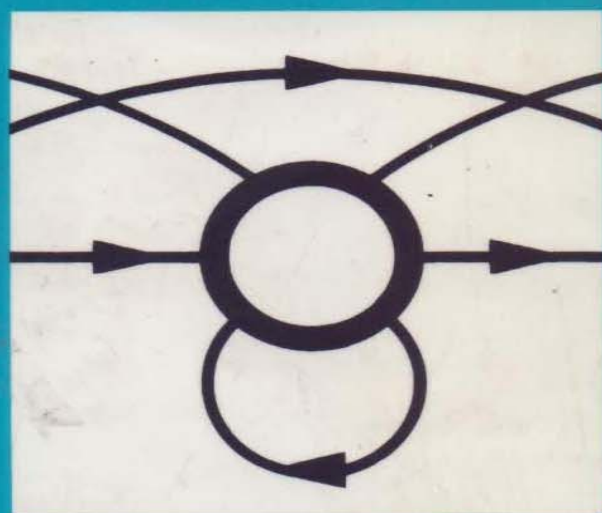


COLLECTION ÉLECTRICITÉ

Traitement de la parole

René Boite, Hervé Bourlard,
Thierry Dutoit, Joël Hancq
et Henri Leich



PRESSES POLYTECHNIQUES ET UNIVERSITAIRES ROMANDES

TABLE DES MATIÈRES

REMERCIEMENTS	V
AVANT-PROPOS.....	VII
CHAPITRE 1 INTRODUCTION.....	1
× 1.1 Traitement de la parole.....	1
1.2 Qu'est-ce que la parole?.....	4
1.2.1 Niveau acoustique	4
1.2.2 Niveau phonétique.....	13
1.2.3 Niveau phonologique	19
1.2.4 Niveau morphologique.....	22
1.2.5 Niveau syntaxique	22
1.2.6 Niveau sémantique	23
1.2.7 Niveau pragmatique (ou niveau du discours).....	24
1.3 Organisation de l'ouvrage	25
1.4 Bibliographie	26
CHAPITRE 2 MODÉLISATION DU SIGNAL DE PAROLE	27
× 2.1 Modélisation d'un signal aléatoire stationnaire.....	27
× 2.1.1 Modélisation d'un signal.....	27
2.1.2 Prédiction linéaire	28
2.1.3 Variance de l'erreur de prédiction.....	28
2.1.4 Estimation des coefficients de prédiction.....	29
2.1.5 Propriété d'orthogonalité.....	31
× 2.1.6 Interprétation de la prédiction linéaire: le modèle autorégressif (modèle AR).....	31
2.1.7 Gain du modèle	32
2.1.8 Filtre inverse.....	33
× 2.1.9 Propriété: fonction d'autocorrélation du signal AR.....	33
2.2 Prédiction avant et prédiction arrière	34
2.2.1 Définitions	34
2.2.2 Variance des erreurs de prédiction.....	35
2.2.3 Relations entre prédiction avant et prédiction arrière	36
2.3 Algorithmes de résolution	37
2.3.1 Méthode de Levinson-Durbin	37
2.3.2 Algorithme de Levinson.....	39
2.3.3 Interprétation de l'algorithme de Levinson: structure en treillis pour le filtre inverse	40

	2.3.4	Méthode de Schur.....	42
	2.3.5	Algorithme de Schur.....	45
	2.3.6	Définitions. Coefficients de corrélation partielle et paramètres LPC.....	46
	2.3.7	Stabilité du modèle AR.....	46
X	2.4	Modélisation ARMA.....	47
X	2.5	Modélisation AR du signal de parole.....	48
	2.5.1	Introduction.....	48
	2.5.2	Energie résiduelle de prédiction.....	49
	2.5.3	Systèmes d'équations.....	50
	2.5.4	Méthode dite de l'autocorrélation.....	51
	2.5.5	Méthode dite de la covariance.....	53
	2.6	Algorithmes de résolution.....	54
	2.6.1	Méthode de l'autocorrélation.....	54
	2.6.2	Méthode de la covariance. Décomposition de Cholesky.....	54
	2.6.3	Algorithme de Cholesky pour la méthode de la covariance.....	55
X	2.6.4	Estimation du modèle AR par la méthode du treillis.....	56
	2.6.5	Algorithme de Burg pour la méthode du treillis....	59
	2.7	Discussion des méthodes de résolution.....	60
	2.7.1	Choix des conditions d'analyse.....	60
X	2.7.2	Complexité des algorithmes.....	62
	2.7.3	Conditionnement des algorithmes de résolution ...	63
	2.7.4	Estimation robuste des coefficients LPC.....	64
	2.8	Bibliographie.....	64
CHAPITRE 3		PROPRIÉTÉS SPECTRALES DU SIGNAL VOCAL.....	65
	3.1	Introduction.....	65
	3.2	Analyse de Fourier court-terme.....	66
	3.2.1	Transformée de Fourier court-terme.....	66
	3.2.2	Calcul par algorithme TFR.....	67
	3.3	Analyse spectrale basée sur la prédiction linéaire.....	70
	3.3.1	Résumé des propriétés de la modélisation AR (cas stationnaire).....	70
	3.3.2	Cas du signal vocal.....	70
	3.3.3	Spectre du modèle.....	71
	3.3.4	Propriété.....	72
	3.3.5	Propriété du logarithme du spectre du modèle AR.....	73
	3.3.6	Comportement asymptotique de la modélisation AR.....	73
	3.3.7	Pondération du spectre par la modélisation AR....	74
	3.3.8	Modélisation AR et analyse spectrale par maximum d'entropie.....	75
	3.3.9	Représentation graphique des spectres court terme	76

✓ 3.4	Analyse spectrale homomorphique	77
× 3.4.1	Principe général. Cepstre complexe	77
× 3.4.2	Propriétés du cepstre complexe	78
× 3.4.3	Définition. Cepstre réel	79
× 3.4.4	Application au signal vocal. Cepstre court terme..	80
× 3.4.5	Estimation correcte de la réponse impulsionnelle.	83
× 3.4.6	Relations avec les coefficients de prédiction	83
× 3.4.7	Exemple d'analyse homomorphique	84
× 3.5	Métrique dans l'espace acoustique	85
× 3.5.1	Distance ou mesure de dissemblance entre vecteurs	85
× 3.5.2	Distance spectrale proprement dite	86
× 3.5.3	Distance cepstrale	86
× 3.5.4	Distance de Itakura-Saito	87
× 3.5.5	Mesures basées sur le rapport de vraisemblance...	88
× 3.5.6	Mesures de dissemblance modifiées	90
3.6	Estimation de la trajectoire des formants	90
3.6.1	Introduction	90
3.6.2	Estimation basée sur les paramètres LPC	91
3.7	Estimation de la période du fondamental et localisation des zones voisées	92
3.7.1	Introduction	92
3.7.2	Méthodes basées sur la fonction d'autocorrélation	93
3.7.3	Méthode basée sur une forme simplifiée de l'autocorrélation (AMDF)	94
3.7.4	Méthode basée sur le filtre inverse (SIFT)	95
× 3.7.5	Méthode du cepstre	95
× 3.7.6	Influence d'un bruit superposé au signal vocal....	96
3.8	Bibliographie	97
CHAPITRE 4	CODAGE	99
4.1	Introduction	99
4.2	Propriétés perceptuelles de l'oreille	101
4.2.1	Seuil d'audition absolu	101
4.2.2	Courbes d'effet de masque	101
4.2.3	Bandes critiques	103
4.3	Théorie de la quantification	105
4.3.1	Quantification instantanée	105
× 4.3.2	Quantification adaptative	116
4.4	Codage dans le domaine temporel	119
4.4.1	Modulation par impulsions codées (MIC-PCM)...	119
4.4.2	Système MIC différentiel (DPCM)	120
4.4.3	Système MIC différentiel adaptatif (ADPCM)	125
4.4.4	Système MIC différentiel adaptatif et prédiction à long terme	129
4.4.5	Mise en forme du bruit de quantification	130

4.5	Codage dans le domaine fréquentiel.....	133
4.5.1	Codage en sous-bandes.....	133
4.5.2	Codage adaptatif de la transformée.....	137
4.6	Codage paramétrique.....	141
4.6.1	Introduction.....	141
4.6.2	Codeur LPC.....	142
4.6.3	Quantification vectorielle.....	143
4.6.4	Codage du résidu (RELP).....	148
4.6.5	Codage par excitation multibandes (MBE).....	151
4.7	Codage par analyse par synthèse.....	156
4.7.1	Excitation de type multiimpulsionnelle (MPE).....	157
4.7.2	Excitation multiimpulsionnelle régulièrement espacée (RPE).....	159
4.7.3	Codage par dictionnaire d'excitations (CELP).....	162
4.8	Débruitage de la parole.....	164
4.8.1	Introduction.....	164
4.8.2	Techniques basées sur l'amplitude spectrale à court-terme.....	165
4.8.3	Modélisation de la parole et filtre de Wiener.....	167
4.8.4	Suppression adaptative du bruit.....	169
4.9	Bibliographie.....	171
CHAPITRE 5 RECONNAISSANCE DE LA PAROLE		
ET DU LOCUTEUR.....		
		175
5.1	Introduction.....	175
5.1.1	Problème et niveaux de complexité.....	175
5.1.2	Approche générale.....	177
5.2	Classification statistique de formes.....	179
5.2.1	Modèles de classification de formes.....	180
5.2.2	Règles simples à retenir.....	181
5.2.3	Bases de la statistique bayésienne.....	182
5.2.4	Classification statistique.....	183
5.2.5	Fonctions discriminantes.....	187
5.2.6	Entraînement.....	189
5.2.7	Algorithme d'estimation et maximisation (EM) ...	191
5.2.8	Classification de données réelles.....	198
5.3	Programmation dynamique.....	200
5.4	Modèles de Markov discrets.....	202
5.4.1	Définition et paramétrisation.....	202
5.4.2	Exemples d'utilisation.....	204
5.4.3	Du modèle de Markov discret au modèle de Markov caché.....	206
5.5	Modèles de Markov cachés (HMM).....	207
5.5.1	Définition.....	207
5.5.2	HMM pour la génération de séquences.....	208
5.5.3	Estimation de la séquence d'états.....	209
5.5.4	Modèles HMM autorégressifs.....	210

5.5.5	Modèles HMM pour la classification de séquences	211
5.5.6	Estimation des paramètres HMM.....	212
X 5.6	Reconnaissance de la parole: schéma-bloc	214
X 5.6.1	Analyse acoustique (acoustic front end)	214
5.6.2	Classification locale et décodage lexical.....	215
5.6.3	Analyse syntaxique et sémantique	216
5.7	Caractérisation du signal de parole.....	217
5.8	Reconnaissance de la parole par DTW.....	219
5.8.1	Introduction	219
5.8.2	Reconnaissance de mots isolés.....	221
5.8.3	Reconnaissance de mots enchaînés	228
5.9	Reconnaissance de la parole par HMM.....	232
5.9.1	Introduction	232
5.9.2	Approche générale.....	233
5.9.3	Modèle acoustique HMM.....	236
5.9.4	Paramétrisation et estimation des probabilités	237
5.9.5	Reconnaissance HMM	243
5.10	Entraînement des modèles HMM.....	245
5.10.1	Introduction	245
5.10.2	Entraînement «avant-arrière» (Baum-Welch)	247
5.10.3	Entraînement Viterbi	253
5.10.4	Estimateurs de probabilités acoustiques locales....	256
5.10.5	Lissage des paramètres.....	259
5.11	Limitations des modèles HMM.....	260
5.11.1	Hypothèses HMM	260
5.11.2	Discrimination	261
5.12	Réseaux de neurones artificiels (ANN).....	266
5.12.1	Description générale.....	267
5.12.2	Entraînement des réseaux de neurones.....	268
5.12.3	Propriétés des réseaux de neurones.....	270
5.12.4	Réseaux de neurones et séquences temporelles	270
5.12.5	Réseaux de neurones et inférence statistique	272
5.13	Modèles hybrides HMM/ANN	278
5.13.1	Motivations.....	278
5.13.2	Paramétrisation et estimation des probabilités	278
5.13.3	Entraînement et reconnaissance par modèles HMM/ANN	282
5.13.4	Pratique.....	285
5.13.5	Autres approches ANN	286
5.14	Contraintes linguistiques et reconnaissance	287
5.14.1	Introduction	287
5.14.2	Modèles phonologiques.....	288
5.14.3	Modèles de langage	291
5.14.4	Reconnaissance avec modèle de langage	295
5.15	Autres considérations	296
5.15.1	Topologies HMM.....	296
5.15.2	Adaptation des systèmes de reconnaissance	297

5.16	Détection de mots clés et niveau de fiabilité	298
5.16.1	Détection de mots clés.....	298
5.16.2	Niveau de fiabilité	299
5.17	Reconnaissance du locuteur.....	301
5.17.1	Introduction	301
5.17.2	Paramètres acoustiques.....	302
5.17.3	Mesures de similarité.....	303
5.17.4	Vérification du locuteur dépendante du texte.....	306
5.17.5	Vérification du locuteur indépendante du texte.....	307
5.17.6	Vérification du locuteur sur base de texte présenté.....	308
5.17.7	Identification, vérification et seuil de décision.....	310
5.18	Performances des systèmes et applications	310
5.18.1	Performances types des systèmes de reconnaissance	310
5.18.2	Applications.....	311
5.19	Résumé et conclusions.....	313
5.20	Bibliographie	314
CHAPITRE 6	INTRODUCTION AU TRAITEMENT DU LANGAGE NATUREL.....	323
6.1	Introduction.....	323
6.2	Généralités	324
6.3	grammaires régulières.....	327
6.3.1	Définition.....	327
6.3.2	Utilisation	329
6.3.3	Analyse régulière.....	331
6.3.4	Inférence régulière.....	332
6.4	Les grammaires hors-contexte.....	333
6.4.1	Définition.....	333
6.4.2	Utilisation	334
6.4.3	Analyse hors-contexte	336
6.4.4	Inférence de grammaires non contextuelles	337
6.5	Extensions des grammaires hors-contexte.....	337
6.6	Lexiques, attributs et notation PATR	339
6.7	Résumé	341
6.8	Bibliographie	342
CHAPITRE 7	SYNTHÈSE DE LA PAROLE À PARTIR D'UN TEXTE.....	345
7.1	Introduction.....	345
7.1.1	Définition.....	345
7.1.2	Applications.....	345
7.1.3	Analyse du problème.....	347
7.1.4	Organisation générale du module de traitement du langage naturel.....	348
7.1.5	Synthèse par règles – synthèse par concaténation.	350

7.2	Analyse morphosyntaxique	351
7.2.1	Prétraitement	351
7.2.2	Analyse morphologique	352
7.2.3	Analyse contextuelle	356
7.2.4	Analyse structurale	363
7.3	Phonétisation automatique.....	364
7.3.1	Analyse du problème.....	364
7.3.2	Phonétisation par règles – phonétisation par dictionnaire.....	366
7.3.3	Module morphophonémique	368
7.3.4	Convertisseur graphème-phonème.....	369
7.3.5	Module de post-traitement phonétique.....	373
7.3.6	Noms propres	374
7.4	Génération de la prosodie.....	374
7.4.1	Introduction	374
7.4.2	Groupes prosodiques	376
7.4.3	Placement automatique des accents	381
7.4.4	Génération de l'intonation.....	382
7.4.5	Génération de la durée.....	388
7.5	Synthèse par règles	389
7.5.1	Description générale.....	389
7.5.2	Synthèse par formants	391
7.5.3	Discussion	394
7.6	Synthèse par concaténation	395
7.6.1	Mise en forme de la base de données de segments	395
7.6.2	Synthèse proprement dite	398
7.6.3	Qualité segmentale	401
7.7	Synthèse par prédiction linéaire	405
7.7.1	Analyse LPC	405
7.7.2	Compression de la base de données paramétrique	405
7.7.3	Modification de la prosodie.....	406
7.7.4	Concaténation.....	406
7.7.5	Filtre de synthèse.....	407
7.7.6	Le choix de σ	408
7.7.7	Qualité segmentale	409
7.8	Synthèse hybride harmonique/stochastique.....	411
7.8.1	Modélisation hybride harmonique/stochastique...	411
7.8.2	Compression de la base de données paramétrique	412
7.8.3	Adaptation de la prosodie.....	412
7.8.4	Concaténation.....	413
7.8.5	Synthèse hybride	414
7.8.6	Qualité segmentale	417
7.9	Synthèse dans le domaine temporel	417
7.9.1	Introduction	417
7.9.2	TD-PSOLA.....	419
7.9.3	MBROLA.....	424
7.9.4	LP-PSOLA	427

	7.10 Résumé	427
	7.11 Perspectives	430
	7.12 Bibliographie	432
CHAPITRE 8	IMPLANTATION DES ALGORITHMES	443
	8.1 Solutions technologiques	443
	8.1.1 Introduction	443
	8.1.2 Approche «implantation sur microprocesseurs»...	444
	8.1.3 Microprocesseurs à usage général	444
	8.1.4 Microprocesseurs spécialisés pour le traitement du signal.....	449
	8.1.5 L'approche «implantation sur structure spécifique»	451
	8.1.6 Composants programmables	452
	8.1.7 Bilan des diverses approches.....	456
	8.1.8 Méthodologies et outils associés	457
	8.2 Programmation des microprocesseurs	457
	8.2.1 Contraintes liées au temps réel	457
	8.2.2 Complexité d'un algorithme. Alternatives «mémoire-puissance de calcul».....	459
	8.2.3 Erreurs de calcul	460
	8.2.4 Traitement en virgule fixe ou en virgule flottante.	462
	8.2.5 Langages de programmation	464
	8.2.6 Exemple d'étude d'implantation: l'analyse LPC ..	464
	8.3 Microprocesseurs spécialisés pour le traitement numérique du signal.....	466
	8.3.1 Structures	466
	8.3.2 Bilan et tendances des microprocesseurs spécialisés pour le traitement du signal	474
	8.4 Bibliographie	481
	GLOSSAIRE DES NOTATIONS	483
	LISTE DES ACRONYMES	487