36
6.5.5	tdcmhmm1P	36
6.5.6	mklogt	36
6.5.7	segmentNew	36
6.5.8	color	37
6.5.9	devoc	37
6.5.10	mklist	37
6.5.11	sed2	37
6.5.12	mktdata	37
6.5.13	WavePara34	37
6.6	実行の仕方	37
6.7	開発メモ	37
6.7.1	original	37
6.7.2	modified	37
6.7.3	Workstation Install Memo	37
6.8	関連研究論文など	37
7	HMM-HARD	41
7.1	作成	41
7.1.1	旧版 (話者適応:codebook mapping)	41
7.1.2	新版 (話者適応: VFS)	41
7.2	機能説明	41
7.2.1	旧版	41
7.2.2	新版	41
7.3	デモソフトウェアの所在	41
7.4	Directory 構造	42
7.5	開発メモ	42
7.5.1	original coding	42
7.5.2	modified coding	42
7.5.3	Workstation Install Memo	42
7.6	関連研究論文など	42
8	SSS-based Speaker-Adapt. Cont. Speech Recognition	45
8.1	作成	45
8.2	機能説明	45
8.3	デモソフトウェアの所在	45
8.4	Directory 構造	45
8.5	Make について	47
8.5.1	Adapt	47
8.5.2	Draw	47
8.5.3	SSS-LR.exe	47
8.5.4	Speaker-select	47
8.5.5	WavePara	47

8.6	実行の仕方	47
8.7	開発メモ	48
8.7.1	original	48
8.7.2	modified	48
8.7.3	Workstation Install Memo	48
8.8	関連研究論文など	48
9	Word-bigram-based Sentence Speech Recognition	49
9.1	作成	49
9.2	機能説明	49
9.3	デモソフトウェアの所在	49
9.4	Directory 構造	49
9.5	Make について	50
9.6	実行の仕方	50
9.7	開発メモ	50
9.7.1	original	50
9.7.2	modified	50
9.7.3	Workstation Install Memo	50
9.8	関連研究論文など	50
10	Word-trigram-based Sentence Speech Recognition	53
10.1	作成	53
10.2	機能説明	53
10.3	デモソフトウェアの所在	53
10.4	Directory 構造	53
10.5	Make について	53
10.6	実行の仕方	53
10.7	開発メモ	54
10.7.1	original	54
10.7.2	modified	54
10.7.3	Workstation Install Memo	54
10.8	関連研究論文など	54
11	Phoneme Recognition using a TDNN	55
11.1	作成	55
11.2	概要説明	55
11.3	デモソフトウェアの所在	55
11.4	Directory 構造	55
11.5	Make について	56
11.6	実行の仕方	56
11.7	開発メモ	56
11.7.1	original	56
11.7.2	modified	56
11.7.3	Workstation Install Memo	56
11.8	関連研究論文など	56
12	Phoneme Spotting using TDNNs	59
12.1	作成	59
12.2	概要説明	59
12.3	デモソフトウェアの所在	59
12.4	Directory 構造	59
12.5	make について	61

12.6 実行の仕方	61
12.7 開発メモ	61
12.7.1 original	61
12.7.2 modified	61
12.7.3 Workstation Install Memo	61
12.8 関連研究論文など	61
13 PD-TDNN for Phoneme Recognition	63
13.1 作成	63
13.2 機能説明	63
13.3 デモソフトウェアの所在	63
13.4 Directory 構造	63
13.5 Make について	64
13.6 実行の仕方	64
13.7 開発メモ	64
13.7.1 original	64
13.7.2 modified	64
13.7.3 Workstation Install Memo	64
13.8 関連研究論文など	64
14 TDNN-LR Cont. Speech Recognition	65
14.1 作成	65
14.2 機能説明	65
14.3 デモソフトウェアの所在	65
14.4 Directory 構造	65
14.5 Make について	66
14.6 実行の仕方	66
14.7 開発メモ	66
14.7.1 original	66
14.7.2 modified	66
14.7.3 Workstation Install Memo	66
14.8 関連研究論文など	66
15 Speaker-Adaptation + TDNN-LR	69
15.1 作成	69
15.2 機能説明	69
15.3 デモソフトウェアの所在	69
15.4 Directory 構造	69
15.5 Make について	70
15.6 実行の仕方	70
15.7 開発メモ	70
15.7.1 original	70
15.7.2 modified	70
15.7.3 Workstation Install Memo	70
15.8 関連研究論文など	70
16 Speaker-Indep. NN-based Cont. Speech Recognition	71
16.1 作成	71
16.2 機能説明	71
16.3 デモソフトウェアの所在	71
16.4 Directory 構造	71
16.5 Make について	72

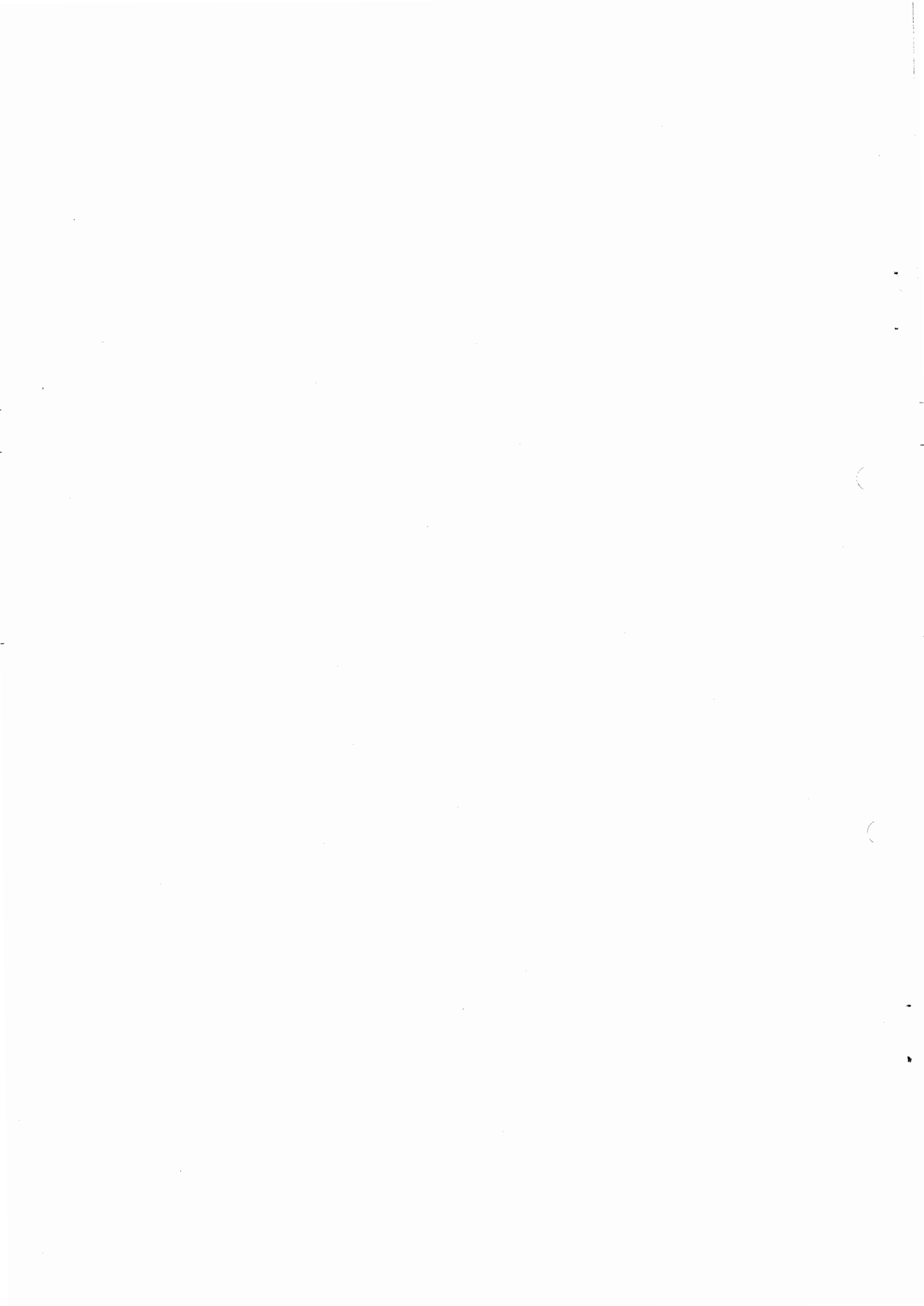
16.6	実行の仕方	72
16.7	開発メモ	72
16.7.1	original	72
16.7.2	modified	72
16.7.3	Workstation Install Memo	72
16.8	関連研究論文など	72
17	FPM-LR	73
17.1	作成	73
17.2	機能説明	73
17.3	Directory 構造	73
17.4	デモソフトウェアの所在	73
17.5	Make について	74
17.5.1	fpm-lr	74
17.5.2	libMyX.a	74
17.5.3	libXKRNL11.a	74
17.5.4	wposition and wait-button	74
17.5.5	SRC-PARA/fpm-lr-p.dec.out	74
17.5.6	SRC-SOCKET/SRC-SERVER/fpm-lr-p.dec.out	74
17.5.7	fpm-lr-client	74
17.6	実行の仕方	74
17.7	開発メモ	74
17.7.1	original	74
17.7.2	modified	74
17.7.3	Workstation Install Memo	75
17.8	関連研究論文など	75
18	A Neural Network for Word Category Prediction	77
18.1	作成	77
18.2	概要説明	77
18.3	デモソフトウェアの所在	77
18.4	Directory 構造	77
18.5	make について	77
18.6	実行の仕方	77
18.7	開発メモ	78
18.7.1	original	78
18.7.2	modified	78
18.7.3	Workstation Install Memo	78
18.8	関連研究論文など	78
19	Noise Reduction Using a Neural Network	79
19.1	作成	79
19.2	機能説明	79
19.2.1	List ファイルのフォーマット	79
19.2.2	起動時の説明	79
19.2.3	Procedure について	80
19.3	デモソフトウェアの所在	80
19.4	Directory 構造	80
19.5	run の仕方	80
19.6	コンパイルの仕方と注意事項	80
19.7	備考	81
19.8	開発メモ	81

19.8.1 original	81
19.8.2 modified	81
19.8.3 Workstation Install Memo	81
19.9 関連研究論文など	81
20 Neural Network Workbench	83
20.1 作成	83
20.2 機能説明	83
20.3 デモソフトウェアの所在	83
20.4 Directory 構造	83
20.5 Make について	84
20.6 実行の仕方	84
20.7 開発メモ	84
20.7.1 original	84
20.7.2 modified	84
20.7.3 Workstation Install Memo	84
20.8 関連研究論文など	84
21 Speech Synthesis Using Non-Uniform Units	85
21.1 作成	85
21.2 機能説明	85
21.3 デモソフトウェアの所在	85
21.4 Directory 構造	85
21.5 Make について	86
21.5.1 nTalk-input	86
21.5.2 nTalk-menu	86
21.5.3 nTalk-select	86
21.5.4 nTalk-spect	86
21.5.5 spect-graph	87
21.5.6 voice, voice-vax	87
21.6 実行の仕方	87
21.7 関連研究論文など	87
21.8 開発メモ	87
21.8.1 original	87
21.8.2 modified	87
21.8.3 Workstation Install Memo	87
22 Voice Conversion Using VQ Codebook Mapping	89
22.1 作成	89
22.2 機能説明	89
22.2.1 List ファイルのフォーマット	89
22.2.2 起動時の説明	89
22.3 デモソフトウェアの所在	90
22.4 Directory 構造	90
22.5 コンパイルの仕方と注意事項	91
22.6 実行の仕方	91
22.7 備考	91
22.8 開発メモ	92
22.8.1 original	92
22.8.2 modified	92
22.8.3 Workstation Install Memo	92
22.9 関連研究論文など	92

23 ATR音声言語翻訳実験システム (ASURA)	93
23.1 作成	93
23.2 機能説明	93
23.3 Directory 構造	94
23.3.1 atrp11	94
23.3.2 as22	98
23.3.3 atrp19	98
23.3.4 atrp17	100
23.4 Make について	100
23.5 実行の仕方 及び、注意	100
23.6 開発メモ	100
23.6.1 original	100
23.6.2 modified	100
23.6.3 Workstation Install Memo	100
23.7 関連研究論文など	100
24 デモ開発用ツール	103
24.1 一般用ツール	103
24.2 研究用ツール	103
24.3 ASURA デモ用ツール	103
24.4 制御用ツール (シェルスクリプト)	104
24.5 プログラム開発用, lib, その他	104
24.6 ファイルの所在	104
25 謝辞	105

表目次

1.1	ソフトウェア早見表 (Hidden Markov Model)	1
1.2	ソフトウェア早見表 (Neural Network Model(1))	1
1.3	ソフトウェア早見表 (Neural Network Model(2))	2
1.4	ソフトウェア早見表 (Speech Synthesis)	2
1.5	デモソフトの各種計算機での動作状況	2
25.1	執筆分担	105



第 1 章

まえがき

本ソフトウェア マニュアルは ATR 自動翻訳電話研究所 音声情報処理研究室で開発された計算機ソフトウェアに関する解説書である。表 1.1 に HMM (Hidden Markov Model) に基づく音声認識系、表 1.2 に NN (Neural Network Model) に基づく音声認識系、表 1.3 に設計ツール、表 1.4 に音声合成系を示す。

表 1.1: ソフトウェア早見表 (Hidden Markov Model)

章	項目	略称	HMM	認識対象	話者	備考
2	11	HMM-LR	離散	文節音声	符号帳写像	Fuzzy VQ (3 codebooks)
3	12	VFS-CHMM	連続	文節音声	VFS	不特定モデルも可
4	14	VFS-DHMM	離散	文節音声	VFS	
5		SI-DHMM	離散	文節音声	不特定	男性話者のみ
6	15	VFS-CHMM-new	連続	文節音声	VFS	項目 12 との差は不明
7	17	HMM-HARD	離散	文節発声, 文	符号帳写像 or VFS	ハードウェア
8	20	SSS-LR	連続 (SSS)	文節音声	VFS	ASURA の認識部
9	18	HMM-bigram	連続	文	不特定	単語 bigram
10	18'	HMM-trigram	連続	文	不特定	単語 trigram
23	19	SL-trans.A	連続	文節発声, 文	VFS	翻訳, 合成と統合

表 1.2: ソフトウェア早見表 (Neural Network Model(1))

章	項目	略称	構造	認識対象	話者	備考
11	1	QUIZ	TDNN	音素	特定	off-line 音声
12	2	SPOT2	TDNN	音素	特定	off-line 音声
13	6	PD-TDNN	PD-TDNN	音素	特定	off-line 音声
14	7	TDNN-LR	TDNN-LR	文節音声	特定	off-line 音声
15	8	AD-TDNN-LR	TDNN-LR	文節音声	話者適応	セグメント話者適応
16	9	IND-TDNN-LR	TDNN-LR	文節音声	不特定	
17	16	FPM-LR	FPM-LR	文節発声, 文	不特定	socket 通信, Multi-CPU

表 1.3: ソフトウェア早見表 (Neural Network Model(2))

章	項目	略称	備考
18	3	NETgram	単語カテゴリー予測
19	4	NR	雑音除去
20	13	NNWB	Neural Network 設計・評価用ツール

表 1.4: ソフトウェア早見表 (Speech Synthesis)

章	項目	略称	備考
21	10	nuuTALK	Non-Uniform Units による合成
22	5	VC	符号帳写像による声質変換

現時点 (93年2月) での各デモソフトウェアの各種計算機・ワークステーションでの動作状況の一覧を以下に示す。

表 1.5: デモソフトの各種計算機での動作状況

略称	項目	DEC	DEC	DEC	HP	SUN
		5000/200	5000/240	Alpha	9000/750	
QUIZ	1	△	○			×
SPOT2	2	△	○			×
NETgram	3	△	○		○	×
NR	4	△	○		○	×
VC	5	△	○		○	×
PD-TDNN	6	△	○			×
TDNN-LR	7	△	○			×
AD-TDNN-LR	8	△	○			×
IND-TDNN-LR	9	△	○		○	×
nuuTALK	10	△	○			×
HMM-LR	11	△	○			×
VFS-CHMM	12	△	○		○	×
NNWB	13	△	○		○	×
VFS-DHMM	14	△	○		○	×
VFS-CHMM-new	15	△	○			×
FPM-LR	16	△	○	?	○	×
HMM-HARD	17	△	○			×
HMM-bigram	18	?	○			×
SL-TRANS	19	?	○			×
SSS-LR	20	△	○	?	○	×
HMM-trigram	22	?	○			×

1.1 デモの起動の仕方 (demox)

ここで説明する demox で起動できるデモは、全て atrq30 で動作するものである。

- 起動シェル名

demox

- ホストマシン名

atrq30 (DEC machine)

- ディレクトリ名

/q30_data/users/demo/

- 操作方法

`demox` を立ちあげるとデモのメニューウインドウが表示されます。メニュー上でマウスのボタンをクリックし選択するとデモが起動します。また、`**quit**` を選択すると demox の環境設定のメニューウインドウが表示され `QuitAll` で終了します。

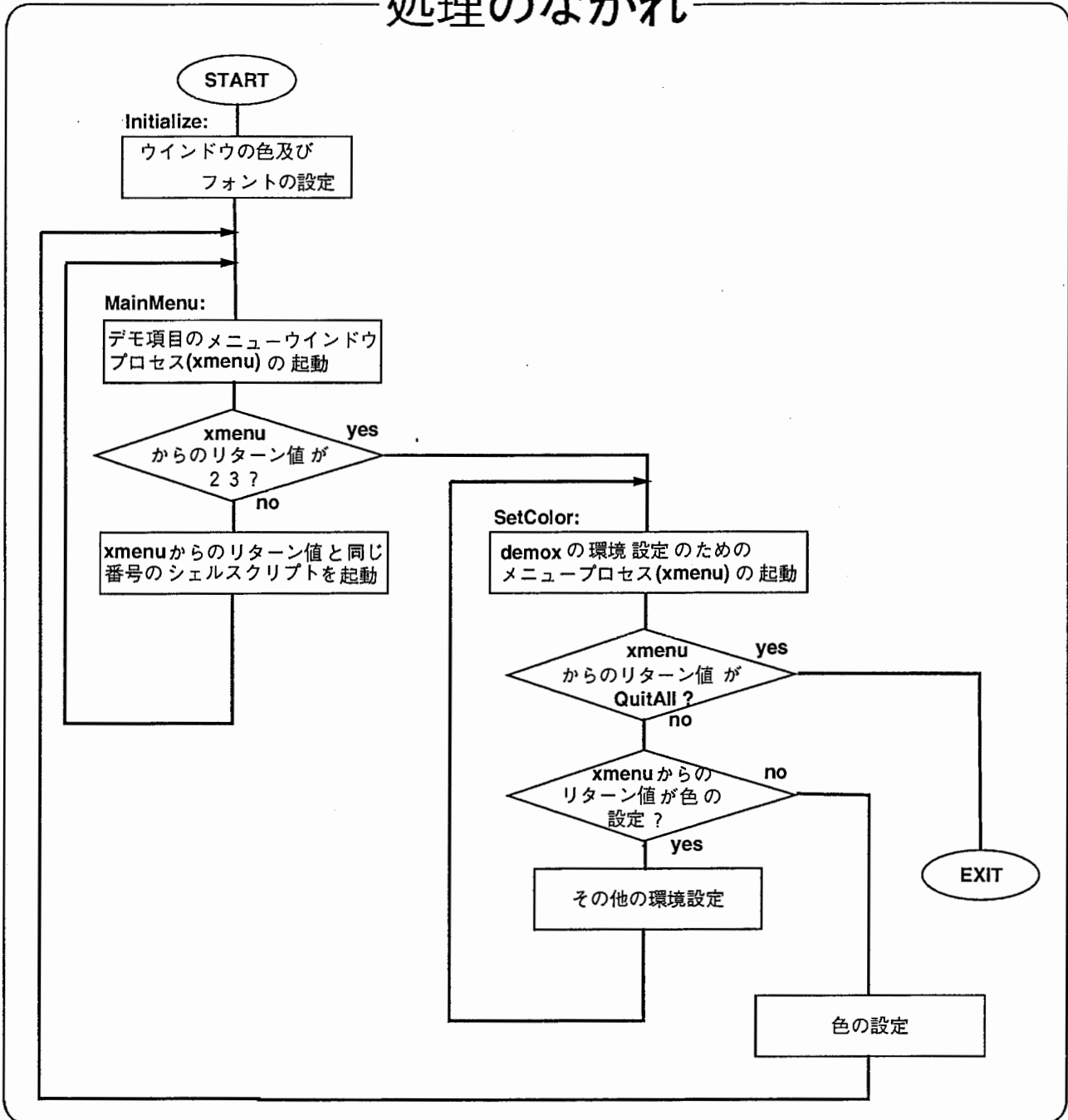
- 概要説明

demox はデモ起動用のシェルスクリプトです。メニューウインドウには、xmenu(tool) を使用しています。メニューで選択されたデモの項目番号と同じ番号のシェルスクリプトを立ちあげることによりデモを起動します。

(例：デモ番号 10 を選択、シェルスクリプト 10 が立ちあがる。)

デモの項目番号と同じ番号のシェルスクリプトには、その番号のデモを起動するためのコマンドが書かれています。

処理のながれ



第 2 章

Speaker-Adapt. Cont. Speech Recognition (CBM)

2.1 作成

作成者 :

オリジナル版: ??

カラー版: 伴?, 高嶋? (1990)

最新版: 大倉

作成日 :

2.2 機能説明

ファジィ VQ に基づく離散 HMM を用いた、文節認識の HMM-LR 連続音声認識システム。話者適応は、コードブックマッピング (適応データは 25 単語)。解析途中の構文仮説を tree 形式で表示する。

2.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

2.4 Directory 構造

```
HMM-LR --- ACP/
|   (C source programs)
|
+-- ADAPT_TREE/
|   (C source programs :Display)
|
+-- Adaptation/
|   (C make up files :SI_SUN)
|
+-- Ctrl/
|   (C source programs)
|
+-- Data/ --- mht/ --- CDBook/
|   (data files) |
|                 |
|                 +-- COR/
|                 |
|                 |
|                 +-- Dur/ --- FVQ/
```



```
|
+-- Hmmlr/
|   (C source programs :Hmmlrfa_Ogv)
|
+-- LR2/
|   (C source programs)
|
+-- Lib/ --- LPC/
|       |   (C source programs :libLPC.a)
|       |
|       +-- XKRNL11/
|           (C source programs :libXKRNL11.a)
|
+-- Log/
|
|
+-- Sound/ --- MANUAL/
|       |   (C source programs)
|       |
|       +-- SEMIAUTO/
|           (C source programs)
|       |
|       +-- BAK_SRC/
|           (C source programs)
|
+-- Sp_Data/
|   (speakers data)
|
+-- TEST/
|
|
+-- Tool/ --- Adapt_25/
|       |
|       |
|       +-- ETC_X11_MY_LIB/
|           (C source programs :libMyX.a)
|       |
|       +-- JXMENU/
|           (C source programs :jxmenu)
|       |
|       +-- JXMENU_BAK/
|           |
|           |
|       +-- <<C source program>>
|           -> menu_cont
|           -> wait_r_ebl
|           -> wait_w_ebl
|           -> wr_cf
|
+-- VQ_CDBOOK/ --- IOlib/
|       |   (C source programs :IOlib.a)
```

```

|
|
|      +-- LPClib/
|      |      (C source programs :libLPC.a)
|      |
|      +-- <<C source programs>>
|      -> WaveCorr
|      -> vq_cdbook
|      -> cdmap_fz_3s
|
|
+-- <<シェルスクリプト>>
|   Start_Hmmlr.csh
|   Demo.csh
|   adapt_menu.csh
|   adapt_menu_new.csh
|   Hmmlr_Only.csh
|   hmmlr_exe.csh
|
+-- <<その他>>
    Makefile    tmp

```

2.5 Make について

HMM-LR/ の下において Makefile を実行すると、必要とするすべての実行イメージが作成される。また、作成された後、directory EXE/ の下へ move されるものもある。作成されるものは、以下のとおりである。また、各ディレクトリーの下にある Makefile を直接、実行することによっても作成できる。

2.5.1 Hmmlr/

directory Hmmlr/ の下において hmmlr.make が実行され、Hmmlrfa_Ogv が作成される。

2.5.2 ADAPT-TREE/

directory ADAPT_TREE/ の下において Makefile が実行され、Display が作成される。

2.5.3 LR2/

directory LR2/ の下において Makefile が実行され、オブジェクトファイルが作成される。

2.5.4 VQ-CDBOOK/Iolib/

directory VQ_CDBOOK/Iolib/ の下において Makefile が実行され、Iolib.a が作成される。

2.5.5 VQ-CDBOOK/LPClib/

directory VQ_CDBOOK/LPClib/ の下において Makefile が実行され、libLPC.a が作成される。

2.5.6 VQ-CDBOOK/

directory VQ_CDBOOK/ の下において、WaveCorr.make が実行され、WaveCorr が作成される。vq_cdbook.make が実行され、vq_cdbook が作成される。cdmap_fz_3s.make が実行され、cdmap_fz_3s が作成される。

2.5.7 Ctrl/

directory Ctrl/ の下において、

```
% make wait_cut_buff_r_ebl
% make wait_cut_buff_w_ebl
% make write_cut_buff
% make read_cut_buff
```

が実行され、wait_cut_buff_r_ebl wait_cut_buff_w_ebl write_cut_buff read_cut_buff が作成される。

2.5.8 Tool/

directory Tool/ の下において、

```
% make menu_cont
% make wait_r_ebl
% make wait_w_ebl
% make wr_cf
```

が実行され、menu_cont wait_r_ebl wait_w_ebl wr_cr が作成される。

2.5.9 Lib/LPC/

directory Lib/LPC/ の下において、Makefile が実行され、libLPC.a が作成される。

2.5.10 Lib/XKRNL11/

directory Lib/XKRNL11/ の下において、Makefile が実行され、libXKRNL11.a が作成される。

2.5.11 オプション

HMM-LR/ の下の Makefile を実行する際に、

```
% make hmmlr
% make display
% make lr2
% make vqcod
% make cut_buf
% make tool
% make library
```

と指定することにより各ディレクトリーの下にある Makefile を実行することができる。

```
% make clear
```

と指定することにより各ディレクトリーの下にあるオブジェクトファイルをすべて消去することができる。

2.6 実行の仕方

<実行形式>

```
% Start_Hmmlr.csh
```

2.7 開発メモ

2.7.1 original

HMM-LR -> 北研二、川端豪、花沢利行

adapt. -> 中村哲

2.7.2 modified

HMM-LR -> 北研二?

adapt. -> 大倉計美

2.7.3 Workstation Install Memo

2.8 関連研究論文など

1. 北研二, 川端豪, 斎藤博昭: "HMM音韻認識と拡張LR構文解析法を用いた連続音声認識," 情報処理学会論文誌, Vol. 31, No. 3, pp. 472-480, March, 1990.
2. 花沢利行, 北研二, 中村哲, 川端豪, 鹿野@@69(音声認識システムの性能評価," 日本音響学会誌, Vol. 46, No. 10, pp. 817-823, October, 1990.
3. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: "HMM Continuous Speech Recognition Using Predictive LR Parsing," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 703-706, May, 1989.
4. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: "Parsing Continuous Speech by HMM-LR Method," *Proceedings of the International Workshop on Parsing Technologies*, pp. 126-131, August, 1989.
5. Toshiyuki Hanazawa, Kenji Kita, Satoshi Nakamura, Takeshi Kawabata, Kiyohiro Shikano: "ATR HMM-LR Continuous Speech Recognition System," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 53-56, April, 1990.
6. 北研二, 川端豪, 斎藤博昭: "HMM音韻認識と予測LRパーザを用いた文節認識," 電子情報通信学会音声研究会, pp. 63-69, October, 1988.
7. 花沢利行, 北研二, 中村哲, 川端豪, 鹿野清宏: "HMM-LR音声認識システムの性能評価," 電子情報通信学会音声研究会, pp. 63-70, December, 1989.
8. 北研二, 川端豪, 斎藤博昭: "HMM音韻認識と予測LRパーザを用いた文節認識," 日本音響学会秋季研究発表会, pp. 259-260, October, 1988.
9. 花沢利行, 北研二, 川端豪, 鹿野清宏: "HMM音韻モデルの文節認識による評価," 日本音響学会春季研究発表会, pp. 81-82, March, 1989.
10. 川端豪, 鹿野清宏, 北研二: "音韻パープレキシティの提案," 日本音響学会春季研究発表会, pp. 93-94, March, 1989.
11. 北研二, 坂野俊哉, 保坂順子, 川端豪: "SL-TRANSにおける文節音声認識 — HMM音韻認識とLR構文解析法による文節音声認識 —," 情報処理学会第39回全国大会, pp. 718-719, October, 1989.
12. 北研二, 江原暉将, 森元逞: "HMM-LR音声認識の大語彙への適用," 情報処理学会第42回全国大会, pp. 2-110-2-111, March, 1991.
13. Toshiyuki Hanazawa, Kenji Kita, Satoshi Nakamura, Takeshi Kawabata, Kiyohiro Shikano: "ATR HMM-LR Continuous Speech Recognition System," *Readings in Speech Recognition*, Waibel, A and Lee, K. F. Eds., Morgan Kaufmann Publishers, pp. 611-614, 1990.

14. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: "GLR Parsing in Hidden Markov Model," *Generalized LR Parsing*, Ed: Tomita, M, Kluwer Academic Publishers, pp. 153-164, 1991.
15. 北 研二: "自動翻訳電話," 第2章 2.3 [2] 隠れマルコフモデルと文脈自由文法を用いた連続音声認識, ATR先端テクノロジーシリーズ, (1993年1月発行予定)
16. Shikano, K., Lee, K-F., and Reddy, R., "Speaker Adaptation through Vector Quantization," Proc. ICASSP86, 49.5, pp.2643-2646.
17. Nakamura, S. and Shikano, K., "Spectrogram normalization using fuzzy vector quantization," Journal of Acoust. Soc. of Japan, Vol.45, No.2, pp.107-114.
18. 花沢, 川端, 鹿野, "Hidden Markov Model を用いた日本語有声破裂音の識別," TR-I-0018.
19. 花沢, 川端, 鹿野, "Duration control methods for HMM phoneme recognition," TR-I-0050.
20. 平田, 川端, 花沢, "HMM-LR 法を用いた文節認識における継続時間長制御パラメータ変換法の検討," TR-I-0076.
21. 北, 川端, 斉藤, "HMM 音韻認識と拡張 LR 構文解析法を用いた連続音声認識," TR-I-0082.
22. 中村, 鹿野, "ファジイベクトル量子化に基づく話者適応化," TR-I-0096.
23. 中村, 花沢, 鹿野, "ベクトル量子化話者適応の HMM 音韻認識への適応," TR-I-0097.
24. 中村, "ベクトル量子化話者適応化の研究," TR-I-0100.
25. 中村, 鹿野, "話者適応化における写像方法の比較," TR-I-0101.
26. 花沢, 川端, 鹿野, "Hidden Markov Model による音韻認識実験の結果," TR-I-0147.
27. 北, "Generalized LR Parsing in Hidden Markov Model," TR-I-0161.
28. 永田, 衛藤, 保坂, "音声認識のための構文規則ガイドブック," TR-I-0240.
29. 保坂, 竹澤, "SL-TRANS における音声認識のための構文規則の拡張," TR-I-0241.
30. 北, "HMM-LR ユーザーズ・マニュアル," TR-I-0246.
31. 北, "HMM-LR ソース・コード," TR-I-0270.
32. 北, "HMM-LR 音声認識プログラム ユーザーズ・マニュアル," TR-I-0271.
33. 北, "A Study on Language Modeling for Speech Recognition," TR-I-0273.

第 3 章

VFS Speaker Adaptation CMD-HMM-LR

3.1 作成

作成者：大倉 計美

作成日：1992年9月

3.2 機能説明

話者適応は、VFS 方式を採用。適応データは、発話様式適応を行なうため、文節発声データを使用(最大20文章)。DASBOXでA/D、WS(HP)上のソフトウェアで音響分析し、話者適応を行なう。HMMは連続混合出力分布型HMM(CDHMM)を用いる。認識時には、適応化された発話者の音素モデル(CDHMM)を用い、入力データを、DASBOXでA/D、WS(HP)上のソフトウェアで音響分析し、HMM-LR(ソフトウェア)によって文節認識を行なう。また、不特定話者(男性)モデルを用いた文節認識も行なえる。

3.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

3.4 Directory 構造

```
Adapt.CHMM/ ---+--- Data/
|          (users data)
|
+--- Dur/  ---+--- mht/  ---+--- 10B/
|          |            |
|          |            +--- 12B/
|          |            +--- 14B/
|          |            +--- 15B/
|          |            +--- 2B/
|          |            +--- 3B/
|          |            +--- 5B/
|          |            +--- 7B/
|          |            +--- IND/
|          |
|          +--- female/
|          +--- male/
|
+--- EXE/  ---+--- Src/  ---+--- ConcatTrain2/
|          |            |          (C source programs)
|          |            |          -> tdcmhmm1DEMO
|          |            |
```

```
|
|
|      +-- Hmmlr/
|      |   (C source programs)
|      |   -> Hmmlr
|      |   -> Hmmlr1200
|      |
|      +-- LOG/
|      |   (Log file)
|      |   (C source programs)
|      |   -> mklogt
|      |
|      +-- NKF/
|      |   (C source program)
|      |   -> nkf
|      |   (その他)
|      |
|      +-- RFILTER/
|      |   (C source program)
|      |
|      +-- Tool/
|      |   (C source programs)
|      |   -> color
|      |   -> color_new
|      |   -> devoc
|      |   -> mklist
|      |   -> sed.2
|      |
|      +-- XMENU/
|      |   (C source program)
|      |   -> Xmenu
|      |
|      +-- XMENU_ORG/
|      |   (XMENU Old Version)
|      |
|      +-- segmentNew/
|      |   (Old Version)
|
+-- Data/
|
+-- List/
|   (Symbolic link to ListF)
|
+-- ListF/
|   (発声リストファイル)
|
+-- ListSB1/
|   (発声リストファイル)
|
+-- BAK_CSH/
|   (Back up file, Old file)
|
```

```

|      +-- Orig/
|      |      (Back up file, Old file)
|      |
|      +-- <<実行イメージ>>
|      |      Hmmlr      Hmmlr1200
|      |      SpeechIn  Xmenu
|      |      color      color_new
|      |      devoc      mklist
|      |      nkf        sed2
|      |      tdcnmhmm1DEMO
|      |
|      +-- <<シェルスクリプト>>
|      |      Analysis   Demo
|      |      Hassei_List.csh
|      |      HmmLR      Lg.csh
|      |      Lg.log.csh  LgING.csh
|      |      Phrase_list.csh
|      |      Selected_Sentence
|      |      Selected_Words
|      |      Speaker_Adaptation
|      |      Speaker_Adaptation_Help
|      |      Word_Registration
|      |      Word_Registration_Help
|      |      list.csh    menu.csh
|      |
|      +-- <<Help Text>>
|      |      SA_Help.doc  WR_Help.doc
|      |
|      +-- <<その他>>
|      |      README.NEW
|
+-- Hmm/  --+-- mht/  --+-- 10B/
|          |          +-- 12B/
|          |          +-- 14B/
|          |          +-- 15B/
|          |          +-- 2B/
|          |          +-- 3B/
|          |          +-- 5B/
|          |          +-- 7B/
|          |
|          +-- female/
|          +-- male/
|
+-- HmmLR/ --+-- Gra/
|           |
|           |
|           +-- TEST/
|           |
|           |

```

```

|           +-- WAVE/
|           |
|           |
|           +-- << * * * * * >>
|           pointer
|
|
+-- MKTDATA/
|   (C source program)
|   -> mktdata
|   (その他)
|
+-- PROG/
|   (C source program)
|   -> WavePara34
|
+-- SPKR/  ---+-- BUF/
|           |
|           |
|           +-- ListF/
|           |
|           |
|           +-- ListSB1/
|           |
|           |
|           +-- <<その他>>
|
+-- TEXT/  ---+-- 11/
|           |
|           |
|           +-- 20/
|           |
|           |
|           +-- <<その他>>
|
+-- <<シェルスクリプト>>
    DEMO

```

3.5 Make について

3.5.1 Hmmlr

directory EXE/Src/Hmmlr/ において Makefile を実行し、実行イメージを EXE/ へ置いてください。

3.5.2 Hmmlr1200

directory EXE/Src/Hmmlr/ において Makefile1200 を実行し、実行イメージを EXE/ へ置いてください。

3.5.3 Xmenu

directory EXE/Src/XMENU/ において xmenu.c をコンパイルし、実行イメージを EXE/ へ置いてください。

3.5.4 color, color-new

directory EXE/Src/Tool/ において color.c もしくは、color_new.c をコンパイルし、実行イメージを EXE/ へ置いてください。ただし、color.c は DEC マシン用で、color_new.c は HP マシン用です。

3.5.5 devoc

directory EXE/Src/Tool/ において devoc.c をコンパイルし、実行イメージを EXE/ へ置いてください。

3.5.6 mklist

directory EXE/Src/Tool/ において mklist.c をコンパイルし、実行イメージを EXE/ へ置いてください。

3.5.7 sed2

directory EXE/Src/Tool/ において sed2.c をコンパイルし、実行イメージを EXE/ の下へ置いてください。

3.5.8 nkf

directory EXE/Src/NKF/ において nkf.c をコンパイルし、実行イメージを EXE/ へ置いてください。

3.5.9 tdcnhmm1DEMO

directory EXE/Src/ConcatTrain2/ において tdcnhmm1DEMO.make を実行し、実行イメージを EXE/ へ置いてください。

3.5.10 mklogt

directory EXE/Src/LOG/ において mklogt.make を実行してください。

3.5.11 mktdata

directory MKTDATA/ において mktdata.c をコンパイルしてください。

3.5.12 WavePara34

directory PROG/ において WavePara34.c をコンパイルしてください。

3.6 実行の仕方

<実行形式>

```
% DEMO
```

3.7 シェルスクリプトの構造

```
DEMO +
  |
  +
  EXE/Demo +
    |
    +
    EXE/menu.csh
      ( from Speaker_Adaptation menu )
      : Voice_Registration
```

```

( from Voice_Registration menu )
: All_Sentence
: Add_Sentence
    EXE/Word_Resistration +
    |
    +
    EXE/list.csh
: Selected_Phrase
    EXE/Selected_Words
: Selected_Sentence
    EXE/Selected_Sentence
: List
    EXE/Hassei_List.csh
: Speaker_Adaptation
    EXE/Analysis
    EXE/Speaker_Adaptation
: Recognition
    EXE/HmmLR
: Help
    EXE/Speaker_Adaptation_Help
(: Change_Speaker )

```

```
no using : Analysis_dummy
```

3.8 開発メモ

3.8.1 original

大倉計美, Adapt.CHMM/EXE/README.NEW, Adapt.CHMM/EXE/menu_new.csh を参照

3.8.2 modified

なし

3.8.3 Workstation Install Memo

3.9 関連研究論文など

1. 大倉計美, 杉山雅英, 嵯峨山茂樹: “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” 連続音声認識シンポジウム, pp.49-50, (1992.02).
2. 大倉計美, 杉山雅英: 移動ベクトル場平滑化話者適応方式を用いた雑音環境下音声認識,” 音学講論, 2-Q-16, pp. 189-190, (1992.3).
3. 大倉計美, 杉山雅英, 嵯峨山茂樹: “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” 音学講論, 2-Q-17, pp. 191-192, (1992.3).
4. 大倉計美, 杉山雅英, 嵯峨山茂樹: “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” 信学技報, SP92-16, pp. 23-28, (1992.6).
5. Kazumi Ohkura, Masahide Sugiyama and Sigeki Sagayama: “Speaker Adaptation Based on Transfer Vector Field Smoothing with Continuous Mixture Density HMMs,” Proc. of ICSLP92, (Oct.1992).

6. 山口耕市, 嵯峨山茂樹: 混合連続分布型 HMM を用いた HMM-LR 連続音声認識, 音学講論, 1-P-5, pp. 113-114, (1992.3).
7. 北 研二, 川端 豪, 斎藤 博昭: “HMM 音韻認識と拡張 LR 構文解析法を用いた連続音声認識,” 情報処理学会論文誌, Vol. 31, No. 3, pp. 472-480, March, 1990.
8. 花沢 利行, 北 研二, 中村 哲, 川端 豪, 鹿野 清宏: “HMM-LR 音声認識システムの性能評価,” 日本音響学会誌, Vol. 46, No. 10, pp. 817-823, October, 1990.
9. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: “HMM Continuous Speech Recognition Using Predictive LR Parsing,” *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 703-706, May, 1989.
10. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: “Parsing Continuous Speech by HMM-LR Method,” *Proceedings of the International Workshop on Parsing Technologies*, pp. 126-131, August, 1989.
11. Toshiyuki Hanazawa, Kenji Kita, Satoshi Nakamura, Takeshi Kawabata, Kiyohiro Shikano: “ATR HMM-LR Continuous Speech Recognition System,” *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 53-56, April, 1990.
12. 北 研二, 川端 豪, 斎藤 博昭: “HMM 音韻認識と予測 LR パーザを用いた文節認識,” 電子情報通信学会音声研究会, pp. 63-69, October, 1988.
13. 花沢 利行, 北 研二, 中村 哲, 川端 豪, 鹿野 清宏: “HMM-LR 音声認識システムの性能評価,” 電子情報通信学会音声研究会, pp. 63-70, December, 1989.
14. 北 研二, 川端 豪, 斎藤 博昭: “HMM 音韻認識と予測 LR パーザを用いた文節認識,” 日本音響学会秋季研究発表会, pp. 259-260, October, 1988.
15. 花沢 利行, 北 研二, 川端 豪, 鹿野 清宏: “HMM 音韻モデルの文節認識による評価,” 日本音響学会春季研究発表会, pp. 81-82, March, 1989.
16. 川端 豪, 鹿野 清宏, 北 研二: “音韻パープレキシティの提案,” 日本音響学会春季研究発表会, pp. 93-94, March, 1989.
17. 北 研二, 坂野 俊哉, 保坂 順子, 川端 豪: “SL-TRANS における文節音声認識 — HMM 音韻認識と LR 構文解析法による文節音声認識 —,” 情報処理学会第 39 回全国大会, pp. 718-719, October, 1989.
18. 研二, 江原 暉将, 森元 逞: “HMM-LR 音声認識の大話彙への適用,” 情報処理学会第 42 回全国大会, pp. 2-110-2-111, March, 1991.
19. Toshiyuki Hanazawa, Kenji Kita, Satoshi Nakamura, Takeshi Kawabata, Kiyohiro Shikano: “ATR HMM-LR Continuous Speech Recognition System,” *Readings in Speech Recognition*, Waibel, A and Lee, K. F. Eds., Morgan Kaufmann Publishers, pp. 611-614, 1990.
20. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: “GLR Parsing in Hidden Markov Model,” *Generalized LR Parsing*, Ed: Tomita, M, Kluwer Academic Publishers, pp. 153-164, 1991.
21. 北 研二: “自動翻訳電話,” 第 2 章 2.3 [2] 隠れマルコフモデルと文脈自由文法を用いた連続音声認識, ATR 先端テクノロジーシリーズ, (1993 年 1 月発行予定)
22. 大倉, 杉山, “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” TR-I-0266
23. 北, “Generalized LR Parsing in Hidden Markov Model,” TR-I-0161.
24. 永田, 衛藤, 保坂, “音声認識のための構文規則ガイドブック,” TR-I-0240.
25. 保坂, 竹澤, “SL-TRANS における音声認識のための構文規則の拡張,” TR-I-0241.
26. 北, “HMM-LR ユーザーズ・マニュアル,” TR-I-0246.

27. 北, "HMM-LR ソース・コード," TR-I-0270.
28. 北, "HMM-LR 音声認識プログラム ユーザーズ・マニュアル," TR-I-0271.
29. 北, "A Study on Language Modeling for Speech Recognition," TR-I-0273.
30. 大倉, 杉山, 嵯峨山, "混合連続分布 HMM 移動ベクトル場平滑化話者適応方式の文節認識性能の評価," TR-I-0284.

第 4 章

Speaker-Adapt. Cont. Speech Recognition

4.1 作成

作成者 : 嵯峨山茂樹
作成日 : 1992 年 7 月

4.2 機能説明

音素モデルとして離散分布型 HMM を用いた HMM-LR 連続音声認識で、話者適応方式は移動ベクトル場平滑化話者適応方式 (VFS) を用いている。

4.3 デモソフトウェアの所在

Machine: atrq30
Directory: /q30_data/users/demo/DEMO_SOFT_DIR

4.4 Directory 構造

```
Adapt.Y
|
+- Data
| |
| +- mht
| | |
| | +- CDBook
| | +- COR
| | +- DAT
| | | |
| | | +- Input
| | | +- Input_Keep
| | |
| | +- Dur
| | +- Hmm
| | +- LAB
| | +- SEQ
| |
| +- si_male
| |
| +- CDBook
| +- COR
```

```

|      +- DAT
|      |  |
|      |  +- Input
|      |  +- Input_Keep
|      |
|      +- Dur
|      +- Hmm
|      +- SEQ
|
+- Drawings
+- EXE
+- FILE
+- Grammer
+- SRC
|  |
|  +- ADAPT_TREE
|  +- CUT_BUF
|  +- FVQ
|  +- Hmmlr
|  +- LPC
|  +- MAP
|  |  |
|  |  +- IO
|  |  +- SCCS
|  |
|  +- VQ_CDBOOK
|
+- sagayama

```

4.5 Make について

4.5.1 Display

Directory Adapt.Y/SRC/ADAPT_TREE/ において Makefile を実行して下さい。実行イメージ Dispaly が作成されます。

4.5.2 libLPC.a

Directory Adapt.Y/SRC/LPC/ において Makefile を実行して下さい。ライブラリー libLPC.a が作成されます。作成された libLPC.a をひとつ上のディレクトリ Adapt.Y/ の下へ置いて下さい。

4.5.3 cut buff

Directory Adapt.Y/SRC/CUT_BUF/ において Makefile を実行して下さい。実行イメージ all_cut_buff と wait_w_ と wr_cf が、それぞれ Directory Adapt.Y/EXE/ に作成されます。

4.5.4 Hmmlrfa-Ogv

Directory Adapt.Y/SRC/Hmmlr/ において hmmlr.make を実行して下さい。実行イメージ Hmmlrfa_Ogv が Directory Adapt.Y/EXE/ に作成されます。

4.5.5 pcm-fzy-hmmlr

Directory Adapt.Y/SRC/FVQ/ において pcm_fzy_hmmlr.make を実行して下さい。実行イメージ pcm_fzy_hmmlr が Directory Adapt.Y/EXE/ に作成されます。

4.5.6 map-3s

Directory Adapt.Y/SRC/MAP/ において map_3s.make を実行して下さい。実行イメージ map_3s が Directory Adapt.T/EXE/ に作成されます。

4.5.7 IOlib.a

Directory Adapt.Y/SRC/MAP/IO/ において IOlib.make を実行して下さい。ライブラリー IOlib.a が Directory ../ に作成されます。

4.5.8 WaveCorr

Directory Adapt.Y/SRC/VQ_CDBOOK/ において WaveCorr.c をコンパイルして下さい。

```
(cc -o ../../EXE/WaveCorr WaveCorr.c)
```

4.6 実行の仕方

<実行形式>

```
% Demo_saga
```

4.7 開発メモ

4.7.1 original coding

服部浩明、高嶋

4.7.2 modified coding

高嶋?、嵯峨山茂樹

4.7.3 Workstation Install Memo

4.8 関連研究論文など

1. 服部浩明, 嵯峨山茂樹: “少数語彙による移動ベクトル場平滑化話者適応方式の文節認識による評価,” 音学講論, 2-Q-15, (1992.03).
2. H. Hattori and S.Sagayama: “Vector Field Smoothing Principle For Speaker Adaptation,” Proc. of ICSLP92, We.fPM.1.4, pp. 381-384, (Oct. 1992).

第 5 章

Speaker-Indep.(male) Cont. Speech Recognition

5.1 作成

作成者 : 大倉計美?
作成日 : 1992 年 7 月

5.2 機能説明

男性の不特定話者を対象とした HMM-LR 連続音声認識で、音素モデルとして離散分布型 HMM を用いている。不特定話者用モデルの作成に際しては、男性話者 10 名分のモデルを学習した。

5.3 デモソフトウェアの所在

Machine: atrq30
Directory: /q30_data/users/demo/DEMO_SOFT_DIR

5.4 Directory 構造

```
Adapt.Y
|
+- Data
| |
| | +- mht
| | |
| | | +- CDBook
| | | +- COR
| | | +- DAT
| | | |
| | | | +- Input
| | | | +- Input_Keep
| | | |
| | | +- Dur
| | | +- Hmm
| | | +- LAB
| | | +- SEQ
| | |
| | +- si_male
| | |
| | | +- CDBook
| | | +- COR
```

```

|      +- DAT
|      | |
|      | +- Input
|      | +- Input_Keep
|      |
|      +- Dur
|      +- Hmm
|      +- SEQ
|
+- Drawings
+- EXE
+- FILE
+- Grammer
+- SRC
| |
| +- ADAPT_TREE
| +- CUT_BUF
| +- FVQ
| +- Hmmlr
| +- LPC
| +- MAP
| | |
| | +- IO
| | +- SCCS
| |
| +- VQ_CDBOOK
|
+- sagayama

```

5.5 Make について

5.5.1 Display

Directory Adapt.Y/SRC/ADAPT_TREE/ において Makefile を実行して下さい。実行イメージ Dispaly が作成されます。

5.5.2 libLPC.a

Directory Adapt.Y/SRC/LPC/ において Makefile を実行して下さい。ライブラリー libLPC.a が作成されます。作成された libLPC.a をひとつ上のディレクトリー Adapt.Y/ の下へ置いて下さい。

5.5.3 cut buff

Directory Adapt.Y/SRC/CUT_BUF/ において Makefile を実行して下さい。実行イメージ all_cut_buff と wait_w_ と wr_cf が、それぞれ Directory Adapt.Y/EXE/ に作成されます。

5.5.4 Hmmlrfa-Ogv

Directory Adapt.Y/SRC/Hmmlr/ において hmmlr.make を実行して下さい。実行イメージ Hmmlrfa_Ogv が Directory Adapt.Y/EXE/ に作成されます。

5.5.5 pcm-fzy-hmmlr

Directory Adapt.Y/SRC/FVQ/ において pcm_fzy_hmmlr.make を実行して下さい。実行イメージ pcm_fzy_hmmlr が Directory Adapt.Y/EXE/ に作成されます。

5.5.6 map-3s

Directory Adapt.Y/SRC/MAP/ において map_3s.make を実行して下さい。実行イメージ map_3s が Directory Adapt.T/EXE/ に作成されます。

5.5.7 IOlib.a

Directory Adapt.Y/SRC/MAP/IO/ において IOlib.make を実行して下さい。ライブラリー IOlib.a が Directory ../ に作成されます。

5.5.8 WaveCorr

Directory Adapt.Y/SRC/VQ_CDBOOK/ において WaveCorr.c をコンパイルして下さい。

```
(cc -o ../../EXE/WaveCorr WaveCorr.c)
```

5.6 実行の仕方

<実行形式>

```
% Demo_si
```

5.7 開発メモ

5.7.1 original coding

服部浩明、高嶋?

5.7.2 modified coding

大倉計美、高嶋?

5.7.3 Workstation Install Memo

5.8 関連研究論文など

不特定話者を対象とした離散分布型 HMM-LR 連続音声認識に関する発表はない。

第 6 章

AdaptCMixHMMLR

6.1 作成

作成者 :

作成日 :

6.2 機能説明

6.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

6.4 Directory 構造

```
AdaptCMixHMMLR
|
+- Data
| |
| | +- MALE
| | |
| | | +- ADAPT
| | | +- HmmO
| | | |
| | | +- MALEmht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | |
| | +- PARA
| | +- WAVE
| | +- list
| |
```

```
| +- ohkura
| | |
| | +- ADAPT
| | +- Hmm154
| | | |
| | | +- ohkuramht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | |
| | +- Hmm154_k100
| | | |
| | | +- ohkuramht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | |
| | +- Hmm23
| | | |
| | | +- ohkuramht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | |
| | +- Hmm48
| | | |
| | | +- ohkuramht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
```

```
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | |
| | +- KAIWA
| | +- PARA
| | +- PHRASE
| | +- WAVE
| | +- list
| |
| +- q
| | |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- taka
| | |
| | +- ADAPT
| | +- Hmm154
| | | |
| | | +- takamht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | |
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- test11
| | |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
+- Data-debug
| |
| +- mau
| | |
| | +- ADAPT
| | +- Hmm102
| | | |
```

```

| | | +- maumht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | | |
| | | +- Hmm13
| | | |
| | | +- maumht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | | |
| | | +- PARA
| | | +- WAVE
| | | +- list
| | | |
| +- mht
| | |
| | | +- HmmO
| | | |
| | | +- mhtmht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | | |
| | | +- list
| | | |
| +- nag
| | |
| | | +- ADAPT
| | | +- PARA
| | | +- WAVE
| | | +- list

```

```
| |
| +- pppp
| |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- qqg
| |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- quiquit
| |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- test
| |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- test1
| |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- test3
| |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- test4
| |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- ttt
```

```
| | |
| | +- ADAPT
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- yama
| | |
| | +- ADAPT
| | +- Hmm23
| | | |
| | | +- yamamht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | |
| | +- PARA
| | +- WAVE
| | +- list
| |
| +- yamafuku
| | |
| | +- ADAPT
| | +- Hmm23
| | | |
| | | +- yamafukumht
| | | |
| | | +- 10B
| | | +- 12B
| | | +- 14B
| | | +- 15B
| | | +- 2B
| | | +- 3B
| | | +- 5B
| | | +- 7B
| | |
| | +- PARA
| | +- WAVE
| | +- list
|
+- Dur
| |
| +- female
| +- mht
| |
```

```
| +- 10B
| +- 12B
| +- 14B
| +- 15B
| +- 2B
| +- 3B
| +- 5B
| +- 7B
|
+- EXE
| |
| +- BAK
| +- BAK2
| +- List
| | |
| | +- old_file
| |
| +- ListF
| | |
| | +- old_file
| |
| +- ListSB1
| +- Src
| | |
| | +- ConcatTrain2
| | +- Hmmlr
| | +- LOG
| | +- Tool
| | +- segmentNew
| |
| +- Work
| +- shell
|
+- Hmm
| |
| +- female
| +- male
| +- mht
| |
| +- 10B
| +- 12B
| +- 14B
| +- 15B
| +- 2B
| +- 3B
| +- 5B
| +- 7B
|
+- HmmlR
| |
| +- Gra
```

```

| +- TEST
| +- WAVE
|
+- MKTDATA
+- PROG
+- SPKR
| |
| +- BUF
| +- ListF
| +- ListSB1
|
+- TEXT
  |
  +- 11
  +- 20

```

6.5 Make について

6.5.1 Hmmlr

directory AdaptCMixHMMLR/EXE/Src/Hmmlr/ において Makefile を実行して下さい。実行イメージ Hmmlr が作成されます。

6.5.2 tdcmhmm1

directory AdaptCMixHMMLR/EXE/Src/ConcatTrain2/ において Makefile (tdcmhmm1.make) を実行して下さい。実行イメージ tdcmhmm1 が作成されます。

6.5.3 tdcmhmm1DEMO

directory AdaptCMixHMMLR/EXE/Src/ConcatTrain2/ において Makefile (tdcmhmm1DEMO.make) を実行して下さい。実行イメージ tdcmhmm1DEMO が作成されます。

6.5.4 tdcmhmm1N

directory AdaptCMixHMMLR/EXE/Src/ConcatTrain2/ において Makefile (tdcmhmm1N.make) を実行して下さい。実行イメージ tdcmhmm1N が作成されます。

6.5.5 tdcmhmm1P

directory AdaptCMixHMMLR/EXE/Src/ConcatTrain2/ において Makefile (tdcmhmm1P.make) を実行して下さい。実行イメージ tdcmhmm1P が作成されます。

6.5.6 mklogt

directory AdaptCMixHMMLR/EXE/Src/LOG/ において Makefile (mklogt.make) を実行して下さい。実行イメージ mklogt が作成されます。

6.5.7 segmentNew

directory AdaptCMixHMMLR/EXE/Src/segmentNew/ において Makefile (segmentNew.make) を実行して下さい。実行イメージ segmentNew が作成されます。

6.5.8 color

directory EXE/Src/Tool/ において color.c をコンパイルし、実行イメージを EXE/ へ置いてください。

6.5.9 devoc

directory EXE/Src/Tool/ において devoc.c をコンパイルし、実行イメージを EXE/ へ置いてください。

6.5.10 mklist

directory EXE/Src/Tool/ において mklist.c をコンパイルし、実行イメージを EXE/ へ置いてください。

6.5.11 sed2

directory EXE/Src/Tool/ において sed2.c をコンパイルし、実行イメージを EXE/ の下へ置いてください。

6.5.12 mktdata

directory MKTDATA/ において mktdata.c をコンパイルしてください。

6.5.13 WavePara34

directory PROG/ において WavePara34.c をコンパイルしてください。

6.6 実行の仕方

<実行形式>

% DEMO

6.7 開発メモ

6.7.1 original

6.7.2 modified

6.7.3 Workstation Install Memo

6.8 関連研究論文など

1. 大倉計美, 杉山雅英, 嵯峨山茂樹: “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” 連続音声認識シンポジウム, pp.49-50, (1992.02).
2. 大倉計美, 杉山雅英: 移動ベクトル場平滑化話者適応方式を用いた雑音環境下音声認識,” 音学講論, 2-Q-16, pp. 189-190, (1992.3).
3. 大倉計美, 杉山雅英, 嵯峨山茂樹: “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” 音学講論, 2-Q-17, pp. 191-192, (1992.3).
4. 大倉計美, 杉山雅英, 嵯峨山茂樹: “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” 信学技報, SP92-16, pp. 23-28, (1992.6).
5. Kazumi Ohkura, Masahide Sugiyama and Sigeki Sagayama: “Speaker Adaptation Based on Transfer Vector Field Smoothing with Continuous Mixture Density HMMs,” Proc. of ICSLP92, (Oct.1992).
6. 山口耕市, 嵯峨山茂樹: 混合連続分布型 HMM を用いた HMM-LR 連続音声認識,” 音学講論, 1-P-5, pp. 113-114, (1992.3).

7. 北 研二, 川端 豪, 斎藤 博昭: “HMM 音韻認識と拡張 LR 構文解析法を用いた連続音声認識,” 情報処理学会論文誌, Vol. 31, No. 3, pp. 472-480, March, 1990.
8. 花沢 利行, 北 研二, 中村 哲, 川端 豪, 鹿野 清宏: “HMM-LR 音声認識システムの性能評価,” 日本音響学会誌, Vol. 46, No. 10, pp. 817-823, October, 1990.
9. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: “HMM Continuous Speech Recognition Using Predictive LR Parsing,” *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 703-706, May, 1989.
10. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: “Parsing Continuous Speech by HMM-LR Method,” *Proceedings of the International Workshop on Parsing Technologies*, pp. 126-131, August, 1989.
11. Toshiyuki Hanazawa, Kenji Kita, Satoshi Nakamura, Takeshi Kawabata, Kiyohiro Shikano: “ATR HMM-LR Continuous Speech Recognition System,” *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 53-56, April, 1990.
12. 北 研二, 川端 豪, 斎藤 博昭: “HMM 音韻認識と予測 LR パーザを用いた文節認識,” 電子情報通信学会音声研究会, pp. 63-69, October, 1988.
13. 花沢 利行, 北 研二, 中村 哲, 川端 豪, 鹿野 清宏: “HMM-LR 音声認識システムの性能評価,” 電子情報通信学会音声研究会, pp. 63-70, December, 1989.
14. 北 研二, 川端 豪, 斎藤 博昭: “HMM 音韻認識と予測 LR パーザを用いた文節認識,” 日本音響学会秋季研究発表会, pp. 259-260, October, 1988.
15. 花沢 利行, 北 研二, 川端 豪, 鹿野 清宏: “HMM 音韻モデルの文節認識による評価,” 日本音響学会春季研究発表会, pp. 81-82, March, 1989.
16. 川端 豪, 鹿野 清宏, 北 研二: “音韻パープレキシティの提案,” 日本音響学会春季研究発表会, pp. 93-94, March, 1989.
17. 北 研二, 坂野 俊哉, 保坂 順子, 川端 豪: “SL-TRANS における文節音声認識 — HMM 音韻認識と LR 構文解析法による文節音声認識 —,” 情報処理学会第 39 回全国大会, pp. 718-719, October, 1989.
18. 研二, 江原 暉将, 森元 逞: “HMM-LR 音声認識の大語彙への適用,” 情報処理学会第 42 回全国大会, pp. 2-110-2-111, March, 1991.
19. Toshiyuki Hanazawa, Kenji Kita, Satoshi Nakamura, Takeshi Kawabata, Kiyohiro Shikano: “ATR HMM-LR Continuous Speech Recognition System,” *Readings in Speech Recognition*, Waibel, A and Lee, K. F. Eds., Morgan Kaufmann Publishers, pp. 611-614, 1990.
20. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: “GLR Parsing in Hidden Markov Model,” *Generalized LR Parsing*, Ed: Tomita, M, Kluwer Academic Publishers, pp. 153-164, 1991.
21. 北 研二: “自動翻訳電話,” 第 2 章 2.3 [2] 隠れマルコフモデルと文脈自由文法を用いた連続音声認識, ATR 先端テクノロジーシリーズ, (1993 年 1 月発行予定)
22. 大倉, 杉山, “混合連続分布 HMM を用いた移動ベクトル場平滑化話者適応方式,” TR-I-0266
23. 北, “Generalized LR Parsing in Hidden Markov Model,” TR-I-0161.
24. 永田, 衛藤, 保坂, “音声認識のための構文規則ガイドブック,” TR-I-0240.
25. 保坂, 竹澤, “SL-TRANS における音声認識のための構文規則の拡張,” TR-I-0241.
26. 北, “HMM-LR ユーザーズ・マニュアル,” TR-I-0246.
27. 北, “HMM-LR ソース・コード,” TR-I-0270.
28. 北, “HMM-LR 音声認識プログラム ユーザーズ・マニュアル,” TR-I-0271.

29. 北, "A Study on Language Modeling for Speech Recognition," TR-I-0273.
30. 大倉, 杉山, 嵯峨山, "混合連続分布 HMM 移動ベクトル場平滑化話者適応方式の文節認識性能の評価," TR-I-0284.

第 7 章

HMM-HARD

7.1 作成

7.1.1 旧版 (話者適応:codebook mapping)

作成者:

HMM-LR の認識制御部: 小野、永井
話者適応部: 荒木、伴、永井、中村 (シャープ: ハードウェア)
漢字 filter+Postfilter: 北、永井、大倉
部分環境依存 LR テーブル: 北、永井
言語翻訳部: 林

作成日 : 1991 年 9 月

7.1.2 新版 (話者適応: VFS)

作成者:

HMM-LR の認識制御部: 小野、永井
話者適応部: 宮沢、永井、小野、伴、大倉、服部
Postfilter: 林 (HMM-LR 出力書式変更に伴う改修)
部分環境依存 LR テーブル: (変更なし)
言語翻訳部: 林

作成日 : 1992 年 10 月

7.2 機能説明

7.2.1 旧版

ファジィ VQ に基づく離散 HMM を用いた文節認識 HMM-LR ハードウェア (三菱電機) と、コードブック・マッピングの話者適応ハードウェア (シャープ) を用いた、高速版の日英音声翻訳実験システム。

7.2.2 新版

話者適応は、VFS 方式を採用。適応データ (10 単語) を HMM-LR ハードウェア上で A/D、音響分析し、WS(HP) で話者適応を行ない、HMM-LR ハードウェアにダウンロードする。話者適応 + ダウンロードで 2 分弱に高速化。

7.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

7.4 Directory 構造

```
HMM-HARD/
  --> Adapt_new.csh
```

7.5 開発メモ

7.5.1 original coding

旧版:

```
HMM-LR HARD  -> 北研二、花沢利行、永井明人(仕様、デバッグデータ)
adapt.HARD   -> 中村哲
```

7.5.2 modified coding

新版:

```
HMM-LR HARD  -> 北研二、永井明人、
                大倉計美(デバッグデータ)、服部浩明(仕様)
adapt.(VFS)  -> 宮沢康永、大倉計美
```

7.5.3 Workstation Install Memo

7.6 関連研究論文など

1. 永井, 北, 花沢, 鈴木, 岩崎, 川端, 中島, 鹿野, 森元, 嵯峨山, 樽松: "HMM-LR 連続音声認識装置の開発と性能評価," 音学講論, 1-5-23, pp.45-46 (1991.10).
2. 中村, 浜口, 永井, 鹿野, 田中, 嵯峨山, 樽松: "ベクトル量子化に基づく話者適応化装置の開発と性能評価," 音学講論, 3-5-6, pp.101-102 (1991.10).
3. Akito NAGAI, Kenji KITA, Toshiyuki HANAZAWA, Tadashi SUZUKI, Tomohiro IWASAKI, Tsuyoshi KAWABATA, Kunio NAKAJIMA, Kiyohiro SHIKANO, Tsuyoshi MORIMOTO, Shigeki SAGAYAMA, Akira KUREMATSU : "Hardware Implementation of Realtime 1000-word HMM-LR Continuous Speech Recognition," ICSLP92 (Canada) (1992.10).
4. 北 研二, 川端 豪, 斎藤 博昭: "HMM 音韻認識と拡張 LR 構文解析法を用いた連続音声認識," 情報処理学会論文誌, Vol. 31, No. 3, pp. 472-480, March, 1990.
5. 花沢 利行, 北 研二, 中村 哲, 川端 豪, 鹿野 清宏: "HMM-LR 音声認識システムの性能評価," 日本音響学会誌, Vol. 46, No. 10, pp. 817-823, October, 1990.
6. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: "HMM Continuous Speech Recognition Using Predictive LR Parsing," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 703-706, May, 1989.
7. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: "Parsing Continuous Speech by HMM-LR Method," *Proceedings of the International Workshop on Parsing Technologies*, pp. 126-131, August, 1989.
8. Toshiyuki Hanazawa, Kenji Kita, Satoshi Nakamura, Takeshi Kawabata, Kiyohiro Shikano: "ATR HMM-LR Continuous Speech Recognition System," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 53-56, April, 1990.
9. 北 研二, 川端 豪, 斎藤 博昭: "HMM 音韻認識と予測 LR パーザを用いた文節認識," 電子情報通信学会音声研究会, pp. 63-69, October, 1988.
10. 花沢 利行, 北 研二, 中村 哲, 川端 豪, 鹿野 清宏: "HMM-LR 音声認識システムの性能評価," 電子情報通信学会音声研究会, pp. 63-70, December, 1989.

11. 北 研二, 川端 豪, 齋藤 博昭: “HMM 音韻認識と予測 LR パーザを用いた文節認識,” 日本音響学会秋季研究発表会, pp. 259-260, October, 1988.
12. 花沢 利行, 北 研二, 川端 豪, 鹿野 清宏: “HMM 音韻モデルの文節認識による評価,” 日本音響学会春季研究発表会, pp. 81-82, March, 1989.
13. 川端 豪, 鹿野 清宏, 北 研二: “音韻パープレキシティの提案,” 日本音響学会春季研究発表会, pp. 93-94, March, 1989.
14. 北 研二, 坂野 俊哉, 保坂 順子, 川端 豪: “SL-TRANS における文節音声認識 — HMM 音韻認識と LR 構文解析法による文節音声認識 —,” 情報処理学会第 39 回全国大会, pp. 718-719, October, 1989.
15. 北 研二, 江原 暉将, 森元 逞: “HMM-LR 音声認識の大語彙への適用,” 情報処理学会第 42 回全国大会, pp. 2-110-2-111, March, 1991.
16. Toshiyuki Hanazawa, Kenji Kita, Satoshi Nakamura, Takeshi Kawabata, Kiyohiro Shikano: “ATR HMM-LR Continuous Speech Recognition System,” *Readings in Speech Recognition*, Waibel, A and Lee, K. F. Eds., Morgan Kaufmann Publishers, pp. 611-614, 1990.
17. Kenji Kita, Takeshi Kawabata, Hiroaki Saito: “GLR Parsing in Hidden Markov Model,” *Generalized LR Parsing*, Ed: Tomita, M, Kluwer Academic Publishers, pp. 153-164, 1991.
18. 北 研二: “自動翻訳電話,” 第 2 章 2.3 [2] 隠れマルコフモデルと文脈自由文法を用いた連続音声認識, ATR 先端テクノロジーシリーズ, (1993 年 1 月発行予定)
19. Shikano, K., Lee, K-F., and Reddy, R., “Speaker Adaptation through Vector Quantization,” Proc. ICASSP86, 49.5, pp.2643-2646.
20. Nakamura, S. and Shikano, K., “Spectrogram normalization using fuzzy vector quantization,” *Journal of Acoust. Soc. of Japan*, Vol.45, No.2, pp.107-114.
21. 服部 浩明, 嵯峨山 茂樹: “少量学習データを用いたコードブックマッピングによる話者適応化,” 音学講論, 1-5-23, pp.49-50 (1991.03).
22. 服部 浩明, 嵯峨山 茂樹: “少数語彙による移動ベクトル場平滑化話者適応方式の文節認識による評価,” 音学講論, 2-Q-15, (1992.03).
23. H. Hattori and S.Sagayama: “Vector Field Smoothing Principle For Speaker Adaptation,” submitted to ICSLP92.
24. H. Hattori and S.Sagayama: “Speaker Adaptation Method Based on Transfer Vector Field Smoothing,” will be submitted to IEICE.
25. H. Hattori and S.Sagayama: “Speaker Adaptation Method Based on Transfer Vector Field Smoothing,” will be submitted to SP.
26. 花沢, 川端, 鹿野, “Hidden Markov Model を用いた日本語有声破裂音の識別,” TR-I-0018.
27. 花沢, 川端, 鹿野, “Duration control methods for HMM phoneme recognition,” TR-I-0050.
28. 平田, 川端, 花沢, “HMM-LR 法を用いた文節認識における継続時間長制御パラメータ変換法の検討,” TR-I-0076.
29. 北, 川端, 齋藤, “HMM 音韻認識と拡張 LR 構文解析法を用いた連続音声認識,” TR-I-0082.
30. 中村, 鹿野, “ファジィベクトル量子化に基づく話者適応化,” TR-I-0096.
31. 中村, 花沢, 鹿野, “ベクトル量子化話者適応の HMM 音韻認識への適応,” TR-I-0097.
32. 中村, “ベクトル量子化話者適応化の研究,” TR-I-0100.
33. 中村, 鹿野, “話者適応化における写像方法の比較,” TR-I-0101.

34. 花沢, 川端, 鹿野, “Hidden Markov Model による音韻認識実験の結果,” TR-I-0147.
35. 北, “Generalized LR Parsing in Hidden Markov Model,” TR-I-0161.
36. 永田, 衛藤, 保坂, “音声認識のための構文規則ガイドブック,” TR-I-0240.
37. 保坂, 竹澤, “SL-TRANS における音声認識のための構文規則の拡張,” TR-I-0241.
38. 北, “HMM-LR ユーザーズ・マニュアル,” TR-I-0246.
39. 北, “HMM-LR ソース・コード,” TR-I-0270.
40. 北, “HMM-LR 音声認識プログラム ユーザーズ・マニュアル,” TR-I-0271.
41. 北, “A Study on Language Modeling for Speech Recognition,” TR-I-0273.
42. 服部, “話者適応および話者個人性の研究,” TR-I-0263.

第 8 章

SSS-based Speaker-Adapt. Cont. Speech Recognition

8.1 作成

作成者 :

作成日 :

8.2 機能説明

8.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

8.4 Directory 構造

```
SSS-LR
|
|-- Adapt
| |
| | +- Seqlist
| | +- Src
| |
|-- Analysis
|-- Data
| |
| | +- hideko
| | |
| | | +- Adapted_HMnet
| | | +- List
| | | +- Para
| | | +- Wave
| | |
| | +- jun
| | |
| | | +- Adapted_HMnet
| | | +- List
| | | +- Para
| | | +- Wave
| |
```

```
| +- kurematu
| | |
| | +- Adapted_HMnet
| | +- List
| | +- Para
| | +- Wave
| |
| +- nagai
| | |
| | +- Adapted_HMnet
| | +- List
| | +- Para
| | +- Wave
| |
| +- saga
| | |
| | +- Adapted_HMnet
| | +- List
| | +- Para
| | +- Wave
| |
| +- singer
| | |
| | +- List
| | +- Para
| | +- Wave
|
+- Draw
| |
| +- data_26phone
| +- data_6cons
|
+- Dur
+- Exe
| |
| +- List
| +- List.org
| |
| +- old_file
|
+- HMnet
| |
| +- FSU
| +- Global
| +- MAU
| +- MHT
| +- MXM
|
+- Reco
| |
| +- Gra
```

```
| +- Src
| | |
| | +- Src_parser
| | +- Src_verify
| |
| +- Wave
|
+- Sp_select
  |
  +- Seqlist
  +- Src
```

8.5 Make について

8.5.1 Adapt

directory SSS-LR/Adapt/Src/ において Makefile を実行して下さい。実行イメージ Adapt が作成されます。

8.5.2 Draw

directory SSS-LR/Draw/ において Makefile を実行して下さい。実行イメージ main が作成されます。

8.5.3 SSS-LR.exe

directory SSS-LR/Reco/Src/ において Makefile を実行して下さい。実行イメージ SSS-LR.exe が作成されます。また、SSS-LR/Reco/Src/Makefile は SSS-LR/Reco/Src/Src_parser/Makefile と SSS-LR/Reco/Src/Src_verify/Makefile を実行しています。

8.5.4 Speaker-select

directory SSS-LR/Sp_select/Src/ において Makefile を実行して下さい。実行イメージ Speaker_select が作成されます。

8.5.5 WavePara

directory SSS-LR/Analysis/ において WavePara34.c をコンパイルして、WavePara34 を作成して下さい。(cc -o WavePara WavePara.c -lm)

8.6 実行の仕方

<実行形式>

% DEMO

8.7 開発メモ

8.7.1 original

8.7.2 modified

8.7.3 Workstation Install Memo

8.8 関連研究論文など

1. 鷹見, 永井, 嵯峨山, “aka(赤)とaki(秋)の /k/ は同じ音? -前後の音素を考慮した高精度音声認識-,” ATR ジャーナル, No.12 (1992.11)
2. 鷹見, “音声認識における HMM とその精度向上のための手法,” 信学技報, SP92-49, pp.17-24 (1992.9).
3. J.Takami, S.Sagayama, “A Successive State Splitting Algorithm for Efficient Allophone Modeling,” ICASSP92 (San Francisco), 66.6 (1992.3).
4. 鷹見, 嵯峨山, “逐次状態分割法 (SSS) により自動生成した隠れマルコフ網の性能評価,” 音講論集, 2-1-8 (1992.3).
5. 鷹見, 嵯峨山, “音素コンテキストと時間に関する逐次状態分割による隠れマルコフ網の自動生成,” 信学技報, SP91-88 (1991.12).
6. 鷹見, 嵯峨山, “逐次状態分割法 (SSS) による隠れマルコフネットワークの自動生成,” 音講論集, 2-5-13, pp.73-74 (1991.10).
7. 鷹見, 宮沢, 永井, 嵯峨山, “話者適応型 SSS-LR 連続音声認識方式における標準話者予備選択の効果,” 音講論集, 2-Q-16 (1993.3) (発表予定).
8. J.Takami, A.Nagai, S.Sagayama, “Speaker Adaptation of the SSS (Successive State Splitting)-Based Hidden Markov Network for Continuous Speech Recognition,” SST92 (Australia) (1992.12)
9. 鷹見, 小坂, 嵯峨山, “話者方向を加えた逐次状態分割法 (SSS) による話者共通隠れマルコフ網の生成,” 音講論集, 3-1-8 (1992.10).
10. 鷹見, 永井, 嵯峨山, “逐次状態分割法 (SSS) と LR パーザを統合した SSS-LR 連続音声認識手法における話者適応の性能評価,” 音講論集, 2-5-5 (1992.10).
11. 鷹見, 嵯峨山, “隠れマルコフ網 (HM-Net) を用いた話者適応,” 音講論集, 1-1-8 (1992.3).
12. 永井明人, 鷹見淳一, 嵯峨山茂樹, “逐次状態分割法 (SSS) と音素コンテキスト依存 LR パーザを統合した SSS-LR 連続音声認識システム,” 信学技報, SP92-33, pp.69-76 (1992.06).
13. Nagai, A., Takami, J. and Sagayama, S., “The SSS-LR Continuous Speech Recognition System: Integrating SSS-derived Allophone Models and a Phoneme-Context-Dependent LR Parser,” Proc. ICSLP92 (Banff), pp.1511-1514.

第 9 章

Word-bigram-based Sentence Speech Recognition

9.1 作成

作成者 : 村上仁一

作成日 : 1992.4.1 (V1), 1992.11.1 (V2)

9.2 機能説明

word bigram を利用した、連続文音声認識 (語彙数 435 単語)

9.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

9.4 Directory 構造

```
HMM-bigram
|
+- language
| |
| +- org
| +- pmarkov1
| +- pmarkov2
| +- pmarkov3
| +- wmarkov1
| +- wmarkov2
| +- wmarkov3
|
+- recog
|
| +- bigram
| |
| | +- INDEPENDENT
| | +- MAU
```

9.5 Make について

directory HMM-bigram/recog/bigram/ において makefile を実行して下さい。実行イメージ main が作成されます。

9.6 実行の仕方

directory HMM-bigram/recog/bigram において

```
% main
```

9.7 開発メモ

9.7.1 original

9.7.2 modified

9.7.3 Workstation Install Memo

9.8 関連研究論文など

1. 荒木, 村上, 池原: “音韻連鎖確率による音声日本文の曖昧さの解消効果について,” 第 35 回情処全大, 1T7 1327-1328 (1987-10).
2. 荒木, 村上, 池原: “2 重音韻マルコフモデルによる日本語の文節認識候補の曖昧さの解消効果,” 信学技報, NLC87-24, (1988-03) .
3. 荒木, 村上, 池原: “2 重音節マルコフモデルによる日本語の文節音節認識候補の曖昧さの解消効果,” 情処論文誌, Vol.30, No.4,pp.467-477 (Apr.1989)..
4. 荒木, 村上, 池原: “2 重音節連鎖確率を用いた日本語音声認識後処理,” NTT R&D, Vol.38, No6, 1989 pp.678-689 (1989). .
5. Araki, Murakami, Ikehara: “Post-processing in Japanese Speech Recognition Using 2nd-order Markov Model of Syllable,” NTT REVIEW Vol.1, No.3, September 1989, pp.96-104..
6. 荒木, 村上, 池原: “m 重マルコフモデルを用いた音節ラティスからの候補絞り込み手法について,” 第 39 回情処全大, 7E-5, pp.575-576..
7. 荒木, 村上, 池原: “音節連鎖確率のタイプと音節認識候補絞り込みについて,” 信学会全大, D-307, pp.6-27 (1989-03)..
8. 荒木, 村上, 池原: “m 重マルコフモデルを用いた音節ラティスからの候補絞り込みアルゴリズム,” 信学技報, CS90-55, pp.13-18 (1990-10)..
9. 村上, 荒木, 池原: “2 重マルコフ連鎖確率モデルを使用した単音節音声入力力の改善,” 信学技報, SP88-29, pp.63-70 (1988-06). .
10. 村上, 坪井: “BIGRAM をもちいた音節 HMM による文節音声認識,” 音学講論, 3-5-15, pp.117-118 (1991-04).
11. 村上, 荒木, 池原: “日本文音節入力に対して 2 重マルコフ連鎖モデルを用いた漢字かな交じり文節候補の抽出精度,” 信学会論文誌, D-2, Vol. J75-D-2, No.1,pp.11-20 (Jan. 1992). .
12. 荒木, 池原, 村上: “音節ラティスに適用するビタビアルゴリズムの評価について,” 第 44 回情処全大, 7N-2, pp. 2.163-164..

13. 村上, 嵯峨山: “連続音声認識における LR パーザと単語 bigram の比較” 音学講論, 1-P-8, pp.189-190 (1992-3)..
14. 村上, 嵯峨山: “単語の trigram を用いた連続音声認識の一アルゴリズム,” 音学講論, 2-Q-7, pp.185-186 (1992-10)..
15. Jin'ichi Murakami and Shigeki Sagayama: “An Efficient Algorithm For Using Word Trigram Models For Continuous Speech Recognition,” SST92, pp. 330-335(1992-12)..

第 10 章

Word-trigram-based Sentence Speech Recognition

10.1 作成

作成者 : 村上仁一

作成日 : 1992.4.1 (V1), 1992.11.1 (V2)

10.2 機能説明

word trigram を利用した、連続文音声認識 (語彙数 435 単語)

10.3 デモソフトウェアの所在

Machine: atr-p14

Directory: /users/demo/DEMO_SOFT_DIR/large/recog/ngram

10.4 Directory 構造

```
atr-p14:/users/demo/DEMO_SOFT_DIR/large/recog/ngram
```

10.5 Make について

directory ngram/ において makefile を実行して下さい。実行イメージ main が作成されます。

10.6 実行の仕方

<実行形式>

```
% main
```

10.7 開発メモ

10.7.1 original

10.7.2 modified

10.7.3 Workstation Install Memo

10.8 関連研究論文など

1. 荒木, 村上, 池原: “音韻連鎖確率による音声日本文の曖昧さの解消効果について,” 第 35 回情処全大, 1T7 1327-1328 (1987-10).
2. 荒木, 村上, 池原: “2 重音韻マルコフモデルによる日本語の文節認識候補の曖昧さの解消効果,” 信学技報, NLC87-24, (1988-03) .
3. 荒木, 村上, 池原: “2 重音節マルコフモデルによる日本語の文節音節認識候補の曖昧さの解消効果,” 情処論文誌, Vol.30, No.4, pp.467-477 (Apr.1989)..
4. 荒木, 村上, 池原: “2 重音節連鎖確率を用いた日本語音声認識後処理,” NTT R&D, Vol.38, No6, 1989 pp.678-689 (1989). .
5. Araki, Murakami, Ikehara: “Post-processing in Japanese Speech Recognition Using 2nd-order Markov Model of Syllable,” NTT REVIEW Vol.1, No.3, September 1989, pp.96-104..
6. 荒木, 村上, 池原: “m 重マルコフモデルを用いた音節ラティスからの候補絞り込み手法について,” 第 39 回情処全大, 7E-5, pp.575-576..
7. 荒木, 村上, 池原: “音節連鎖確率のタイプと音節認識候補絞り込みについて,” 信学会全大, D-307, pp.6-27 (1989-03)..
8. 荒木, 村上, 池原: “m 重マルコフモデルを用いた音節ラティスからの候補絞り込みアルゴリズム,” 信学技報, CS90-55, pp.13-18 (1990-10)..
9. 村上, 荒木, 池原: “2 重マルコフ連鎖確率モデルを使用した単音節音声入力 of 改善,” 信学技報, SP88-29, pp.63-70 (1988-06). .
10. 村上, 坪井: “BIGRAM をもちいた音節 HMM による文節音声認識,” 音学講論, 3-5-15, pp.117-118 (1991-04).
11. 村上, 荒木, 池原: “日本文音節入力に対して 2 重マルコフ連鎖モデルを用いた漢字かな交じり文節候補の抽出精度,” 信学会論文誌, D-2, Vol. J75-D-2, No.1, pp.11-20 (Jan. 1992). .
12. 荒木, 池原, 村上: “音節ラティスに適用するビタビアルゴリズムの評価について,” 第 44 回情処全大, 7N-2, pp. 2.163-164..
13. 村上, 嵯峨山: “連続音声認識における LR パーザと単語 bigram の比較” 音学講論, 1-P-8, pp.189-190 (1992-3)..
14. 村上, 嵯峨山: “単語の trigram を用いた連続音声認識の一アルゴリズム,” 音学講論, 2-Q-7, pp.185-186 (1992-10)..
15. Jin'ichi Murakami and Shigeki Sagayama: “An Efficient Algorithm For Using Word Trigram Models For Continuous Speech Recognition,” SST92, pp. 330-335(1992-12)..

第 11 章

Phoneme Recognition using a TDNN

11.1 作成

作成者 : 宮武? (北岡 [DEC])

作成日付 : 1988.12

11.2 概要説明

モジュール構成 TDNN による音素識別のデモ。認識対象とする音素サンプルは単語波形データファイルから読み込む。単語をクイズ形式で発音する。単語の極一部の音を発音し、それがどのような単語に含まれる音であるかを、クイズとして質問し、一定時間が経過したのち、答えとして単語を発音する。また、音声の形を画面に表示する。

11.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

11.4 Directory 構造

```
QUIZ/ --- SRC/ -----< C ソース及びヘッダファイル>
|
|           Makefile           dspdefs.h
|           quiz_x11.c         net_disp.h
|           snet_disp_x11.c   net.h
|           set_color.c       c.h
|           areadshort.c      QuizColordef.h
|           fft.c             Mn.h
|           ham.c             Nnp.h
|           Nn.c
|
|
+--- TRAIN/ -----< * * * * * >
|           consall.disp      consall.net
|           consall.voc       consall.wght
|
|
+--- WORK/ -----< * * * * * >
|           demo.rand3
|           demo.rand3."920601"
|           demo.rand2.sav
|
```


4. A. Waibel, 沢井秀文, 鹿野清宏, "Phoneme Recognition by Modular Construction of Time-Delay Neural Networks," 音講論, 2-P-12, pp.225-226 (1988.10).
5. 沢井秀文, A. Waibel, 宮武正典, 鹿野清宏, "モジュール構成ニューラルネットワークのスケールアップによる音韻認識," 音声研究会資, SP88-105, pp.73-80 (1988.12).
6. A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, K.J. Lang, Phoneme Recognition Using Time -Delay Neural Networks, Trans., ASSP-37, No.3, pp.328-339 (MAR. 1989).
7. A. Waibel, H. Sawai, K. Shikano, "Consonant Recognition by Modular Construction of Large Phonemic Time-Delay Neural Networks," Proc. of ICASSP89, (1989.05).
8. A. Waibel, H. Sawai, K. Shikano, "Modularity and Scaling in Large Phonemic Neural Networks," IEEE Trans. Acoust., Speech, Signal Processing, Vol.37, pp.1888-1898 (1989.12).
9. テクニカルレポート : TR-I-0034, TR-I-0058

第 12 章

Phoneme Spotting using TDNNs

12.1 作成

作成者 : 宮武 (北岡 [DEC])
作成日付 : 1988.10

12.2 概要説明

TDNN を用いて単語発声の音素スポッティング (スキャンニング) を行なうデモ。認識は単語波形データファイルを読み込んで行なう。データとしてもっている単語が、どのような音で構成されているかを、表示する。また、音の形も表示する。

12.3 デモソフトウェアの所在

Machine: atrq30
Directory: /q30_data/users/demo/DEMO.SOFT_DIR

12.4 Directory 構造

```
SPOT2/  --+--- SPOT2_SRC/
          |   <C ソース及びヘッダファイル>
          |
          |
          +--- Data.fft.100/
          |   < * * * * * >
          |
          |
          +--- Data.phon.100/
          |   < * * * * * >
          |
          |
          +--- MKS_SRC/
          |   <C ソース及びヘッダファイル>
          |
          |
          +--- nnwidget/
          |   < * * * * * >
          |
          |
```


12.5 make について

ディレクトリー SPOT_SRC/ において、Makefile を実行すると、実行イメージ Spot2 が、SPOT2/ のしたに作成される。

12.6 実行の仕方

<実行形式>

```
% Spot2 List.100 -f Data.fft.100 -o Data.phon.100 -m 0.05 -h -v
```

<注意>

実行時に Data.fft.100 と、Data.phon.100 のしたに、それぞれ データファイルが 作成される。

12.7 開発メモ

12.7.1 original

12.7.2 modified

12.7.3 Workstation Install Memo

12.8 関連研究論文など

1. 沢井, 他, 音講論, 2-P-11, pp.223-224 (1988.10).
2. 沢井, 他, 音講論, 2-P-25, pp.279-280 (1989.03).
3. 宮武, 他, 音講論, 2-P-24, pp.277-278 (1989.03).
4. 宮武正典, 沢井秀文, 鹿野清宏, “時間遅れニューラルネットワークを用いた音韻スポッティング法,” 信学会, SA-1-4, Vol.1, pp.335-336 (1989.03).
5. 宮武正典, 沢井秀文, 鹿野清宏, “連続音声中の音韻スポッティングのための TDNN 構成法,” 音声研究会資, SP89-32, pp.63-68(1989.06).
6. 宮武正典, 沢井秀文, 鹿野清宏, “時間遅れ神経回路網 (TDNN) を用いた音韻スポッティングの改良,” 音講論, 1-1-25, pp.49-50(1989.10).
7. 沢井秀文, 宮武正典, A. Waibel, 鹿野清宏, “連続音声認識のための時遅れ神経回路網を用いた音韻/音節スポッティング,” 信学論, vol.J72-D-II, No.8, pp.1151-1158 (1989.08).
8. H.Sawai, et al., “Spotting Japanese CV-Syllables and Phonemes Using Time-Delay Neural Networks,” ICASSP89,(1989.05).
9. H.Sawai, et al., “Spotting Phonemes and Syllables for Continuous Speech Recognition Using Time-Delay Neural Networks,” Systems and Computers in Japan, vol.21, No.9, pp.71-79 (1990.09).
10. テクニカルレポート : TR-I-0089, TR-I-0090, TR-I-0103


```

|
+-- (実行イメージ)
    PD_TDNN_Demo

```

13.5 Make について

ディレクトリー SRC/ において、Makefile を実行すると、実行イメージ PD_TDNN_Demo が、PD_TDNN/ のしたに作成される。

13.6 実行の仕方

<実行形式>

```
% PD_TDNN_Demo
```

13.7 開発メモ

13.7.1 original

13.7.2 modified

13.7.3 Workstation Install Memo

13.8 関連研究論文など

1. 鷹見, 嵯峨山, 対判定型 TDNN による音素認識, 音声研究会資料, SP90-10, pp.9-16 (1990-06).
2. 鷹見, 嵯峨山, 対判定型 TDNN における中間値学習の効果, 音響学会講演論文集, 2-P-15, pp.159-160 (1990-09).
3. J.Takami, A.Kai, S.Sagayama, "A Pairwise Discriminant Approach using Artificial Neural Networks for Continuous Speech Recognition," JASJ(E), Vol.13, No.6, pp.411-418 (1992.11).
4. 鷹見, 嵯峨山, 甲斐, "対判定型 TDNN と予測 LR パーザの結合による連続音声認識," 音講論集, 1-5-22, pp.43-44 (1991.10).
5. J.Takami, S.Sagayama, A.Kai, "Speech Recognition by Combining Pairwise Discriminant Time-Delay Neural Networks and Predictive LR-Parser," NNSP91 (Princeton), pp.327-336 (1991.9).
6. J.Takami, S.Sagayama, "A Pairwise Discriminant Approach to Robust Phoneme Recognition by Time-Delay Neural Networks," ICASSP91 (Toronto), 8.S2.13 (1991.5).
7. 鷹見, 嵯峨山, "分類型ニューラルネットワークを用いた音声認識におけるロバストネスの問題について," HMM-NN シンポジウム (強羅) (1990.2).
8. J.Takami, S.Sagayama, "Phoneme recognition by pairwise discriminant TDNN," ICSLP90 (Kobe), S16.5.1, pp.677-680 (1990.11).
9. 鷹見, 嵯峨山, "対判定型 TDNN における中間値学習の効果," 音講論集, 2-P-15, pp.159-160 (1990.9).
10. 鷹見, 嵯峨山, "対判定型 TDNN による音素認識," 信学技報, SP90-10, pp.9-16 (1991.6).

第 14 章

TDNN-LR Cont. Speech Recognition

14.1 作成

作成者：小森(伴、高島)

作成日付：1991.09

14.2 機能説明

学習に、Neural Fuzzy 学習法を使用した TDNN と LR パーザを組み合わせた TDNN-LR 特定話者文節認識のデモ。認識は文節波形データファイルを読み込んで行なう。また、デモの中で従来の Discrete タイプの TDNN による文節認識も行ない両者を比較している。

14.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

14.4 Directory 構造

```
TDNN_LR/ --+--- SRC/
|      (Cソース及びヘッダファイル)
|
+--- List/
|      (リストファイル)
|
+--- FILE/
|      ( * * * * * )
|
+--- wave1/ --+--- MAU/ --+--- LBL/ --+--- DO/
|              |              |      (ラベルファイル)
|              |              |
|              |              +--- DSB/
|              |              (ラベルファイル)
|              |
|              +--- WAV/ --+--- DO/
|                          |      (音声データ)
|                          |
|                          +--- DSB/
|                          (音声データ)
|
+--- SPOT_RES/ --+--- FZ/
```

```

|           | ( * * * * * )
|           |
|           +--- HD/
|           ( * * * * * )
|
+--- lr0/ -----+----- Src/
| ( * * * * * ) (Cソース、ヘッダファイル)
|
+--- <<実行イメージ>>
|   TDNN-LR_FZ
|
+--- <<シェルスクリプト>>
|   new_tdnn.sh
|
+--- <<その他>>
|   old_TDNN-LR/   test.sh

```

14.5 Make について

ディレクトリ SRC/ において、Makefile を実行すると、実行イメージ TDNN-LR_FZ が、TDNN_LR/ の下に作成される。

14.6 実行の仕方

<実行形式>

```
% new_tdnn.sh MAU
```

<注意>

実行時に、SPOT_RES/ の下にデータファイルが作成される。

14.7 開発メモ

14.7.1 original

14.7.2 modified

14.7.3 Workstation Install Memo

14.8 関連研究論文など

1. 南, 他, TDNN 音韻スポットティングと予測 LR パーザを用いた大語彙単語音声認識, TR-I-0144 (1990-02).
2. Miyatake M., Sawai H., Minami and K. Shikano, : "Integrated Training for Spotting Japanese Phonemes Using Phonemic Time-Delay Neural Networks," Proc. of ICASSP, S8.10, pp.449-452 (Apr. 1990).
3. Hidefumi Sawai, "The TDNN-LR Large-Vocabulary and Continuous Speech Recognition System," Proc. of ICSLP90, pp.1349-1352 (1990-11).
4. Hidefumi Sawai, Yasuhiro Minami, Masanori Miyatake, Alex Waibel and Kiyohiro Shikano, "Connectionist Approaches to Large Vocabulary Continuous Speech Recognition," Special Issue on Continuous Speech Recognition and Understanding, IEICE Trans. Vol.E 74, No.7, PP 1834-1844, July 1991.
5. 小森 康弘, 嵯峨山 茂樹, A.H.Waibel: "音素識別ニューラルネットワークにおけるファジィ学習法," 音学講論, 1-5-15, pp.33-34 (1991.03).

6. 小森 康弘, 嵯峨山 茂樹, A.Waibel: “ニューラル・ファジー学習による TDNN-LR 連続音声認識システムの性能向上,” 信学技報, SP91-24, pp.49-56(1991.6).
7. 小森 康弘, 福沢 圭二, 杉山 雅英, A.H.Waibel, 嵯峨山 茂樹: “ニューラル・ファジー学習の連続音声認識における効果,” 音学講論, 2-5-11, pp.69-70(1991.10).
8. Yasuhiro Komori, “Neural Fuzzy Training Approach for Continuous Speech Recognition Improvement”, Proc. ICASSP92, Vol.1 ,pp.405-408, 1992-03.
9. テクニカルレポート : TR-I-0085, TR-I-0144

第 15 章

Speaker-Adaptation + TDNN-LR

15.1 作成

作成者：福沢 (伴、高島)

作成日：1992-02

15.2 機能説明

セグメントベース話者適応ニューラルネットワークと TNN-LR を組み合わせた実音声入力による文節認識のデモ。

入力話者のマイク入力された文節発声をセグメントベース話者適応ニューラルネットワークにより標準話者の音声特徴量へ変換し、標準話者により学習された TDNN-LR に入力することにより文節単位の認識を行なう。適応学習には時間がかかる為 (1 日～2 日)、入力話者による単語発声データ (50～100 単語) を用いてあらかじめ学習しておく。TDNN の学習には、Neural Fuzzy 学習法を使用。

15.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

15.4 Directory 構造

```
AD_TDNN-LR/ --- SRC/
|          (C ソース 及びヘッダファイル)
|
+--- LRO/
|      ( * * * * * )
|
+--- ADAPT_NET/
|      (NETWORK FILE)
|
+--- DURATION_NET/
|      (NETWORK FILE ,DURATION FILE)
|
+--- WAVE_DATA/
|      ( * * * * * )
|
+--- SPOT_RES/
|      ( * * * * * )
|
+--- WEIGHT/ --- ADP_WEIGHT/
```

```

|           | (WEIGHT FILE)
|           |
|           +--- TDNN_WEIGHT/
|           (WEIGHT FILE)
|
+--- <<実行イメージ>>
|   ad_tdnn_lr   SpeechIn
|
+--- <<シェルスクリプト>>
|   at_tdnn_lr.sh

```

15.5 Make について

ディレクトリー SRC/ において、Makefile を実行すると、実行イメージ ad_tdnn_lr が、AD_TDNN_LR/ のしたに作成される。

15.6 実行の仕方

<実行形式>

```
% ad_tdnn_lr.sh 5
```

<注意>

実行時に、SPOT_RES/ と WAVE_DATA/ の下に、それぞれ データファイルが作成される。

15.7 開発メモ

15.7.1 original

15.7.2 modified

15.7.3 Workstation Install Memo

15.8 関連研究論文など

1. 福沢圭二, 沢井秀文, 杉山雅英, “ニューラルネットワークによる恒等写像を用いた話者適応,” 音学講論, 秋季研究発表会, 1-8-16, pp.31-32, (1990.09).
2. 福沢圭二, 小森康弘, 沢井秀文, 杉山雅英, “セグメントベース話者適応ニューラルネットワークを用いた文節音声認識,” 音学講論, 秋季研究発表会, 3-5-10, (1991.10).
3. 福沢圭二, 小森康弘, 杉山雅英, “TDNN-LR 連続音声認識における不特定話者 TDNN と話者適応ニューラルネットワークの性能比較,” 音学講論, 春季研究発表会, 2-Q-21, (1992.03).
4. 福沢圭二, 小森康弘, 沢井秀文, 杉山雅英, “セグメントベース話者適応ニューラルネットワークと TDNN-LR を用いた文節音声認識,” 音声研究会資料, SP91-105, (1992.01).
5. Keiji Fukuzawa, Hidehumi Sawai, Masahide Sugiyama, “Segment-based Speaker Adaptation by Neural Network,” pp.442-451, Proc. NNSP'91, (1991.09).
6. Keiji Fukuzawa, Yasuhiro Komori, Hidehumi Sawai, Masahide Sugiyama, “A Segment-based Speaker Adaptation Neural Network Applied to Continuous Speech Recognition,” 55.1, Proc. ICASSP92, vol.1, pp.433-436; (1992.03).

第 16 章

Speaker-Indep. NN-based Cont. Speech Recognition

16.1 作成

作成者 : 福沢 (伴、高島)

作成日 : 1992-02

16.2 機能説明

不特定話者の音素識別を行なう TDNN と LR パーザを組み合わせた TDNN-LR システムによる実音声入力による不特定話者文節認識のデモ。

16.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

16.4 Directory 構造

```
IND_TDNN_LR/ --- SRC/
|   (Cソース 及びヘッダファイル)
|
+--- LRO/
|   ( * * * * * )
|
+--- DURATION_NET/
|   (NETWORK FILE ,DURATION FILE)
|
+--- WAVE_DATA/
|   ( * * * * * )
|
+--- SPOT_RES/
|   ( * * * * * )
|
+--- WEIGHT/
|   (WEIGHT FILE)
|
+--- ETC_LIB_SRC/ --- ETC_X11_MT_LIB/
|                   |   (Cソース)
|                   |
|                   +--- XKRNL11/
|                   |   (Cソース及びヘッダファイル)
```

```

|
|
|          +-- <<ライブラリー>>
|          |
|          |          libMYX.a   libXKRNL11.a
|          |
|          +-- <<実行イメージ>>
|          |
|          |          ind_tdnn-lr   SpeechIn
|          |
|          +-- <<シェルスクリプト>>
|          |
|          |          ind_tdnn-lr.sh

```

16.5 Make について

ディレクトリ SRC/ において、Makefile を実行すると、実行イメージ ind_tdnn-lr が、IND_TDNN-LR/ の下に作成される。

16.6 実行の仕方

<実行形式>

```
% ind_tdnn-lr.sh 5
```

<注意>

実行時に、SPOT_RES/ と WAVE_DATA/ の下に、それぞれ データファイルが作成される。

16.7 開発メモ

16.7.1 original

16.7.2 modified

16.7.3 Workstation Install Memo

16.8 関連研究論文など

1. 福沢圭二, 小森康弘, 杉山雅英, "TDNN-LR 連続音声認識における不特定話者 TDNN と話者適応ニューラルネットワークの性能比較," 音学講論, 春季研究発表会, , 2-Q-21, (1992.03).
2. 沢井, 中村, 福沢, 杉山, ニューラルネットワークによる不特定話者音声認識へのアプローチ法について、音響学会講演論文集, 1-5-17, pp.37-38 (1991-03).
3. 中村, 沢井, 杉山, TDNN を用いた不特定話者音素認識, 音響学会講演論文集, 2-5-8, 63-64 (1991-10).
4. Satoru Nakamura, Hidefumi Sawai and Masahide Sugiyama, Speaker-Independent Phoneme Recognition Using Large-Scale Neural Networks, Proc. of ICASSP92 (1992).
5. K.Fukuzawa, Y.Kato, M.Sugiyama, A Fuzzy Patition Model(FPM) Neural Network Architecture for Speaker Independent Continuous Speech Recognition, ICSLP92, Th.PM.P.14, pp.1383-1386 (1992-10).

第 17 章

FPM-LR

17.1 作成

作成者 : 伴、加藤

作成日 : 1992年8月

17.2 機能説明

Fuzzy Partition Model ニューラルネットワークアーキテクチャを用いた不特定話者連続音声認識のデモ

17.3 Directory 構造

```
FPM-LR
|
+- DURATION_NET
+- ETC_LIB_SRC
| |
| +- ETC_X11_MY_LIB
| +- XKRNL11
|
+- LRO
+- SPOT_RES
+- SRC
+- SRC_ETC
+- SRC_PARA
+- SRC_SOCKET
| |
| +- SRC_CLIENT
| +- SRC_SERVER
| +- include
|
+- WAVE_DATA
+- WEIGHT
```

17.4 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

17.5 Make について

17.5.1 fpm-lr

directory FPM-LR/SRC/ において Makefile を実行して下さい。一つ上の directory FPM-LR/ の下に実行イメージ fpm-lr が作成されます。(Makefile には DecMotif 用, Hpmotif 用, dwt 用が用意されています。)

17.5.2 libMyX.a

directory FPM-LR/ETC_LIB_SRC/ETC_X11_MY_LIB/ において Makefile を実行して下さい。一つ上の directory FPM-LR/ETC_LIB_SRC/ の下にライブラリー libMyX.a が作成されます。

17.5.3 libXKRNL11.a

directory FPM-LR/ETC_LIB_SRC/XKRNL11/ において Makefile を実行して下さい。一つ上の directory FPM-LR/ETC_LIB_SRC/ の下にライブラリー libXKRNL11.a が作成されます。

17.5.4 wposition and wait-button

directory FPM-LR/SRC_ETC/ において Makefile を実行して下さい。実行イメージ wposition と wait_button が一つ上の directory FPM-LR/ に作成されます。

17.5.5 SRC-PARA/fpm-lr-p.dec.out

directory FPM-LR/SRC_PARA/ において Makefile を実行して下さい。実行イメージ fpm-lr_p.dec.out が作成されます。(Makefile には DecMotif 用, HpMotif 用が用意されています。)

17.5.6 SRC-SOCKET/SRC-SERVER/fpm-lr-p.dec.out

directory FPM-LR/SRC_SOCKET/SRC_SERVER/ において Makefile を実行して下さい。実行イメージ fpm-lr_p.dec.out 二つ上の directory FPM-LR/ に作成されます。(Makefile には DecMotif 用, HpMotif 用が用意されています。)

17.5.7 fpm-lr-client

directory FPM-LR/SRC_SOCKET/SRC_CLIENT/ において makefile を実行して下さい。実行イメージ fpm-lr_client 二つ上の directory FPM-LR/ に作成されます。

17.6 実行の仕方

<実行形式>

```
% fpm-lr_para.sh
```

17.7 開発メモ

17.7.1 original

1992年8月

17.7.2 modified

パラレル処理	: 1992年9月
サーバ & クライアント方式	: 1992年10月
文認識組込み	: 1992年11月

17.7.3 Workstation Install Memo

17.8 関連研究論文など

1. 加藤, 杉山, 多入力素子をもつニューラルネットワークを用いた連続音声認識, 音響学会講演論文集, 3-1-1, pp.71-72 (1992-03).
2. 福沢, 加藤, 杉山, FPM-LR を用いた不特定話者連続音声認識, 音響学会講演論文集, 3-7-3, pp.167-168 (1992-10).
3. 加藤, 杉山, Fuzzy Partition Model を用いた連続音声認識, SP92-??, pp. (1992-06).
4. 加藤, 杉山, 多入出力素子をもつニューラルネットワークを用いた連続音声認識, 電子情報通信学会「マルコフモデル・ニューラルネットワークを包含する新しい音声認識手法」時限研究専門委員会, SPREC-91-2, pp.47-48 (1992-02).
5. Y.Kato, M.Sugiyama, Fuzzy Partition Models and Their Effect in Continuous Speech Recognition, NNSP92, pp.111-120 (1992-8).
6. K.Fukuzawa, Y.Kato, M.Sugiyama, A Fuzzy Partition Model(FPM) Neural Network Architecture for Speaker Independent Continuous Speech Recognition, ICSLP92, Th.PM.P.14, pp.1383-1386 (1992-10).
7. Y.Kato, M.Sugiyama, Fuzzy Partition Models and Their Incremental Training for Continuous Speech Recognition, ASJ(E), Vol.13, No.6, pp.411-418 (1992 Nov.).

第 18 章

A Neural Network for Word Category Prediction

18.1 作成

作成者 : 中村 雅巳、北岡
作成日付 : 1988 年 10 月

18.2 概要 説明

単語品詞予測モデルをニューラルネットワークでどのように実現しているかを紹介するデモ。

18.3 デモソフトウェアの所在

Machine: atrq30
Directory: /q30_data/users/demo/DEMO_SOFT_DIR

18.4 Directory 構造

```
NETgram/ --- SRC/  
|   <C ソース 及びヘッダファイル>  
|  
+---data/  
|   <データ ファイル>  
|  
+--- <リスト ファイル>  
|   List.19  
|  
+--- <実行イメージ>  
|   spot19  
|  
+--- BAKS/
```

18.5 make について

ディレクトリー SRC/ において、Makefile を実行すると、実行イメージ spot19 が、NETgram/ のしたに作成される。

18.6 実行の仕方

<実行形式>

% spot19 List.19

18.7 開発メモ

18.7.1 original

18.7.2 modified

18.7.3 Workstation Install Memo

18.8 関連研究論文など

1. 中村, 鹿野, コネクショニストモデルによる単語列予測の検討, 音学講論, 3-P-8, pp.243-244 (1988-03).
2. 中村, 鹿野, ニューラルネットによる N-gram 単語列予測モデルの検討, 音学講論, 2-P-2, pp.205-206 (1988-10).
3. 丸山, 中村, 川端, 鹿野, HMM 音韻認識と NETgram を用いた単語音声認識, 音響学会講演論文集, 2-P-8, pp.145-146 (1989-10).
4. M.Nakamura, K.Maruyama, T.Kawabata, K.Shikano, Neural Network Approach to Word Category Prediction for English Texts, COLING'90, Helsinki (1990-08).

第 19 章

Noise Reduction Using a Neural Network

19.1 作成

作成者：田村 震一

作成日付：

19.2 機能説明

ニューラルネットワークを用いて雑音の混入した音声から雑音抑圧を行うデモ指定した List ファイルから、OUTPUT ファイル名と INPUT ファイル名を読みとり波形を表示する。

19.2.1 List ファイルのフォーマット

(ファイル数の MAX は def.h で定義。MAXFC 50)

```
[display_name]  [file_1]  [file_2]
:               :       :
:               :       :
```

```
file_1 --> noisy file name.
file_2 --> noise-free file name.
```

* example (Training data) name of file must begin at 't' !!

```
TRAINING_1      t1_o   t1_i
TRAINING_2      t1_o   t1_i
:               :       :
```

* example (Non-training data)

```
NON_TRAINING_1  n1_o   n1_i
NON_TRAINING_2  n1_o   n1_i
:               :       :
```

19.2.2 起動時の説明

まず、一番目のデータが読みこまれ、2つの波形が表示されます。その際ファイル名の頭文字が "t" ならば << TRAININGDATA >>、"t" 以外ならば << NON-TRAINING DATA >> というタイトルが表示されます。

画面のメニューは、マウスの左ボタンで操作します。

Next : 次のデータファイルを表示する。

Prev : 一つ前のデータファイルを表示する。
 Select : データファイルを選択する。
 D/A : 音声出力 (daout) を行う。
 Procedure : 波形の拡大。(後述)
 Exit : 終了。

19.2.3 Procedure について

Procedure メニューで左ボタンをクリックすると、波形データの 15 msec が拡大されて、新たなウィンドウに表示されます。この時、一番下についているスクロールバーを操作することにより、波形データをスクロールさせて任意の 15 msec 部分を表示させることができます。

Return : Procedure ウィンドウを消して元のウィンドウに戻る。

19.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

19.4 Directory 構造

```
NR/ --- SRC/
    | (Cソース及びヘッダファイル)
    |
    +--- DATA/
        | (データファイル)
        |
        +--- <<リストファイル>>
            | List
            |
            +--- <<実行イメージ>>
                | nnn
                |
                +--- BAKS/
```

19.5 run の仕方

<実行形式>

```
program List -f data_directory
```

```
----> " nnn List -f DATA "
```

19.6 コンパイルの仕方と注意事項

NR/SRC/ の下において Makefile を実行してください。コンパイルには、/usr/lib/ の下に libX11.a libdwt.a /usr/include/ の下に X11/ が、必要です。

19.7 備考

波形出力はユーティリティを使用しています。

(Transform.c Transform.h TransformP.h Wave.c Wave.h WaveP.h)

19.8 開発メモ

19.8.1 original

19.8.2 modified

19.8.3 Workstation Install Memo

19.9 関連研究論文など

1. 田村震一, 波形入出力による雑音抑圧ニューラルネットワークの解析, 音学講論, 2-P-18, pp.237-238 (1988-10).
2. S.Tamura, An Analysis of a Noise Reduction Neural Network, Proc. ICASSP89, Vol.3, pp.2001-2004 (1989-05).
3. 田村, 中村, 波形入出力による雑音抑圧ニューラルネットワークの改良, 音響学会講演論文集, 3-4-18, pp.303-304 (1990-03).
4. 大倉, 杉山, 波形入出力による雑音抑圧ニューラルネットの音声認識への応用, 音響学会講演論文集, 1-8-3, pp.5-6 (1990-09).
5. 田村, 中村, 波形入出力による雑音抑圧ニューラルネットワークの最適化, 信学技法, SP89-120, pp.9-14, 1990.
6. 田村, 中村, 波形入出力による雑音抑圧ニューラルネットワークの改良, 春季音学講論, 3-4-18, pp.303-304, 1990.
7. S.Tamura, M.Nakamura, Improvements to the noise reduction neural network, ICASSP90, Albuquerque, pp.825-828, 1990.

第 20 章

Neural Network Workbench

20.1 作成

作成者 : 中村雅己 (村上、北岡 [DEC])
作成日 : 1989.10

20.2 機能説明

ニューラルネットワークの研究用支援プログラムのデモ。

ワークステーション上にネットワークワーク構成をグラフィック表示し、ネットワークの設計・構築、学習状況のモニタ、学習後のネットワークの解析、テストランを行なう。ニューラルネットワークの学習プログラムとして DCP2 を使用。

20.3 デモソフトウェアの所在

Machine: atrq30
Directory: /q30_data/users/demo/DEMO_SOFT_DIR

20.4 Directory 構造

```
NNWB/ --+--- SRC/ --+--- DEC_UTILS/
|
|
|          +--- MOTIF_UTILS/
|          |
|          |
|          +--- <<C source programs>>
|              -> nnwb
|
+--- DCP2/ --+--- bdc_samp/
|
|
|          +--- src/
|              (C source programs)
|              -> dcp2
|
|          +--- <<data file>>
|              dcp2.additional
|
+--- Sample/
|
```

```
|  
+-- TRAIN/  
|  
|  
+-- bdg/  
|  
|  
+-- <<実行イメージ>>  
    dcp2   nnwb
```

20.5 Make について

Directory NNWB/SRC/ において、Makefile を実行してください。実行イメージ nnwb が作成されます。Directory NNWB/DCP2/src/ において、Makefile を実行してください。実行イメージ dcp2 が作成されます。また、作成された実行イメージは、Directory NNWB/ の下へ置いてください。

20.6 実行の仕方

<実行形式>

```
% nnwb
```

20.7 開発メモ

20.7.1 original

20.7.2 modified

20.7.3 Workstation Install Memo

20.8 関連研究論文など

1. 中村, 田村, 他, “ニューラルネット開発用ワークベンチシステム-ネットワークエディタおよびモニタ機能について,” 音響学会講演論文集, 2-P-26, pp.181-182 (1989-10).
2. 中村, 鹿野: ニューラルネットワークによる N-gram 単語予測モデルの検討, 音響学会講演論文集, 2-P-2, pp.205-206 (1988-10).
3. 中村, 田村, 他: ニューラルネット開発用ワークベンチシステム-ネットワークエディタおよびモニタ機能について-, TR-I-0113 (1990-09).
4. 中村, 鹿野: ニューラルネットワークにおけるバックプロパゲーション学習の効率化方法, TR-I-0113 (1990-09).

第 21 章

Speech Synthesis Using Non-Uniform Units

21.1 作成

作成者 : 道尻、海木、岩橋、三村

作成日 : 平成 4 年 1 0 月

21.2 機能説明

種々の複合音声単位を選択的に用いる音声規則合成実験システムの紹介デモ。本音声規則合成実験システムは、ルールによる韻律情報の生成、ケプストラム距離に基づいた 4 つの単位選択基準を用いた最適な複合音声単位の選択、およびケプストラムパラメータによる素片接続・編集・合成といった処理を統合化したものである。スペクトル生成手法および韻律決定手法に関するデモ、任意文章のキーボード入力による合成のデモなどを含む。

21.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

21.4 Directory 構造

```
newTALK/ --+-- CURSOR/
|      (C source program)
|
+-- DEMO_VOICE/
|
|
+-- INPUT_SRC/
|      (C source program :nTalk_input)
|
+-- MENU_SRC/
|      (C source program :nTalk_menu)
|
+-- PATH_DATA/
|
|
+-- SELECT_SRC/
|      (C source program :nTalk_select)
|
+-- SENTENCE/
|
|
```

```

+-- SPECT_GRAPH/
|   (C source program :spect_graph)
|
+-- SPECT_SRC/
|   (C source program :nTalk_spect)
|
+-- SPECT_WAV/
|
|
+-- VOICE_SCR/
|   (C source program :voice ,voice_vax)
|
+-- XINPUT_C/
|   (C source program)
|
+-- XMENU_C/
|   (C source program)
|
+-- <<実行イメージ>>
|   nTalk_input      nTalk_menu
|   nTalk_select     nTalk_spect
|   voice            voice_vax
|   spect_graph
|
+-- <<シェルスクリプト>>
|   demo_conversation1.sh
|   demo_conversation1_vax.sh
|   demo_conversation2.sh
|   demo_conversation2_vax.sh
|   demo_outline.sh
|   demo_outline_vax.sh
|   voice_out.sh
|   spect_wav_vax.sh
|
+-- <<データファイル>>

```

21.5 Make について

21.5.1 nTalk-input

directory INPUT_SRC/ において Makefile を実行してください。newTALK/ の下に nTalk_input が作成されます。

21.5.2 nTalk-menu

directory MENU_SRC/ において Makefile を実行してください。newTALK/ の下に nTalk_menu が作成されます。

21.5.3 nTalk-select

directory SELECT_SRC/ において Makefile を実行してください。newTALK/ の下に nTalk_select が作成されます。

21.5.4 nTalk-spect

directory SPECT_SRC/ において Makefile を実行してください。newTALK/ の下に nTalk_spect が作成されます。

21.5.5 spect-graph

directory SPECT_GRAPH/ において Makefile を実行してください。newTALK/ の下に sprct_graph が作成されます。

21.5.6 voice, voice-vax

direcory VOICE_SRC/ において C source program voice.c 及び voice_vax.c をコンパイルし、newTALK/ のしたへ move してください。

21.6 実行の仕方

<実行形式>

```
% nTalk_menu
```

21.7 関連研究論文など

21.8 開発メモ

21.8.1 original

21.8.2 modified

21.8.3 Workstation Install Memo

第 22 章

Voice Conversion Using VQ Codebook Mapping

VQ 符号帳の写像に基づく声質変換 (Voice Conversion) 法のデモ

ベクトル量子化手法を用いて話者の音声空間を複数個のベクトルで表現し、話者ベクトルの集合間の対応関係に基づき、ベクトルを入れ換えることにより、話者の声質変換を実現している。

22.1 作成

作成者 : 阿部 匡伸 (現在 NTT HI 研究所)

連絡先: ave@nttspch.ntt.jp

作成日付 :

22.2 機能説明

日本語の単語 (3つ) と文章 (2つ) と英語の文章 (3つ) の声質変換デモ。画面上でクリックした単語または文章について、定められた順序で声質変換の daout を行う。単語と文章はあらかじめ決まっているので、List ではそのデータのディレクトリとファイル名を指定する。

22.2.1 List ファイルのフォーマット

DIR_1: VQ/DATA/demo2/mau	--> 男声データのディレクトリ。
DIR_2: VQ/DATA/demo2/mau_P	--> pitch 変換データの "
DIR_3: VQ/DATA/demo2/mau_s	--> spectrum 変換データの "
DIR_4: VQ/DATA/demo2/mau_fsk	--> pitch & spectrum 変換データの "
DIR_5: VQ/DATA/demo2/fsk	--> 女声データの "
TAN_1: 124	--> " パイプ " の音声データファイル名。
TAN_2: 130	--> " びょうしゃ "
TAN_3: 132	--> " びょうどう "
BUoN_1: SB3_02	--> " こちらは 第一回 通訳 電話 国際 …… "
BUN_2: SC3_16	--> " 京都駅から京都国際会議場まで時間 …… "
ENG_1: mit_1	--> " My name is MITalk. "
ENG_2: mit_7	--> " May I have your name ? "
ENG_3: mit_6	--> " Can I help you ? "

22.2.2 起動時の説明

画面に 1~5 までの円 (上記のディレクトリに値する) と、単語と文章がそれぞれ表示され、選んだ単語または文章について、定められた順序で daout が行われる。(daout している時、その番号の円が Flush する) Exit にて終

了。

日本語の単語 (3つ) の場合 : 1 ~ 5 まで daout する。
 日本語の文章 (2つ) の場合 : 1 と 4 を daout する。
 英語の文章 (3つ) の場合 : 1 と 4 を daout する。

Training Sample : DEMO の 7 番を実行する。
 (/usr/users/speech/ の下に LVC_sample* が必要)

22.3 デモソフトウェアの所在

Machine: atrq30

Directory: /q30_data/users/demo/DEMO_SOFT_DIR

22.4 Directory 構造

```

VQ/ --- SRC/
|   (Cソース及びヘッダファイル)
|
+--- DATA/ --- AD/
|           |
|           +--- CEP/
|           |
|           +--- PMOD8/ --- SYN_03/
|           |           +--- SYN_07/
|           |           +--- <<データファイル>>
|           |           *****
|           |
|           +--- TIME/
|           |
|           +--- demo/ --- fsk/
|           |           +--- mau/
|           |           +--- mau_fsk/
|           |           +--- temp1/
|           |           +--- temp2/
|           |           +--- <<リストファイル>>
|           |           list list.BAK
|           |
|           +--- demo2/ --- FVQmau_fsk/
|           |           +--- FVQmau_mau/
|           |           +--- fsk/
|           |           +--- mau/
|           |           +--- mau_fsk/
|           |           +--- mau_mht/
|           |           +--- mau_mnm/
|           |           +--- mau_p/
|           |           +--- mau_s/
|           |           +--- mht/
|           |           +--- mht_mau/
|           |           +--- mnm/
|           |           +--- mnm_mau/

```

```

|           |           +-- <<リストファイル>>
|           |           list list.BAK
|           |
|           |           +-- demo3/ ---+-- fsk/
|           |           +-- mit/
|           |           +-- mit_fsk/
|           |
|           |           +-- select/ ---+-- fsk/
|           |           +-- fsk_fts/
|           |           +-- fts/
|           |           +-- mau/
|           |           +-- mau_fsk/
|           |
|           |           +-- <<X image>>
|           |           VC.img
|
+-- fsk/
|   (データファイル)
|
+-- mit/
|   (データファイル)
|
+-- <<Traning Sample File>>
|   LVC_sample
|
+-- <<リストファイル>>
|   List
|
+-- <<実行イメージ>>
|   qvox.eng
|
+-- BAKS/
|   BAKS_2/

```

22.5 コンパイルの仕方と注意事項

VQ/SRC/ の下において Makefile を実行してください。コンパイルには /usr/lib/ の下に libX11.a libdwt.a と /usr/include/ の下に X11/ が必要です。

22.6 実行の仕方

<実行形式>

```
% qvox.eng List
```

22.7 備考

daout の長さは、x_io.eng.c の da_flush() という関数の中の sleep(n) で調節しています。円を Flush させる回数は、Qvox.eng.h の中で定義しています。daout に渡すファイル名の長さは 40 文字までとなっています。それをこえると "Invalid Number of Channel" というエラーがでますので、注意してください。

22.8 開発メモ

22.8.1 original

当時の派遣アルバイトの方 (丸顔の女の方, 名前は忘れてしまいました) に書いてもらったものです。(by 阿部)

22.8.2 modified

22.8.3 Workstation Install Memo

22.9 関連研究論文など

1. 阿部 匡伸, 中村 哲, 鹿野 清宏, 桑原 尚夫, ベクトル量子化による声質変換, 日本音響学会秋季研究発表会, (1987-10).
2. 阿部 匡伸, 中村 哲, 鹿野 清宏, 桑原 尚夫, ベクトル量子化による声質変換の評価, 日本音響学会春季研究発表会, (1988-3).
3. 阿部 匡伸, Cross-language voice conversion, 日本音響学会春季研究発表会, (1990-3).
4. 阿部 匡伸, 中村 哲, 鹿野 清宏, 桑原 尚夫, ベクトル量子化による声質変換, 電子通信情報学会 音声研究会, (1988-2).
5. 阿部 匡伸, 中村 哲, 鹿野 清宏, 桑原 尚夫, 言語間にわたる声質変換, 電子通信情報学会 音声研究会 (1990-2).
6. M.Abe, S.Nakamura, K.Shikano, H.Kuwabara, Voice conversion through vector quantization, ICASSP88 (1988).
7. M.Abe, K.Shikano, H.Kuwabara, Cross-language voice conversion, ICASSP90 (1990).
8. M.Abe, S.Nakamura, K.Shikano, H.Kuwabara, Voice conversion through vector quantization, The Journal of the Acoustical Society of Japan, Vol.11, No.2, pp.71-76 (Mar. 1990).
9. M.Abe, K.Shikano, H.Kuwabara, Statistical analysis of bilingual speaker's speech for cross-language voice conversion, The Journal of the Acoustical Society of America, Vol.90, No.1, pp.76-82 (July 1991).

第 23 章

ATR 音声言語翻訳実験システム (ASURA)

23.1 作成

作成者：竹沢寿幸

作成日：1992年11月30日

23.2 機能説明

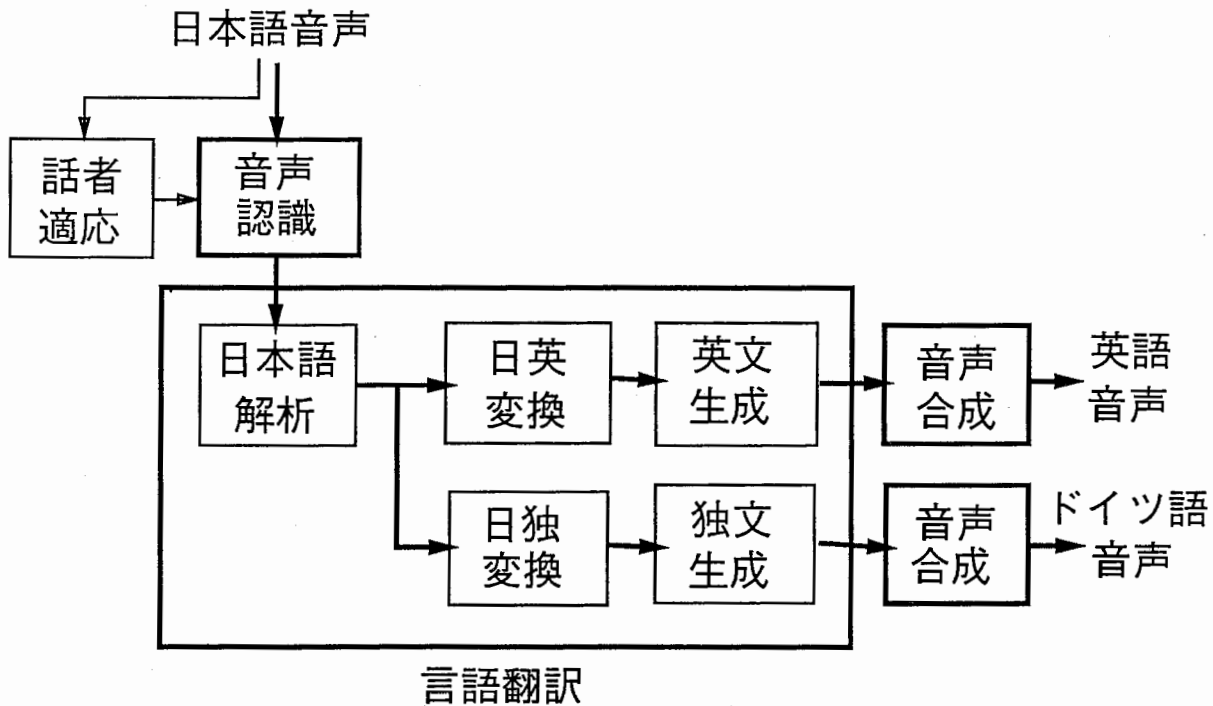
[概要] 異なる言語を話す人とのコミュニケーションが円滑に行なえることを目指して、音声言語翻訳実験システム ASURA を作成し、その評価実験を行ないました。音声認識部は ATR で提案・開発した話者適応機能付きの連続音声認識方式を採用しており、多数話者に対して高い認識率を達成しています。言語翻訳部は、ATR で提案・開発した話し言葉翻訳手法を採用しており、英語の他に新たにドイツ語も出力するように拡張しています。

[特長]

1. 音声認識、言語翻訳および音声合成を一貫して結合したもので、特に、音声認識と言語処理との間の音声言語処理において、それらの間の相互作用を円滑に行なう構成となっています。
2. 音声認識における基本単位である音素モデルは、前後の音素の影響を詳細に反映した精度の高いモデルです。少量の学習データでの話者適応が可能なので、多数話者に対しても高い音声認識率を達成しています。
3. 言語翻訳では、日本語の話し言葉に頻出する省略表現や、間接的な依頼のような多種多様な表現を扱うことができます。処理系と規則を完全に分離し、言語表現の取り扱いを一般的にしていますので、拡張性に優れています。

[性能] 国際会議に関する問合せを実験タスクとし、音声認識部、言語翻訳部ともに約 1,500 語の語彙を扱っています。言語翻訳部では日本語話し言葉の標準的な言いまわしの約 90% を扱うことができます。

[システム構成]



23.3 Directory 構造

23.3.1 atrp11

```

atrp11:/users/cstar/DEMO_R1.0/
|
+- COMMU/
| |
| +- ASURA_MAIN_SUB/
| | |
| | +- 93.1.12/
| | +- 93.1.14/
| | +- 93.1.20/
| | +- P_SSS_LR/
| |
| +- EXE/
| | |
| | +- ALL_LOG/
| | +- CMU_LOG/
| | +- GER_LOG/
| | +- LIST/
| | +- OLD_csh/
| | +- TEST_LOG/
| | +- sh.old/
| |
| +- MODEM/
| +- OPTION_WINDOW/
| | |
| | +- BACK930125/
| |
|

```

```
| +- PLOT_WINDOW/
| | |
| | +- BACK930113/
| | +- BACK930125/
| |
| +- X25/
|
+- FONTS/
| |
| +- HELV.bdf/
| +- HP.snf/
| +- SUN.snf/
| +- TBGM.bdf/
|
+- RECOG/
| |
| +- Adapt/
| | |
| | +- Seqlist/
| |
| +- BINHP -> SRC/BINHP/
| +- BINSUN -> SRC/BINSUN/
| +- Data/
| | |
| | +
| |
| +- Dur/
| +- ETC/
| +- Exe/
| +- Exe_C-STAR/
| | |
| | +- BAK.930119/
| | +- LOGS/
| |
| +- Gra/
| +- HMnet/
| | |
| | +- FMS/
| | +- FTK/
| | +- FYM/
| | +- MAU/
| | +- MHT/
| | +- MXM/
| |
| +- Reco/
| +- SRC/
| | |
| | +- ADAPT/
| | +- ADDHEADER/
| | +- ALLOC_DSP/
| | +- BINHP/
```

```

| | +- BINHP.org/
| | +- BINSUN/
| | +- DEMOCANDX2/
| | | |
| | | +- TMP/
| | |
| | +- DSP/
| | +- ENDIAN/
| | +- FZK_NGRAM/
| | +- GKILL/
| | +- INCLUDE/
| | +- LPC/
| | +- MAKE_SEQLIST/
| | +- MAKE_WORD_SAMPLE/
| | +- PARAIN/
| | +- PARAIN.930111/
| | +- PARAIN.930201/
| | +- POSTFILTER/
| | +- RMHEADER/
| | +- SOCKET/
| | +- SP_SELECT/
| | +- SSS/
| | | |
| | | +- Src_parser/
| | | +- Src_verify/
| | |
| | +- SSSLIKE/
| | +- THROUGH/
| | +- VAR_CONV/
| | +- WAIT_CUT_BUFF/
| | +- XEXTRACT2/
| | | |
| | | +- BAK.930115/
| | | | |
| | | | +- DISP/
| | | | |
| | | | +- BAK.930108/
| | | | +- BAK.930113/
| | | | |
| | | | +- FONT_TEST/
| | | | |
| | | | +- DISP/
| | | | +- Exe/
| | | | +- PARENT/
| | | | |
| | | | +- PARENT/
| | | | |
| | | | +- BAK.930113/
| | | | |
| | | | +- TRANS_DISP/
| | | |

```

```
| | | +- Ver1.1.930115/
| | | |
| | | +- DISP/
| | | +- Exe/
| | | +- PARENT/
| | | +- TRANS/
| | | +- TRANS_PARENT/
| | |
| | +- gomi/
| | +- inc/
| |
| +- Sp_select/
| | |
| | +- Seqlist/
| | |
| +- Unseg/
|
+- TOOL/
| |
| +- BITMAP_BUTTON/
| +- ENDIAN/
| +- NKF/
| +- P-JANUS/
| | |
| | +- DEMO_MODEM/
| | +- MODEM/
| |
| +- RFILTER/
| +- SpeechIn/
| | |
| | +- BINHP/
| | +- DOC/
| | +- SRC/
| | |
| | +- .tmp/
| | | |
| | | +- adin_as28.c/
| | | +- d_as28.c/
| | | +- daout_as28.c/
| | | +- dasbox.h/
| | | +- udm.h/
| | |
| | +- RCS/
| |
| +- SpeechIn.930111/
| | |
| | +- BINHP/
| | +- DOC/
| | +- SRC/
| | |
| | +- .tmp/
```

```

| | +- RCS/
| |
| | +- WAIT_BUTTON/
| | +- XINPUT_SAGA/
| | +- XMENU_SAGA/
| |
+- TRANS/
|
+- filterdemo/
|
+- resultfile/
+- translation.C-STAR/
|
+- TEST/
+- TRANS_SRC/
| |
| | +- BAKS/
| |
+- TRANS_SRC.930119/

```

23.3.2 as22

as22 (SUN machine) においては、atrp11:/users/cstar/DEMO_R1.0/ が as22:/home/cstar/DEMO_R1.0/ に、リンクされています。

23.3.3 atrp19

atrp19:/users/project/asura/demo/cstar/demo/

```

|
+- cstar/
| |
| | +- .vue/
| | |
| | +- apps/
| | +- atrp19:0/
| | |
| | +- current/
| | |
| | +- palettes/
| | +- types/
| | |
| | +- tools/
| |
+- english/
| | |
| | +- com-interface/
| | |
| | +- bacup.921215/
| | |
| | +- emorph_dict/
| | +- src/
| | |
| | |

```

```
| | | +- talk/
| | |
| | +- world/
| | |
| | +- mk-acp/
| | +- mk-nlp/
| |
| +- eval/
| | |
| | +- inputfile/
| | +- outputfile/
| | +- translation/
| |
| +- german/
| | |
| | +- Data/
| | | |
| | | +- irdata/
| | |
| | +- LOG/
| | +- com-interface -> ../english/com-interface/
| | +- gmorph_dict/
| | +- tmp/
| | +- world/
| | |
| | +- mk-acp/
| | +- mk-nlp/
| |
| +- tmp/
|
+- filterdemo/
  |
  +- asura_disp/
  +- asura_disp.bak/
  | |
  | +- bak/
  | +- bin/
  |
  +- inputfile/
  +- postfilter/
  +- resultfile/
  +- speechsynthesis/
  +- translation/
  +- translation.C-STAR/
    |
    +- NEW_TRANS/
    +- OLD_TRANS/
    +- SOC_TEST/
```

23.3.4 atrp17

```
atrp17:/users/cstar/C_STAR/
|
+- WV/
+- job_loop/
+- work/
```

23.4 Make について

テクニカルレポートを参照してください。

23.5 実行の仕方 及び、注意

[注意]

現在 ASURA を実行するには、以下のマシンが必要です。また、それぞれをネットワークでつなぐ必要があります。

- as22 (SUN) : 通信制御、及び通信オペレーション用
- atrp11 (HP) : 音声認識、及び話者オペレーション用
- atrp19 (HP) : 翻訳用
- atrp17 (HP) : 音声合成用

さらに、実行時には、モデムをつなぐ相手側のシステムが必要です。

[実行形式]

as22 (SUN machine) のディレクトリー /home/cstar/DEMO_R1.0/COMMU/EXE/ において以下のコマンドを実行して下さい。(注意: ディレクトリー /home/cstar/DEMO_R1.0/ は、atrp11(HP machine) にリンクされています。)

```
% ASURA
```

23.6 開発メモ

23.6.1 original

23.6.2 modified

23.6.3 Workstation Install Memo

23.7 関連研究論文など

1. S.Sagayama, M.Sugiyama, etal, ATREUS: Continuous Recognition Systems at ATR Interpreting Telephony Research Laboratories, SST92 (1992), 採録決定.
2. 「ATR 音声言語翻訳実験システム ASURA」竹沢寿幸、森元暹、谷戸文廣、鈴木雅実、嵯峨山茂樹、樽松明 (ATR 自動翻訳電話研究所) 情報処理学会第 46 回全国大会 (1993 年 3 月発表予定)
3. 竹沢寿幸, 森元暹, 樽松明: “日英音声言語翻訳実験システム SL-TRANS 2 における音声対話処理”, 1991 年電子情報通信学会秋季全国大会, D-53 (1991-09).
4. 竹沢寿幸, 大倉計美, 森元暹, 嵯峨山茂樹, 樽松明: “日英音声言語翻訳実験システム SL-TRANS 2”, 日本音響学会平成 3 年度秋季研究発表会講演論文集 (1991-10).

5. Tsuyoshi Morimoto, Masami Suzuki, Toshiyuki Takezawa, Gen'ichiro Kikui, Masaaki Nagata, Mutsuko Tomokiyo: "A Spoken Language Translation System: SL-TRANS2", *Proceedings of the fifteenth International Conference on Computational Linguistics (COLING-92)*, pp. 1048-1052 (1992-08).
6. Tsuyoshi Morimoto, Toshiyuki Takezawa, Kazumi Ohkura, Masaaki Nagata, Fumihito Yato, Shigeki Sagayama, Akira Kurematsu: "Enhancement of ATR's Spoken Language Translation System: SL-TRANS2", *Proceedings of the International Conference on Spoken Language Processing (ICSLP-92)* (1992-10).
7. 嵯峨山、他: 「自動翻訳電話実験システム: ASURA」日本音響学会平成5年度春季研究発表会講演論文集(1993-3) 発表予定.
8. S.Sagayama, M.Sugiyama, etal: "ATREUS: Continuous Recognition Systems at ATR Interpreting Telephony Research Laboratories," *Proc SST92 (Brisbane)*, pp. 324-329 (1992).

第 24 章

デモ開発用ツール

デモソフトウェアを効率的に開発するために以下に示すようなツールを作成・使用しています。簡単にツール名称・使用方法を以下の形式で述べます。

ソースディレクトリー名: ファイル名
簡単な機能紹介

24.1 一般用ツール

1. COLORVIEW: colorviewer
X display 上で、表現可能なカラーを検索する。
2. FONTLIST: fontlist
DEC station 上の、X 用 DECfont の検索と表示。
3. HARDCOPY: hardcopy X display の、ハードコピーをとる。
4. XGRAPH1: xgraph1
簡易、グラフ表示ツール。

24.2 研究用ツール

1. SPAC_N: spac_n
音声データの、分析表示ツール。???

24.3 ASURA デモ用ツール

1. DEMOCANDX2: democandx2, srv-democandx2
音声認識候補、表示 window
2. XEXTRACT2: cli_xextract2, srv_xextract2
cli_trans, srv_trans
翻訳候補、結果、表示 window
3. XMATRIX: xmatrix
likelihood map その他の、
カラーグラデュエイト表示 window

24.4 制御用ツール (シェルスクリプト)

1. XINPUT_SAGA: xinput
Key 入力要求 window
2. XMENU_SAGA: xmenu
メニュー選択要求 window
3. XW_BUTTON: wait_button
ボタンプレス待ちプロセス。

24.5 プログラム開発用, lib, その他

1. TOOLKIT_WIDGETS
2. NnfpWidget: wbtNn
ニューラルネットの発火パターン表示用、Xtoolkit Widget
3. TRANS_COORDINATES
4. XG: XG.o
5. XGKS: libXGKS.a
6. XKRNL11: libXKRNL11.a
7. DR_SPECT: libDr.Spect.a, libcXgks.a
GKS 風座標変換を使った、X の基本描画のための、ライブラリー。

24.6 ファイルの所在

ファイルの所在¹を以下に記します。

Location of source

hostname: atrq01
directory: /q01/users/local_xtools/src

Location of binary

hostname: 使用されているマシン
directory: /usr/local/bin or demo/bin

¹ライブラリーについては、デモ用ソフトと共に置かれています。

第 25 章

謝辞

編集にあたり御執筆協力いただいた方々に心よりお礼を申し上げます。執筆分担は以下の通りです。本解説を作成した時期に既にデモソフトウェアの作成担当者・開発責任者がいない場合があり、説明が不十分な部分もあり得るます。その点について御容赦願うこととします。

表 25.1: 執筆分担

項目	担当者	略称	デモ項目
11	永井明人	HMM-LR	Speaker-Adapt. Cont. Speech Recognition (CBM)
12	宮沢康永	VFS-CHMM	VFS Speaker Adaptation CMD-HMM-LR
14	山口耕市	VFS-DHMM	Speaker-Adapt. Cont. Speech Recognition
	山口耕市	SI-DHMM	Speaker-Indep.(male) Cont. Speech Recognition
15	山口耕市	VFS-CHMM-new	
17	宮沢康永	HMM-HARD	
20	鷹見淳一	SSS-LR	SSS-based Speaker-Adapt. Cont. Speech Recognition
18	村上仁一	HMM-bigram	Word-bigram-based Sentence Speech Recognition
18'	村上仁一	HMM-trigram	Word-trigram-based Sentence Speech Recognition
1	福沢圭二	QUIZ	Phoneme Recognition using a TDNN
2	福沢圭二	SPOT2	Phoneme Spotting using TDNNs
3	加藤喜永	NETgram	A Neural Network for Word Category Prediction
4	加藤喜永	NR	Noise Reduction Using a Neural Network
6	鷹見淳一	PD-TDNN	PD-TDNN for Phoneme Recognition
7	福沢圭二	TDNN-LR	TDNN-LR Cont. Speech Recognition
8	福沢圭二	AD-TDNN-LR	Speaker-Adaptation + TDNN-LR
9	福沢圭二	IND-TDNN-LR	Speaker-Indep. NN-based Cont. Speech Recognition
13	福沢圭二	NNWB	Neural Network Workbench
16	加藤喜永	FPM-LR	Speaker-Indep. NN-based Cont. Speech Recognition
10	岩橋直人	nuuTALK	Speech Synthesis Using Non-Uniform Units
5	杉山雅英	VQ	Voice Conversion Using VQ Codebook Mapping
19	竹沢寿幸	SL-trans.A	

