Modélisation Bayésienne d'une boucle perception-action : application à la lecture et à l'écriture

Estelle Gilet

▶ **To cite this version:**

Estelle Gilet. Modélisation Bayésienne d'une boucle perception-action : application à la lecture et à l'écriture. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 2009. Français. NNT : . tel-00430866

HAL Id: tel-00430866

https://theses.hal.science/tel-00430866

Submitted on 10 Nov 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE JOSEPH FOURIER - GRENOBLE

N^o attribué par la bibliothèque
|_|_|_|_|_|_|_|_|_|_|

Thèse

pour obtenir le grade de

DOCTEUR DE L'UJF

Spécialité : *Mathématiques et Informatique*

préparée au laboratoire LIG et l'INRIA Rhône-Alpes, dans le cadre de l'École doctorale *Mathématiques, Sciences et Technologies de l'Information, Informatique*

présentée et soutenue publiquement par

Estelle GILET

le 02 octobre 2009

Titre :

Modélisation Bayésienne d'une boucle perception-action : application à la lecture et à l'écriture

Co-directeurs de thèse :

Pierre Bessière et Julien Diard

Composition du jury :

M.	Jean-Luc Schwartz	Président
M.	Gabriel Baud-Bovy	Rapporteur
M.	Jacques Droulez	Rapporteur
M.	Pierre Baraduc	Examinateur
M.	Pierre Bessière	Directeur de thèse
M.	Julien Diard	Co-directeur de thèse

Remerciements

Je souhaite tout d'abord remercier mes directeurs de thèse, Pierre Bessière et Julien Diard pour leur soutien, leurs conseils et leur amitié. Je remercie aussi les membres du Jury pour l'intérêt qu'ils ont porté à mon travail et leurs précieux commentaires sur le document. Je n'oublie pas tous les membres de l'équipe E-Motion, qui m'ont accueilli durant ce travail et avec qui j'ai pu avoir des discussions enrichissantes. Enfin, j'adresse un grand merci à Christophe, à ma famille et mes amis qui m'ont soutenu durant tout ce temps.

Merci à tous.

Résumé

Comment notre système moteur peut-il influencer notre perception ? Quels sont les mécanismes qui permettent de réaliser un même mouvement avec différents effecteurs ?

Afin d'apporter des éléments de réponse à ces deux questions issues des sciences cognitives, nous avons étudié un système sensori-moteur : celui mis en oeuvre pour les tâches de lecture et écriture de lettres cursives isolées. Notre contribution est un modèle formel de cette boucle sensori-motrice, qui tient compte des observations et théories issues de la littérature.

L'hypothèse centrale est qu'il existe un codage interne des lettres commun à la perception et à l'action, et, de plus, indépendant de l'effecteur. Ce codage, pivot entre la perception et l'action, consiste en des séquences de points de contrôle, géométriquement remarquables, identifiables à partir de la trace des lettres, et permettant de récupérer une partie de l'information cinématique.

Le formalisme utilisé est celui de la *programmation bayésienne*. La définition d'un programme bayésien se fait par l'écriture d'une distribution de probabilité conjointe.

Connaissant cette probabilité conjointe, le modèle est capable de résoudre plusieurs tâches cognitives. Chacune de ces tâches correspond à une *question probabiliste* qui est résolue automatiquement par inférence bayésienne :

- la lecture,
- la lecture en simulant le geste d'écriture,
- la copie de traces,
- la reconnaissance du scripteur,
- l'écriture avec différents effecteurs.

Nous avons simulé informatiquement chacune de ces inférences et comparé leurs résultats aux observations issues des expériences de psychophysique. Le modèle permet d'effectuer des prédictions et ainsi de suggérer de nouveaux protocoles expérimentaux.

Table des matières

1 Introduction **15**
- 1.1 Contribution . . . 16
  - 1.1.1 Description des sous-modèles . . . 17
  - 1.1.2 Simulation interne du mouvement . . . 18
  - 1.1.3 Suggestion de protocoles expérimentaux . . . 18
- 1.2 Plan de lecture . . . 19
- 1.3 Cadre de travail . . . 20

2 Architecture du modèle BAP **21**
- 2.1 Une représentation interne unique ? . . . 22
  - 2.1.1 Théories de la perception de la parole . . . 22
  - 2.1.2 BAP : La représentation interne des lettres pour la lecture et l'écriture . 24
- 2.2 Quel est l'espace de la représentation interne de lettres ? . . . 24
  - 2.2.1 Observations expérimentales . . . 25
  - 2.2.2 L'équivalence motrice dans les modèles d'écriture . . . 27
  - 2.2.3 BAP : Dans quel espace sont représentées les lettres ? . . . 28
- 2.3 Comment les lettres sont-elles codées ? . . . 29
  - 2.3.1 Revue des modèles de perception . . . 29
  - 2.3.2 Revue des modèles de production . . . 34
  - 2.3.3 BAP : La représentation interne des lettres . . . 38
- 2.4 Simulation interne du mouvement . . . 41
  - 2.4.1 Etudes expérimentales . . . 41
  - 2.4.2 Simulation du mouvement dans la perception . . . 43
- 2.5 Résumé . . . 45

3 Programmation Bayésienne **47**
- 3.1 Formalisme Bayésien . . . 47
  - 3.1.1 Probabilités et représentation des connaissances . . . 47
  - 3.1.2 Notion de Programmation Bayésienne . . . 49
  - 3.1.3 Implémentation . . . 49
- 3.2 Exemple introductif . . . 49
  - 3.2.1 Variables Probabilistes . . . 50
  - 3.2.2 Distribution conjointe et décomposition . . . 50
  - 3.2.3 Question . . . 58
- 3.3 Résumé . . . 61

4 Modèle BAP **63**
4.1 Activation des sous-modèles . . . 64
4.1.1 Cohérence . . . 64
4.1.2 Cas particulier 1 : activation ou inhibition des sous-modèles . . . 64
4.1.3 Cas particulier 2 : court-circuit . . . 65
4.2 Construction du modèle . . . 67
4.2.1 Représentation interne pivot . . . 67
4.2.2 Représentations internes sensorielles et motrices . . . 69
4.2.3 Modélisation de l'effecteur . . . 70
4.2.4 Activation des sous-modèles . . . 71
4.2.5 Simulation motrice . . . 73
4.3 Résumé . . . 77

5 Perception **79**
5.1 Modèle de perception . . . 79
5.1.1 Variables . . . 80
5.1.2 Termes de la décomposition . . . 80
5.1.3 Formes paramétriques . . . 80
5.2 Lecture . . . 80
5.2.1 Question et inférence . . . 81
5.2.2 Reconnaissance mono-scripteur . . . 83
5.2.3 Reconnaissance multi-scripteur . . . 91
5.3 Reconnaissance du scripteur . . . 94
5.3.1 Question et inférence . . . 94
5.3.2 Taux de reconnaissance . . . 94
5.3.3 Lettre inconnue . . . 95
5.4 Conclusion . . . 95
5.4.1 Résumé . . . 95
5.4.2 Suggestion d'un protocole expérimental . . . 96

6 Action **99**
6.1 Contrôle moteur dans la littérature . . . 99
6.1.1 Etudes expérimentales . . . 99
6.1.2 Modèles de génération de trajectoires . . . 100
6.1.3 Génération de trajectoires à partir de la théorie de contrôle optimal . . . 102
6.1.4 Modèles dans l'espace cartésien . . . 102
6.1.5 Modèles dans l'espace articulaire . . . 109
6.1.6 Confrontation des modèles . . . 114
6.1.7 Contrôle optimal dans notre modèle . . . 116
6.2 Définition du modèle . . . 117
6.2.1 Modèle de génération de trajectoires . . . 118
6.2.2 Modèle de l'effecteur . . . 120
6.3 Tâche d'écriture . . . 122
6.3.1 Question et inférence . . . 122
6.3.2 Résultats . . . 124
6.3.3 Analyse d'une trace . . . 124
6.3.4 Variabilité intra-scripteur . . . 127

6.3.5 Variabilité inter-scripteurs . . . 128
6.4 Equivalence motrice . . . 130
6.4.1 Bras à deux articulations . . . 131
6.4.2 Robot holonome . . . 132
6.4.3 Résultats . . . 136
6.5 Résumé . . . 139

7 Etude du lien perception-action 141
7.1 Copie . . . 141
7.1.1 Copie de traces . . . 141
7.1.2 Copie de lettres . . . 143
7.1.3 Comparaison des deux types de copie . . . 147
7.1.4 Résumé . . . 148
7.1.5 Suggestion de protocole expérimental : Etude de l'influence de la représentation interne sur la copie de lettres . . . 149
7.2 Définition du modèle de simulation motrice . . . 150
7.2.1 Variables probabilistes . . . 150
7.2.2 Termes de la décomposition . . . 151
7.3 Lecture avec simulation motrice . . . 152
7.3.1 Question et inférence . . . 152
7.3.2 Analyse d'une trace . . . 153
7.3.3 Taux de reconnaissance . . . 154
7.3.4 Comparaison avec et sans simulation motrice . . . 155
7.3.5 Reconnaissance du scripteur . . . 157
7.3.6 Lettres tronquées . . . 158
7.4 Reproduction d'expérience . . . 162
7.4.1 Analyse et modélisation . . . 162
7.4.2 Résultats . . . 168
7.5 Résumé . . . 172

8 Conclusion 173
8.1 Hypothèses . . . 173
8.2 Modélisation . . . 174
8.3 Résultats . . . 174
8.4 Discussion et perspectives . . . 179
8.4.1 Espace visuel . . . 180
8.4.2 Représentation interne des lettres . . . 181
8.4.3 Rôle de la variable λ . . . 183
8.4.4 Allographe . . . 183
8.4.5 Influence de la base de données . . . 184
8.4.6 Apprentissage du contrôle moteur . . . 184
8.4.7 Sources de variabilité . . . 185
8.4.8 Geste simulé . . . 185
8.4.9 Lecture avec ou sans simulation motrice . . . 187
8.4.10 Dédoublement du modèle de perception . . . 188
8.4.11 Apprentissage . . . 188
8.4.12 Modélisation des mots . . . 189

8.4.13 Application à d'autres domaines 190

Glossaire **191**

Table des figures

1.1 Architecture du modèle 16
1.2 Distribution de probabilité conjointe du modèle BAP 17

2.1 Architecture du modèle BAP 21
2.2 La représentation interne des lettres est le pivot entre la perception et l'action . 25
2.3 Similarités d'écriture avec différents effecteurs 26
2.4 Modèle de l'écriture de Meulenbroek 27
2.5 Equivalence motrice 28
2.6 Ecriture en ligne et hors-ligne 30
2.7 Classification selon la difficulté de reconnaissance 30
2.8 Caractéristiques locales 32
2.9 Caractéristiques globales 32
2.10 Modèle de Markov 33
2.11 Modèle masse-ressort 35
2.12 Trace résultante du modèle masse-ressort 36
2.13 Formes de base 37
2.14 Séquence de points de contrôle 37
2.15 Optimisation de la position des points de contrôle 38
2.16 Points de contrôle 40
2.17 Profils de vitesse en x et y 40
2.18 Etude sur les contraintes dynamiques 42
2.19 Modèle perception-action intégrant la boucle de simulation du mouvement . . 44

3.1 Extrait du modèle global. 50
3.2 Distribution de probabilité de la position en x 53
3.3 Lettres normalisées 54
3.4 Profils de vitesse en x et y 54
3.5 Vitesses filtrées 55
3.6 Distribution de probabilité de la position en x du troisième point de contrôle . 56
3.7 Deux des allographes de la lettre l possibles. 56
3.8 Distributions de probabilité de la position des points aux indices 2 à 5 57
3.9 Ensemble des points de contrôle dont on cherche à connaître la lettre. 60
3.10 Trace dont on a extrait les points de contrôle. 60

4.1 Structure générale du modèle BAP 63
4.2 Illustration d'un court-circuit 66
4.3 Représentation interne des lettres pivot entre perception et action 68

4.4 Les représentations internes sensorielles et motrices sont distinctes. 69
4.5 Structure globale du modèle. 71
4.6 Activation ou inhibition des sous-modèles. 72
4.7 Simulation motrice de l'écriture 74
4.8 Simulation motrice de l'écriture : dédoublement de la boucle 75
4.9 Notation des variables dans le schéma global du modèle 76

5.1 Architecture du modèle de perception 79
5.2 Architecture du modèle impliquée dans la tâche de lecture 81
5.3 Exemple de trace dont on cherche à connaître la lettre 84
5.4 Ensemble des points définissant la trace dont on cherche à reconnaître la lettre. 84
5.5 En bleu, points de contrôle de la trace perçue. 85
5.6 Matrice de confusion. 86
5.7 Erreur de classification 87
5.8 Probabilité en fonction du nombre de points de contrôle 88
5.9 Probabilité selon les positions et vitesses 89
5.10 Distribution de probabilité de la position en y du cinquième point de contrôle . 89
5.11 Allographe de la lettre g le plus rencontré lors de l'apprentissage. 90

6.1 Trajectoire décrite par la main et profil de vitesse 100
6.2 Minimum-jerk 104
6.3 Comparaison entre le minimum-jerk et les traces expérimentales 105
6.4 Courbes prédites par le modèle de minimum-jerk 107
6.5 Mouvement point à point à profil d'accélération bang-bang 108
6.6 Tâches de pointage 109
6.7 Bras robotique à deux articulations 110
6.8 Traces prédites par le modèle de minimum-torque 111
6.9 Trajectoires générées par les sujets 112
6.10 Trajectoires prédites par le modèle de minimum-variance 114
6.11 Comparaison avec les données expérimentales 115
6.12 Comparaison des modèles 116
6.13 Trajectoires générées par le modèle minimum-accélération. 118
6.14 Modèle de minimum-accélération 119
6.15 Modélisation du bras composé de deux articulations (l'épaule et le coude). . . 120
6.16 Accélérations articulaires 125
6.17 Vitesses articulaires 125
6.18 Positions articulaires 126
6.19 Trace produite par le modèle BAP 126
6.20 Des traces produites par le modèle BAP 127
6.21 Variabilité intra-scripteur 128
6.22 Variabilité inter-scripteurs 129
6.23 Variabilité inter-scripteurs de la lettre a 129
6.24 Variabilité inter-scripteurs pour les lettres n, v, e et w 130
6.25 Bras robotique utilisé par le modèle BAP. 131
6.26 Trace écrite par le bras simulé 132
6.27 Présentation du robot holonome 132
6.28 Vitesse réelle de la roue composée de la vitesse contrôlée et non contrôlée. . . . 133

6.29 Vitesse réelle du robot en fonction des vitesses contrôlées et non contrôlées . . . 134
6.30 Repère global du robot et repère centré sur une roue. 134
6.31 Trace « a » réalisée par le bras simulé, robotique et le robot holonome 136
6.32 Traces réalisées par le robot holonome 136
6.33 Reproduction des styles d'écriture 137
6.34 Reproduction des styles d'écriture avec les trois effecteurs 138
6.35 Modélisation des trois effecteurs du modèle BAP 139

7.1 Architecture du modèle BAP pour la copie de traces 143
7.2 Trace à recopier 143
7.3 Trace recopiée par le modèle BAP 144
7.4 Ensemble de traces recopiées par le modèle BAP 144
7.5 Architecture du modèle où on a représenté les deux variables scripteurs. 145
7.6 Architecture du modèle pour la copie de lettres 146
7.7 Lettre à recopier 147
7.8 Lettre recopiée par plusieurs scripteurs 147
7.9 Prototype de la lettre w 148
7.10 Copie de lettres 148
7.11 Architecture du modèle BAP pour la simulation motrice 151
7.12 Trace fournie au modèle pour la reconnaissance avec simulation motrice. 153
7.13 Trace simulée générée par le modèle 153
7.14 Matrice de confusion de perception avec avec simulation motrice 154
7.15 Trace correspondant à la lettre d reconnue par le modèle comme une lettre a. . 155
7.16 Trajectoire générée à partir des points de contrôle de la figure 7.15. 156
7.17 Architecture du modèle comprenant les deux variables scripteurs W_L et W_S. . . 156
7.18 Lettre tronquée. 158
7.19 Simulation motrice de la lettre tronquée. 159
7.20 Trace o tronquée 159
7.21 Trace u tronquée 159
7.22 Lettre découpée en huit secteurs 166
7.23 Lettre découpée en trois secteurs 166
7.24 Résultats de l'expérience : pourcentage de réponses correctes 168
7.25 Résultat sans simulation motrice 169
7.26 Résultats avec simulation motrice 170
7.27 Produit de distribution de probabilité 170
7.28 Variation du paramètre α 171

8.1 Architecture du modèle BAP 175
8.2 Ordre de présentation de la discussion 179
8.3 Modèle sensoriel 180
8.4 Projection dans le colliculus supérieur 181
8.5 Points de contrôle en accélération nulle 182
8.6 Comparaison entre les différents types de points de contrôle 182
8.7 Boucle de perception simulée. 186
8.8 Boucle de perception simulée incluant la simulation du geste. 186

Chapitre 1

Introduction

Si je vous demande d'écrire votre nom, vous allez effectuer cette tâche très facilement. Vous pouvez même la réaliser en regardant ailleurs ou en pensant à autre chose. Mais que se passe-t-il si je vous demande d'écrire votre nom avec votre pied dans le sable ou la neige, par exemple ? La trace que vous produirez va être légèrement déformée par rapport à votre écriture habituelle mais pourtant, vous êtes capable de réaliser cette tâche et ce, sans aucun entraînement. Cet effet, connu sous le nom d'*équivalence motrice* (Bernstein, 1967), soulève une première interrogation :

Quels sont les mécanismes qui permettent de réaliser un même mouvement avec différents effecteurs ?

Lorsque des ballerines observent d'autres ballerines en train de danser, leur cerveau semble simuler les actions comme si c'était elles qui les effectuaient. Ce qui n'est probablement pas le cas chez vous, sauf si vous êtes expert en danse classique, bien sûr. Il se pourrait que, afin de mieux appréhender les actions des autres, notre cerveau mette en œuvre une *simulation* des gestes perçus (Calvo-Merino *et al.*, 2004). Pour qu'il puisse réaliser ces simulations, il faut évidemment avoir une bonne connaissance des actions effectuées. Lorsque nous regardons un ballet, nous ne sommes donc pas tous égaux, notre perception dépend de notre capacité à reproduire les gestes réalisés. Cette constatation soulève une seconde interrogation :

Comment notre système moteur peut-il influencer notre perception ?

Si nous partons de l'hypothèse que nous simulons l'action lorsque nous observons, il semble normal de prédire que nous sommes plus performants lorsque nous observons nos propres actions plutôt que celles des autres. Si on vous présente un point se déplaçant le long d'une trace que vous avez vous même écrite, vous serez capable de prédire la suite de la trajectoire : le mouvement continue-t-il ou s'arrête-t-il ? A l'inverse, si on vous présente un point se déplaçant le long d'une trajectoire que vous n'avez pas écrite, vous ne serez plus capable de faire cette prédiction. Et pourtant, cet effet persiste même lorsque l'on s'assure que les trajectoires sont identiques du point de vue géométrique. Cette expérience réalisée par Knoblich *et al.* (2002) renforce l'idée que l'action est simulée pendant la perception. Plus l'action observée ressemble à celle que vous effectuez vous-même, et plus la simulation sera précise.

1.1 Contribution

Afin d'apporter des éléments de réponse aux deux questions précédentes, nous avons modélisé un système perception-action : le modèle BAP (Bayesian Action Perception). Plus précisément, nous nous plaçons dans le cadre de la lecture et de l'écriture. Ce système a pour but de traiter les mécanismes qui permettent de réaliser un même mouvement avec différents effecteurs et d'étudier l'influence de l'action lors de la perception.

Nous avons tout d'abord défini sa structure globale en tenant compte des observations et théories issues de la littérature. Cela nous a permis de définir un ensemble d'hypothèses fondatrices du modèle. Comme nous pouvons le voir figure 1.1, le modèle BAP est composé de trois sous-modèles : le modèle de représentation interne des lettres, le modèle de perception et le modèle d'action. Les principales hypothèses du modèle BAP liées à l'influence du système

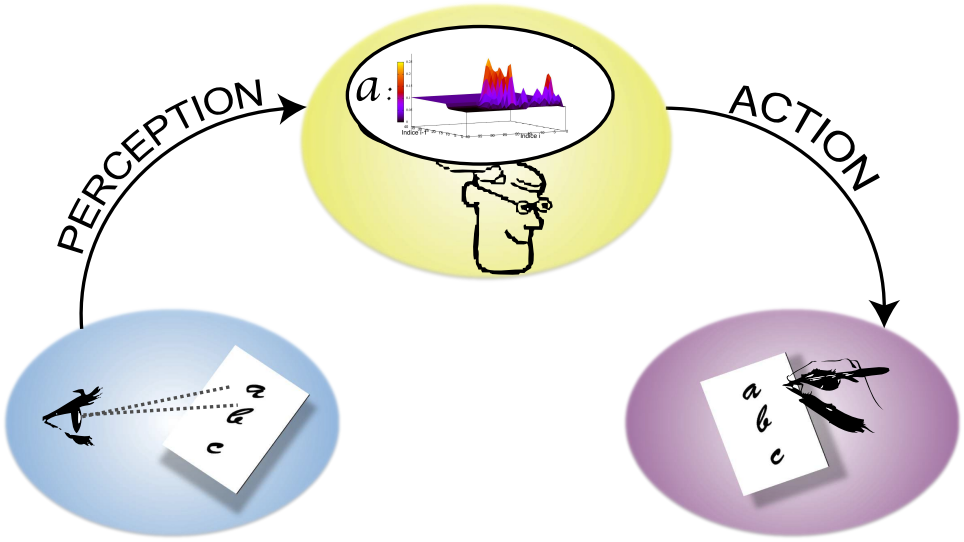


FIG. 1.1 – Architecture du modèle : la représentation des lettres (en jaune) est le pivot entre la perception et l'action. Le modèle de perception est représenté en bleu et le modèle de production en violet.

moteur lors de la perception sont les suivantes :

- la représentation interne des lettres joue le rôle de pivot entre la perception et l'action,
- l'action peut être simulée lors de la perception.

Nous avons utilisé le formalisme bayésien pour décrire mathématiquement le modèle BAP. La modélisation bayésienne s'appuie sur les probabilités pour représenter des connaissances. Dans cette thèse, nous utilisons le formalisme de la Programmation Bayésienne qui donne un cadre pour leur organisation. La formulation d'un modèle sous forme bayésienne se fait par l'écriture d'une distribution de probabilité conjointe. La figure 1.2 exprime la relation entre les variables du modèle BAP sous forme mathématique (à gauche) et sous forme graphique (à droite).

Connaissant cette probabilité conjointe, le modèle est capable de résoudre plusieurs tâches cognitives. Chacune de ces tâches correspond à une question probabiliste qui est résolue au-

$$P\left(\begin{array}{cccccccc} C_{LV} & C_{LP} & C_{LS} & C_V & \lambda_p & \lambda_v & \lambda_s & \lambda_l \\ E & C_P & C_S & V & P & L & W & S \end{array}\right)$$

$$= \left(\begin{array}{l} P(C_{LV} \mid L\ W)\, P(C_{LP} \mid L\ W) \\ P(L)\, P(W)\, P(C_{LS} \mid L\ W) \\ P(\lambda_V \mid C_{LV}\ C_V)\, P(\lambda_P \mid C_{LP}\ C_P) \\ P(C_V \mid V)\, P(P \mid C_P)\, P(C_S \mid S) \\ P(\lambda_l \mid C_{LV}\ C_{LP})\, P(S \mid P)\, P(P) \\ P(E \mid P)\, P(\lambda_S \mid C_{LS}\ C_S) \end{array}\right)$$

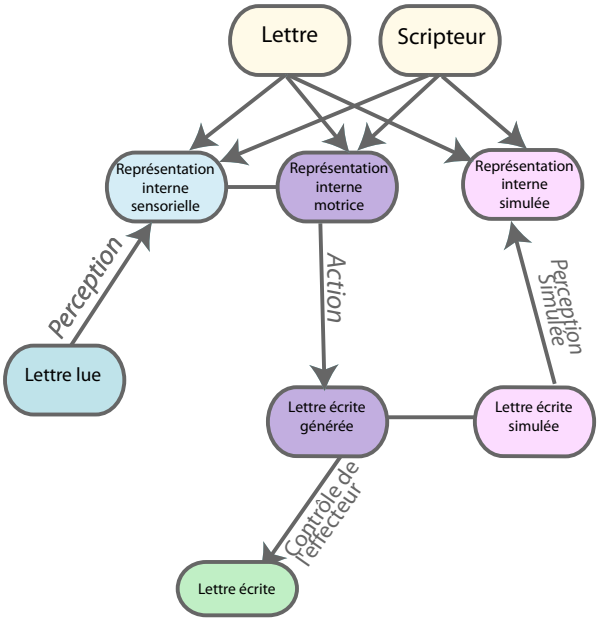


FIG. 1.2 – A gauche, distribution de probabilité conjointe du modèle BAP et à droite, son architecture.

tomatiquement par inférence bayésienne.

Dans la suite de ce chapitre nous allons décrire chacune des sous-parties du modèle BAP. Nous présenterons aussi les études que nous avons réalisées liées à la simulation du mouvement et une des utilisations possible du modèle BAP : la suggestion de protocoles expérimentaux.

1.1.1 Description des sous-modèles

Comme nous l'avons vu précédemment, notre modèle est constitué de 3 sous-modèles : le modèle de représentation interne des lettres, le modèle de perception et le modèle d'action. Dans la suite de ce chapitre, nous allons présenter la fonction de chacun des ces trois sous-modèles.

1.1.1.1 Modèle de représentation interne des lettres

La représentation interne des lettres joue le rôle de pivot entre la perception et la production. Elle intervient donc logiquement dans ces deux processus. La représentation choisie doit alors être aisément utilisable dans les mécanismes impliqués aussi bien dans la perception que l'action. De plus, afin d'assurer l'équivalence motrice, nous avons fait l'hypothèse que la représentation interne des lettres est indépendante de l'effecteur : elle est donc dans l'espace cartésien, et non dans un espace articulaire.

Dans le modèle BAP, nous avons supposé que chaque lettre est représentée par une séquence de points particuliers de la trajectoire : *les points de contrôle*. Ces points sont définis aux points de tangence remarquable (tangente horizontale ou verticale).

1.1.1.2 Modèle de perception

La trace perçue, entrée du modèle de perception, est une séquence de points représentant un caractère de type cursif. Le modèle de perception correspond à un système d'extraction des points de contrôle à partir d'une trace.

Une fois la représentation interne et le modèle de perception définis, le modèle BAP est utilisé pour résoudre différentes tâches de perception : lecture (reconnaissance de caractères cursifs isolés), reconnaissance du scripteur, détection de caractères inconnus.

1.1.1.3 Modèle d'action

Le modèle d'action est composé de deux sous-modèles : le modèle de génération de trajectoire et le modèle de l'effecteur. Le modèle de génération de trajectoire détermine une trajectoire à partir de la représentation interne des lettres (séquence de points de contrôle). Il est basé sur la théorie du contrôle optimal et plus précisément sur la minimisation des accélérations. La trajectoire est générée dans le même espace que celui de la représentation des lettres : l'espace cartésien. Le modèle de l'effecteur est composé d'un modèle cinématique (conversion de la trajectoire générée dans l'espace de l'effecteur) et d'un modèle de contrôle de l'effecteur (génération des commandes motrices de l'effecteur). Dans BAP, nous avons modélisé trois effecteurs : un bras à deux articulations simulé, un bras robotique à deux articulations et un robot holonome.

Une fois la représentation interne et le modèle d'action définis, le modèle BAP est utilisé pour résoudre la tâche d'écriture.

1.1.2 Simulation interne du mouvement

Nous avons également étudié l'influence des connaissances motrices sur la perception. Nous avons fait l'hypothèse que simuler le geste d'écriture aide à la reconnaissance des lettres. Dans le modèle BAP, lors de la reconnaissance de lettres, le processus d'action intervient pour simuler la trajectoire et aider le processus de perception.

Nous avons tout d'abord testé cette hypothèse sous plusieurs conditions de tâche de lecture : lecture mono-scripteur, lecture multi-scripteur, lecture de traces incomplètes, etc. Nous l'avons également testé lors de la reproduction d'une expérience de psychologie cognitive basée sur la prédiction des lettres à venir (Kandel *et al.*, 2000a).

1.1.3 Suggestion de protocoles expérimentaux

Le modèle BAP est capable de résoudre plusieurs tâches cognitives : la lecture avec ou sans simulation interne du mouvement, l'écriture avec différents effecteurs, la prédiction de lettres, la copie de lettres, etc. Nous pouvons donc simuler le modèle sous plusieurs conditions, ce qui nous a permis de réaliser des prédictions expérimentales. Cependant, les données expérimentales correspondantes n'étant pas existantes, nous avons dû nous contenter de suggérer de nouveaux protocoles expérimentaux.

Il est important de noter que nous considérons ces suggestions comme des amorces de protocoles. En effet, la rédaction d'un protocole expérimental est très délicate. Il faut que sa description soit suffisamment claire pour que l'expérience puisse être reproduite à l'identique et il doit faire l'objet d'une analyse critique pour notamment détecter d'éventuels biais. Or,

cela ne fait pas partie de notre domaine de compétences. C'est pourquoi nous ne fournissons uniquement que des pistes de protocoles.

1.2 Plan de lecture

Chapitre 2 Le chapitre 2 présente l'architecture du modèle du modèle BAP. Nous y étudions quatre questions, à la lumière de la littérature pertinente, qui nous amènent à énoncer nos hypothèses fondamentales :

- Y a-t-il une représentation unique des lettres pour la perception et l'action ou s'agit-il de représentations distinctes ?
- Quel est l'espace dans lequel les représentations internes sont définies ?
- Comment sont définies les représentations internes des lettres ?
- Une simulation de l'action influence t-elle la perception ?

Nous analysons tout d'abord la relation entre la perception et l'action dans le cadre de la perception de mouvements biologiques. Nous exposerons les diverses théories présentes dans la littérature avant d'analyser le lien entre la perception et l'action dans le cadre de la perception de traces manuscrites.

Afin de définir l'espace dans lequel est décrite la représentation interne des lettres, nous avons traité les observations issues de l'équivalence motrice : nous avons effectué une brève revue des observations et défini des hypothèses. Ce chapitre décrit donc les bases du modèle BAP en définissant sa structure globale.

Chapitre 3 Le chapitre 3 introduit le formalisme utilisé, la Programmation Bayésienne, en s'appuyant sur un exemple, extrait du modèle BAP. Une fois le modèle décrit mathématiquement, nous l'utilisons pour illustrer la résolution d'une tâche cognitive par inférence bayésienne. Cet exemple décrit la relation entre la représentation interne des lettres, les lettres et les scripteurs.

Chapitre 4 Le chapitre 4 décrit l'intégralité du modèle BAP dans le formalisme bayésien. Nous revenons sur chacune des hypothèses réalisées dans le chapitre 2 pour donner son équivalence mathématique. Cela nous permet donc, à la fin du chapitre, d'obtenir une description complète du modèle BAP ; les chapitres suivants seront consacrés à son utilisation. Dans le chapitre 4, nous décrivons également un outil : les variables de cohérence. Cet outil, exprimé également dans le formalisme bayésien, nous permet d'activer ou de désactiver les sous-modèles. Il sera notamment utile pour comparer la tâche de lecture avec la tâche de lecture impliquant la simulation de l'action.

Chapitre 5 Dans le chapitre 5, nous présentons le modèle de perception avec les résultats concernant la tâche de lecture en perception pure (la simulation de l'action n'est pas activée). Cela nous permet de déterminer les principales confusions entre lettres réalisées par le modèle BAP. Nous simulons le modèle dans plusieurs conditions : reconnaissance mono-scripteur (un seul scripteur connu), multi-scripteurs (plusieurs scripteurs sont identifiés) et omni-scripteurs (les scripteurs ne sont pas tous connus de la base de données). Nous verrons également dans ce chapitre que le modèle BAP est capable de reconnaître le scripteur à l'origine de la trace présentée.

Chapitre 6 Le chapitre 6 définit le système de génération de trajectoires. Nous présentons tout d'abord une revue des modèles existants. Puis, nous modélisons celui que nous employons au sein du modèle BAP. Ce chapitre décrit également la modélisation des effecteurs. Dans le modèle BAP, nous avons modélisé un bras simulé, un bras robotique à deux articulations et un robot à trois roues. L'utilisation de ces trois effecteurs nous permet de valider les hypothèses liées à l'équivalence motrice. Enfin, ce chapitre expose les résultats liés à la tâche d'écriture. Nous étudierons notamment les variabilités intra-scripteur et inter-scripteurs.

Chapitre 7 Dans le chapitre 7, nous définissons le modèle de simulation interne des mouvements ainsi que les différents résultats liés à la simulation motrice. Etant donné que nous pouvons utiliser le modèle pour simuler des tâches cognitives dans plusieurs conditions, nous pouvons comparer les résultats de la lecture avec et sans simulation motrice. De plus, nous abordons la notion de copie et distinguons deux tâches cognitives : la copie de traces et la copie de lettres. Nous reproduisons également une expérience de psychologie cognitive basée sur la prédiction des lettres à venir. Nous verrons qu'avec le modèle BAP nous pouvons simuler l'expérience et comparer les résultats à ceux des sujets.

Chapitre 8 Enfin, le dernier chapitre présente les pistes de recherche et discute nos résultats.

1.3 Cadre de travail

Dans la suite de ce document, lorsque nous parlerons de lecture, il s'agira de *reconnaissance de caractères cursifs isolés*. Nous sommes conscients qu'il ne s'agit que d'une petite partie de la tâche cognitive « lecture » (nous ne traitons pas les lettres en script et les liaisons entre les caractères par exemple). Cependant, afin d'alléger le texte dans la suite de ce document, nous nous autorisons à utiliser cet abus de langage. Il en est de même pour le terme « écriture » que nous utiliserons à la place de *génération de trajectoires correspondant à des caractères cursifs isolés*. En résumé, dans le reste de ce document, à chaque fois, que le mot lecture ou écriture sera employé, il faudrait le remplacer par :

- **lecture :** reconnaissance de caractères cursifs isolés,
- **écriture :** génération de trajectoires correspondant à des caractères cursifs isolés.

Nous n'abordons pas dans cette thèse l'étude de l'apprentissage de la lecture et de l'écriture. Nous ne traitons donc pas dans ce document de la phase d'apprentissage de la lecture et de l'écriture (nous supposons la représentation interne des lettres et le contrôle moteur comme acquis). De plus, cette thèse n'a pas pour but de développer un système de reconnaissance d'écriture industriel. Nous ne cherchons donc en aucun cas à obtenir des taux de reconnaissance à 100%. Notre but est d'obtenir un modèle qui puisse offrir un cadre explicatif, et d'analyse et de prédictions de la performance humaine.

Chapitre 2

Architecture du modèle BAP

Dans ce chapitre, nous étudions les hypothèses fondamentales du modèle BAP par le biais d'une revue sur les boucles perception-action. Plus précisément, nous nous intéressons à la représentation interne des lettres : nous déterminons l'espace de codage de cette représentation ainsi que les codes des lettres et ce, à partir des travaux sur les modèles de lecture et d'écriture. Enfin, nous abordons la notion de boucle perception-action à un plus bas niveau : l'influence de l'action lors de la perception.

Afin de préciser notre cadre de travail, nous avons défini une première ossature du modèle présentée figure 2.1.

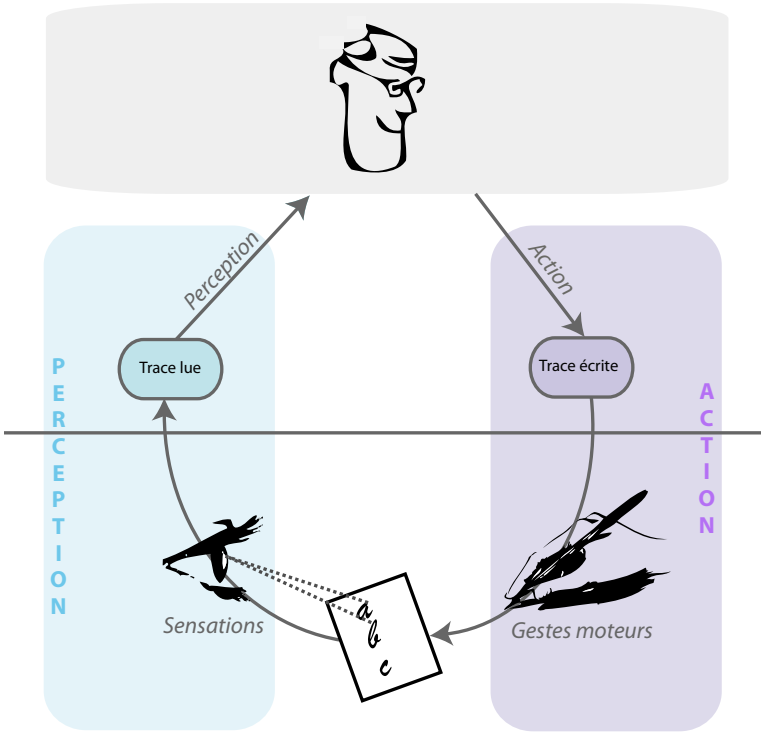


FIG. 2.1 – Architecture du modèle BAP composée d'un modèle de perception et d'un modèle d'action.

Sur la figure 2.1, nous faisons la distinction entre le *modèle de perception* et le *modèle sensoriel*. Le modèle sensoriel correspond au modèle qui projette l'image de la feuille dans l'espace visuel. Il s'agit du « modèle de l'œil ». Afin de simplifier cette projection nous faisons l'hypothèse, dans le cadre de ce travail, que l'espace visuel correspond à l'espace cartésien. Dans le modèle BAP, le modèle sensoriel correspond donc au modèle identité. Dans la suite de ce travail, afin d'alléger les futurs schémas, nous n'indiquerons plus le modèle sensoriel. Le modèle de perception, quant à lui, extrait la représentation interne des lettres à partir de la trajectoire fournie par le modèle sensoriel.

Nous effectuons également une distinction entre le *modèle d'action* et le *modèle moteur* : le modèle d'action génère une trajectoire à partir de la représentation interne des lettres et le modèle moteur contrôle l'effecteur. Cependant, dans le modèle moteur, nous ne modélisons pas le stylo (son poids, les frottements, etc.). Nous fournissons uniquement les commandes motrices appliquées à l'effecteur qui sont fonctions de la trajectoire à effectuer. Le contrôle de l'effecteur est réalisé en boucle ouverte : les informations sensorielles ne sont pas utilisées pour corriger d'éventuelles dérives sur la trajectoire. Nous ne traitons pas non plus, dans ce manuscrit, de la notion de copie d'efférence.

Ce chapitre est axé autour de quatre questions :

- **Y a-t-il une représentation unique des lettres pour la perception et l'action ou s'agit-il de représentations distinctes ?**
- **Quel est l'espace dans lequel les représentations internes sont définies ?**
- **Comment sont définies les représentations internes des lettres ?**
- **Une simulation de l'action est-elle nécessaire pour la perception ?**

Pour chaque question, nous présentons une revue des théories puis nous définissons les hypothèses effectuées pour le modèle BAP.

2.1 Des représentations distinctes pour la perception et la production des lettres ou une représentation commune ?

2.1.1 Théories de la perception de la parole

Les principales théories articulant perception et action ont émergé du domaine de la parole. La perception de la parole est l'ensemble des mécanismes permettant aux systèmes sensoriels (visuels, acoustiques, etc.) et au système nerveux central de décoder le message émis par le locuteur. Plusieurs théories s'opposent dans ce domaine :

- les théories auditives,
- la théorie motrice de la perception de la parole,
- les théories de l'interaction locuteur-auditeur,
- la théorie de la perception pour le contrôle de l'action.

La caractéristique majeure des théories auditives est que la perception y est étudiée indépendamment de l'articulation (c'est-à-dire de l'action). Les théories de l'interaction locuteur-auditeur font l'hypothèse que les propriétés du langage émergent des conditions de communication (Lindblom, 1990). Dans la suite de cette section, nous détaillons uniquement la théorie motrice de la perception et la théorie de la perception pour le contrôle de l'action car elles définissent notre cadre de travail.

2.1.1.1 Théorie motrice de la perception

Liberman (1957) a développé la théorie motrice de la perception en partant du constat que des signaux acoustiques similaires avec des mouvements articulatoires sous-jacents différents sont perçus comme des signaux différents. Il a développé l'idée que la perception de la parole correspond à identifier des objets moteurs. Ces objets correspondent aux invariants phonétiques des commandes motrices des muscles. Cependant, étant donné que les gestes implémentés dans le système vocal sont déformés par la co-articulation[1], il est nécessaire de distinguer les intentions de gestes du locuteur des mouvements réels qu'il effectue (Liberman et Mattingly, 1985). Liberman identifie donc les intentions de gestes, et non les actions réelles, comme les objets fondamentaux de la perception de la parole.

De plus, il s'est appuyé sur la notion de *module* (Fodor, 1983), c'est-à-dire de systèmes qui réalisent des tâches spécifiques : par exemple, la localisation de son pour l'audition ou la perception de profondeur ou de couleur pour la vision. Le mécanisme de perception de la parole est donc vu comme un module : *le module phonétique* (il est à distinguer du module du système auditif). La théorie motrice de la perception propose de regrouper la perception et la production de la parole dans un même module. Cela permet d'expliquer entre autre comment la représentation acquise par la perception est disponible pour la production. Dans le cas contraire (si l'on avait considéré des modules différents), il faudrait alors expliquer comment la même structure évolue conjointement dans ces deux modules.

Liberman a résumé la théorie motrice de la parole avec ces deux affirmations :

(a) Percevoir la parole c'est percevoir les intentions des gestes vocaux réalisés par le locuteur.

(b) Le système moteur de la parole joue un rôle dans la perception.

La théorie motrice de la perception a été étudiée dans d'autres domaines que la parole telle que les mouvements du bras. Rizzolatti et Arbib (1998) ont trouvé des neurones dans le cortex prémoteur du singe (aire F5) qui répondent à la fois lorsque le singe effectue une action de prise d'un objet et lorsqu'il perçoit un autre singe ou humain qui effectue cette même action.

2.1.1.2 Théorie de la Perception pour le Contrôle de l'Action

Dans la théorie motrice de la perception, percevoir la parole c'est percevoir les intentions des gestes du locuteur. La théorie de la perception pour le contrôle de l'action (Schwartz, 2001; Schwartz *et al.*, 2002) est légèrement différente : l'ensemble des processus perceptifs (auditifs, visuels, etc.) fournissent un ensemble de représentations qui permettent à la fois de contrôler ses propres actions et de suivre les actions des autres. Dans ce cadre théorique les représentations perceptives et motrices ne sont ni indépendantes (comme elles le sont souvent dans les théories dites « auditives » de la parole), ni confondues (comme dans les théories « motrices »), mais bien interdépendantes et, en réalité, « co-construites » au cours de l'apprentissage. Cette théorie a été modélisée dans le cadre de l'apprentissage de la parole sur un *bébé androïde* (Serkhane, 2005). La croissance du conduit vocal n'est pas uniforme et les

[1] co-articulation : selon les phonèmes qui l'entourent dans une phrase, un phonème n'est pas prononcé de la même manière. Ceci est dû au fait que le conduit vocal, comme tout système physique, a une certaine inertie qui l'empêche de « passer » instantanément d'un phonème à un autre. La modification de la configuration du conduit vocal pour passer d'un phonème à un autre se fait de façon progressive et les deux phonèmes subissent une distorsion. Ce phénomène n'est pas audible, ni le locuteur ni l'auditeur ne s'en rendent compte, néanmoins la coarticulation est primordiale à l'intelligibilité de la phrase.

capacités motrices de l'enfant évoluent au cours du temps. La question est de savoir comment la relation perception-production intervient dans le développement.

2.1.2 BAP : La représentation interne des lettres pour la lecture et l'écriture

> Nous l'avons vu, dans le cadre de la parole, plusieurs théories discutent de l'unicité des représentations sensorielles et motrices. La théorie motrice de la perception considère une unique représentation motrice alors que la théorie pour le contrôle de l'action considère deux représentations distinctes. Cependant, ces deux représentations évoluent conjointement.
> Dans le cadre de l'écriture, il est possible de se poser la même question : *La représentation interne sensorielle est-elle distincte de la représentation interne motrice ?*
> La parole et l'écriture partagent plus que le simple fait que ce sont deux formes de communications que l'on peut générer ou percevoir. Ce sont deux systèmes sensori-moteur qui dépendent du locuteur/scripteur, et qui ont une grande variabilité. Il ne serait donc pas surprenant que les systèmes de reconnaissance et de production de la parole et les systèmes de reconnaissance et de production de l'écriture aient des points communs (Huckvale, 1992).
> Dans cette thèse, nous nous plaçons dans le cadre de la théorie de la perception pour le contrôle de l'action. Dans le modèle BAP, nous faisons l'hypothèse que l'espace interne dans lequel est codé la perception est le même que celui dans lequel est codé la production. Nous avons donc des représentations de même nature qui sont le pivot entre la perception et la production des lettres. Les représentations sensorielles et motrices ne sont pas confondues, elles sont co-construites au cours de l'apprentissage. La figure 2.2 présente le schéma de notre modèle : la représentation interne des lettres joue le rôle de pivot entre la perception et la production.
> Ces deux représentations sont interdépendantes et basées sur le même espace : c'est pourquoi nous les avons liées dans notre modèle.
> De plus, nous faisons l'hypothèse que le scripteur et la lettre sont indépendants : nous les avons donc modélisés de façon distincte dans la figure 2.2. A chaque lettre et à chaque scripteur est donc associée une représentation interne.

2.2 Quel est l'espace de la représentation interne de lettres ?

Si on vous demande d'écrire votre nom sans contraintes spécifiques, vous considérez cela comme une tâche très facile. Vous pouvez même réaliser cette tâche dans de nombreuses circonstances, comme en pensant à autre chose ou en regardant ailleurs. Mais qu'en est-il d'écrire votre nom avec votre pied dans le sable ou la neige par exemple ? La trace réalisée sera peut-être un peu déformée par rapport à votre écriture classique mais même sans aucun apprentissage d'écriture au pied, votre nom sera lisible.

Cette effet est connu sous le nom d'*équivalence motrice* (Bernstein, 1967). Dans cette section, nous allons tout d'abord étudier les observations expérimentales résultantes de cet effet (dans le domaine de l'écriture mais aussi dans d'autres domaines), puis nous analyserons un modèle particulier de l'écriture permettant d'écrire avec différents effecteurs. Enfin, nous verrons comment nous intégrons la notion d'équivalence motrice au sein du modèle BAP.

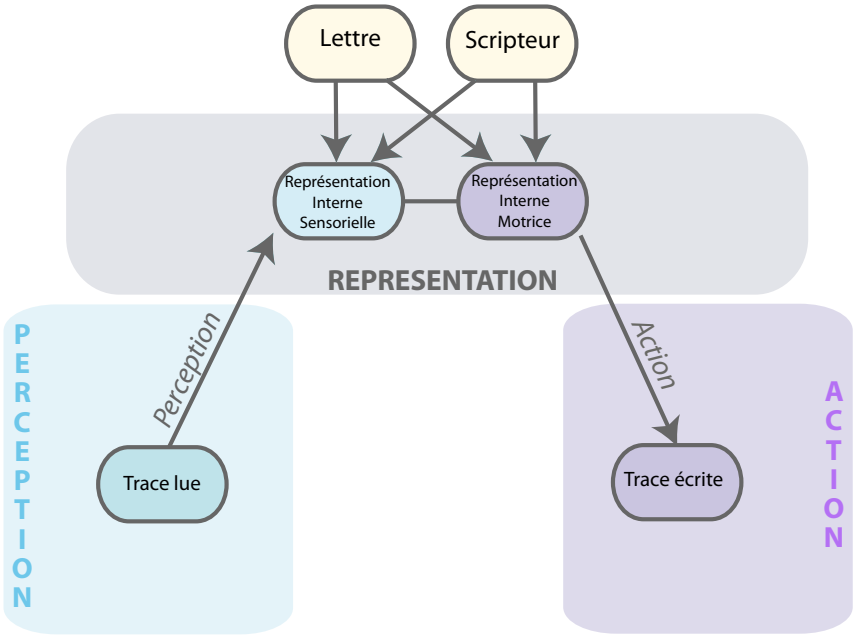


FIG. 2.2 – La représentation interne des lettres est le pivot entre la perception et la production des lettres. Le modèle de *perception* détermine la *Représentation interne sensorielle* à partir de la *Trace lue*. A l'inverse, le modèle d'*action* détermine la *Trace écrite* à partir de la *Représentation interne motrice*.

2.2.1 Observations expérimentales

Hors du domaine de l'écriture, il existe plusieurs exemples d'équivalence motrice.

2.2.1.1 Domaine de la parole

Dans le cadre de la parole, il est toujours possible de parler si l'on bloque une partie de l'effecteur. Par exemple, lors d'une expérience réalisée par Savariaux *et al.* (1995), un tube labial de 20 mm de diamètre imposant ainsi une grande aire labiale a été insérée entre les lèvres de 11 locuteurs. Les auteurs leur ont ensuite demandé de produire dans ces conditions la voyelle [u], usuellement produite avec une faible ouverture labiale. Ils ont alors recueilli des données articulatoires et acoustiques (signal sonore produit). Les stratégies développées pour faire face à la perturbation ont été ainsi analysées. Les résultats ont confirmé que la compensation était possible grâce à une complète réorganisation articulatoire. Ils ont aussi pu mettre en évidence que le comportement des locuteurs est fortement basé sur leur volonté d'améliorer leurs résultats acoustiques : la cible auditive est donc clairement présente dans la tâche du locuteur.

2.2.1.2 Tâches de pointage

Un autre exemple d'équivalence motrice concerne les tâches de pointage. Plusieurs études ont été réalisées (Bortolami *et al.*, 2008) et en particulier celle d'Adamovich *et al.* (2001), dans laquelle il s'agit de gestes de pointage de la main vers l'avant. Lors d'une expérience, les

auteurs ont montré que les sujets sont plus rapides dans les tâches de pointage combinant un mouvement du bras et du tronc que dans les tâches de pointage où le tronc est bloqué. Ils suggèrent donc que dans le second cas, il est nécessaire de compenser le blocage du tronc pour pouvoir effectuer le mouvement.

2.2.1.3 Domaine de l'écriture

Dans le cadre de l'écriture, selon Serratrice et Habib (1993), le transfert entre les effecteurs est immédiat et concerne tous les segments corporels (bouche, main non dominante, pied). C'est-à-dire qu'il est possible d'écrire avec n'importe quelle partie de son corps de façon instantanée. La figure 2.3 présente les traces réalisées par un sujet avec sa main dominante, sa main non dominante, avec un stylo bloqué entre les dents et avec son pied. On observe

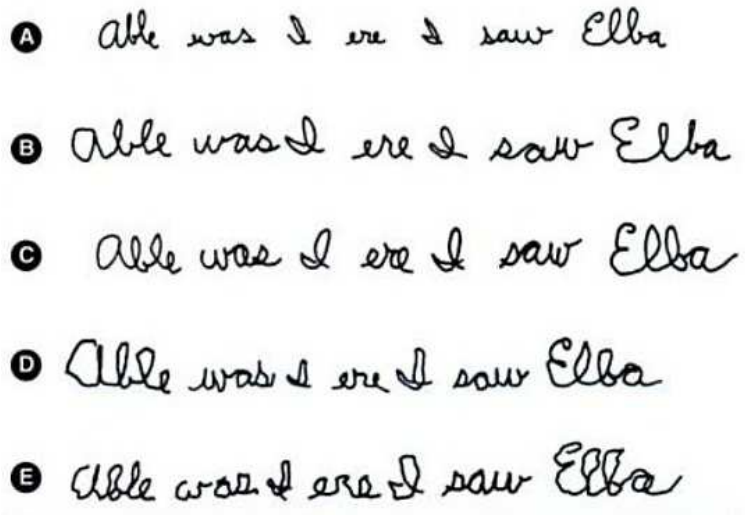


FIG. 2.3 – Similarités d'écriture avec différents effecteurs. Traces réalisées par un sujet avec sa main dominante (ligne A), avec le poignet bloqué (ligne B), avec sa main non dominante (ligne C), avec le stylo entre les dents (ligne D), avec le stylo accroché au pied (ligne E) (figure tirée de Raibert (1977)).

que les formes des lettres sont similaires quelque soit l'effecteur utilisé (Wright, 1990). Lors de plusieurs expériences, certaines invariances ont été mises en évidence : la forme globale des lettres ainsi que l'inclinaison des lettres restent inchangées. Cependant, Castiello et Stelmach (1993) ont montré que la similitude n'est pas parfaite : la longueur des segments et les profils de vitesse diffèrent. Mais après un apprentissage ces différences ont tendance à disparaître, ce qui laisse supposer que ces différences sont principalement dues au contrôle de l'effecteur lui-même.

L'équivalence motrice a également été étudiée en imagerie cérébrale et Wing (2000) en a fait une revue : les études ont pu mettre en évidence les différentes zones motrices activées lors de l'écriture. Ils ont pu distinguer les aires correspondantes à chacun des effecteurs mais aussi les aires motrices activées quelque soit l'effecteur utilisé. Ils ont conclu en suggérant que l'écriture est encodée dans un espace indépendant de toute référence motrice.

2.2.2 L'équivalence motrice dans les modèles d'écriture

L'*équivalence motrice* suggère que la représentation interne des mouvements est indépendante de l'effecteur généralement utilisé pour effectuer le mouvement. Dans le cadre de l'écriture, cela suggère que la représentation interne des lettres est indépendante du bras et de la main utilisée pour écrire (Wing, 2000; Berkenblit et Feldman, 1988; Keele *et al.*, 1995).

Meulenbroek *et al.* (1996) ont utilisé ce constat pour développer un modèle de l'écriture. Dans leur modèle, la représentation interne des lettres est indépendante de l'effecteur et est définie dans un espace cartésien. Cette représentation est ensuite convertie selon l'effecteur utilisé. Les auteurs utilisent un bras avec trois articulations : une au coude, une à l'épaule et une au poignet.

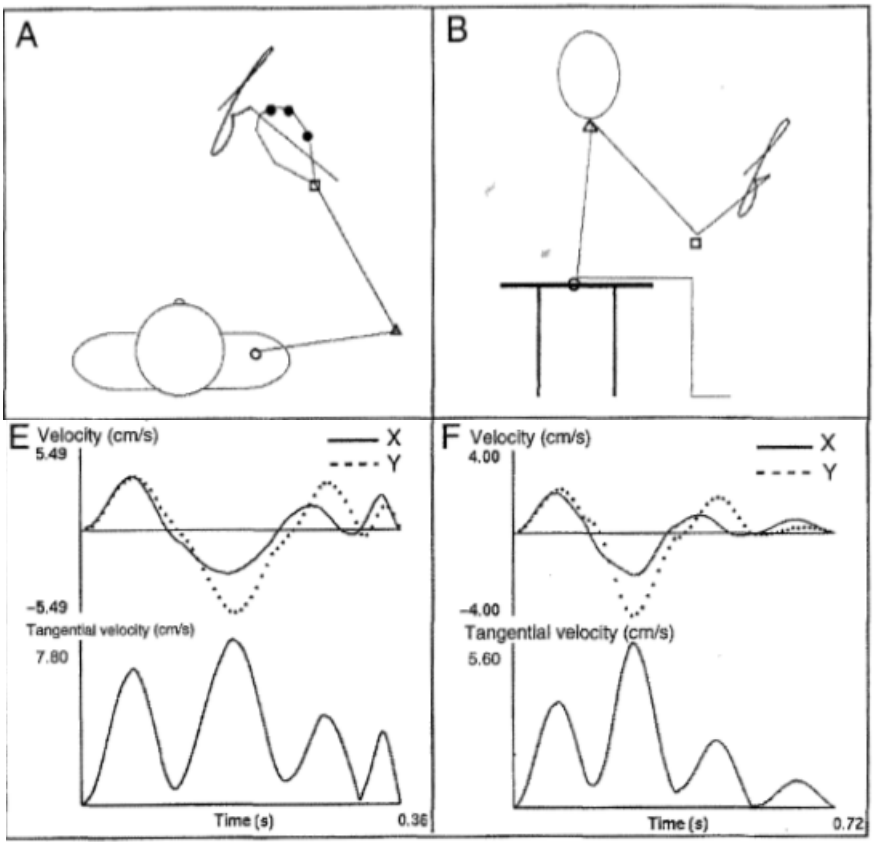


FIG. 2.4 – Résultats du modèle de Meulenbroek (figure tirée de Meulenbroek *et al.* (1996)) A : trajectoire générée dans le plan horizontal par les articulations de l'épaule, du coude et du poignet. B : trajectoire générée dans le plan sagittal par les articulaires de la hanche, de l'épaule et du coude. E et F : profils de vitesse du bout du stylo dans l'espace cartésien générés par les deux effecteurs.

La figure 2.4 présente une lettre « f » générée par deux effecteurs modélisés : la lettre de gauche a été générée dans le plan horizontal par les articulations de l'épaule, du coude et du poignet alors que la lettre de droite a été générée dans le plan sagittal, par les articulations de la hanche, de l'épaule et du coude. Nous observons que quelque soit l'effecteur utilisé la trace

est lisible et que les profils de vitesse dans l'espace cartésien sont similaires (E et F figure 2.4).

Ce modèle utilise une représentation des lettres indépendante de l'effecteur : des points dans l'espace cartésien. Cette étude montre donc que quelque soit l'effecteur utilisé, leur modèle peut générer une trace à partir d'une représentation indépendante de l'effecteur.

2.2.3 BAP : Dans quel espace sont représentées les lettres ?

> Afin de respecter l'équivalence motrice dans le modèle BAP, nous faisons une nouvelle hypothèse : la *représentation interne des lettres utilise le même espace que celui de la tâche*, c'est-à-dire dans notre cas un espace cartésien. Nous nous plaçons dans un espace à deux dimensions (x, y) qui correspond au plan de la feuille. Nous ne tenons pas compte de la composante verticale qui correspond à la pression appliquée sur la feuille.
> La figure 2.5 présente le modèle BAP avec la modélisation de l'effecteur et la précision des espaces (en rouge pour espace cartésien et en vert pour espace de l'effecteur). Les représentations interne motrice et perceptive sont définies dans l'espace cartésien ainsi que la trace lue. Le contrôle de l'effecteur se fait quant à lui dans l'espace de l'effecteur.

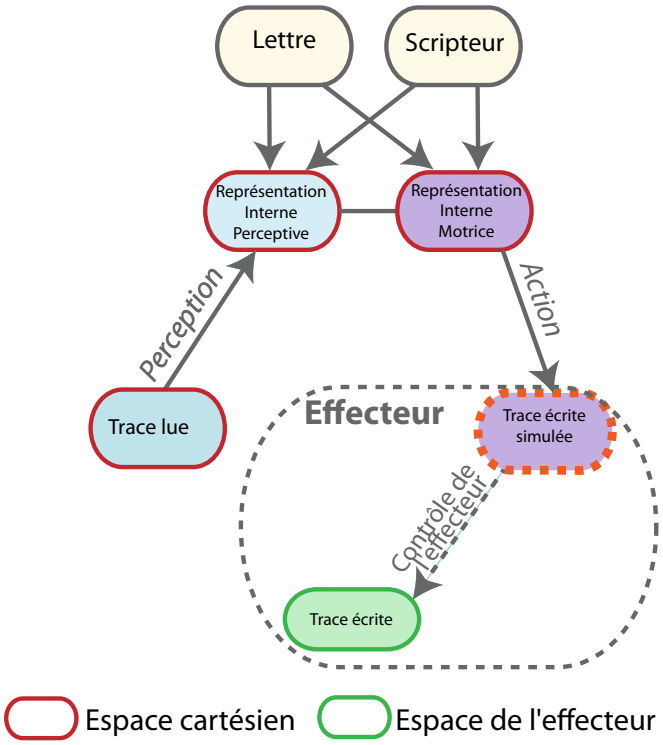


FIG. 2.5 – Equivalence motrice : la représentation interne des lettres est indépendante de l'effecteur.

BAP : Espace de la trajectoire simulée

A partir de la représentation interne motrice, le modèle de production génère une trace que nous nommerons *trace générée*. Cette trace n'est pas réalisée par un effecteur. Elle est seulement planifiée. A ce stade du modèle, nous ne pouvons pas décider de l'espace dans lequel cette trace est générée (en orange sur la figure). C'est pourquoi nous avons représenté en pointillé l'effecteur sur la figure et les liens entre les différentes représentations. La trace générée peut être planifiée dans l'espace cartésien ou dans l'espace de l'effecteur. Le modèle de l'effecteur se divise en deux modules : le *modèle cinématique de l'effecteur* et le modèle de *contrôle de l'effecteur*. Le premier modèle correspond à convertir la représentation interne ou la trajectoire générée dans l'espace de l'effecteur. Le contrôle de l'effecteur détermine les variables de commandes motrices réellement appliquées sur l'effecteur. L'agencement des sous-modèles *modèle cinématique de l'effecteur* et *contrôle de l'effecteur* dépend de l'espace dans lequel la trace générée est planifiée. Par exemple, dans le cas où la trace générée est planifiée dans l'espace cartésien, il faut la convertir dans l'espace de l'effecteur à l'aide d'un modèle cinématique, puis modéliser le contrôle de l'effecteur. Par contre, dans le cas où la trace générée est planifiée dans l'espace de l'effecteur, il faut d'abord convertir la représentation interne motrice dans l'espace de l'effecteur avant d'appliquer le modèle de génération de trajectoire. Une fois la *trace générée* planifiée, il ne reste alors plus qu'à contrôler l'effecteur. Le choix de l'espace dans lequel est planifiée la trajectoire générée sera effectué dans le chapitre 6, à l'aide d'un état de l'art sur les modèles de génération de trajectoires.

2.3 Comment les lettres sont-elles codées ?

Dans le modèle BAP, nous avons fait l'hypothèse que la représentation interne est commune à la perception et à la production des lettres. Nous avons également fait l'hypothèse qu'elle est indépendante de l'effecteur et qu'elle est définie dans l'espace de feuille (espace cartésien). Cependant, nous n'avons pas décrit sa nature exacte. Afin de choisir un représentation adéquate pour notre modèle, nous nous sommes intéressés aux représentations des lettres utilisées dans les modèles de production et de reconnaissance d'écriture présents dans la littérature. Dans cette section, nous allons donc effectuer une revue des modèles de la lecture puis de l'écriture en notant tout particulièrement la représentation des lettres choisie.

2.3.1 Revue des modèles de perception

Depuis les années 90, de nombreux systèmes de reconnaissance d'écriture ont vu le jour. Tout d'abord, nous pouvons distinguer la reconnaissance en ligne des systèmes de reconnaissance hors-ligne (par exemple, reconnaissance optique de caractères (OCR) (Impedovo *et al.*, 1991)). Dans le cas hors-ligne, il s'agit de reconnaître les caractères à partir d'une image discrète constituée d'un ensemble de pixels (voir figure 2.6 à droite). Dans ce cas, il n'y a a priori aucune information temporelle du tracé. En revanche, dans la reconnaissance en ligne, les informations recueillies sont constituées d'une suite ordonnée de points (définis par leurs coordonnées) échantillonnée généralement à cadence fixe (voir figure 2.6 à gauche).

Nous distinguons, de plus, la lecture des caractères d'imprimerie de la lecture de l'écriture cursive. Dans le cas de l'écriture cursive, il faut aussi distinguer les caractères liés des caractères isolés. En effet, la lecture des symboles isolés ne nécessite pas de segmentation alors que

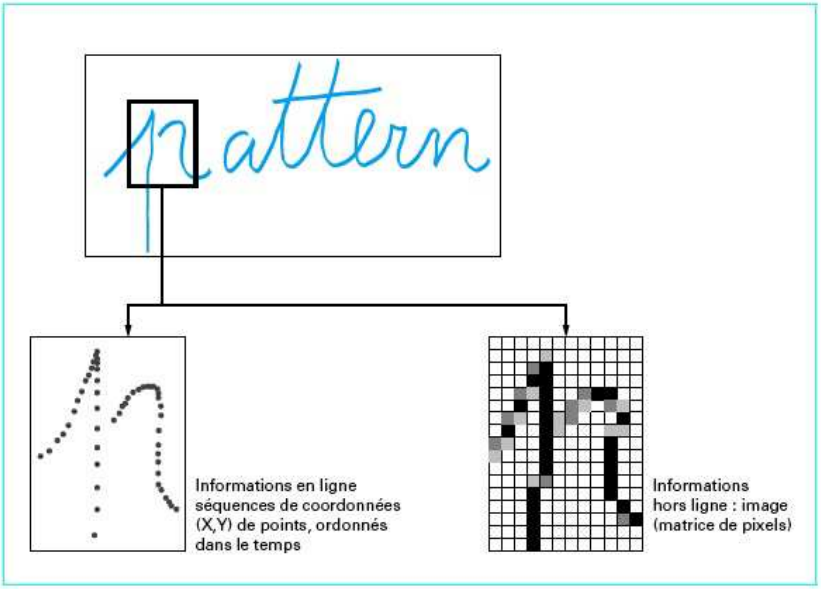


FIG. 2.6 – Ecriture en ligne et hors-ligne (figure extraite de Crettez et Lorette (1998)).

la lecture de caractères liés implique une segmentation en différentes entités. Cette étape constitue une des difficultés majeures de la reconnaissance de l'écriture cursive.

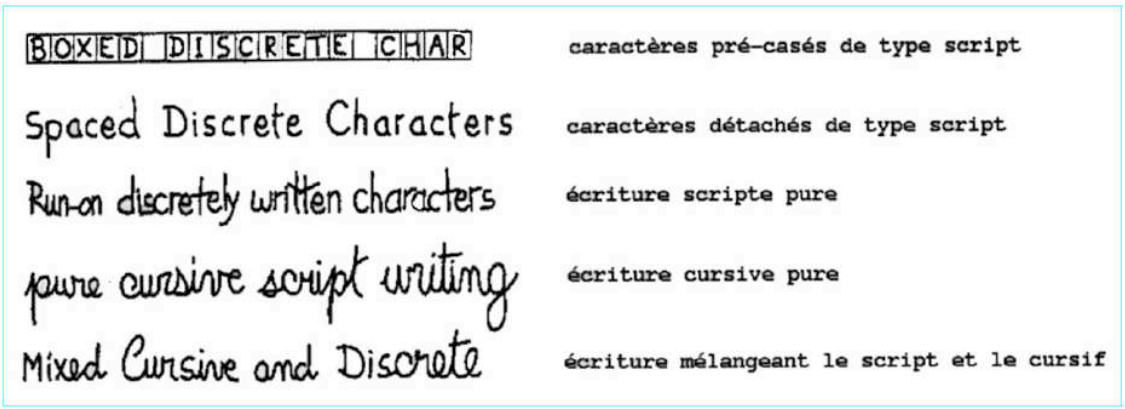


FIG. 2.7 – Classification selon la difficulté de reconnaissance (figure extraite de Tappert *et al.* (1990)).

La figure 2.7 présente une classification par ordre de difficulté croissante de reconnaissance effectuée par Tappert *et al.* (1990) : caractères pré-casés de type script, caractère détachés de type script, écriture script pure, écriture cursive pure et écriture mélangeant le script et le cursif.

BAP : Type de lettres modélisées

Dans ce travail, nous définissons un modèle de reconnaissance **en ligne** de **caractères isolés** de types **cursifs**. En effet, dans le modèle BAP, nous nous intéressons à l'influence des connaissances motrices sur la perception. Nous prenons donc en compte les informations cinématiques. C'est pour cette raison que nous nous intéressons uniquement aux systèmes en ligne de reconnaissance de caractères de types cursifs (l'écriture cursive est plus liée qu'une écriture de type script). Enfin, nous ne nous intéressons pas aux ligatures entre les lettres. C'est pourquoi, nous étudions seulement des caractères isolés.

Dans la suite de cette section, nous allons nous intéresser aux principales méthodes utilisées pour la reconnaissance de caractères isolés (Crettez et Lorette, 1998; Lorette, 1999; Vuori, 2002) : méthodes structurelles, markoviennes, neuro-mimétiques. Les modèles plus récents correspondent en général à une fusion de ces différentes techniques. Les performances de chacun de ces modèles sont testées sur différentes bases de données et donc difficiles à comparer. En globalité, les modèles (indépendants du scripteur et basés sur un alphabet de 26 lettres) ont un taux d'erreur de reconnaissance des caractères entre 5 et 10%. Pour chaque méthode, nous nous intéressons aussi à la représentation des lettres choisie : cette dernière doit être suffisamment invariante pour être indépendante du scripteur mais suffisamment signifiante pour être reconnue (Teulings, 1993). Par exemple, les méthodes de reconnaissance par comparaison globale (*pattern matching*) ne sont pas adaptées à la variabilité de l'écriture et sont donc très peu employées.

2.3.1.1 Méthodes markoviennes

Modèle Les modèles de Markov cachés (HMM, Hidden Markov Models) sont des techniques largement utilisées dans la reconnaissance de l'écriture. A chaque lettre L est associée un modèle de Markov caché Λ qui émet un signal prenant la forme de la trace écrite de la lettre (notée O). La trace écrite O d'une lettre est formée par la séquence des observations o_i émises successivement par chacun des états S_i du modèle (ou par chacune des transitions successives entre états). Il est possible de calculer la probabilité pour qu'une séquence d'observations o_i soit générée par chaque modèle de lettre Λ. La reconnaissance d'une lettre consiste à trouver parmi tous les modèles Λ possibles celui avec la probabilité la plus élevée.

Résultats Ces méthodes de reconnaissance ont été appliqués à la reconnaissance en ligne de caractères manuscrits isolés et elles donnent de bons résultats à condition d'avoir des bases de données d'apprentissage de très grandes tailles (Anquetil et Lorette, 1995). Les taux de reconnaissance obtenus s'échelonnent entre 80 et 95% selon les lettres (avec un apprentissage réalisé sur un ensemble de 1857 lettres et la reconnaissance effectuée sur une autre ensemble de 1032 lettres (tous scripteurs confondus)).

Comparés à d'autres approches de reconnaissance (que nous détaillerons par la suite), les modèles de Markov cachés sont efficaces pour modéliser différentes sources de connaissance (Yacoubi, 1996; Mahjoub et Ellouze, 1998; Anquetil et Lorette, 1995).

Représentation des lettres Les observations o_i correspondent aux primitives des lettres, c'est-à-dire à des portions de tracé. Les primitives modélisent à la fois les aspects statiques et dynamiques de l'écriture. Les primitives locales (pics, boucles, bosses, levés de stylo) et leurs

orientations modélisent l'aspect dynamique de l'écriture : elles représentent la trajectoire de la pointe du stylo. Ces primitives locales sont présentées figure 2.8.

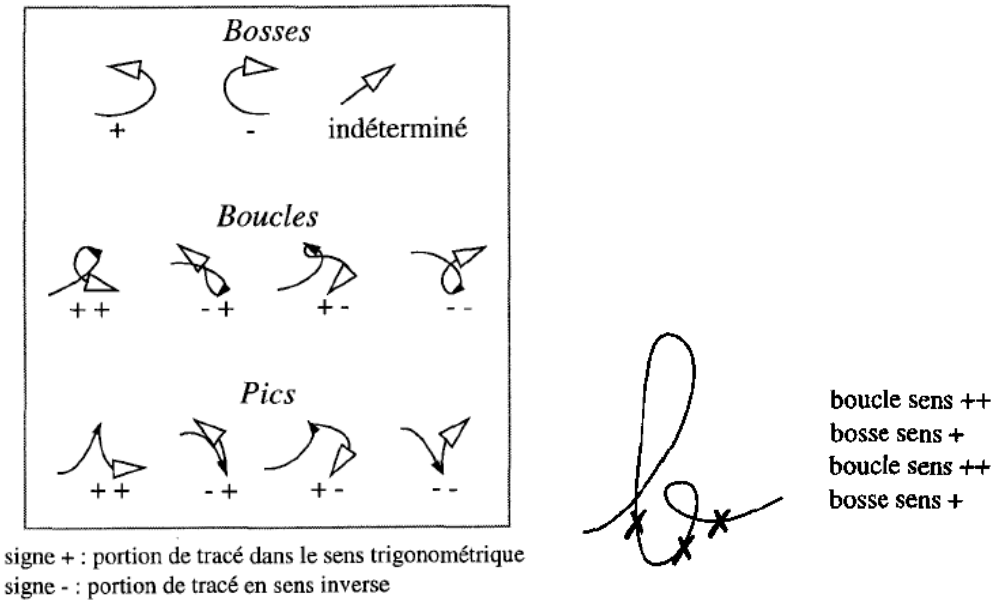


FIG. 2.8 – A gauche, la typologie des caractéristiques associées aux primitives locales et à droite la segmentation en primitives locales (figures extraites de Anquetil et Lorette (1995)).

Les primitives globales reflètent quant à elles les aspects géométriques du tracé. La figure 2.9 présente la position du centre de gravité de la lettre, les proportions du rectangle englobant la lettre et les positions relatives du début et de la fin du tracé par rapport au rectangle englobant.

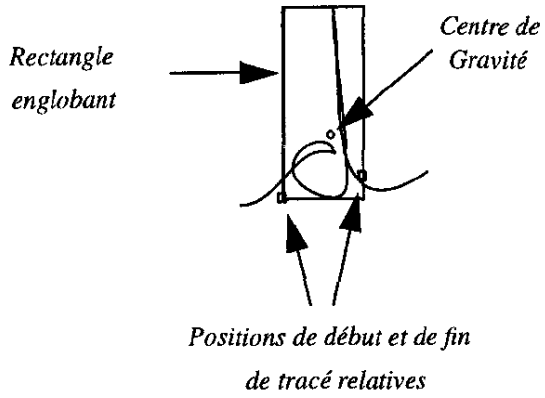


FIG. 2.9 – Aspect géométrique de la lettre (figure extraite de Anquetil et Lorette (1995)).

La figure 2.10 représente la modélisation associée à la classe de chaque lettre (modèle HMM gauche-droite de type Bakis). Cette architecture permet de tenir compte de l'absence, de l'insertion ou de la substitution de primitives dans la lettre. Par exemple, un saut d'états

représente une absence de primitive : c'est le cas du chemin qui correspond à la transition entre les états 0 et 4. Le chemin qui correspond aux transitions 0–3 et 3–4 modélise l'émission d'une primitive o_i et du point de segmentation s_i qui lui est associé. A chaque lettre est donc associé un modèle de ce type permettant de modéliser des primitives.

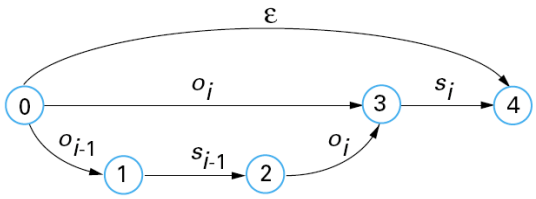


FIG. 2.10 – Modèle de Markov gauche-droite associé à chaque lettre (figure extraite de Crettez et Lorette (1998)).

2.3.1.2 Méthodes neuro-mimétiques

Modèle Les approches basées sur des réseaux de neurones artificiels ont été largement utilisées en reconnaissance d'écriture. Les réseaux de neurones sont des mémoires associatives dans lesquelles l'information d'écriture contenue dans la phase d'apprentissage est stockée implicitement dans le poids des neurones du réseaux. En effet, les réseaux de neurones mettent en œuvre le principe d'apprentissage par expérience. Ils infèrent un système de décision qui est fonction du nombre de cas d'apprentissages rencontrés. Ils ne fournissent pas toujours de règle exploitable par un humain. Ils restent souvent une boîte noire qui fournit une réponse quand on lui présente une donnée. Grâce à leur capacité de classification et de généralisation, les réseaux de neurones sont souvent utilisés dans la reconnaissance de caractères.

Résultats Ces systèmes obtiennent des taux de reconnaissance de mots de l'ordre de 90%. L'avantage de tels systèmes est que tout le procédé de reconnaissance de l'écriture, (normalisation des données, extraction des informations, classification) peut être effectué à partir des réseaux de neurones.

Représentation des lettres Dans les réseaux de neurones, les lettres sont représentées par un vecteur de valeurs numériques calculées à partir de la trace. Le nombre de caractéristiques est de l'ordre de 100 à 300. Par exemple, Schomaker *et al.* (1993) ont développé un système de reconnaissance en ligne d'écriture basé sur deux principes :

- la modélisation des traits,
- la modélisation des caractères.

Les traits sont représentés par 14 caractéristiques qui sont déterminées à l'aide des minimum locaux de vitesse. 9 d'entre elles sont des caractéristiques angulaires, 3 des caractéristiques cartésiennes et 2 des caractéristiques globales liées à la structure de la trace. De plus, chaque caractère correspond à une concaténation de 1 à 6 traits.

2.3.1.3 Méthodes structurelles

Modèle Les méthodes statistiques (approches markoviennes, réseau de neurones, etc.) considèrent une description comme une modélisation mathématique de l'objet. Les approches structurelles quant à elle, considèrent une description comme une représentation du contenu de l'objet. Il s'agit non pas d'extraire des mesures statistiques, mais des informations sémantiques relatives à la topologie de la forme. En effet, prises individuellement, les valeurs mesurées relatives aux caractéristiques (ou leurs présences/absences) ne fournissent pas dans tous les cas suffisamment d'éléments pour classifier les différentes formes à reconnaître. Des informations sur la relation de ces caractéristiques sont alors nécessaires.

Résultats Cette technique a été appliquée à la reconnaissance de chiffres hors-ligne (Suen *et al.*, 1992). Avec une base d'apprentissage comprenant 200 fois chaque chiffre (2000 traces), le taux de reconnaissance obtenu par le modèle est supérieur à 90%.

Représentation des lettres Dans les méthodes structurelles, les lettres sont donc représentées par une description sous forme de composants et de relations entre ces composants. La première phase de la reconnaissance consiste à ramener la donnée brute à un ensemble de composants, c'est-à-dire de la segmenter en fragments, suivant un ou plusieurs niveaux de décomposition. Par exemple, les méthodes basées sur les parcours de formes détectent la structure de la trace (Delalandre, 2005). On distingue pour ce type de méthodes deux traitements : le suivi de traits et la détection de jonctions. Une supervision de ces traitements permet alors d'organiser la détection de chaque trait de la forme et d'en établir les relations pour en construire le graphe de structure.

2.3.2 Revue des modèles de production

Dans cette section, nous présentons plusieurs modèles de production de l'écriture de la littérature : le modèle masse-ressort, le modèle d'Edelman, le modèle VITE. Nous présentons pour chacun de ces modèles, la technique utilisée pour produire des lettres, et la représentation interne adoptée.

2.3.2.1 Modèle masse-ressort

Modèle Un système masse-ressort est un système mécanique à un degré de liberté qui est constitué d'une masse m accrochée à un ressort d'une constante de raideur k et d'un coefficient d'amortissement f. La masse est contrainte de se déplacer dans une seule direction. Son mouvement est dû à deux forces :

- une force de rappel $F_R = -kx$,
- une force de frottements fluides $F_A = -f\frac{dx}{dt}$,

D'après la relation fondamentale de la dynamique, on obtient :

$$m\frac{d^2x}{dt^2} + f\frac{dx}{dt} + kx = 0. \tag{2.1}$$

C'est une équation différentielle du second ordre. Nous définissons deux paramètres :

- la pulsation naturelle du système : $\omega_0 = \sqrt{\frac{k}{m}}$,
- le taux d'amortissement : $\zeta = \frac{f}{2\sqrt{km}}$.

Nous obtenons donc :

$$\frac{d^2x}{dt^2} + 2\zeta\omega_0\frac{dx}{dt} + \omega_0^2 x = 0. \tag{2.2}$$

Dans le cas où le coefficient d'amortissement est nul (force de frottements nulle), on obtient :

$$\frac{d^2x}{dt^2} + \omega_0^2 x = 0. \tag{2.3}$$

Les solutions de l'équation différentielle sont de la forme $x = x_0\cos(\omega_0 t + \phi)$, ce qui est caractéristique d'un oscillateur harmonique.

Représentation des lettres Le modèle masse-ressort décrit l'écriture comme le résultat de deux composantes orthogonales : une horizontale et une verticale. Pour produire l'effet de translation vers la droite, un coefficient c est ajouté à la vitesse selon l'axe des abscisses (v_x). Les profils de vitesse des oscillateurs vertical et horizontal, incluant la translation sont :

$$\begin{aligned} v_x &= a\sin(\omega_x(t - t_0) + \phi_x) + c, \\ v_y &= b\sin(\omega_y(t - t_0) + \phi_y), \end{aligned} \tag{2.4}$$

où t est le temps écoulé depuis le début du mouvement (t_0). Les coefficients a (resp. b) et ϕ_x (resp. ϕ_y) sont l'amplitude et la phase de l'oscillateur horizontal (resp. vertical). ω_x et ω_y sont leurs pulsations. Chaque forme graphique est caractérisée en termes de valeurs de phase relative et d'amplitude relative spécifique. La figure 2.11 présente des cycloïdes obtenus avec les paramètres $a = b = 2c$ et des déphasages de 90° à -30° . Une lettre serait alors composée

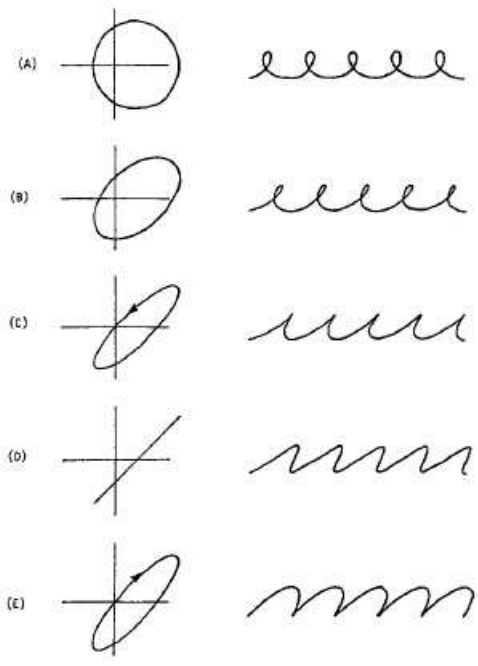


FIG. 2.11 – Résultats obtenus avec les coefficients $a = b = 2c$ et pour des changements de phases de A : 90°, B : 60°, C : 30°, D : 0°, E : -30° (figure tirée de Hollerbach (1981)).

de l'enchaînement d'une ou plusieurs relations spécifiques entre les oscillateurs.

Résultats Le modèle masse-ressort a été proposé par Hollerbach (1981); Hollerbach et Flash (1982); Sallagoïty (2004); Mermelstein et Eden (1964); Eden (1962). La figure 2.12 présente un exemple de mot formé à partir de la concaténation de plusieurs formes graphiques.

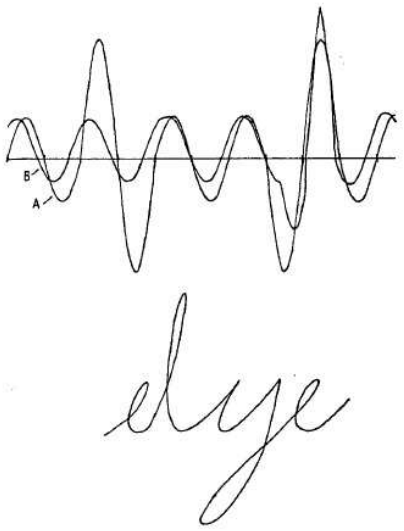


FIG. 2.12 – En haut, l'amplitude et la modulation utilisée pour produire la séquence *elye* et en bas le résultat obtenu (figure tirée de Hollerbach (1981)).

Plamondon a proposé un modèle similaire basé sur une représentation intrinsèque (Plamondon, 1987; Plamondon et Maarse, 1989; Plamondon *et al.*, 1989). Le système d'écriture est représenté mécaniquement sans référence à des axes fixes. Au lieu de décrire la vitesse du bout du stylo en coordonnées cartésiennes (v_x, v_y), on le décrit dans un référentiel curviligne (repère de Frenet). La génération de l'écriture est donc réduite à la production de deux types de déplacements : un déplacement linéaire du bout de l'effecteur le long de l'abscisse curviligne et un déplacement angulaire. L'avantage de cette représentation est que la production ne dépend d'aucun axe fixe et est donc robuste à la rotation.

2.3.2.2 Modèle avec formes de base

Modèle Edelman et Flash (1987) ont proposé un modèle où les lettres correspondent à une concaténation de mouvements. L'hypothèse est que les mouvements obtenus lors de l'écriture sont planifiés segments par segments. Un répertoire de segments de base a été composé afin d'obtenir n'importe quelle lettre.

Représentation des lettres La production des mouvements de ces formes de base est basée sur l'optimisation de la dynamique et du coût de la dérivée de l'accélération (modèle de minimum-jerk que nous détaillerons section 6.1.4). Afin de réaliser les segments, les auteurs utilisent un algorithme d'interpolation en imposant une vitesse au point intermédiaire de ces formes. Une fois les formes de bases réalisées, il suffit de les combiner pour obtenir n'importe quel caractère. Les segments de base générés par le modèle sont présentés figure 2.13.

FIG. 2.13 – Les quatre segments de base : le crochet, la coupe, le gamma et l'ovale (figure tirée de Edelman et Flash (1987)).

Résultats Edelman et Flash (1987) a défini une structure hiérarchique de la production de caractères cursifs. Son modèle traite l'écriture comme tout autre mouvement du bras : il utilise le principe général d'optimisation dynamique. Le principe du modèle de Morasso et Mussa (1982) est similaire : une lettre est considérée comme une série de mouvements de base. Ces mouvements de base réalisés par des B-splines[2] sont ensuite enregistrés pour former un catalogue.

2.3.2.3 Modèle VITE

Modèle Le modèle Vector-Integration To Endpoint (VITE) (Bullock et Grossberg, 1988; Bullock *et al.*, 1993) est fondé sur un modèle de réseau de neurones artificiel.

Représentation des lettres Il utilise une représentation vectorielle basée sur la direction et la longueur du mouvement. Le réseau de neurones génère un vecteur entre la position courante et la position finale à partir de la somme pondérée d'une population de vecteurs.

Résultats Avec ce type de modèle il est possible de générer des trajectoires rectilignes avec un profil de vitesse de type courbe en cloche, ce qui est important car c'est ce que l'on observe expérimentalement chez l'humain (voir chapitre 6.1). Cependant, ce modèle n'explique pas les mouvements courbes.

2.3.2.4 Modèle de Wada

Représentation des lettres Le modèle de Wada et Kawato (2004) représente une lettre comme une séquence de points de contrôle. Les auteurs choisissent comme points de contrôle les minimums et maximums locaux de vitesse (points d'accélération nulle). La figure 2.14 présente la séquence de points de contrôle obtenus pour le mot « hello ». Les points de contrôle sont ensuite reliés en utilisant le modèle de minimum-torque (ce modèle est présenté section 6.1.5.1).

FIG. 2.14 – Séquence de points de contrôle obtenus pour le mot « hello » (figure tirée de Wada et Kawato (2004)).

[2]combinaison linéaire de courbes de Bezier.

Résultats Afin d'optimiser le nombre et la position des points de contrôle, Tokunaga et Wada (2004) ont proposé un procédé d'acquisition de ces points de contrôle : plusieurs sujets tracent une trajectoire de référence ce qui permet de déterminer une série de points de contrôle qui sont les minimum et maximum locaux de vitesse (étape 1). Le modèle de minimum-torque est ensuite appliqué pour générer la trajectoire à partir de ces points de contrôle (étape 2). Puis, la trajectoire générée est évaluée par rapport à la trajectoire de référence ce qui permet d'ajuster le nombre et la position des points de contrôle (étape 3). Les étapes 2 et 3 sont répétées afin d'obtenir une trajectoire optimale.

La Figure 2.15 présente la trajectoire générée et les points de contrôle déterminés par cette procédure.

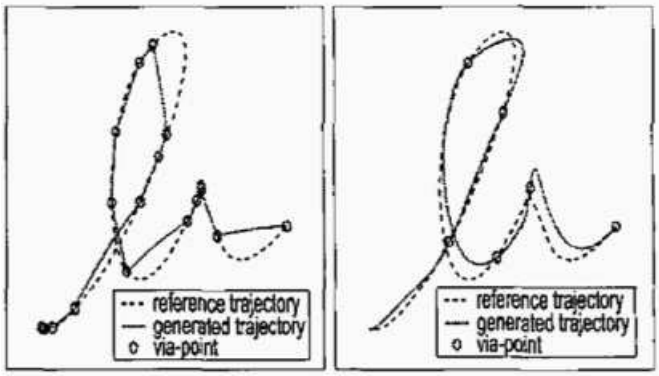


FIG. 2.15 – La trajectoire et les points de contrôle générés à la première phase du modèle (à gauche) et à la dernière phase (à droite) (figure tirée de Tokunaga et Wada (2004)).

2.3.3 BAP : La représentation interne des lettres

Avant d'effectuer un résumé des représentations internes des modèles de perception et de production de l'écriture présents dans la littérature, nous rappelons les contraintes liées aux hypothèses que nous avons réalisées précédemment :

1. La représentation interne des lettres est commune à la perception et à la production des lettres.
2. La représentation interne des lettres est indépendante des effecteurs.

Le premier point suggère que la représentation doit être à la fois pertinente du point de vue de la perception et de l'action : le modèle doit être capable à partir d'une trace d'extraire les informations correspondantes à la représentation, mais aussi à partir de la représentation choisie de construire une nouvelle trace.
Le second point, quant à lui, nous indique que la représentation des lettres doit être définie dans un espace indépendant de l'effecteur. **Nous avons donc choisi de définir la représentation dans l'espace de la feuille, c'est-à-dire à un espace cartésien à deux dimensions**.

Dans les modèles de reconnaissance d'écriture, les représentations des lettres peuvent être classées en deux catégories : celles basées sur les caractéristiques locales et celles basées sur les caractéristiques globales. Nous pouvons directement écarter les caractéristiques globales de notre choix, car même si elles sont très efficaces pour représenter les lettres lors de la reconnaissance, elles ne sont d'aucune utilité lors de la production de l'écriture. Par exemple, connaître le barycentre de la trace ne nous permet pas de définir la trajectoire correspondante à la trace. Les caractéristiques locales quant à elles représentent des points particuliers de la trajectoire ou des morceaux de trajectoire.
Du point de vue des modèles de production de l'écriture, les lettres sont représentées de diverses manières : dans les modèles *masse-ressort*, les formes graphiques sont caractérisées en terme de phase et d'amplitude des oscillateurs. Les lettres sont alors la concaténation de plusieurs formes graphiques. L'avantage d'une telle représentation est que la production ne dépend d'aucun repère fixe et est donc robuste à la rotation. Cependant, il serait difficile d'utiliser une telle représentation interne pour pouvoir reconnaître les caractères. En effet, il n'est pas aisé à partir d'une trace de déterminer la phase et l'amplitude d'un système masse-ressort ayant généré cette trace.
Le modèle de production avec des *formes de base* permet de produire des lettres en concaténant des segments de base. Il est important de noter que les formes de base sont réalisées à partir de séquence de points qui définissent les formes de base, et donc que ce type de modèle correspond à une représentation basée sur des points de contrôle.
Le dernier modèle de production des lettres que nous avons étudié (modèle de Wada) est également basé sur des points de contrôle. Les points sont placés sur des maximums et minimums locaux de vitesse.
Le tableau présenté ci-dessous résume les différentes représentations des lettres utilisées dans les modèles de la littérature. Pour chaque type de représentation, nous évaluons sa pertinence du point de vue de la perception et de la production de lettres.

Nom du modèle	Type de représentation	Production	Perception
Méthodes markoviennes	primitives locales	-	++
	primitives globales	- -	++
Réseaux de neurones	caractéristiques locales	-	++
	caractéristiques globales	- -	++
Méthodes structurelles	relation entre caractéristiques	-	++
Modèle masse-ressort	phase et amplitude	++	- -
Modèle avec formes de base	points de contrôle	++	-
Modèle VITE	vecteur	++	-
Modèle de Wada	points de contrôle	++	-

Dans le modèle BAP, nous avons cherché à avoir une représentation interne des lettres qui soit à la fois exploitable pour produire et reconnaître des lettres. Nous faisons l'hypothèse que les lettres sont représentées par des points de contrôle. Nous avons vu qu'une telle représentation est pertinente du point de vue de la production. Afin qu'elle le soit également pour la perception, nous choisissons un critère pour déterminer la position des points qui soit géométriquement identifiable.
Nous avons donc choisi de placer les points de contrôle :
– aux points de tangence remarquable (lorsque la tangente est horizontale ou verticale),
– aux points de rebroussement.
En effet, il est facile de détecter ces points sur la trajectoire. La figure 2.16 présente une lettre et ses points de contrôle déterminés de manière géométrique.

Il s'avère que ces points spécifiques de la trajectoire sont aussi placés à l'endroit où au moins une des dérivées s'annule :

$$\frac{dx}{dt}(t) = 0 \ \vee \ \frac{dy}{dt}(t) = 0. \tag{2.5}$$

La figure 2.17 présente la position des points de contrôle par rapport au profil de vitesse et par rapport à la lettre. Nous avons donc choisi des points de contrôle pertinents pour la perception (géométriquement identifiables dans l'espace visuel) mais aussi pertinents vis-à-vis de l'action, du moins du point de vue cinématique (passage à zéro dans l'espace des vitesses).

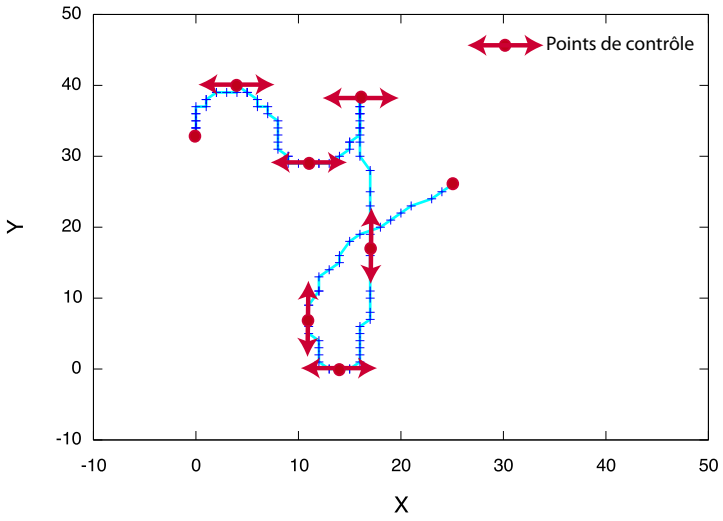


FIG. 2.16 – Les points de contrôle placés aux points de tangence remarquable (horizontaux ou verticaux) ou aux points de rebroussement.

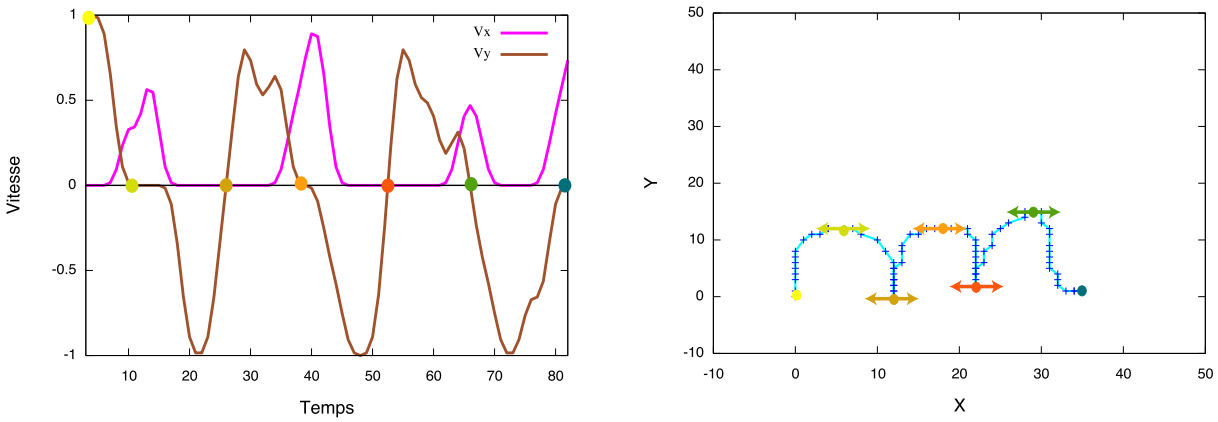


FIG. 2.17 – A gauche, les profils de vitesses ($\dot{x}$ en rose et $\dot{y}$ en marron), avec les positions des points de contrôle. A droite, les positions des points de contrôle dans la lettre.

2.4 Une simulation motrice de l'action est-elle nécessaire pour la perception ?

Une des idées principales du *Sens du Mouvement* de Berthoz (1998) est que « la perception n'est pas seulement une interprétation de messages sensoriels : elle est contrainte par l'action, elle est simulation interne de l'action, elle est jugement et prise de décision, elle est anticipation des conséquences de l'action. » Dans cette section, nous allons voir que la perception est *contrainte par l'action* et est *simulation interne de l'action*. Tout d'abord, nous effectuons une revue des observations expérimentales. Puis, nous présentons la théorie de la simulation.

2.4.1 Etudes expérimentales

De nombreuses études expérimentales montrent l'implication du système moteur dans la perception. C'est ce qu'ont montré Kerzel et Bekkering (2000) dans l'une de leurs expériences. Dans cette étude, les sujets devaient prononcer les syllabes /*ba*/ ou /*da*/ (indiquées à l'écran) alors qu'ils voyaient un film représentant une bouche articulant /*ba*/ ou /*da*/ et inversement. Les résultats montrent la présence d'interférence entre la perception et la production : les sujets répondent plus vite la syllabe /*ba*/ lorsque le film présente une vidéo d'un mouvement /*ba*/ plutôt que /*da*/. Cette expérience suggère que percevoir les gestes active un programme moteur qui interagit avec le programme moteur activé simultanément pour la même action.

Galantucci *et al.* (2006), lors d'une revue, ont classé en trois catégories les implications du système moteur dans la perception :

- les contraintes motrices liées à l'anatomie du corps,
- les contraintes motrices liées à la dynamique,
- les contraintes motrices liées aux différences entre individus.

Dans la suite de ce paragraphe, c'est avec cette classification que nous allons illustrer le lien entre perception et action.

2.4.1.1 Contraintes anatomiques

Durant une expérience de Shiffrar et Freyd (1990), les sujets devaient regarder des paires de photographies de personnes dans différentes positions et déterminer le chemin du mouvement pris entre ces positions. Il s'avère que la réponse des sujets est cohérente avec les mouvements normaux du corps humain (ce qui ne correspond pas toujours au plus court chemin). Ces mouvements qui ont été choisis préservent les limitations des articulations et la solidité du corps. Ce résultat semble indiquer que le système de perception possède des connaissances sur les limitations physiques des différents corps. En effet, le système perceptif ne considère que des chemins biologiquement plausibles.

2.4.1.2 Contraintes dynamiques

Plusieurs études expérimentales ont montré que la perception d'un objet en mouvement est sensible aux lois qui spécifient le mouvement de systèmes moteurs biologiques. Par exemple, la conformité de la trajectoire avec la loi de la puissance 2/3 affecte la perception de la trace (Viviani et Stucchi, 1992). Ces résultats ont été montrés chez l'adulte (Kandel *et al.*, 2000b), les nouveaux-nés (Méary *et al.*, 2007) et aussi sur les poussins (Vallortigara *et al.*, 2005). Dans cette dernière expérimentation, Vallortigara *et al.* (2005) ont montré que des poussins nouveau-nés sans expérience visuelle et motrice sont sensibles aux mouvements de locomotion

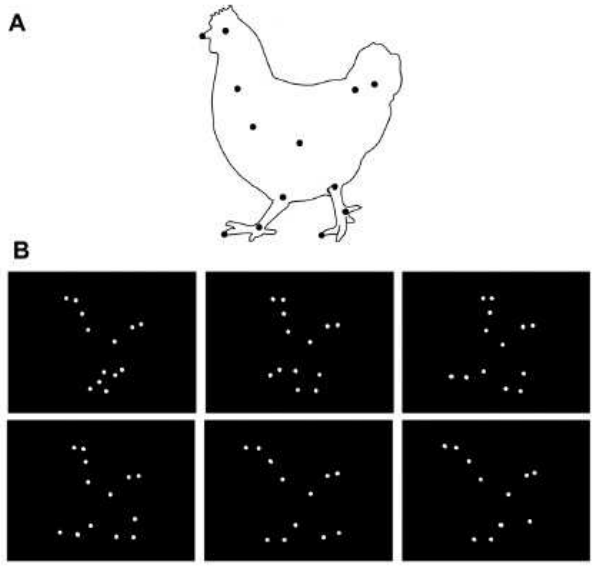


FIG. 2.18 – A : Disposition des points lumineux sur la poule. B : Six images extraites de l'animation d'une poule en mouvement (figure tirée de Vallortigara *et al.* (2005)).

d'une poule. Ils préfèrent regarder l'écran sur lequel ils peuvent observer des points lumineux représentant le déplacement d'une poule plutôt que des mouvements rigides. La figure 2.18 présente les points lumineux correspondant au déplacement d'une poule. Les résultats indiquent que les poussins préfèrent les mouvements d'une poule marchant plutôt qu'aux mouvements rigides (60%) et qu'aux mouvements aléatoires (60%). Les auteurs ont aussi montré que les poussins étaient sensibles au déplacement d'un chat, déplacement qui ne pourra jamais être exécuté par le poussin. Cependant, d'autres études sont en contradiction avec ces résultats. Par exemple, Ildei-Bidet (2006) a montré chez les enfants et les patients parkinsoniens des décalages entre les compétences motrices et les compétences perceptives. Ces résultats semblent suggérer que la perception visuelle des mouvements humains n'implique pas systématiquement une intervention du système moteur.

2.4.1.3 Différences entre individus

Les compétences motrices peuvent aussi inclure des contraintes propres à un individu (Reid *et al.*, 2005). Knoblich et Flach (2006) ont résumé les preuves récentes suggérant que les connaissances motrices forment aussi l'identité de la personne. Lors d'expériences, les sujets ont tendance à mieux reconnaître leur mouvement que celui des autres. Par exemple, dans le cadre de l'écriture, Knoblich et Prinz (2001) ont testé si les sujets pouvaient reconnaître la cinématique de leur propre écriture : ils observaient deux cinématiques (points bougeant le long d'une trajectoire) de la production du même symbole et devaient déterminer quel était le leur. Les résultats montraient que les participants pouvaient reconnaître leur propre écriture, avec seulement des informations cinématiques.

2.4.2 Simulation du mouvement dans la perception

La théorie de la simulation motrice de Jeannerod (2001), qui consiste à simuler les états mentaux d'autrui, décrit que le système moteur est un système de simulation qui est activé pour réaliser des actions et aussi pour reconnaître les actions des autres. Les actions ne sont pas exécutées mais le moyen d'atteindre le but et les conséquences de l'action sont simulés. Ces hypothèses sont en accord avec les études montrant le rôle des compétences motrices lors de la perception visuelle des mouvements humains et les études montrant l'activation neurale des structures motrices durant la perception des mouvements humain. En effet, Raos *et al.* (2007) ont constaté que l'observateur se place lui-même à la place de la personne observée pour prédire ses actions.

Il est important de noter que ce que nous appelons simulation de l'action n'est pas une action volontaire d'un sujet mais un processus inconscient. Lors de la simulation de l'action, on simule le point de vue de l'autre, et on utilise le résultat pour comprendre l'autre et prédire son comportement. Cependant, il est nécessaire de distinguer trois cas concernant la simulation de l'action (Proust, 2000) :

- Simulation égocentrique : ce cas consiste à transformer les informations de perception à la première personne et donc à se mettre soi-même à la place de la personne que l'on souhaite simuler.
- Simulation projective : ce cas implique que le simulateur se projette à la place du simulé.
- Simulation neutre : simuler n'implique pas de transfert de perspective. Le contenu est seulement celui de la situation simulée.

2.4.2.1 Simulation du mouvement dans le cadre de la lecture et de l'écriture

Bien que lorsque l'on observe des traces manuscrites, on n'observe pas directement le mouvement effectué par le scripteur pour produire ces traces, nous proposons qu'il est possible de simuler mentalement l'action que le scripteur a effectué pour obtenir cette trace. Il est donc possible d'appliquer la théorie de la simulation dans le cadre de la lecture-écriture. Cette idée est soutenue par plusieurs expériences.

Tout d'abord, Longcamp (2003) a montré que lorsque nous reconnaissons les lettres, nous nous référons implicitement aux mouvements que nous produirions pour écrire ces mêmes lettres. L'auteur a mesuré, par Imagerie par Résonance Magnétiques fonctionnelle, l'activité cérébrale de sujets adultes qui lisaient ou qui écrivaient des lettres. Il a observé qu'une zone corticale prémotrice qui est activée pendant l'écriture l'est aussi pendant la lecture (alors que les sujets sont immobiles).

Durant une autre étude expérimentale (Knoblich *et al.*, 2002), les sujets devaient, à partir de la cinématique de leurs propres productions et des productions d'un autre sujet, déterminer s'ils étaient à l'origine du tracé. Lorsque l'écriture n'était pas contrainte, les sujets répondaient correctement uniquement pour leurs propres productions. Lorsque les cinématiques des productions étaient normalisées par des contraintes spatiales, les prédictions étaient très précises dans les deux cas (reconnaissance de leurs traces et reconnaissance des traces des autres sujets). Il semble donc que des mécanismes particuliers, basés sur la simulation par le sujet de son propre mouvement, permettent de reconnaître les traces.

Dans le cadre de l'écriture et de la lecture, il s'agirait de simuler l'action d'écriture pour mieux reconnaître la lettre. Cela consiste à se placer à la place du scripteur et à effectuer mentalement l'action qu'il aurait effectuée pour obtenir cette trace.

2.4.2.2 BAP : Simulation de l'écriture pour la lecture

Il est possible de s'intéresser à l'intégration de cette théorie de la simulation dans les modèles sensori-moteurs et notamment dans notre modèle perception-action. Nous avons vu précisément que dans le cadre de l'écriture, cela reviendrait à simuler l'action d'écriture pour mieux reconnaître la lettre. Nous avons modélisé cette simulation en ajoutant une nouvelle dépendance (**perception simulée**) entre la *trace écrite simulée* et la *représentation interne motrice*, comme présenté sur la figure 2.19.
Dans le modèle BAP, la lecture avec simulation du mouvement peut être décrite de la façon suivante : lors de la perception d'une lettre, la *trace lue* est convertie dans l'espace de représentation interne. Etant donné que l'espace de représentation interne est le même pour la perception et pour la production, il est possible de simuler l'action à partir de cette représentation interne pour obtenir une *trace écrite simulée*. Le processus de perception est une nouvelle fois exécuté pour convertir cette *trace écrite simulée* dans l'espace de représentation interne. La reconnaissance de la lettre se fait donc à partir de la représentation interne perceptive et de la représentation interne motrice.

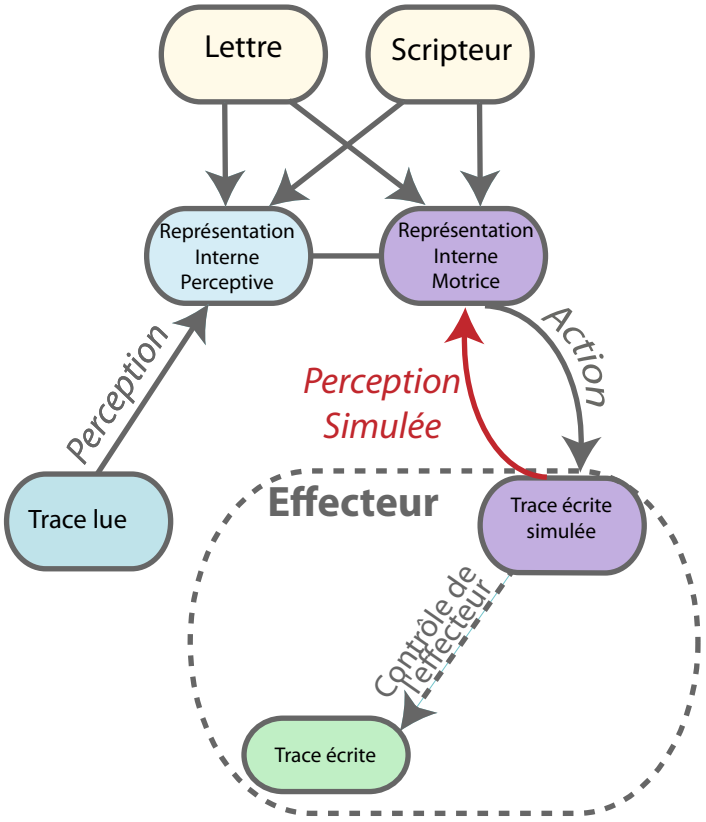


FIG. 2.19 – Modèle perception-action intégrant la boucle de simulation du mouvement

2.5 Résumé

Dans ce chapitre, nous avons défini une liste d'hypothèses issues de la littérature. Ces hypothèses sont les suivantes :

- Représentation interne des lettres :
  - elle est le pivot entre la perception et l'action,
  - son espace de codage est indépendant de l'effecteur,
  - elle est encodée par des séquences de points de contrôle.
- La perception ne joue aucun rôle lors de la production de lettres, nous ne prenons pas en compte les retours sensoriels (boucle ouverte).
- La production de mouvements est simulée lors de la perception de lettres.

Dans le chapitre suivant, nous allons présenter le formalisme de la Programmation Bayésienne qui permet d'exprimer mathématiquement ces hypothèses.

Chapitre 3

Programmation Bayésienne

Nous utilisons tout au long de ce travail un unique formalisme qui nous permet à la fois de définir le modèle et de le mettre en œuvre par la simulation de tâches. La méthodologie de ce formalisme est la Programmation Bayésienne. Dans un premier temps vient la description de ce formalisme. Nous illustrons ensuite l'outil de programmation bayésienne à travers la modélisation de la représentation interne des lettres.

3.1 Formalisme Bayésien

La modélisation Bayésienne s'appuie sur les probabilités pour représenter des connaissances (Bessière *et al.*, 1998a,b). Nous utilisons le formalisme de la programmation Bayésienne, qui donne un cadre pour leur organisation (Lebeltel, 1999). Cette section est adaptée de Colas (2006); Diard (2003); Koike (2005); Bessière (2003).

3.1.1 Probabilités et représentation des connaissances

Nous souhaitons construire un modèle à partir d'un certain nombre de connaissances préalables et d'observations. Nous considérons le calcul des probabilités comme une généralisation de la logique classique : la logique est étendue à des propositions dont la vérité n'est pas connue avec certitude. Les probabilités permettent alors de représenter la plausibilité de propositions logiques. La modélisation bayésienne manipule des variables selon les règles des probabilités à l'image d'un raisonnement logique. Nous présentons ici les probabilités et les règles d'inférence.

3.1.1.1 Variables et distributions

Il n'est pas rare de ne pas connaître avec précision la valeur v d'une variable V de domaine D_V. Dans ce cas, notre connaissance de la valeur de cette variable s'exprime sous la forme d'une distribution de probabilité sur son domaine. Une distribution de probabilité sur V est une fonction P de D_V vers $\mathbb{R}^+$ d'intégrale (ou somme) unité sur D_V :

$$\int_{D_V} P(V) = 1. \tag{3.1}$$

Par la suite, nous utiliserons la notation $P(V)$ pour désigner une distribution de probabilité sur une variable discrète V. Nous définissons donc $P([V = v])$ la probabilité que la variable V soit égale à v.

Distribution conjointe Etant données deux variables A et B de domaines respectifs D_A et D_B, on s'intéresse à leur conjonction $A \wedge B$. Afin d'alléger les notations, la distribution de probabilité conjointe $P(A \wedge B)$ sera notée :

$$P(A\ B). \tag{3.2}$$

Distribution conditionnelle Nous nous intéressons maintenant à la valeur d'une première variable supposant connaître exactement la valeur d'une seconde. La distribution de probabilité de A sachant que B a pour valeur b ($b \in D_B$) est notée :

$$P(A \mid [B = b]). \tag{3.3}$$

3.1.1.2 Inférence

Deux règles fondamentales sont à la base de toute inférence probabiliste : la règle de Bayes et la règle de normalisation.

Règle de Bayes La règle de Bayes (ou règle du produit) met en relation la distribution conjointe avec les distributions conditionnelles :

$$P(A\ B) \ = \ P(A \mid B)\ P(B) \ = \ P(B \mid A)\ P(A). \tag{3.4}$$

On observe que cette règle est symétrique en A et B du fait de la commutativité de la conjonction.

Règle de normalisation La seconde règle nécéssaire à l'inférence est la marginalisation :

$$\sum_B P(A\ B) \ = \ P(A). \tag{3.5}$$

Cette règle découle de la normalisation des distributions (équation 3.1) et de la règle de Bayes (équation 3.4).

3.1.1.3 Indépendance

Indépendance marginale Lorsque l'on a :

$$\forall\ b \ \in \ D_B, \ \ P(A \mid [B = b]) \ = \ P(A), \tag{3.6}$$

on dit que A et B sont indépendantes. Nous pouvons alors décomposer la distribution de probabilité conjointe de la façon suivante :

$$P(A\ B) \ = \ P(A)\ P(B). \tag{3.7}$$

Indépendance conditionnelle Lorsque l'on a :

$$P(A \mid B\ C) \ = \ P(A \mid B), \tag{3.8}$$

on dit que A est indépendante de C sachant B. On notera que l'indépendance conditionnelle n'implique pas l'indépendance marginale, et réciproquement.

3.1.2 Notion de Programmation Bayésienne

Il existe plusieurs formalismes de modélisation bayésienne qui ont pour point commun de s'intéresser à la distribution conjointe. La programmation bayésienne s'appuie sur une écriture algébrique de la distribution conjointe. Elle s'attache à fournir un cadre pour spécifier des connaissances probabilistes afin de pouvoir calculer la distribution conjointe. Cela permet de répondre à une question, c'est-à-dire calculer une distribution portant sur les variables ou un partie des variables de la conjointe. La définition de la conjointe est un processus en trois parties :

- définition des variables sur lesquelles raisonner ;
- décomposition de la distribution conjointe en produit de facteurs plus simples ;
- spécification de chacun des facteurs de la décomposition ;

La décomposition est une étape de la spécification de la distribution conjointe portant sur la structure du modèle. Le but est de la simplifier. Pour cela, on utilise la règle de Bayes pour l'écrire sous la forme d'un produit. Le programmeur fait des hypothèses d'indépendance conditionnelle pour simplifier les facteurs de la conjointe : on obtient alors une décomposition de la distribution conjointe. On spécifie ensuite le moyen de calculer chaque distribution, facteur de cette décomposition. En général, il s'agit soit d'une forme paramétrique telle une loi uniforme, ou une loi gaussienne et des valeurs des paramètres de cette loi, telle la moyenne et l'écart-type pour la gaussienne, soit un mécanisme pour calculer ces valeurs au vu d'un ensemble d'observations expérimentales (apprentissage). De cette manière, on obtient une expression calculable de la distribution conjointe.

Durant la phase d'exploitation, nous utilisons cette description afin de répondre à une question. Poser une question, consiste à calculer la distribution de probabilité d'un certain nombre de variables de la description connaissant les valeurs d'autres variables. L'inférence bayésienne permet de répondre à n'importe quelle question à partir de la spécification de la conjointe.

3.1.3 Implémentation

Dans ce document nous ne discutons pas des mécanismes pour réaliser informatiquement les inférences. Nous disposons du moteur d'inférence bayésienne ProBT
(http ://www.probayes.com/) qui implémente ces algorithmes. Du point du vue du travail présenté dans ce document, l'inférence est un problème résolu par ce moteur.

3.2 Exemple introductif

Dans cette section, nous fournissons un exemple détaillé de Programmation Bayésienne. Il concerne la relation entre les représentations internes des lettres, les lettres et les scripteurs. La figure 3.1 décrit cette relation (c'est un extrait du modèle global que nous définirons dans le chapitre suivant).

Dans un premier temps, nous allons définir les variables probabilistes, puis nous définirons la distribution conjointe en produit de facteurs plus simples. Et enfin, nous spécifierons chacun de ces facteurs.

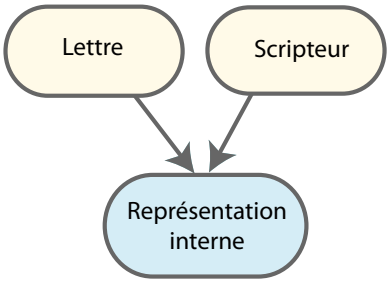


FIG. 3.1 – Extrait du modèle global.

3.2.1 Variables Probabilistes

Nous notons L la variable correspondant aux lettres. Cette variable peut prendre 22 valeurs distinctes (nous ne considérons que les lettres sans levé de crayon) : a, b, c, d, e, f, g, h, k, l, m, n, o, p, q, r, s, u, v, w, y, z.

La variable correspondant aux scripteurs est W. Elle peut prendre 4 valeurs possibles : *Julienne*, *Jean-Louis*, *Christophe*, *Estelle*.

Nous notons $C_L^{0:N}$ la variable correspondant à la représentation interne d'une lettre et, plus précisément, à la séquence de ses points de contrôle. Nous avons fixé la valeur de N à 15, ce qui correspond à limiter à 16 le nombre de points de contrôle dans une lettre. $C_L^{0:N}$ correspondent à l'ensemble $\{C_{Lx}^{0:N}\ C_{Ly}^{0:N}\ C_{L\dot{x}}^{0:N}\ C_{L\dot{y}}^{0:N}\}$, variables de position en x et y et variables de vitesse $\dot{x}$ et $\dot{y}$. La variable $C_L^{0:N}$ est en fait elle-même une conjonction de 64 variables ($16*4$) : $C_{Lx}^{0:N}\ C_{Ly}^{0:N}\ C_{L\dot{x}}^{0:N}\ C_{L\dot{y}}^{0:N}$.

Nous notons C_{Lx}^n la variable correspondant à l'abscisse du $n^{\text{ème}}$ point de contrôle de la séquence (avec $n \in [0, N] \cap \mathbb{N}$). La variable C_{Lx}^n est une variable entière à valeur dans $[0, 40] \cap \mathbb{N}$. Il en est de même pour la variable C_{Ly}^n, ordonnée du $n^{\text{ème}}$ point de contrôle de la séquence. En d'autres termes, les points de contrôle sont placés sur une grille de $41 * 41$ positions possibles.

La vitesse en x du $n^{\text{ème}}$ point de contrôle de la séquence est notée $C_{L\dot{x}}^n$. Elle peut prendre 9 valeurs possibles (entier compris entre -4 et 4). Nous avons fortement discrétisé l'espace des vitesses car cette résolution est suffisante pour notre application tout en permettant d'avoir une dimension minimale pour l'espace de représentation des lettres.

3.2.2 Distribution conjointe et décomposition

La description sur les variables que nous venons d'effectuer a pour but la définition de la distribution conjointe :

$$P(C_L^{0:N}\ L\ W). \tag{3.9}$$

La règle du produit nous permet de décomposer cette expression, en l'exprimant sous forme d'un produit de distributions. Avec la définition des variables données ci-dessus, la décomposition de la distribution de probabilité conjointe prend la forme suivante :

$$\begin{aligned} P(C_L^{0:N}\ L\ W) &= P(C_L^{0:N} \mid L\ W)\ P(L\ W) && (3.10)\\ &= P(C_L^{0:N} \mid L\ W)\ P(L \mid W)\ P(W) && (3.11)\\ &= P(C_L^{0:N} \mid L\ W)\ P(L)\ P(W). && (3.12) \end{aligned}$$

Les équations 3.10 et 3.11 sont obtenues en appliquant la règle de Bayes. De plus, nous faisons l'hypothèse que la connaissance du scripteur (W) n'apporte aucune information sur la connaissance de la lettre (L). L'équation 3.12 est obtenue en appliquant cette hypothèse d'indépendance entre les variables L et W.

Le terme correspondant à la représentation des lettres dans la décomposition de la distribution de probabilité 3.12 est :

$$P(C_L^{0:N} \mid L\ W). \tag{3.13}$$

Il exprime donc la distribution de probabilité de la séquence des points de contrôle connaissant la lettre et le scripteur. En d'autres termes, pour chaque lettre et chaque scripteur est associé une distribution de probabilité de la séquence des points de contrôle.

Nous faisons une hypothèse d'indépendance entre les quatre dimensions x, y, $\dot{x}$ et $\dot{y}$. La distribution de probabilité des points de contrôle sachant la lettre et le scripteur est donc égale aux produits des distributions de probabilité de leurs positions et de leurs vitesses :

$$P(C_L^{0:N} \mid L\ W) = \begin{pmatrix} P(C_{Lx}^{0:N} \mid L\ W)\ P(C_L^{0:N} \mid L\ W) \\ P(C_{L\dot{x}}^{0:N} \mid L\ W)\ P(C_{L\dot{y}}^{0:N} \mid L\ W) \end{pmatrix}. \tag{3.14}$$

De plus, nous faisons l'hypothèse que la position (ou vitesse) d'un point de contrôle ne dépend que de la position (ou vitesse) du point de contrôle précédent. En conséquence, sur la dimension x, nous obtenons :

$$P(C_{Lx}^{0:N} \mid L\ W) = P(C_{Lx}^0 \mid L\ W) \prod_{n=1}^{N} P(C_{Lx}^n \mid C_{Lx}^{n-1}\ L\ W). \tag{3.15}$$

La décomposition est similaire sur les trois autres dimensions, ce qui nous permet de décomposer la distribution de probabilité des points de contrôle de la façon suivante :

$$= \begin{pmatrix} P(C_{Lx}^0 \mid L\ W)\ P(C_{Ly}^0 \mid L\ W) \\ P(C_{L\dot{x}}^0 \mid L\ W)\ P(C_{L\dot{y}}^0 \mid L\ W) \end{pmatrix} \prod_{n=1}^{N} \begin{pmatrix} P(C_{Lx}^n \mid C_{Lx}^{n-1}\ L\ W)\ P(C_{Ly}^n \mid C_{Ly}^{n-1}\ L\ W) \\ P(C_{L\dot{x}}^n \mid C_{L\dot{x}}^{n-1}\ L\ W)\ P(C_{L\dot{y}}^n \mid C_{L\dot{y}}^{n-1}\ L\ W) \end{pmatrix}. \tag{3.16}$$

Le terme $P(C_L^{0:N} \mid L\ W)$ est une distribution de probabilité qui n'est pas facile à spécifier directement sous cette forme à 64 dimensions. Ces hypothèses d'indépendances permettent donc de passer d'une distribution sur un espace de dimension 64 et de taille $41^{32} \times 9^{32}$ à un produit de 64 (16×4) distributions beaucoup plus simples sur des espaces de dimension 1 et de taille 41 (position) ou 9 (vitesse).

3.2.2.1 Formes paramétriques

Nous définissons les formes paramétriques des termes de la décomposition.

- $P(L)$ est une distribution de probabilité de loi uniforme car nous n'avons aucun a priori sur les lettres. Une loi uniforme est une forme paramétrique à un paramètre qui est le nombre de cas de la variable sur laquelle porte cette distribution. Cette loi indique l'équiprobabilité de tous les cas ; elle sera choisie en général pour modéliser notre ignorance sur un phénomène. Dans notre formalisme, la définition des variables est statique, donc cette loi n'évoluera pas en fonction des données expérimentales. Soit L définie par

son domaine $D_L = \{0, 1, ..., (k_l - 1)\}$ et son cardinal k_l, la distribution de probabilité de loi uniforme se note :

$$P(L) = U_{k_l}(L)$$
$$\forall l \in D_L, \quad P([L = l]) = \frac{1}{k_l} = \frac{1}{22}.$$

- $P(W)$: A ce terme est associée une distribution de loi uniforme, car nous n'avons pas d'a priori sur le scripteur. Nous définissons $D_W = \{0, 1, ..., (k_w - 1)\}$ le domaine de W et k_w son cardinal. Nous avons alors :

$$P(W) = U_{k_w}(W)$$
$$\forall w \in D_w, \quad P([W = w]) = \frac{1}{k_w} = \frac{1}{4}.$$

- $P(C_L \mid L\ W)$: Il nous reste à définir les formes paramétriques de chacun des termes du produit suivant :

$$\begin{pmatrix} P(C^0_{Lx} \mid L\ W)\ P(C^0_{Ly} \mid L\ W) \\ P(C^0_{L\dot{x}} \mid L\ W)\ P(C^0_{L\dot{y}} \mid L\ W) \end{pmatrix} \prod_{n=1}^{N} \begin{pmatrix} P(C^n_{Lx} \mid C^{n-1}_{Lx}\ L\ W)\ P(C^n_{Ly} \mid C^{n-1}_{Ly}\ L\ W) \\ P(C^n_{L\dot{x}} \mid C^{n-1}_{L\dot{x}}\ L\ W)\ P(C^n_{L\dot{y}} \mid C^{n-1}_{L\dot{y}}\ L\ W) \end{pmatrix}. \tag{3.17}$$

Pour ne pas nous interdire des cas potentiels non observés durant la phase d'apprentissage, nous souhaitons qu'aucune probabilité ne soit nulle. De plus, nous ne voulons pas imposer de forme à priori sur les distributions de probabilité. C'est pourquoi, nous faisons l'hypothèse que la distribution de probabilité de chacun des termes du produit est une loi de succession de Laplace :

$$P([C^n_{Lx} = v] \mid [C^{n-1}_{Lx} = c]\ [L = l]\ [W = w]) = \frac{p_i + \alpha}{N + k\alpha}. \tag{3.18}$$

La variable p_i correspond au nombre de fois où $[C^n_{Lx} = v]$ dans les données durant la phase d'apprentissage. Le paramètre N correspond au nombre total d'observations effectuées.
Etant donné que l'on souhaite n'avoir aucune probabilité nulle, nous fixons un *prior* : chacun des cas a été vu au moins une fois durant la phase d'apprentissage. Ce *prior* s'exprime mathématiquement en fixant la variable α à 1. La variable k correspond au cardinal de la variable C^n_{Lx}. Lorsque l'on n'a jamais rencontré d'observations pour un cas donné ($N = 0$), la distribution de probabilité obtenue est donc de type uniforme. A l'inverse, les cas pour lesquels on dispose de beaucoup de données expérimentales et donc $p_i >> k\alpha$, la distribution revient à un histogramme $\frac{p_i}{N}$.
Intuitivement, ce choix de loi de succession de Laplace correspond donc à construire un histogramme avec la garantie mathématique de ne pas avoir de probabilité nulle pour les cas non observés (ce qui est une probabilité souhaitable pour la réalisation informatique de l'inférence).

3.2.2.2 Apprentissage

Base de données d'apprentissage L'apprentissage est effectué sur une base de données de 3520 traces. Notre base de donnée comprend 4 scripteurs et 22 lettres possibles :

a, *b*, *c*, *d*, *e*, *f*, *g*, *h*, *k*, *l*, *m*, *n*, *o*, *p*, *q*, *r*, *s*, *u*, *v*, *w*, *y*, *z*.

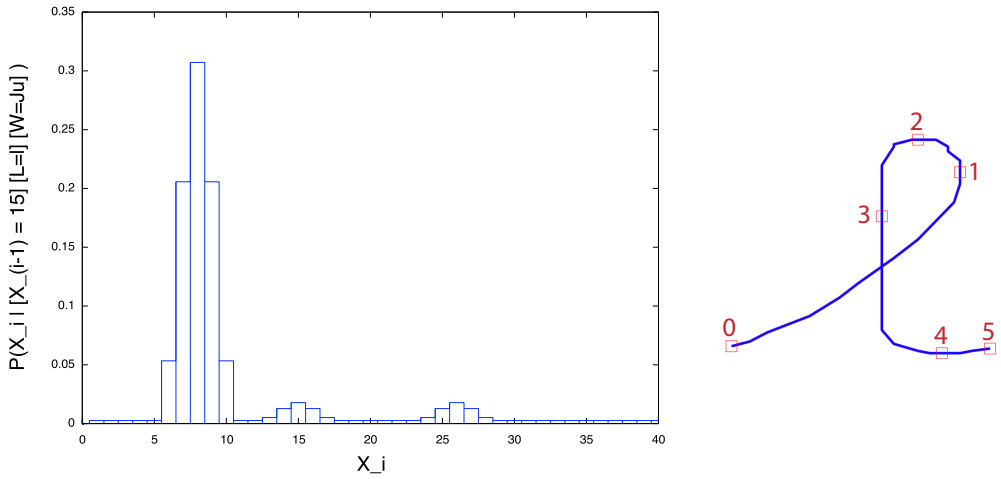


FIG. 3.2 – A gauche, distribution de probabilité de la position en x du troisième point de contrôle de la lettre l pour le scripteur $Julienne$: $P(C^3_{Lx} \mid [C^2_{Lx} = 15]\ [L = l]\ [W = Julienne])$. A droite, une trace l produite par le scripteur $Julienne$, avec les différents points de contrôle.

Nous avons supprimé de notre alphabet les lettres comprenant un levé de stylo telles que le i ou le t.

En pratique, l'acquisition des traces est faite à partir d'un stylo (Intuos 3, Grip Pen) et d'une tablette graphique (Wacom, Intuos 3) d'une précision de 0.25 mm et avec un débit maximum de 200 points par seconde.

Identification Après avoir spécifié les connaissances préalables, nous avons fixé les formes paramétriques associées aux termes. Il nous reste maintenant à déterminer les paramètres des formes paramétriques, en l'occurrence des lois de succession de Laplace. Ces paramètres sont définis lors d'une phase d'apprentissage de notre modèle. Pour chaque lettre et chaque scripteur et chaque point de contrôle C^{n-1}_{Lx}, nous déterminons le nombre d'observations p_i pour chaque point de contrôle C^n_{Lx}.

La figure 3.2 présente la distribution de probabilité obtenue pour :

$$P(C^3_{Lx} \mid [C^2_{Lx} = 15]\ [L = l]\ [W = Julienne]). \tag{3.19}$$

Nous observons que lorsque la position en x du second point de contrôle vaut 15, la position en x du troisième point de contrôle a une forte probabilité de valoir entre 7 et 9. On a représenté à droite une lettre l avec les différentes positions des points de contrôle. Le troisième point de contrôle du l est celui dans la partie descendante de la trace. On observe que la distribution de probabilité de la position en x est centrée sur la valeur 8, ce qui est normal étant donné que le point 3 est à gauche du point 2. De plus, nous observons que quelque soit la valeur c^3_{lx} de C^3_{Lx}, la probabilité de $P([C^3_{Lx} = c^3_{lx}] \mid [C^2_{Lx} = 15]\ [L = l]\ [W = Julienne]))$ n'est jamais nulle.

Traitement de la trace Pour pouvoir exploiter les traces, il faut au préalable effectuer un ensemble de traitements. Il faut en particulier normaliser les traces pour pouvoir comparer les lettres quelque soit leur taille. Il est aussi nécessaire de filtrer les vitesses des traces qui

seront nécessaires pour la détermination des points de contrôle. Ceci permet d'éliminer l'effet escalier dû à la discrétisation de la tablette graphique.

Normalisation Une fois la trace acquise, on normalise sa taille, en conservant sa proportion hauteur - largeur. La trace est normalisée dans un carré de 40 pixels par 40 pixels. La figure 3.3 présente un "e" et "l" avant et après normalisation.

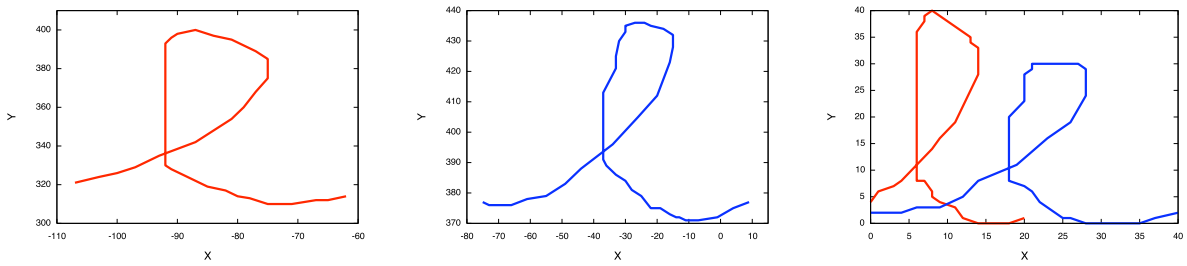


FIG. 3.3 – A gauche et au centre : une trace représentant la lettre "l" et la lettre "e" avant normalisation. A droite : les mêmes lettres après normalisation.

Filtre La figure 3.4 présente la vitesse en x et en y pour une lettre m.

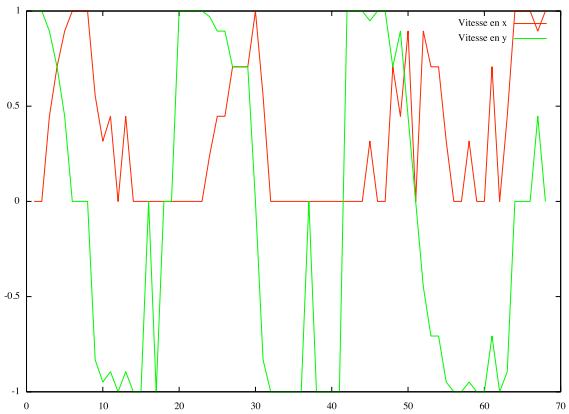


FIG. 3.4 – En rouge, la vitesse en x et en vert, la vitesse en y pour une lettre m.

Nous observons que la vitesse fluctue énormément ce qui est dû à notre mode d'acquisition des données. En effet, la trace produite par le sujet est fortement discrétisée par la tablette graphique ce qui résulte en un bruit de haute fréquence. Pour réduire ce bruit, nous allons filtrer les vitesses par un filtre passe-bas. Nous déduirons de la trace filtrée les différents points de contrôle.

Nous utilisons un filtre à réponse impulsionnelle finie de taille $2N + 1$ défini par :

$$f(t) = \sum_{k=-N}^{N} h_N(k) g(t+k), \tag{3.20}$$

avec $h_N(k)$ les coefficients du filtre et g le signal d'entrée. Nous choisirons par la suite un filtre binomial dont les coefficients $h_N(k)$ correspondent aux coefficients binomiaux (générés par le

triangle de Pascal) :

$$h_N(k) = \frac{C_{2N}^{k+N}}{2^{2N}} \qquad \text{avec} \quad k \in [-N; N]. \tag{3.21}$$

Il s'agit donc d'un filtre passe-bas qui a pour but de lisser la courbe et de limiter les brusques variations de vitesse.

Dans notre cas, nous utilisons un filtre de taille 13 ($N = 6$) dont les coefficients sont :

$$\frac{1}{4096} \{1, 12, 66, 220, 495, 792, 924, 792, 495, 220, 66, 12, 1\}.$$

La figure 3.5 présente la vitesse en x et en y avant filtrage (à gauche) et après filtrage (à droite) pour une lettre m.

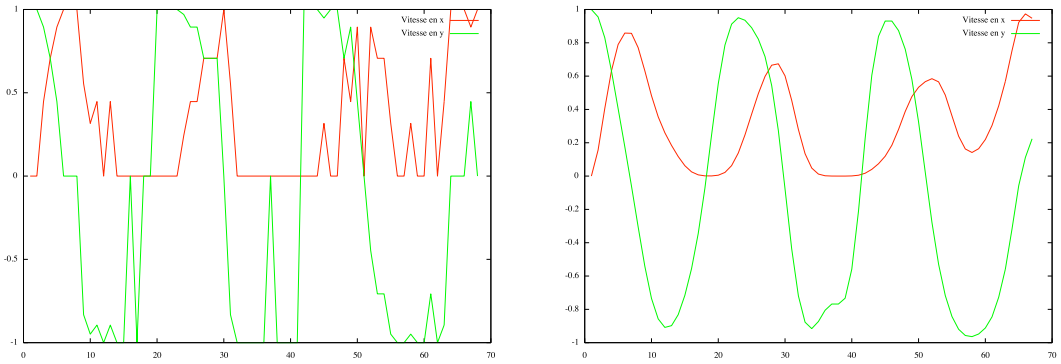


FIG. 3.5 – En rouge, la vitesse en x et en vert, la vitesse en y avant filtrage (à gauche) et après filtrage (à droite).

Lettre uniforme Nous avons défini une lettre supplémentaire, notée $\mathscr{U}$, dans notre répertoire correspondant à une lettre dont on ne sait rien. Nous avons donc défini chacun des termes $P(C_L^n \mid C_L^{n-1}\ [L = \mathscr{U}]\ W)$ comme des distributions de probabilité uniforme. Cette nouvelle lettre de notre répertoire nous permet de classer les symboles inconnus dans cette catégorie lors de la reconnaissance des caractères. La variable L peut donc prendre 23 valeurs possibles : a, b, c, d, e, f, g, h, k, l, m, n, o, p, q, r, s, u, v, w, y, z,$\mathscr{U}$.

Analyse des distributions de probabilité Dans ce paragraphe, nous analysons les distributions de probabilité une fois les paramètres des lois de successions de Laplace identifiés.

La figure 3.6 présente la distribution de probabilité de la position en x du troisième point de contrôle de la lettre l pour le scripteur 2 en fonction des différentes positions en x à l'indice précédent. On observe que la position de x à l'indice 3 dépend de la position de x à l'indice 2. La figure 3.2 présentée précédemment est un extrait de cette figure. On la retrouve lorsque la position en x à l'indice $i - 1$ vaut 15. On observe sur la figure 3.6 une probabilité forte juste au dessus de la diagonale ce qui implique que quelque soit la position en x à l'indice 2, il y a une forte probabilité que la position en x de l'indice 3 soit légèrement inférieure : le point de contrôle à l'indice 3 est plus à gauche que celui à l'indice 2. De plus, nous observons sur cette figure un autre mode qui n'est pas sur la diagonale : lorsque la valeur de x à l'indice

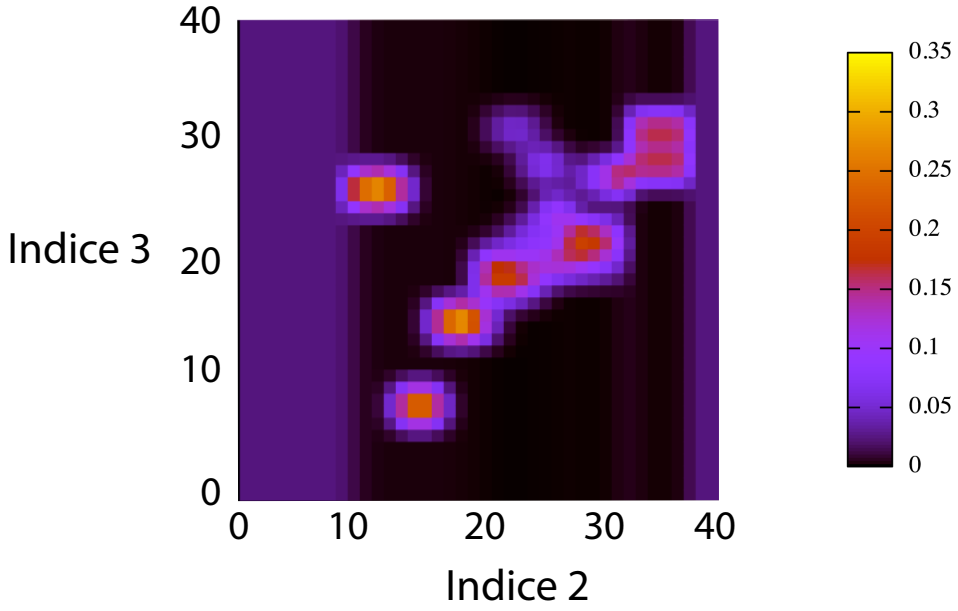


FIG. 3.6 – Distribution de probabilité de la position en x du troisième point de contrôle de la lettre l pour le scripteur *Julienne* : $P(C^3_{Lx} \mid C^2_{Lx}\ [L = l]\ [W = Julienne])$ en fonction des différentes position en x à l'indice précédent.

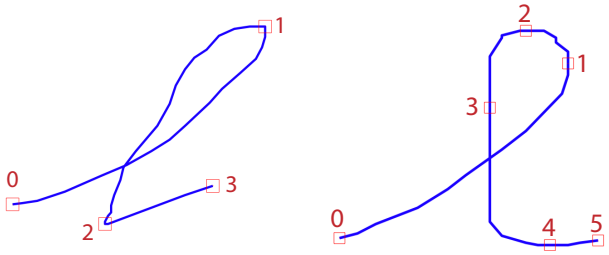


FIG. 3.7 – Deux des allographes de la lettre l possibles.

2 vaut 12, il y a une forte probabilité que la valeur de x à l'indice 3 soit égale à 26. Ceci vient du fait que l'on n'a pas un unique allographe pour chaque lettre. En effet, il peut y avoir des insertions ou suppressions de points de contrôle entre les différents allographes. Les distributions de probabilité reflètent cette variété d'allographes à travers les différents modes. Nous avons vu figure 3.2 à droite, un allographe possible pour la lettre l. La figure 3.7 présente un autre allographe qui est à l'origine du mode qui n'est pas placé sur la diagonale.

Un exemple plus complet La figure 3.8 présente les distributions de probabilité de la position en y des points de contrôle à l'indice 2, 3, 4 et 5 en fonction des points de contrôle précédents. On observe que ces distributions de probabilité reflètent deux allographes possibles de la lettre l.

Pour chacune des distributions, nous avons entouré en rouge les modes représentant les points de contrôle de l'allographe de gauche et en vert ceux représentant l'allographe de droite.

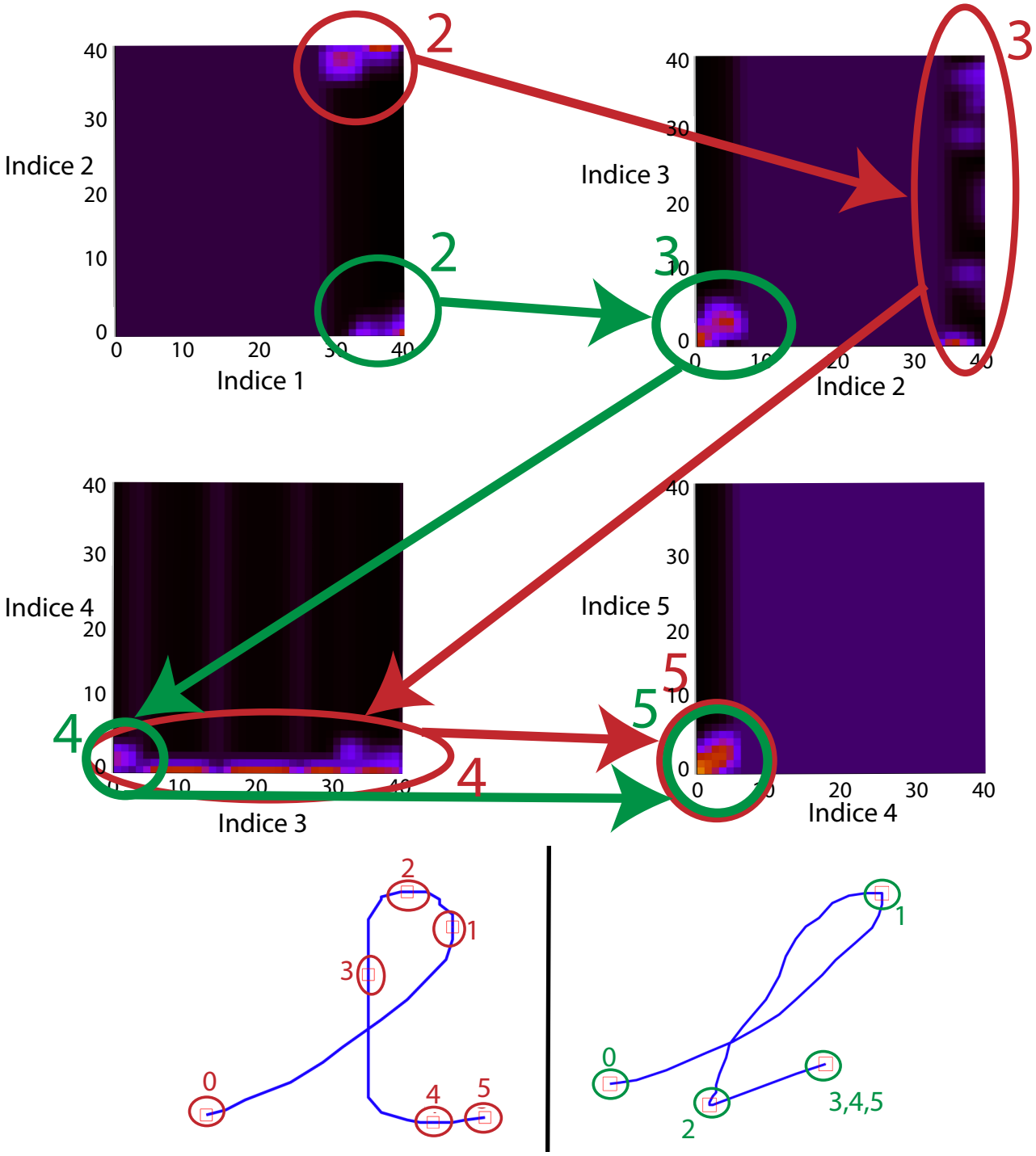


FIG. 3.8 – En haut, distributions de probabilité de la position en y des points de contrôle à l'indice 2, 3, 4 et 5 en fonction des points de contrôle précédents pour la lettre L et le scripteur *Julienne*. En bas, les deux allographes de la lettre l qui ont été encodés par les distributions de probabilité.

Les quatre distributions de probabilité représentent :

- en haut à gauche : $P(C^2_{Ly} \mid C^1_{Ly}\ [L = l]\ [W = Julienne])$. On observe deux modes : celui entouré de rouge est centré en $(38, 0)$ et celui entouré de vert est centré en $(38, 35)$. Ces deux modes correspondent aux distributions de probabilité des deux points d'indice 2 des lettres l comme on peut le voir sur le bas de la figure. En rouge, le point de contrôle est placé en haut de la trace et en vert il est situé en bas.
- en haut à droite : $P(C^3_{Ly} \mid C^2_{Ly}\ [L = l]\ [W = Julienne])$. On observe un mode centré en $(2, 2)$, (entouré en vert). Il correspond à la position du dernier point de contrôle de l'allographe situé en bas à droite. La zone entourée en rouge correspond à la distribution de probabilité de la position en y du point de contrôle de la partie descendante du l. Suivant les traces, ce point est placé plus ou moins haut. Nous n'observons donc pas de pic de probabilité mais une distribution de probabilité quasiment uniforme.
- en bas à gauche : $P(C^4_{Ly} \mid C^3_{Ly}\ [L = l]\ [W = Julienne])$. On observe que quelque soit la valeur de la position à l'indice précédent, la distribution de probabilité de la nouvelle position est centrée en 0. En d'autres termes, quelque soit la position du point de contrôle d'indice 3, le point de contrôle d'indice 4 est toujours situé en bas de la trace. Pour l'allographe vert, nous avons reproduit la distribution de probabilité à l'indice précédent.
- en bas à droite : $P(C^5_{Ly} \mid C^4_{Ly}\ [L = l]\ [W = Julienne])$. Nous observons un unique mode centré en $(2, 3)$. Quelque soit l'allographe, il s'agit de la position du dernier point de contrôle.

Nous observons qu'avec ce type de forme paramétrique (les lois de successions de Laplace), il est possible d'encoder implicitement plusieurs allographes pour une lettre donnée. Cela nous permettra par la suite de reconnaître une plus grande variété de lettres et de produire différents allographes.

3.2.3 Question

Jusqu'à présent, dans l'exemple introductif, nous avons décrit les variables probabilistes, la distribution de probabilité conjointe, sa décomposition, les formes paramétriques de chaque terme de la décomposition et les paramètres de ces formes paramétriques. Les phases de spécification et d'identification sont terminées, nous disposons d'une description complètement définie.

Durant la phase d'exploitation, nous utilisons cette description afin de répondre à une question. Poser une question consiste à calculer la distribution de probabilité d'un certain nombre de variables de la description connaissant les valeurs d'autres variables. Nous posons la question suivante :

Quelle est la lettre correspondant à la séquence de points de contrôle donnée ?

$$P(L \mid [C_L^{0:N} = c_l^{0:N}]\ [W = w]). \tag{3.22}$$

Cette question correspond à déterminer la distribution de probabilité de la lettre connaissant une séquence de points de contrôle donnée $(c_l^{0:N})$ et le scripteur (w).

3.2.3.1 Inférence

L'inférence bayésienne permet de répondre à n'importe quelle question à partir de la spécification de la distribution de probabilité conjointe. Tout d'abord, en utilisant la règle de

Bayes, nous obtenons :

$$P(L \mid [C_L^{0:N} = c_l^{0:N}]\ [W = w]) = \frac{P(L\ [C_L^{0:N} = c_l^{0:N}]\ [W = w])}{P([C_L^{0:N} = c_l^{0:N}]\ [W = w])}. \tag{3.23}$$

Nous voyons que le dénominateur est une constante, indépendante de la variable L. Nous renommons cette constante de normalisation Z_1 :

$$P(L \mid [C_L^{0:N} = c_l^{0:N}]\ [W = w]) = \frac{1}{Z_1} P(L\ [C_L^{0:N} = c_l^{0:N}]\ [W = w]). \tag{3.24}$$

Nous appliquons la décomposition de la distribution de probabilité conjointe :

$$P(L \mid [C_L^{0:N} = c_l^{0:N}]\ [W = w]) = \frac{1}{Z_1} P([C_L^{0:N} = c_l^{0:N}] \mid L\ [W = w])\ P(L)\ P([W = w]). \tag{3.25}$$

Le terme $P([W = w])$ suit une loi uniforme. Sa valeur numérique est donc toujours la même, quelque soit la valeur de w. Il peut donc être intégré dans la constante de normalisation. Il en est de même pour le terme $P(L)$ qui suit également une distribution de probabilité de loi uniforme. La constante de normalisation devient Z_2 :

$$P(L \mid [C_L^{0:N} = c_l^{0:N}]\ [W = w]) = \frac{1}{Z_2} P([C_L^{0:N} = c_l^{0:N}] \mid L\ [W = w]). \tag{3.26}$$

Poser cette question revient donc à comparer les points de contrôle avec les points de contrôle de chacune des lettres mémorisées. Cette comparaison est effectuée au niveau de chaque point de contrôle pour la position en x et y et pour les vitesses $\dot{x}$ et $\dot{y}$:

$$\begin{aligned} &P(L \mid [C_L^{0:N} = c_l^{0:N}]\ [W = w]) \\ &\propto \begin{pmatrix} P([C_{Lx}^0 = c_{lx}^0] \mid L\ [W = w]) \\ P([C_{Ly}^0 = c_{ly}^0] \mid L\ [W = w]) \\ P([C_{L\dot{x}}^0 = c_{l\dot{x}}^0] \mid L\ [W = w]) \\ P([C_{L\dot{y}}^0 = c_{l\dot{y}}^0] \mid L\ [W = w]) \end{pmatrix} \\ &\prod_{n=1}^{N} \begin{pmatrix} P([C_{Lx}^n = c_{lx}^n] \mid [C_{Lx}^{n-1} = c_{lx}^{n-1}]\ L\ [W = w]) \\ P([C_{Ly}^n = c_{ly}^n] \mid [C_{Ly}^{n-1} = c_{ly}^{n-1}]\ L\ [W = w]) \\ P([C_{L\dot{x}}^n = c_{l\dot{x}}^n] \mid [C_{L\dot{x}}^{n-1} = c_{l\dot{x}}^{n-1}]\ L\ [W = w]) \\ P([C_{L\dot{y}}^n = c_{l\dot{y}}^n] \mid [C_{L\dot{y}}^{n-1} = c_{l\dot{y}}^{n-1}]\ L\ [W = w]) \end{pmatrix}. \end{aligned} \tag{3.27}$$

Chacun de ces termes correspond à la distribution de probabilité de C_L sachant L et W. Or, dans notre cas, on connaît les valeurs de C_L et W et on veut obtenir une distribution de probabilité sur les valeurs de la variable L. La distribution de probabilité correspondant à la question est donc calculée par « inversion » de ces termes.

3.2.3.2 Résultat

Nous allons analyser le résultat pour la question suivante :

$$P(L \mid [C_L^{0:N} = c_l^{0:N}]\ [W = Estelle]), \tag{3.28}$$

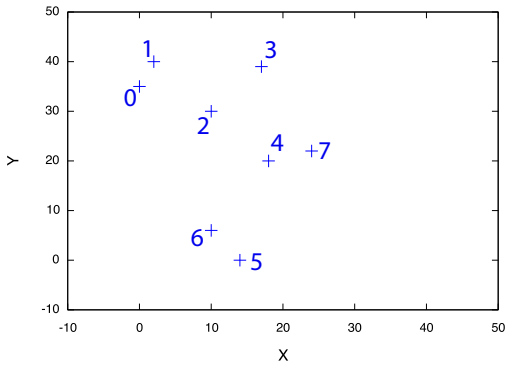


FIG. 3.9 – Ensemble des points de contrôle dont on cherche à connaître la lettre.

avec $c_l^{0:N}$ l'ensemble des points de contrôle du scripteur *Estelle* défini figure 3.9.

La distribution de probabilité résultante à cette question est la suivante :

	a	b	c	d	e	f	g	h	k	l	m	n	o	p	q	r	s	u	v	w	y	z	∅
Probabilité	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0

Les points de contrôle ont été reconnus avec certitude comme appartenant à une lettre y. Ce résultat est cohérent étant donné qu'ils avaient été extrait de la lettre présentée figure 3.10 (un « y »).

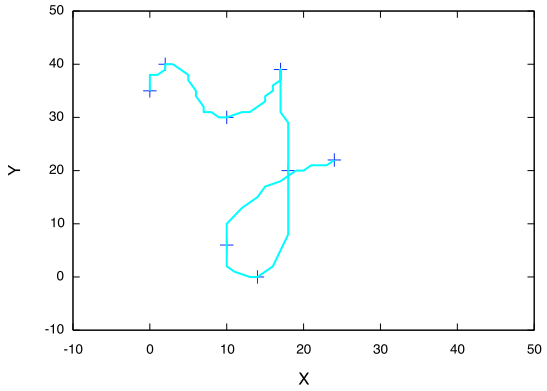


FIG. 3.10 – Trace dont on a extrait les points de contrôle.

La question $P(L \mid [C_L^{0:N} = c_l^{0:N}]\ [W = w])$ nous a permis d'illustrer le modèle de représentation interne des lettres. Nous avons décrit son inférence et présenté le résultat pour une séquence donnée de points de contrôle. Il est important de noter que le modèle permet, grâce au mécanisme d'inférence, de répondre à bien d'autres questions :

- $P(L \mid [C_L^{0:N} = c_l^{0:N}])$: quelle est la lettre correspondant à cette séquence de points de contrôle ne connaissant pas le scripteur qui l'a produite ?
- $P(W \mid [C_L^{0:N} = c_l^{0:N}]\ [L = l])$: qui est le scripteur à l'origine de cette séquence de points de contrôle correspondant à la lettre l ?
- $P(W \mid [C_L^{0:N} = c_l^{0:N}])$: qui est le scripteur à l'origine de cette séquence de points de contrôle ne sachant pas à quelle lettre elle correspond ?
- $P(C_L \mid [L = l]\ [W = w])$: quels sont les points de contrôle correspondant à la lettre l

pour le scripteur w ?
– etc.

Toutes ces questions serons présentées chapitre 5.

3.3 Résumé

Dans ce chapitre, nous avons tout d'abord présenté notre cadre de modélisation : la programmation bayésienne. Nous l'avons ensuite illustré au travers d'un exemple simplifié : le modèle de représentation des lettres. Nous avons défini les variables du modèle, sa distribution de probabilité conjointe ainsi que sa décomposition. Nous avons défini les formes paramétriques associées à chacun des termes et identifié les paramètres. Nous avons ensuite utilisé notre modèle en implantant une question. Cette question correspond à reconnaître la lettre à partir de la représentation interne et donc de la séquence de points de contrôle.

Chapitre 4

Modèle BAP

L'exemple précédent illustre l'application de la programmation bayésienne à la modélisation de la représentation interne des lettres. Nous proposons, dans ce chapitre, de définir le modèle complet de la boucle perception-action appliquée à la lecture et à l'écriture. Pour cela, nous utilisons la structure générale du schéma que nous avons défini chapitre 2 et que nous représentons figure 4.1.

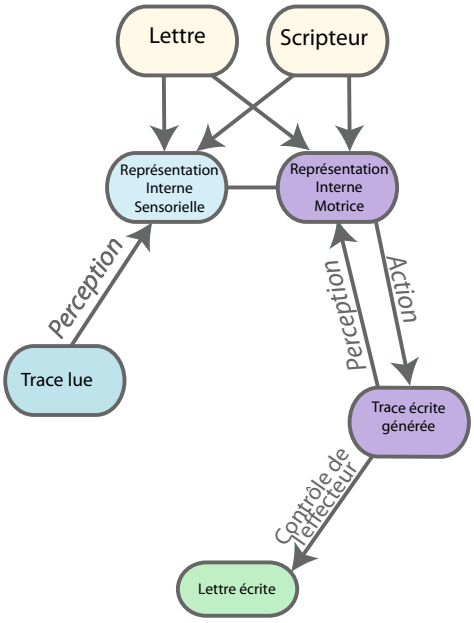


FIG. 4.1 – Structure générale du modèle BAP : la représentation interne des lettres est le pivot entre la perception et la production.

Nous observons que le modèle BAP est composé de sous-modèles (sous-modèles de perception, d'action, de représentation interne des lettres et de simulation motrice). La première section de ce chapitre est consacrée à la présentation d'un outil bayésien, qui permet d'activer ou de désactiver explicitement ces sous-modèles lors de l'inférence. Lors de la seconde section, nous fournirons l'équivalence mathématique du schéma présenté figure 4.1.

4.1 Activation des sous-modèles

4.1.1 Cohérence

Nous présentons dans ce paragraphe un modèle de relation entre variables : les modèles de cohérence. Nous le présentons tel qu'il a été décrit par Pradalier *et al.* (2003); Pradalier (2004) avant de donner un exemple dans notre modèle.

4.1.1.1 Définition : Variable de cohérence

Une variable de cohérence λ est une variable booléenne qui indique si la variable A est cohérente avec la variable B.

4.1.1.2 Définition : Modèle de cohérence

Un modèle de cohérence sur A et B est constitué d'une variable de cohérence λ et d'une distribution conjointe $P(\lambda\ A\ B)$ se décomposant en

$$P(\lambda\ A\ B)\ =\ P(A)\ P(B)\ P(\lambda \mid A\ B). \tag{4.1}$$

Les variables A et B apparaissent comme des variables indépendantes dans cette distribution. Cela signifie deux choses :

- D'une part, toute la connaissance sur la relation entre A et B est contenue dans la distribution $P(\lambda \mid A\ B)$.
- D'autre part, sans savoir si A et B sont cohérentes l'une avec l'autre (c'est-à-dire λ non spécifiée), il n'y a aucune raison a priori pour que ces variables dépendent l'une de l'autre.

4.1.2 Cas particulier 1 : activation ou inhibition des sous-modèles

Les variables de cohérence ont déjà été appliquées dans plusieurs cas. Par exemple, Koike (2005) les a utilisées pour la fusion de commandes, Le Hy (2007) pour effectuer de la fusion par cohérence (pour fusionner un nombre important de distributions de probabilité) et Pradalier (2004) les a utilisées pour traduire la cohérence entre les observations réelles et son modèle. Dans notre modèle, nous allons les utiliser dans un autre cas particulier.

Loi Dirac Nous définissons le terme $P(\lambda \mid A\ B)$ par une loi Dirac[1] :

$$P([\lambda = 1] \mid [A = a]\ [B = b]) = \begin{cases} 1 \text{ si } a = b \\ 0 \text{ sinon.} \end{cases} \tag{4.2}$$

Inférence Dans le cas où la variable de cohérence n'est pas spécifiée, il n'y a aucune dépendance entre les variables A et B. Il est donc important de noter que la distribution de probabilité $P([A = a])$ n'est pas égale à la distribution de probabilité $P([A = a] \mid [\lambda = 1])$. En utilisant l'inférence bayésienne et la définition du terme $P(\lambda \mid A\ B)$, nous allons décomposer ces deux termes et analyser les dépendances des termes :

[1] Loi Dirac : cette forme paramétrique possède un paramètre qui est une valeur particulière de la variable sur laquelle elle porte : pour cette valeur donnée, la probabilité est 1 ; elle est 0 pour toutes les autres valeurs.

- $P([A = a])$: c'est un terme de la décomposition. Il ne dépend donc pas de B.
- $P([A = a] \mid [\lambda = 1])$: nous appliquons tout d'abord la règle de Bayes :

$$P([A = a] \mid [\lambda = 1]) \tag{4.3}$$
$$= \frac{P([A = a]\ [\lambda = 1])}{P([\lambda = 1])}. \tag{4.4}$$

Puis, nous appliquons la règle de normalisation :

$$= \frac{\sum_B P([A = a]\ [\lambda = 1]\ B)}{P([\lambda = 1])}. \tag{4.5}$$

Nous utilisons la décomposition définie à l'équation 4.1

$$= \frac{\sum_B P([A = a])\ P(B)\ P([\lambda = 1] \mid [A = a]\ B)}{P([\lambda = 1])}. \tag{4.6}$$

Nous avons vu précédemment que $P([\lambda = 1] \mid [A = a]\ B)$ ne vaut 1 que lorsque la valeur de A est égale à la valeur de B et vaut 0 pour tous les autres cas. Par conséquent, il n'y a qu'un seul terme de la somme sur B non nul (pour $B = A$). Nous obtenons donc :

$$P([A = a] \mid [\lambda = 1]) \tag{4.7}$$
$$= \frac{P([A = a])\ P([B = a])\ P([\lambda = 1] \mid [A = a]\ [B = a])}{P([\lambda = 1])} \tag{4.8}$$
$$= \frac{P([A = a])\ P([B = a])}{P([\lambda = 1])} \tag{4.9}$$
$$\propto P([A = a])\ P([B = a]). \tag{4.10}$$

Nous observons que $P(A \mid [\lambda = 1])$ dépend de $P([A = a])$ et de $P([B = a])$.

Interprétation Lorsque l'on ne définit pas la variable λ, le terme $P([A = a])$ est indépendant de la valeur de la variable B. La partie du modèle contenant la variable B est donc inhibée. A l'inverse, lorsque nous définissons la variable λ à 1, nous forçons B à prendre la valeur de A (dû à la définition de la loi Dirac). Nous activons donc la partie du modèle contenant la variable B. C'est cette propriété des variables de cohérence que nous allons utiliser dans cette thèse pour activer ou désactiver des sous-modèles.

4.1.3 Cas particulier 2 : court-circuit

Dans cette section, nous allons illustrer un second cas particulier des variables de cohérence. Nous utilisons ces variables dans le but de court-circuiter un sous-modèle. Nous nous plaçons dans la configuration présentée figure 4.2. La décomposition de la distribution de probabilité conjointe correspondante est :

$$P(\lambda\ A\ B\ C) = P(A \mid C)\ P(B \mid C)\ P(\lambda \mid A\ B)\ P(C). \tag{4.11}$$

Loi Dirac Nous définissons, comme précédemment, le terme $P(\lambda \mid A\ B)$ par une loi Dirac :

$$P([\lambda = 1] \mid [A = a]\ [B = b]) = \begin{cases} 1 \text{ si } a = b \\ 0 \text{ sinon.} \end{cases} \tag{4.12}$$

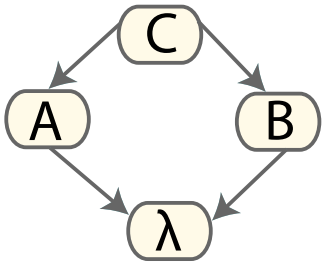


FIG. 4.2 – Modèle définissant la relation entre les variables A, B, C et λ.

Inférence Nous allons étudier les deux questions suivantes :
- $P(B \mid [\lambda = 1]\ [A = a])$,
- $P(B \mid [A = a])$.

Pour chacune de ces deux cas, nous allons détailler l'inférence et analyser la signification des questions.

- $P(B \mid [\lambda = 1]\ [A = a])$: nous appliquons tout d'abord la règle de Bayes :

$$\begin{aligned} & P(B \mid [\lambda = 1]\ [A = a]) \\ = \ & \frac{P(B\ [\lambda = 1]\ [A = a])}{P([\lambda = 1]\ [A = a])}. \end{aligned} \tag{4.13}$$

Puis, nous appliquons la règle de normalisation :

$$= \frac{\sum_C P(B\ [\lambda = 1]\ [A = a]\ C)}{\sum_{CB} P([\lambda = 1]\ [A = a])}. \tag{4.14}$$

Nous utilisons la décomposition définie équation 4.11 :

$$= \frac{\sum_C P([A = a] \mid C)\ P(B \mid C)\ P([\lambda = 1] \mid [A = a]\ B)\ P(C)}{\sum_{CB} P([A = a] \mid C)\ P(B \mid C)\ P([\lambda = 1] \mid [A = a]\ B)\ P(C)} \tag{4.15}$$

Nous obtenons alors :

$$\begin{aligned} & P([B = a] \mid [\lambda = 1]\ [A = a]) \\ = \ & \frac{\sum_C P([A = a] \mid C)\ P([B = a] \mid C)\ P(C)}{\sum_C P([A = a] \mid C)\ P([B = a] \mid C)\ P(C)} \\ = \ & 1 \end{aligned} \tag{4.16}$$

et

$$\begin{aligned} & P([B \neq a] \mid [\lambda = 1]\ [A = a]) \\ = \ & 0 \end{aligned} \tag{4.17}$$

Nous avons vu précédemment que $P([\lambda = 1] \mid [A = a]\ B)$ ne vaut 1 que lorsque la valeur de A est égale à la valeur de B et vaut 0 pour tous les autres cas. Par conséquent, il n'y a qu'une seule valeur pour B telle que la probabilité $P(B \mid [\lambda = 1]\ [A = a])$ soit non nulle : $B = a$. Nous observons donc que les termes $P(A \mid C)$ et $P(B \mid C)$ n'interviennent plus dans l'inférence : nous les avons donc court-circuités.

- $P(B \mid [A = a])$: nous appliquons tout d'abord la règle de Bayes :

$$\begin{aligned} & P(B \mid [A = a]) \\ = \quad & \frac{P(B\ [A = a])}{P([A = a])}. \end{aligned} \tag{4.18}$$

Puis, nous appliquons la règle de normalisation :

$$= \quad \frac{\sum_{C\lambda} P(B\ \lambda\ [A = a]\ C)}{P([A = a])}. \tag{4.19}$$

Nous utilisons la décomposition définie équation 4.11 :

$$\begin{aligned} = \quad & \frac{\sum_{C\lambda} P([A = a] \mid C)\ P(B \mid C)\ P(\lambda \mid [A = a]\ B)\ P(C)}{P([A = a])} \\ \propto \quad & \sum_{C} P([A = a] \mid C)\ P(B \mid C). \end{aligned} \tag{4.20}$$

Etant donné qu'il y a une sommation sur λ le terme $P(\lambda \mid [A = a]\ B)$ vaut 1. Il n'intervient donc plus dans l'inférence. Nous observons que la valeur de B dépend cette fois du produit des termes $P([A = a] \mid C)\ P(B \mid C)$.

Interprétation Lorsque nous définissons la variable λ à 1, le terme $P(B \mid A\ [\lambda = 1])$ est indépendant des modèles $P(A \mid C)$ et $P(B \mid C)$. Nous les avons donc court-circuités. A l'inverse, lorsque nous ne définissons pas λ dans $P(B \mid A)$ le modèle de C est activé.

4.2 Construction du modèle

Dans le chapitre 2, nous avons effectué différentes hypothèses sur l'architecture de notre modèle. Ces hypothèses nous ont permis de construire, pas à pas, l'ossature du modèle. Dans cette section, nous allons reprendre chaque étape de construction du modèle pour donner l'équivalence mathématique dans le formalisme Bayésien.

4.2.1 Représentation interne pivot

Nous avons dans un premier temps fait l'hypothèse que la représentation interne des lettres est le pivot entre la perception et la production. La figure 4.3 (à gauche) présente la structure du modèle.

4.2.1.1 Variables probabilistes

Les notations des variables correspondant à la *lettre* (L), au *scripteur* (W) et à la *représentation interne* $C_L^{0:N}$ ont été présentées section 3.2.1. Afin d'alléger les notations dans les formules que nous manipulerons dans ce chapitre, nous omettons l'exposant $^{0:N}$ des variables de représentation interne des lettres. La *trace lue* est notée V et la *trace produite* P. La figure 4.3 (à droite) présente le modèle avec les noms des variables définis ci-dessus.

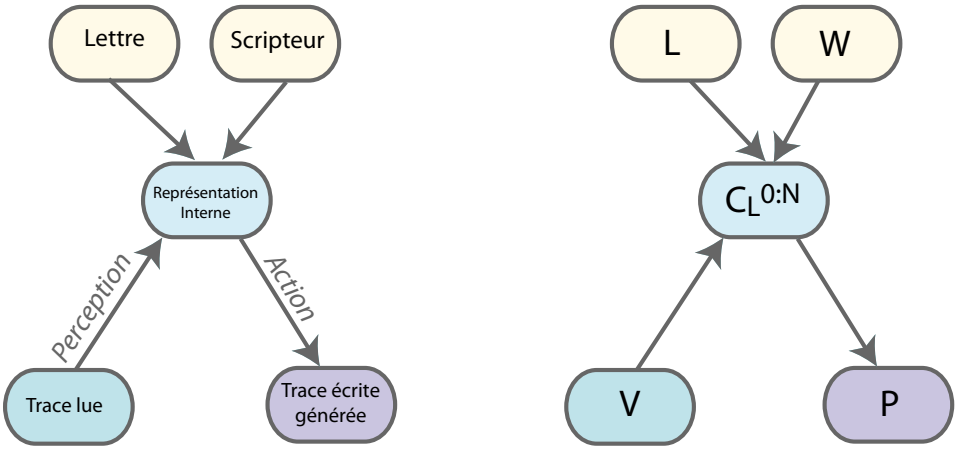


FIG. 4.3 – La représentation interne des lettres est le pivot entre la perception et la production.

4.2.1.2 Distribution conjointe et décomposition

La description des variables que nous venons d'effectuer a pour but la définition de la distribution conjointe :

$$P\left(\; C_L \;\; V \;\; P \;\; L \;\; W \; \right). \tag{4.21}$$

En appliquant plusieurs fois la règle de Bayes, nous obtenons :

$$\begin{aligned} & P\left(C_L \; V \; P \; L \; W\right) \\ & \quad = P(V \mid C_L \; L \; W \; P) P(C_L \; L \; W \; P) \\ & \quad = P(V \mid C_L \; L \; W \; P) P(P \mid C_L \; L \; W) P(C_L \; L \; W). \end{aligned} \tag{4.22}$$

Nous simplifions chacun des termes par des hypothèses d'indépendances conditionnelles. Cela permet de réduire les dimensions des termes sur lesquelles ces hypothèses portent :

- $P(C_L \; L \; W)$: ce terme a déjà été traité section 3.2.2 . Nous avions déterminé sa décomposition :

$$P(C_L \; L \; W) = P(C_L \mid L \; W) P(L) P(W). \tag{4.23}$$

- $P(P \mid C_L \; L \; W)$: nous faisons l'hypothèse que si l'on connaît la valeur de la représentation interne (C_L), la valeur de la trace écrite (P) est indépendante de la lettre (L) et du scripteur (W). Cette hypothèse d'indépendance se traduit mathématiquement par la simplification suivante :

$$P(P \mid C_L \; L \; W) = P(P \mid C_L). \tag{4.24}$$

- $P(V \mid C_L \; L \; W \; P)$: nous faisons l'hypothèse que la valeur de la trace lue (V) est indépendante de la lettre (L), du scripteur (W) et de la trace écrite (P) si l'on connaît la valeur de la représentation interne (C_L). Nous obtenons la simplification suivante :

$$P(V \mid C_L \; L \; W \; P) = P(V \mid C_L). \tag{4.25}$$

Etant données nos hypothèses d'indépendance, la décomposition de la distribution de probabilité conjointe s'écrit donc :

$$\begin{aligned} &P\left(C_L\ V\ P\ L\ W\right)\\ &\quad = P(V \mid C_L)P(P \mid C_L)P(C_L \mid L\ W)P(L)P(W). \end{aligned} \tag{4.26}$$

4.2.2 Représentations internes sensorielles et motrices

Nous avons fait l'hypothèse section 2.1.2 que les représentations internes sensorielles et motrices ne sont pas confondues. Nous avons donc deux objets distincts représentés figure 4.4 (à gauche). Il est important de noter que ces deux représentations sont interdépendantes et sont donc liées.

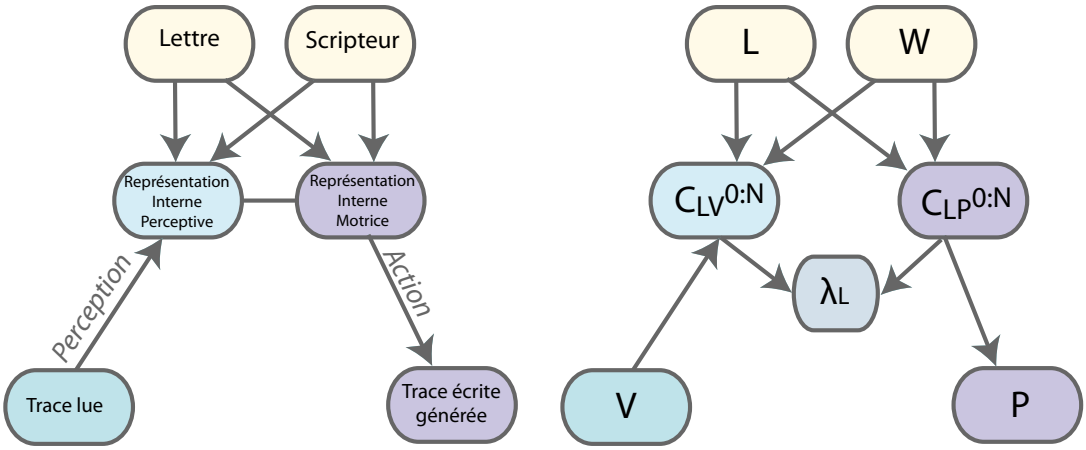


FIG. 4.4 – Les représentations internes sensorielles et motrices sont distinctes.

4.2.2.1 Variables probabilistes

Nous notons la représentation interne sensorielle C_{LV} et la représentation interne motrice C_{LP}. Leurs définitions sont identiques à celle de C_L. Le lien entre les variables de représentation interne sensorielle et motrice est défini par une variable de cohérence notée λ_L. Cette variable joue le rôle de court-circuit du modèle des lettres. Nous verrons son utilisation en détail section 7.1. La figure 4.4 (à droite) présente l'architecture du modèle avec la notation des variables définie ci-dessus.

4.2.2.2 Distribution conjointe et décomposition

La distribution conjointe est :

$$P\left(\ \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ \right). \tag{4.27}$$

En appliquant la règle de Bayes, nous obtenons la décomposition suivante :

$$\begin{aligned} &P\left(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\right)\\ &\quad = P\left(\lambda_L \mid C_{LV}\ C_{LP}\ V\ P\ L\ W\right) P\left(C_{LV}\ C_{LP}\ V\ P\ L\ W\right). \end{aligned} \tag{4.28}$$

Nous simplifions chacun des termes :

- $P(\lambda_L \mid C_{LV}\ C_{LP}\ V\ P\ L\ W)$: par construction, la valeur λ_L est indépendante de la lettre (L), du scripteur (W), de la trace lue (V) et de la trace écrite (P) si l'on connaît la valeur de la représentation interne sensorielle (C_{LV}) et la valeur de la représentation interne motrice (C_{LP}). Nous obtenons la simplification suivante :

$$P(\lambda_L \mid C_{LV}\ C_{LP}\ V\ P\ L\ W) = P(\lambda_L \mid C_{LV}\ C_{LP}). \tag{4.29}$$

- $P(C_{LV}\ C_{LP}\ V\ P\ L\ W)$: nous avons précédemment défini le terme $P(C_L\ V\ P\ L\ W)$. En remplaçant C_L par $C_{LV}\ C_{LP}$, nous obtenons la décomposition suivante :

$$\begin{aligned} &P(C_{LV}\ C_{LP}\ V\ P\ L\ W) \\ &\quad = P(V \mid C_{LV}\ C_{LP})P(P \mid C_{LP}\ C_{LV})P(C_{LV}\ C_{LP} \mid L\ W)P(L)P(W). \end{aligned} \tag{4.30}$$

  Nous faisons l'hypothèse que la trace lue ne dépend pas de la représentation interne motrice :

$$P(V \mid C_{LV}\ C_{LP}) = P(V \mid C_{LV}). \tag{4.31}$$

  et que la trace écrite ne dépend pas de la représentation interne sensorielle :

$$P(P \mid C_{LP}\ C_{LV}) = P(P \mid C_{LP}). \tag{4.32}$$

  Pour le terme $P(C_{LV}\ C_{LP}|\ L\ W)$, nous appliquons la règle de Bayes :

$$P(C_{LV}\ C_{LP}|\ L\ W) = P(C_{LV} \mid C_{LP}\ L\ W)P(C_{LP} \mid L\ W). \tag{4.33}$$

  et nous faisons l'hypothèse que la représentation interne sensorielle est indépendante de la représentation interne motrice si l'on connaît les valeurs de L et W. Nous obtenons en appliquant cette hypothèse d'indépendance conditionnelle :

$$P(C_{LV} \mid C_{LP}\ L\ W) = P(C_{LV} \mid L\ W). \tag{4.34}$$

Si l'on regroupe toutes ces simplifications, nous obtenons la décomposition de la distribution de probabilité conjointe suivante :

$$\begin{aligned} &P\left(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\right) \\ &\quad = \left(\begin{array}{l} P(\lambda_L \mid C_{LV}\ C_{LP})P(C_{LV} \mid L\ W)P(C_{LP} \mid L\ W) \\ P(L)P(W)P(V \mid C_{LV})P(P \mid C_{LP}) \end{array} \right). \end{aligned} \tag{4.35}$$

Comme nous l'avons vu précédemment, le terme $P(\lambda_L \mid C_{LV}\ C_{LP})$ est défini par une distribution de probabilité de loi Dirac :

$$P([\lambda_L = 1] \mid [C_{LV} = c_{lv}]\ [C_{LP} = c_{lp}]) = \left\{ \begin{array}{l} 1 \text{ si } c_{lv} = c_{lp} \\ 0 \text{ sinon.} \end{array} \right. \tag{4.36}$$

4.2.3 Modélisation de l'effecteur

4.2.3.1 Variables probabilistes

Nous notons E l'ensemble des variables correspondant à l'effecteur. La figure 4.5 présente le schéma global du modèle sur lequel on a représenté l'effecteur.

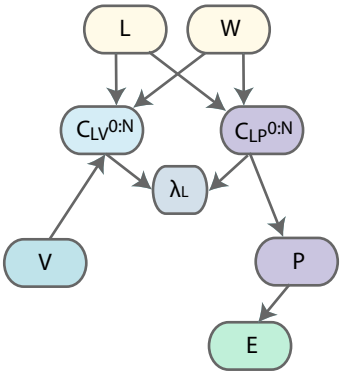


FIG. 4.5 – Structure globale du modèle.

4.2.3.2 Distribution conjointe et décomposition

Nous appliquons la règle de Bayes sur la distribution de probabilité conjointe :

$$\begin{aligned} &P(E\ \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W) \\ &\quad = P(E \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W)P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W). \end{aligned} \tag{4.37}$$

Le second terme peut être décomposé comme à l'équation 4.35. Pour le premier terme, nous faisons l'hypothèse que la valeur des variables de l'effecteur (E) ne dépend que de la valeur des variables de la trace écrite (P). En appliquant cette hypothèse d'indépendance conditionnelle, nous obtenons :

$$P(E \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W) = P(E \mid P). \tag{4.38}$$

La distribution de probabilité conjointe peut donc être décomposée de la façon suivante :

$$\begin{aligned} &P\,(E\ \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W) \\ &\quad = \left(\begin{array}{l} P(\lambda_L \mid C_{LV}\ C_{LP})P(C_{LV} \mid L\ W)P(C_{LP} \mid L\ W) \\ P(L)P(W)P(V \mid C_{LV})P(P \mid C_{LP})P(E \mid P) \end{array} \right). \end{aligned} \tag{4.39}$$

4.2.4 Activation des sous-modèles

Au début de ce chapitre, nous avons analysé un outil mathématique permettant d'activer ou de désactiver des sous-modèles : les variables de cohérence. Afin de pouvoir inhiber ou activer certaines parties du modèle, nous avons rajouté des variables de cohérence. Pour cela, nous avons dédoublé les variables représentation interne et ajouté des variables de cohérence pour les lier. C'est ce que présente la figure 4.6 (à gauche) avec les deux variables de cohérence λ_V et λ_P.

4.2.4.1 Variables probabilistes

Nous notons C_V et C_P les représentations internes sensorielles et motrices dupliquées. Leurs définitions sont identiques à celles de C_{LV} et C_{LP}. La figure 4.6 (à droite) présente l'architecture du modèle avec la notation des variables définie ci-dessus.

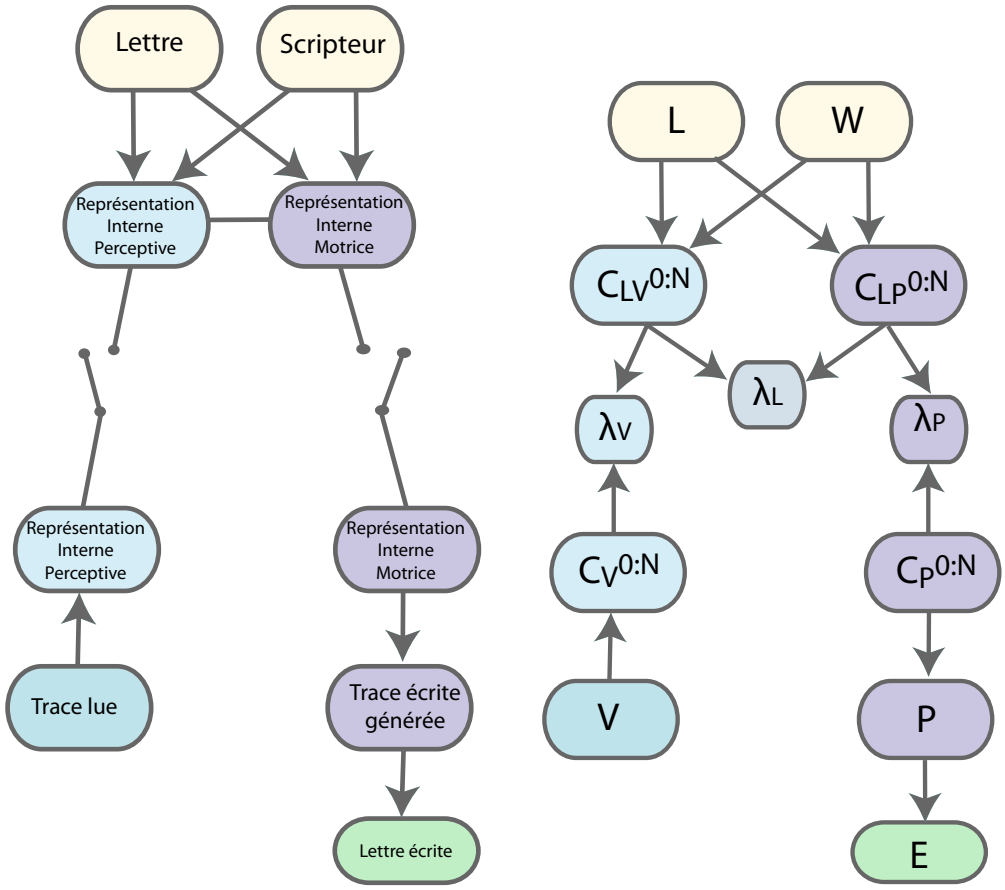


FIG. 4.6 – Activation ou inhibition des sous-modèles.

4.2.4.2 Distribution conjointe et décomposition

La distribution conjointe est :

$$P\left(\begin{array}{cccccccccccc}\lambda_L & C_{LV} & C_{LP} & V & P & L & W & \lambda_V & \lambda_P & C_V & C_P & E\end{array}\right). \tag{4.40}$$

En appliquant la règle de Bayes, nous obtenons la décomposition suivante :

$$\begin{aligned} & P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P) \\ & \quad = \left(\begin{array}{l} P(\lambda_V \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_P) \\ P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_P) \end{array}\right) \\ & \quad = \left(\begin{array}{l} P(\lambda_V \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_P) \\ P(\lambda_P \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P) \\ P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P) \end{array}\right). \end{aligned} \tag{4.41}$$

Nous simplifions chacun des termes :

- $P(\lambda_V \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_P)$: par construction, la valeur λ_V est indépendante de la lettre (L), du scripteur (W), de la trace lue (V), de la trace écrite (P), de l'effecteur (E), des représentations internes motrices (C_P et C_{LP}) et des variables

de cohérence λ_P, λ_L si l'on connaît les représentations internes sensorielles C_V et C_{LV}. Nous obtenons la simplification suivante :

$$P(\lambda_V \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_P) = P(\lambda_V \mid C_{LV}\ C_V). \tag{4.42}$$

- $P(\lambda_P \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P)$: de la même façon, nous pouvons simplifier ce terme :

$$P(\lambda_P \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P) = P(\lambda_P \mid C_{LP}\ C_P). \tag{4.43}$$

- $P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P)$: ce terme peut être décomposé de la façon suivante (la décomposition est similaire à celle présentée équation 4.44 : les variables C_{LV} et C_{LP} ont été dupliquées) :

$$\begin{aligned} &P\,(E\ \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ C_V\ C_P) \\ &\quad = \left(\begin{array}{l} P(\lambda_L \mid C_{LV}\ C_{LP})P(C_{LV} \mid L\ W)P(C_{LP} \mid L\ W) \\ P(L)P(W)P(C_V \mid V)P(P \mid C_P)P(E \mid P)P(V)P(C_P) \end{array} \right). \end{aligned} \tag{4.44}$$

Si l'on regroupe ces simplifications, nous obtenons la décomposition de la distribution de probabilité conjointe suivante :

$$\begin{aligned} &P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P) \\ &\quad = \left(\begin{array}{l} P(\lambda_V \mid C_{LV}\ C_V)\ P(\lambda_P \mid C_{LV}\ C_V) \\ P(\lambda_L \mid C_{LV}\ C_{LP})P(C_{LV} \mid L\ W)P(C_{LP} \mid L\ W) \\ P(L)P(W)P(C_V \mid V)P(P \mid C_P)P(E \mid P)P(V)P(C_P) \end{array} \right). \end{aligned} \tag{4.45}$$

Comme nous l'avons vu précédemment, les termes $P(\lambda_V \mid C_{LV}\ C_V)$ et $P(\lambda_P \mid C_{LP}\ C_P)$ sont définis par des distributions de probabilité de loi Dirac :

$$P([\lambda_V = 1] \mid [C_{LV} = c_{lv}]\ [C_V = c_v]) = \left\{ \begin{array}{l} 1 \text{ si } c_{lv} = c_v \\ 0 \text{ sinon.} \end{array} \right. \tag{4.46}$$

$$P([\lambda_P = 1] \mid [C_{LP} = c_{lp}]\ [C_P = c_p]) = \left\{ \begin{array}{l} 1 \text{ si } c_{lp} = c_p \\ 0 \text{ sinon.} \end{array} \right. \tag{4.47}$$

4.2.5 Simulation motrice

Nous avons fait l'hypothèse section 2.4.2 que pour mieux reconnaître la lettre l'action d'écriture était simulée. Cette hypothèse se traduit schématiquement par une boucle de *perception simulée* au niveau de l'action. Ce schéma est présenté figure 4.7.

Nous observons que nous avons une boucle entre la représentation interne motrice et la trace écrite générée. Or, pour des contraintes algébriques sur la décomposition de la conjointe, le modèle ne doit pas contenir de boucle. Afin de répondre à ce critère, nous avons modifié le graphe 4.7 en dédoublant la variable *représentation interne motrice*. Afin de garder une symétrie à notre modèle nous avons aussi dédoublé la variable *lettre écrite générée* et la variable de cohérence λ_P qui devient λ_S pour obtenir le nouveau graphe 4.8.

4.2.5.1 Variables probabilistes

Nous notons C_{LS} et C_S les représentation internes simulées. Leurs définitions sont identiques aux autres variables de représentations internes (C_{LV}, C_{LP}, C_V, C_P). La trace écrite simulée est notée S. La figure 4.9 présente le schéma global du modèle en respectant la notation définie ci-dessus.

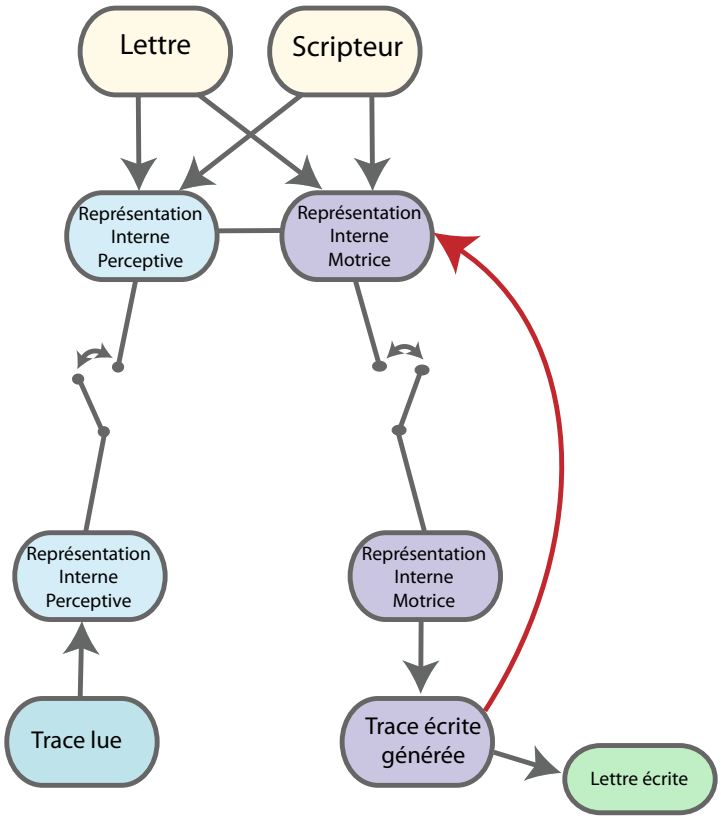


FIG. 4.7 – Simulation motrice de l'écriture

4.2.5.2 Distribution conjointe et décomposition

La distribution de probabilité conjointe est :

$$P\Big(\ C_{LV}\ \ C_{LP}\ \ C_{LS}\ \ C_V\ \ \lambda_p\ \ \lambda_v\ \ \lambda_s\ \ \lambda_l\ \ E\ \ C_P\ \ C_S\ \ V\ \ P\ \ L\ \ W\ \ S\ \Big). \quad (4.48)$$

Nous appliquons la règle de Bayes sur la distribution de probabilité conjointe :

$$\begin{aligned}
&P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ \lambda_S\ C_S\ C_{LS}\ S)\\
&\quad= \begin{pmatrix} P(C_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ \lambda_S\ S\ C_{LS})\\ P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ \lambda_S\ S\ C_{LS}) \end{pmatrix}\\
&\quad= \begin{pmatrix} P(C_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ \lambda_S\ S\ C_{LS})\\ P(\lambda_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ C_{LS})\\ P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ C_{LS}) \end{pmatrix}\\
&\quad= \begin{pmatrix} P(C_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ \lambda_S\ S\ C_{LS})\\ P(\lambda_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ C_{LS})\\ P(C_{LS} \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S)\\ P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ ) \end{pmatrix}\\
&\quad= \begin{pmatrix} P(\lambda_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ C_{LS}\ C_S)\\ P(C_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ C_{LS})\\ P(C_{LS} \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S)\\ P(S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P)\\ P(\lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P) \end{pmatrix}. \quad (4.49)
\end{aligned}$$

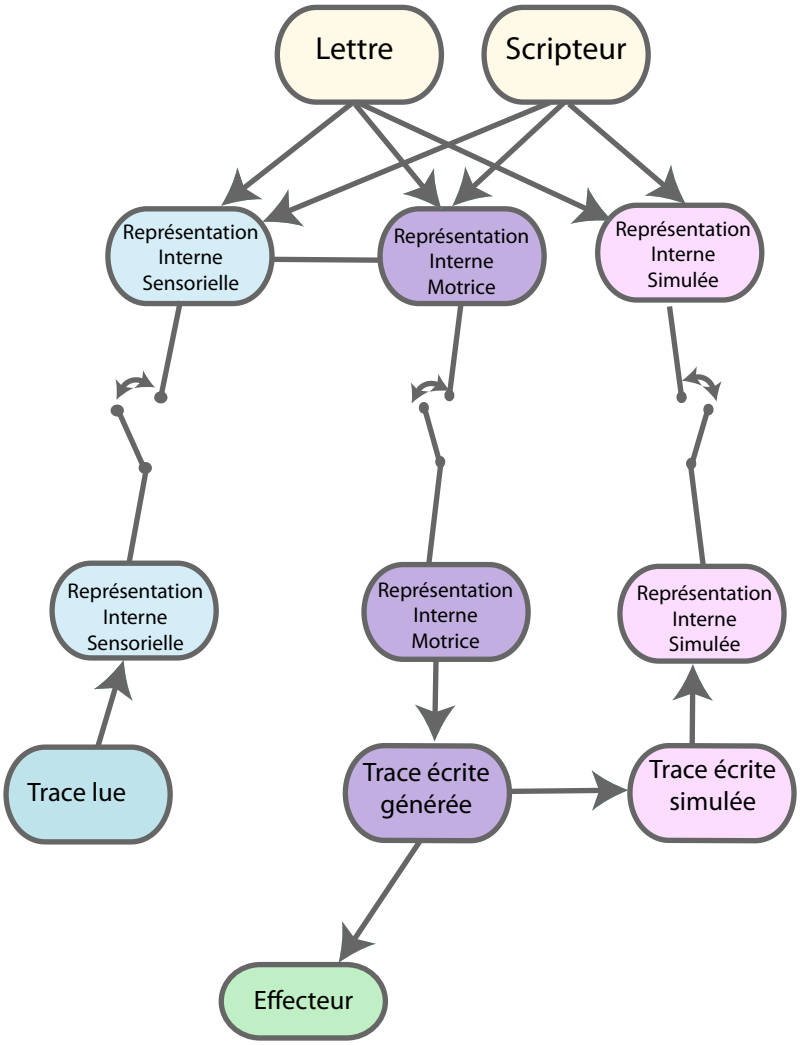


FIG. 4.8 – Simulation motrice de l'écriture : dédoublement de la boucle

Le dernier terme peut être décomposé comme à l'équation 4.45. Pour les autres termes, nous utilisons des hypothèses d'indépendances conditionnelles pour les simplifier :

- $P(\lambda_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ C_{LS}\ C_S)$: par construction, la variable de cohérence ne dépend que de C_{LS} et de C_S :

$$P(\lambda_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ C_{LS}\ C_S) = P(\lambda_S \mid C_{LS}\ C_S). \quad (4.50)$$

- $P(C_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ \ S\ C_{LS})$: nous faisons l'hypothèse que la représentation interne sensorielle C_S, si l'on connaît la valeur de la trace écrite simulée (S), est indépendante de toutes les autres variables. En appliquant cette hypothèse d'indépendance conditionnelle, nous obtenons la simplification suivante :

$$P(C_S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P\ S\ C_{LS}) = P(C_S \mid S). \quad (4.51)$$

- $P(S \mid \lambda_L\ C_{LV}\ C_{LP}\ V\ P\ L\ W\ E\ C_V\ C_P\ \lambda_V\ \lambda_P)$: nous faisons l'hypothèse que la valeur de la trace écrite simulée ne dépend que de la valeur de la trace écrite générée. Nous

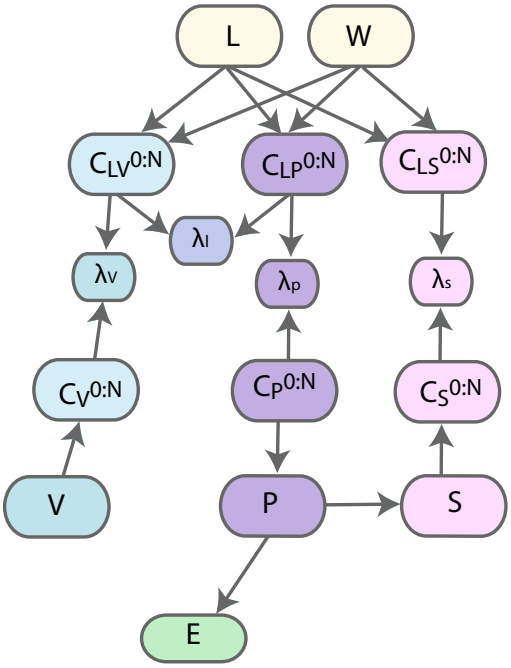


FIG. 4.9 – Notation des variables dans le schéma global du modèle

obtenons :

$$P(S \mid P). \tag{4.52}$$

En appliquant toutes ces simplifications, nous obtenons la décomposition de la distribution de probabilité conjointe suivante :

$$P\left(C_{LV}\ C_{LP}\ C_{LS}\ C_V\ \lambda_p\ \lambda_v\ \lambda_s\ \lambda_l\ E\ C_P\ C_S\ V\ P\ L\ W\ S \right)$$

$$= \begin{pmatrix} P(C_{LV} \mid L\ W)\ P(C_{LP} \mid L\ W)\ P(C_{LS} \mid L\ W) \\ P(L)\ P(W) \\ P(\lambda_V \mid C_{LV}\ C_V)\ P(\lambda_P \mid C_{LP}\ C_P)\ P(\lambda_S \mid C_{LS}\ C_S) \\ P(C_V \mid V)\ P(P \mid C_P)\ P(C_S \mid S) \\ P(\lambda_l \mid C_{LV}\ C_{LP})\ P(S \mid P)\ P(V) \\ P(E \mid P) \end{pmatrix}. \tag{4.53}$$

Nous rappelons la spécification de chacune des variables de la décomposition :

- L : lettre,
- W : scripteur,
- C_{LV} et C_V : représentations internes sensorielles,
- C_{LP} et C_P : représentations internes motrices,
- C_{LS} et C_S : représentations internes simulées,
- λ_V, λ_P, λ_S, λ_L : variables de cohérence,

- V : trace perçue,
- P : trace écrite générée (la trace n'est pas écrite mais seulement planifiée),
- S : trace écrite simulée,
- E : effecteur.

4.3 Résumé

Dans ce chapitre nous avons montré le schéma global de notre modèle de la lecture et de l'écriture. Celui-ci s'appuie sur une liste d'hypothèses issues de la littérature et les implante au sein du cadre théorique de la programmation bayésienne. En particulier, dans ce chapitre, nous avons mis l'accent sur l'importance de l'expression des hypothèses de notre modèle pour la décomposition de la distribution de probabilité conjointe.

Dans les chapitres suivants, nous allons montrer comment se comporte le modèle dans différentes situations qui peuvent se présenter dans la boucle sensori-motrice de la lecture et de l'écriture.

Chapitre 5

Perception

Ce chapitre concerne le sous-modèle relatif à la perception. Par perception, nous entendons le processus de recueil et de traitement de l'information sensorielle. Dans notre cas, l'information sensorielle provient uniquement de la vue. Dans ce travail, le modèle de vision est associé à un modèle identité : la trace résultante du modèle de vision est identique à la trace vue par la rétine. En effet, nous faisons l'hypothèse simplificatrice que l'espace de la rétine est l'espace cartésien. Le modèle de vision correspond donc à une projection de l'espace de la feuille (espace cartésien) dans l'espace cartésien. Le rôle du modèle de perception consiste alors à extraire les points de contrôle de cette trace.

La première partie de ce chapitre est consacrée à la description des variables et relations impliquées dans ce sous-modèle. La seconde partie de ce chapitre correspond aux premiers résultats liés à la tâche de lecture : nous verrons comment grâce aux seuls modèles de perception et de représentation interne des lettres, le modèle permet de résoudre la tâche de lecture.

5.1 Modèle de perception

La figure 5.1 présente les variables et relations impliquées dans le modèle de perception.

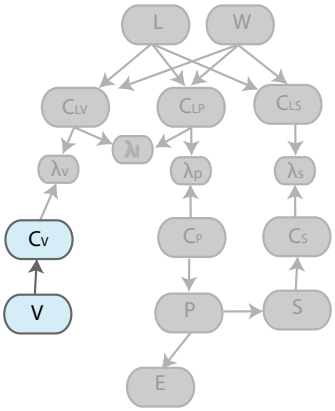


FIG. 5.1 – Architecture du modèle où l'on a grisé les variables non impliquées dans le sous-modèle de perception.

5.1.1 Variables

La variable correspondant à la trace perçue, V, est définie comme un ensemble de variables de positions $\{V_x^{0:M}, V_y^{0:M}\}$. M correspond au nombre de points maximal dans la trace. Dans notre cas, il varie entre 20 (pour une trace représentant la lettre c) et 97 (pour une trace représentant la lettre m). Les variables V_x^m et V_y^m correspondent à la position en x et y (dans un repère cartésien) du $m^{\text{ème}}$ point de la trajectoire. Ce sont des variables entières à valeur dans $[0, 40] \cap \mathbb{N}$.

5.1.2 Termes de la décomposition

Les termes de la décomposition en relation avec la perception des traces sont :

$$P(\ V\ )\ \text{ et }\ P(\ C_V\ \mid\ V\ ). \tag{5.1}$$

Le terme $P(V)$ correspond à la distribution de probabilité de la trace perçue. Il est associé une loi uniforme, car nous n'avons pas d'a priori sur la répartition des positions des points de la trace. Le terme de la décomposition correspondant à l'extraction des points de contrôle de la trace est :

$$P(C_V \mid V). \tag{5.2}$$

Dans le chapitre précédent, nous avons vu que C_V correspondait à l'ensemble des quatre variables $\{C_{Vx}^{0:N}, C_{Vy}^{0:N}, C_{V\dot{x}}^{0:N}, C_{V\dot{y}}^{0:N}\}$. Le terme de la décomposition correspondant à la perception des traces est donc :

$$P(C_{Vx}^{0:N}\ C_{Vy}^{0:N}\ C_{V\dot{x}}^{0:N}\ C_{V\dot{y}}^{0:N} \mid V_x^{0:M}\ V_y^{0:M}). \tag{5.3}$$

5.1.3 Formes paramétriques

Le terme $P(C_{Vx}^{0:N}\ C_{Vy}^{0:N}\ C_{V\dot{x}}^{0:N}\ C_{V\dot{y}}^{0:N} \mid V_x^{0:M}\ V_y^{0:M})$ est associé à une loi Dirac qui vaut 1 pour l'ensemble de valeur $\{c_{vx}^{0:N}, c_{vy}^{0:N}, c_{v\dot{x}}^{0:N}, c_{v\dot{y}}^{0:N}\}$ donnée par une fonction déterministe f qui prend en entrée l'ensemble des points d'une trace.

$$P\left(\begin{array}{cc|c} [C_{Vx}^{0:N} = c_{vx}^{0:N}] & [C_{Vy}^{0:N} = c_{vy}^{0:N}] & [V_x^{0:M} = v_x^{0:M}] \\ [C_{V\dot{x}}^{0:N} = c_{v\dot{x}}^{0:N}] & [C_{V\dot{y}}^{0:N} = c_{v\dot{y}}^{0:N}] & [V_y^{0:M} = v_y^{0:M}] \end{array}\right) = \left\{\begin{array}{l} 1 \text{ si } \left\{\begin{array}{c} c_{vx}^{0:N}, c_{vy}^{0:N} \\ c_{v\dot{x}}^{0:N}, c_{v\dot{y}}^{0:N} \end{array}\right\} = f(v_x^{0:M}, v_y^{0:M}) \\ 0 \text{ sinon.} \end{array}\right. \tag{5.4}$$

Cette fonction f est définie par un algorithme d'extraction des points de contrôle basé sur la définition des points de contrôle que nous avons effectué section 2.3.3. En d'autres termes, f identifie dans une trace les annulations des vitesses de V_x et V_y, après lissage de la trace.

5.2 Lecture

Etant donné que nous avons défini dans le chapitre précédent la représentation interne des lettres (en vert sur la figure 5.2) et dans ce chapitre le modèle de perception (en rose sur la figure 5.2), nous pouvons déjà analyser les résultats pour la tâche de lecture. Nous rappelons que par le terme « lecture », nous entendons reconnaissance de caractères cursifs isolés.

Dans la suite de ce chapitre, nous allons traiter les tâches suivantes :

– lecture en supposant le scripteur connu,

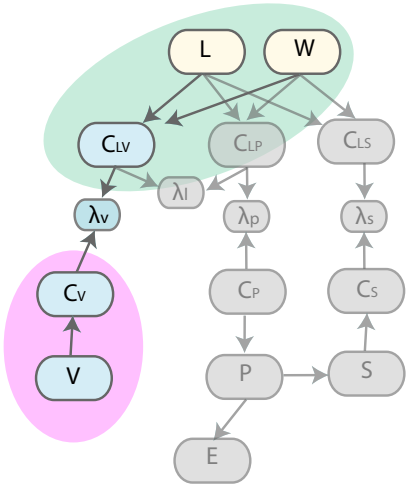


FIG. 5.2 – Architecture du modèle où l'on a grisé les variables non impliquées dans la tâche de lecture. En vert, le modèle de représentation interne des lettres et en rose le modèle de perception.

- lecture sans connaître le scripteur à l'origine des traces,
- reconnaissance du scripteur à partir d'une trace,
- reconnaissance du scripteur à partir d'une trace dont on connaît la lettre.

5.2.1 Question et inférence

La question relative à la lecture est la suivante :

Quelle est la lettre correspondant à la trace que je perçois et qui a été écrite par un scripteur que je connais ?

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]). \tag{5.5}$$

En d'autres termes cette question correspond à : quelle est la lettre correspondant à la trace $\{v_x^{0:M}, v_y^{0:M}\}$ écrite par le scripteur w ? L'inférence bayésienne permet de répondre à cette question à partir de la décomposition présentée équation 4.53. Tout d'abord, en utilisant la règle de Bayes, nous obtenons :

$$\begin{aligned} &P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ &= \frac{P(L\ [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1])}{P([V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1])}. \end{aligned} \tag{5.6}$$

Nous voyons que le dénominateur est une constante, indépendante de la variable L. Nous renommons cette constante de normalisation Z_1 :

$$\begin{aligned} &P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ &= \frac{1}{Z_1} P(L\ [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]). \end{aligned} \tag{5.7}$$

Puis, nous appliquons la règle de normalisation au numérateur :

$$\begin{aligned} & P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ & = \frac{1}{Z_1} \sum_A P \left(\begin{array}{l} L\ [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]\ C_{LV}\ C_V \\ C_{LP}\ C_{LS}\ \lambda_p\ \lambda_s\ \lambda_l\ E\ C_P\ C_S\ P\ S \end{array} \right), \end{aligned} \tag{5.8}$$

avec $A = \{C_{LV}\ C_{LP}\ C_{LS}\ C_V\ \lambda_p\ \lambda_s\ \lambda_l\ E\ C_P\ C_S\ P\ S\}$.

Nous appliquons la décomposition de la distribution de probabilité conjointe :

$$\begin{aligned} & P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ & = \frac{1}{Z_1} \sum_A \left(\begin{array}{l} P(C_{LV} \mid L\ [W = w])\ P([\lambda_V = 1] \mid C_{LV}\ C_V)\ P(L)\ P([W = w]) \\ P(C_V \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}])\ P([V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]) \\ P(C_{LP} \mid L\ W)\ P(C_{LS} \mid L\ W)\ P(\lambda_P \mid C_{LP}\ C_P)\ P(\lambda_S \mid C_{LS}\ C_S) \\ P(P \mid C_P)\ P(C_S \mid S)\ P(\lambda_l \mid C_{LV}\ C_{LP})\ P(S \mid P)\ P(P)\ P(E \mid P) \end{array} \right). \end{aligned} \tag{5.9}$$

Nous avons vu section 4.1.3 qu'étant donné que λ_P, λ_S et λ_L ne sont pas définis, les parties motrices et de simulation du modèle sont désactivées. Nous obtenons donc :

$$\begin{aligned} & P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ & = \frac{1}{Z_1} \sum_B \left(\begin{array}{l} P(C_{LV} \mid L\ [W = w])\ P([\lambda_V = 1] \mid C_{LV}\ C_V)\ P(L)\ P([W = w]) \\ P(C_V \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}])\ P([V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]) \end{array} \right), \end{aligned} \tag{5.10}$$

avec $B = \{C_{LV}\ C_V\}$.

Etant donné que le terme $P(C_V \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}])$ est défini par une loi Dirac, il n'y a qu'une seule valeur de C_V pour laquelle ce terme est non nul. La sommation sur cette variable disparaît donc.

$$\begin{aligned} & P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ & = \frac{1}{Z_1} \sum_{C_{LV}} \left(\begin{array}{l} P(C_{LV} \mid L\ [W = w])\ P([\lambda_V = 1] \mid C_{LV}\ [C_V = f(v_x^{0:M}, v_y^{0:M})])\ P(L)\ P([W = w]) \\ P([V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]) \end{array} \right). \end{aligned}$$

Les termes $P([W = w])$ et $P([V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}])$ suivent une loi uniforme. Leur valeur numérique est donc toujours la même, quelque soit la valeur de w, $v_x^{0:M}$ et $v_y^{0:M}$. Ils peuvent donc être sortis de la somme et intégrés dans la constante de normalisation qui devient Z_2 :

$$\begin{aligned} & P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ & \quad = \frac{1}{Z_2} \sum_{C_{LV}} P(C_{LV} \mid L\ [W = w])\ P([\lambda_V = 1] \mid C_{LV}\ [C_V = f(v_x^{0:M}, v_y^{0:M})])\ P(L). \end{aligned} \tag{5.11}$$

D'après la définition du terme $P([\lambda_V = 1] \mid C_{LV}\ [C_V = f(v_x^{0:M}, v_y^{0:M}))$ vue section 4.1.3, nous obtenons :

$$\begin{aligned} & P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ & \quad = \frac{1}{Z_2} \left(P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P(L) \right). \end{aligned} \tag{5.12}$$

Comme le terme $P(L)$ est défini par une distribution de probabilité de loi uniforme, il peut également être intégré dans la constante de normalisation :

$$\begin{aligned} &P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ &\quad = \frac{1}{Z_3} P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]). \end{aligned} \tag{5.13}$$

Poser la question de lecture revient donc à comparer les points de contrôle de la trace lue avec les points de contrôle de chacune des lettres mémorisées (ce qui nous ramène à l'exemple introductif). Cette comparaison est effectuée au niveau de chaque point de contrôle pour la position en x et y et pour la vitesse en $\dot{x}$ et $\dot{y}$:

$$\begin{aligned} &P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ &\propto \begin{pmatrix} P([C^0_{LVx} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C^0_{LVy} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C^0_{LV\dot{x}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C^0_{LV\dot{y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \end{pmatrix} \\ &\quad \prod_{n=1}^{N} \begin{pmatrix} P([C^n_{LVx} = f(v_x^{0:M}, v_y^{0:M})] \mid [C^{n-1}_{LVx} = f(v_x^{0:M}, v_y^{0:M})]\ L\ [W = w]) \\ P([C^n_{LVy} = f(v_x^{0:M}, v_y^{0:M})] \mid [C^{n-1}_{LVy} = f(v_x^{0:M}, v_y^{0:M})]\ L\ [W = w]) \\ P([C^n_{LV\dot{x}} = f(v_x^{0:M}, v_y^{0:M})] \mid [C^{n-1}_{LV\dot{x}} = f(v_x^{0:M}, v_y^{0:M})]\ L\ [W = w]) \\ P([C^n_{LV\dot{y}} = f(v_x^{0:M}, v_y^{0:M})] \mid [C^{n-1}_{LV\dot{y}} = f(v_x^{0:M}, v_y^{0:M})]\ L\ [W = w]) \end{pmatrix}. \end{aligned} \tag{5.14}$$

5.2.2 Reconnaissance mono-scripteur

Dans cette partie, nous allons détailler les résultats pour la question :

Sachant cette trace que je vois et qui a été écrite par Estelle, à quelle lettre correspond-elle ?

$$P(L \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [W = Estelle]\ [\lambda_V = 1]). \tag{5.15}$$

Cette question correspond à : étant donnée la trace $\{v_x^{0:M}, v_y^{0:M}\}$ et étant donné que le scripteur est *Estelle*, quelle est la lettre ?

Nous analyserons tout d'abord les résultats pour une trace puis nous les détaillerons pour un ensemble de traces. Enfin, nous analyserons la raison des confusions entre les lettres lors de la reconnaissance.

5.2.2.1 Analyse d'une trace

Nous allons analyser le résultat pour la question suivante :

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = Estelle]\ [\lambda_V = 1]), \tag{5.16}$$

avec $v_x^{0:M}, v_y^{0:M}$ l'ensemble de points définis figure 5.3. Nous indiquons donc au modèle le scripteur (ici *Estelle*) à l'origine de cette trace.

La figure 5.4 présente l'ensemble des points $v_y^{0:M}$ en fonction de $v_x^{0:M}$ (en rouge). Afin de mieux visualiser l'ensemble des points, nous les avons reliés (en bleu). Cette représentation nous permet de visualiser la lettre à reconnaître (ici un *l*).

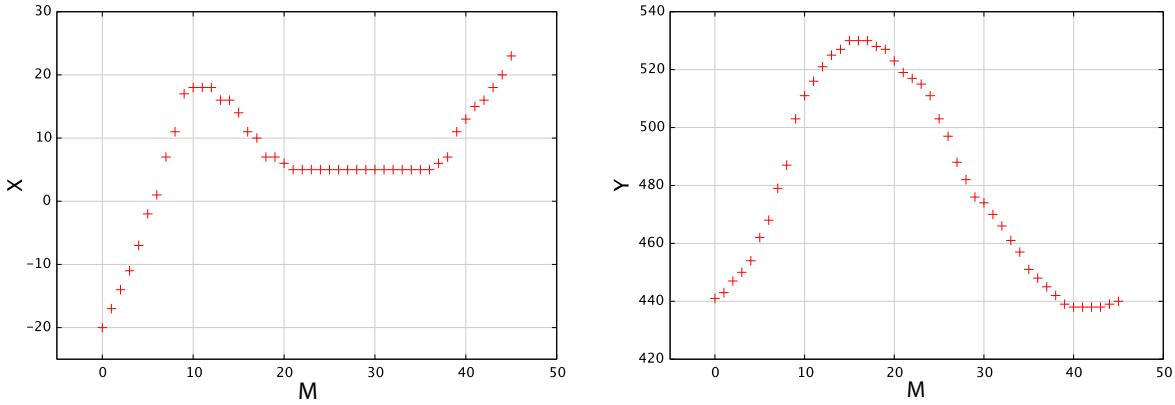


FIG. 5.3 – Ensemble des points définissant la trace dont on cherche à reconnaître la lettre. A gauche, v_x en fonction de l'indice dans la séquence ($m \in [0:M] \cap \mathbb{N}$) et à droite v_y en fonction du même indice ($m \in [0:M] \cap \mathbb{N}$).

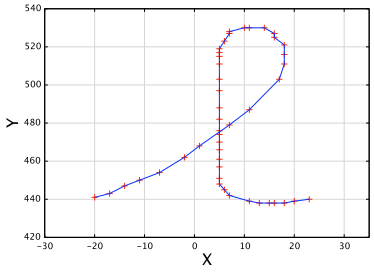


FIG. 5.4 – Ensemble des points définissant la trace dont on cherche à reconnaître la lettre.

Nous avons vu précédemment que :

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = Estelle]\ [\lambda_V = 1]) \propto P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = Estelle]). \quad (5.17)$$

La reconnaissance de la lettre est donc uniquement influencée par la position et la vitesse aux points de contrôle. La figure 5.5 présente la position et la vitesse des points de contrôle de la trace 5.4, extraits par la fonction f.

La distribution de probabilité résultante à cette question est la suivante :

	a	b	c	d	e	f	g	h	k	l	m	n	o	p	q	r	s	u	v	w	y	z	$\mathscr{W}$
Probabilité	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0

Nous noterons que les 0 dans cette distribution de probabilité correspondent réellement à 10^{-40} et le 1 à $1 - (22 * 10^{-40})$.

La trace a été reconnue avec certitude comme étant une lettre l (ce qui est en accord avec ce que le sujet avait tracé) :

$$P([L = l] \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = Estelle]\ [\lambda_V = 1]) = 1. \quad (5.18)$$

5.2.2.2 Taux de reconnaissance

Nous posons maintenant cette même question sur un ensemble de traces produites par le scripteur *Estelle*, afin d'évaluer le taux de reconnaissance de chaque lettre. Il est important

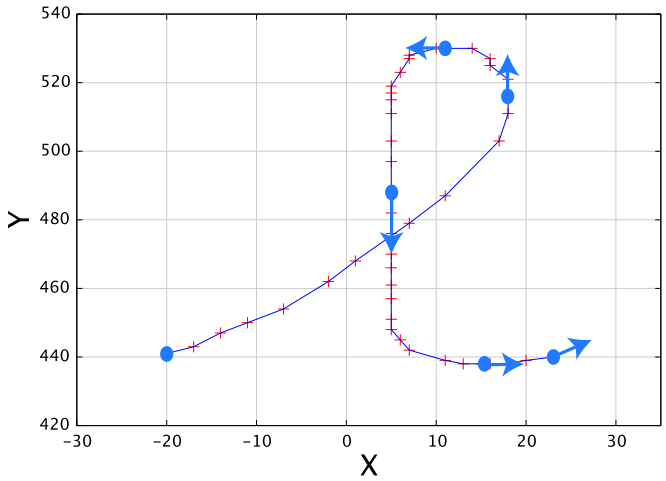


FIG. 5.5 – En bleu, points de contrôle de la trace perçue.

de noter que ces traces ne font pas partie de la base de données d'apprentissage que nous avons utilisée pour identifier les paramètres des formes paramétriques du modèle. Nous avons utilisé une validation croisée en K blocs pour l'échantillonnage de nos données. Parmi les K blocs, un seul est retenu pour la phase de test et les $(K-1)$ autres blocs sont utilisés pour l'apprentissage des paramètres des formes paramétriques. La validation croisée est répétée K fois de telle façon à ce que chaque bloc soit utilisé une fois pour la phase de test. La moyenne des K résultats est calculée pour produire une unique estimation.

Nous obtenons la matrice de confusion[1] présentée figure 5.6.

Chaque ligne correspond au taux de reconnaissance d'une lettre. Par exemple, si l'on prend la première ligne de résultat, nous obtenons que la lettre a a une probabilité de 0.95 d'être reconnue comme un a et une probabilité de 0.05 d'être reconnue comme n'étant pas une lettre (la lettre est reconnue comme étant la lettre « inconnue », notée $\mathscr{U}$).

Le taux de reconnaissance global est de 93,36% : les résultats sont satisfaisants étant donné qu'il n'y a que très peu de confusions. Les principales confusions sont soit entre des lettres similaires comme le g et le q soit avec la lettre « uniforme » (ce dernier cas indique que la trace n'a pas été reconnue comme étant une lettre). Dans la section suivante, nous allons analyser un cas de confusion en détail.

[1]La matrice de confusion, est un outil servant à mesurer la qualité d'un système de classification. Chaque colonne de la matrice représente la probabilité d'une classe estimée, tandis que chaque ligne représente la probabilité d'une classe réelle (probabilité de la classe égale à 1).

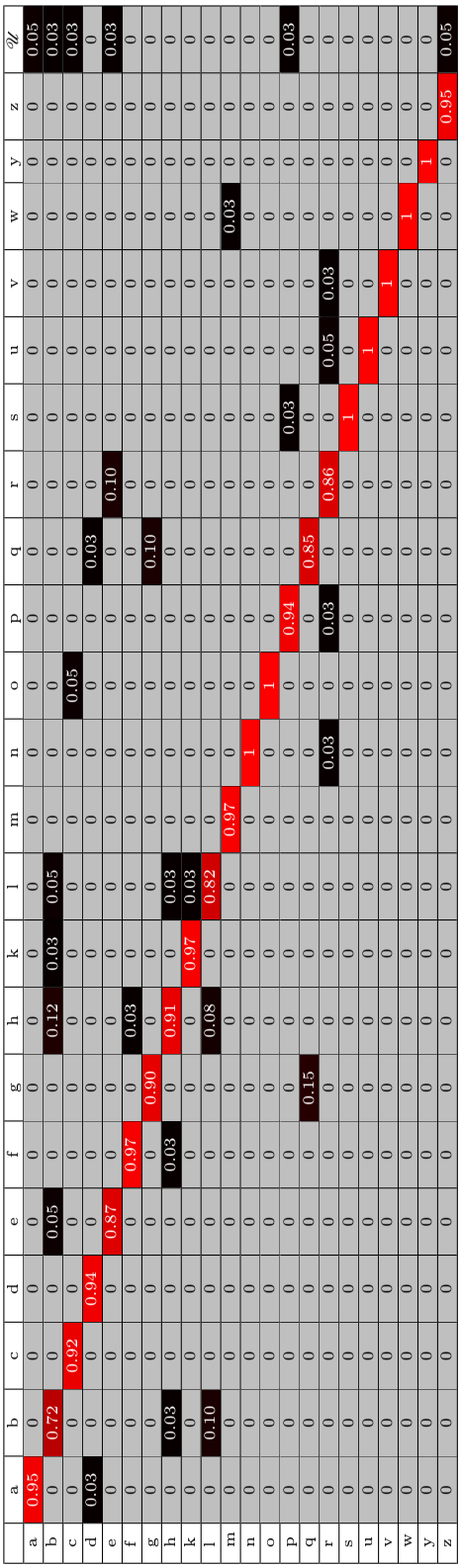


FIG. 5.6 – Matrice de confusion.

5.2.2.3 Analyse de la confusion des lettres g et q

Nous allons analyser, dans cette section, la raison de la confusion entre les lettres g et q. La trace présentée figure 5.7, qui représente une lettre g, a été classifiée par le modèle comme étant une lettre q avec une probabilité de 1. C'est cette trace que nous allons analyser en détail. Sur cette même figure, nous pouvons également voir la position des points de contrôle de la trace lue (carrés rouges).

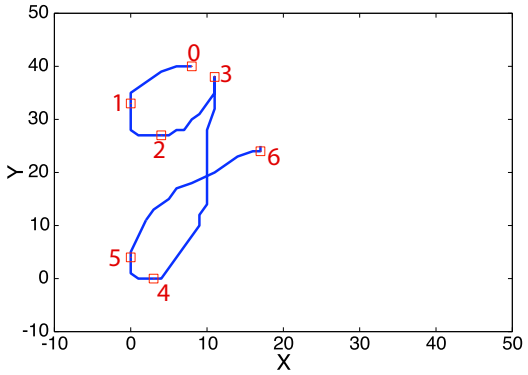


FIG. 5.7 – En bleu, la trace perçue qui est classifiée comme étant un q et en rouge les points de contrôle perçus.

Nous avons vu précédemment que poser la question de « lecture » revient à comparer la position et la vitesse des points de contrôle de la trace lue avec la position et la vitesse des points de contrôle de chacune des lettres mémorisées. Afin de comprendre les raisons de cette erreur de classification, nous avons tout d'abord déterminé les points de contrôle qui sont les plus probables d'appartenir à la séquence des points de contrôle de la lettre q. Pour cela, nous posons les questions suivantes au modèle :

$$P(L \mid [C_{LVx}^{0:I} = c_{lvx}^{0:I}]\ [C_{LVy}^{0:I} = c_{lvy}^{0:I}]\ [C_{LV\dot{x}}^{0:I} = c_{lvx}^{0:I}]\ [C_{LV\dot{y}}^{0:I} = c_{lvx}^{0:I}]\ [W = Estelle]\ [\lambda_V = 1]), \quad (5.19)$$

avec I variant entre 0 et N. Cela correspond à poser une première fois la question avec uniquement le premier point de contrôle, une seconde fois avec le premier et le second point de contrôle, et ainsi de suite. La figure 5.8 présente la probabilité que la lettre L soit égale à g ou q en fonction de I.

On observe que jusqu'au quatrième point de contrôle (indice 3), le modèle n'arrive pas à classifier la lettre comme étant un g ou un q : les distributions de probabilité oscillent entre 0.26 et 0.74. Ces résultats sont cohérents étant donné qu'il s'agit de la première partie de la trace (le « rond ») qui est commune aux deux lettres.

Nous allons maintenant regarder les raisons de l'erreur de classification au niveau du cinquième point de contrôle (indice 4). Etant donné que nous avons effectué une hypothèse d'indépendance sur les quatre composantes x, y, $\dot{x}$, $\dot{y}$, nous pouvons regarder pour chacune d'elle la quantité d'information apportée en ce qui concerne le quatrième point de contrôle. Pour cela, nous allons poser les quatre questions suivantes au modèle :

$$P([L = g] \mid [C_{LVx}^{0:4} = c_{lvx}^{0:4}]\ [C_{LVy}^{0:3} = c_{lvy}^{0:3}]\ [C_{LV\dot{x}}^{0:3} = c_{lv\dot{x}}^{0:3}]\ [C_{LV\dot{y}}^{0:3} = c_{lv\dot{y}}^{0:3}]\ [W = Estelle]\ [\lambda_V = 1]) \quad (5.20)$$

$$P([L = g] \mid [C_{LVx}^{0:3} = c_{lvx}^{0:3}]\ [C_{LVy}^{0:4} = c_{lvy}^{0:4}]\ [C_{LV\dot{x}}^{0:3} = c_{lv\dot{x}}^{0:3}]\ [C_{LV\dot{y}}^{0:3} = c_{lv\dot{y}}^{0:3}]\ [W = Estelle]\ [\lambda_V = 1]) \quad (5.21)$$

$$P([L = g] \mid [C_{LVx}^{0:3} = c_{lvx}^{0:3}]\ [C_{LVy}^{0:3} = c_{lvy}^{0:3}]\ [C_{LV\dot{x}}^{0:4} = c_{lv\dot{x}}^{0:4}]\ [C_{LV\dot{y}}^{0:3} = c_{lv\dot{y}}^{0:3}]\ [W = Estelle]\ [\lambda_V = 1]) \quad (5.22)$$

$$P([L = g] \mid [C_{LVx}^{0:3} = c_{lvx}^{0:3}]\ [C_{LVy}^{0:3} = c_{lvy}^{0:3}]\ [C_{LV\dot{x}}^{0:3} = c_{lvx}^{0:3}]\ [C_{LV\dot{y}}^{0:4} = c_{lv\dot{y}}^{0:4}]\ [W = Estelle]\ [\lambda_V = 1]). \quad (5.23)$$

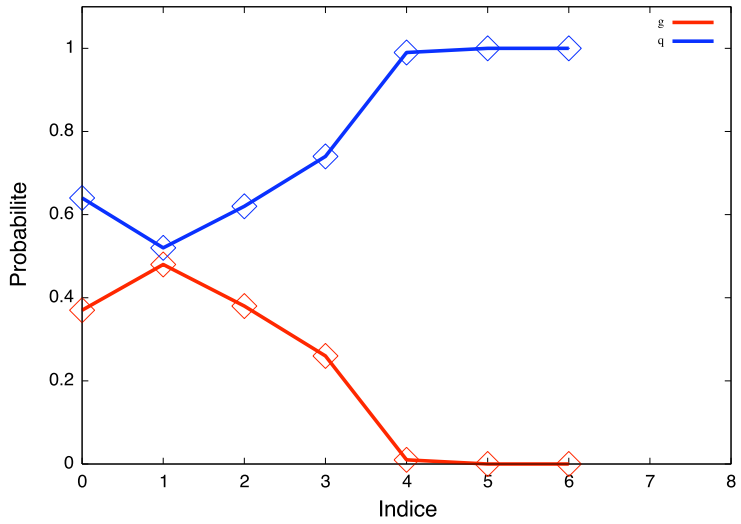


FIG. 5.8 – Probabilité que le trace soit un g en rouge ou un q en bleu en fonction du nombre de points de contrôle présentés au modèle.

La figure 5.9 présente les résultats de ces quatre questions. En marron, nous avons la probabilité de la lettre g sachant les positions et vitesses des quatre premiers points de contrôle et la position en x du cinquième point. En vert, nous avons la probabilité de la lettre g sachant les positions et vitesses des quatre premiers points de contrôle et la position en y du cinquième point. En rose (resp. en bleu) est représentée la probabilité de la lettre g sachant les positions et vitesses des quatre premiers points de contrôle et la vitesse en x (resp. en y) du cinquième point. Nous observons que l'erreur de classification vient de la position en y du cinquième point de contrôle. En effet, la probabilité de la lettre g a fortement chuté selon la position en y alors qu'elle est restée stable avec les informations de vitesse et a même fortement augmenté avec la position en x.

Nous allons maintenant analyser les distributions de probabilité de la position en y du cinquième point de contrôle des lettres g et q que nous avons déterminés lors de la phase d'apprentissage et nous allons les comparer figure 5.10 avec la position du point de contrôle de la trace lue. Nous observons figure 5.7 que la position en y du quatrième point de contrôle est 38 et que la position du cinquième est 0. Nous avons entouré en vert sur la figure 5.10, la valeur de la probabilité qui nous intéresse. Nous observons que la probabilité de la position en y du cinquième point de contrôle est plus élevé pour la lettre q (0.5) que la lettre g (0.2). Ces résultats sont tout a fait cohérents pour la lettre q étant donné qu'il s'agit du dernier point de contrôle de la lettre, donc le plus bas (position à 0). En revanche, pour la lettre g, cela indique que nous n'avons quasiment jamais vu de tel allographe lors de la phase d'apprentissage. Le plus souvent c'est l'allographe présenté figure 5.11 que nous avons rencontré avec un point supplémentaire entre le quatrième et le cinquième point. La position la plus probable du cinquième point se situe entre 30 et 40.

Discussion Au niveau du cinquième point de contrôle, nous avons vu que si l'on considère uniquement la probabilité selon la position en x, la lettre g est plus probable que la lettre q. Pour améliorer la reconnaissance, il faudrait supprimer l'hypothèse d'indépendance entre les

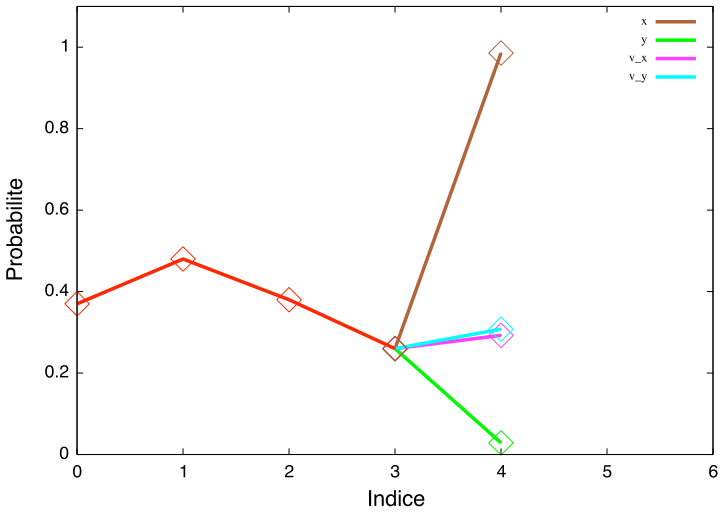


FIG. 5.9 – Probabilité de reconnaître la lettre g selon les informations de position et de vitesse du cinquième point de contrôle. En rouge, probabilité de reconnaître la lettre g sachant les informations de position et de vitesse du point de contrôle de l'indice et des précédents. En marron, probabilité de reconnaître la lettre g en fonction des informations de position en x du cinquième point et des points précédents. En vert, en fonction de la position y et des précédents. En violet, en fonction de la vitesse en x et des précédents et en bleu de fonction de la vitesse en y et des précédents.

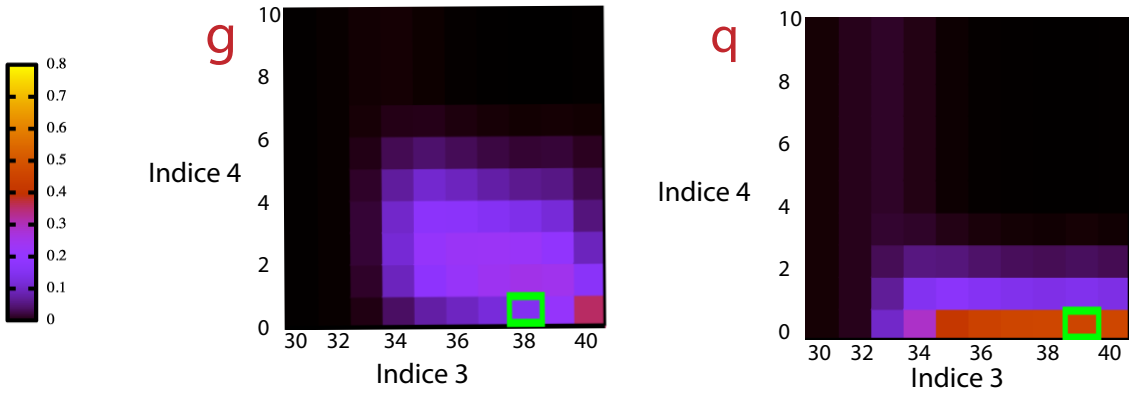


FIG. 5.10 – Distribution de probabilité de la position en y du cinquième point de contrôle des lettres g et q. En vert, à gauche : $P([L = g] \mid [C^4_{Lx} = 0]\ [C^3_{Lx} = 38])$ et en vert à droite : $P([L = q] \mid [C^4_{Lx} = 0]\ [C^3_{Lx} = 38])$

dimensions x, y, $\dot{x}$ et $\dot{y}$ des points de contrôle que nous avons réalisé précédemment. Dans ce cas, la probabilité du quadruplet $< x, y, \dot{x}, \dot{y} >$ du cinquième point de contrôle pour la lettre q serait faible étant donné que la position en x ne correspond pas à l'allographe enregistré. Nous aurions alors une probabilité faible pour les lettres g et q et la lettre serait alors reconnue comme la lettre uniforme. Etant donné que nous n'avons jamais rencontré cet allographe lors de l'apprentissage, il serait alors normal de classifier cette trace comme n'étant pas une lettre connue.

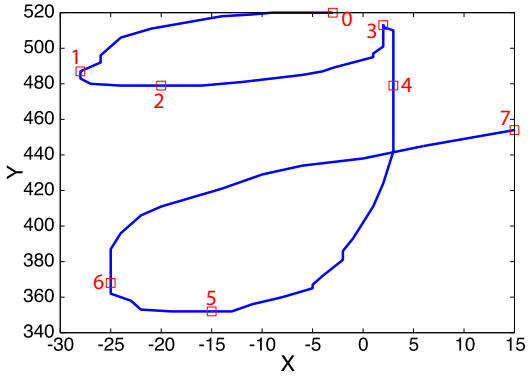


FIG. 5.11 – Allographe de la lettre g le plus rencontré lors de l'apprentissage.

Cependant, supprimer l'hypothèse d'indépendance entre les dimensions x, y, $\dot{x}$ et $\dot{y}$ des points de contrôle serait assez coûteux en temps de calcul notamment pour le modèle de production. Nous verrons par la suite que dans l'inférence de la tâche d'écriture, nous avons une sommation sur les points de contrôle. Dans notre cas, cela revient à réaliser une sommation sur un espace à 100 dimensions ($40+40+10+10$). Dans le cas où nous supprimons l'hypothèse d'indépendance, la sommation se ferait sur un espace à $40 * 40 * 10 * 10$ dimensions.

5.2.2.4 Suggestion de protocole expérimental

La matrice de confusion présente les différentes erreurs faites par le modèle lors de la reconnaissance des lettres. Nous observons que les confusions se font principalement entre des lettres qui sont visuellement similaires telles que le b, le h et le l ou le q et le g. Or, nous n'avons aucune comparaison avec des résultats chez l'humain. Afin de valider ces prédictions expérimentales réalisées par notre modèle, nous suggérons donc un protocole expérimental.

Hypothèses	Les confusions entre les lettres sont principalement entre des lettres visuellement similaires.
Matériel	Points lumineux se déplaçant le long de la trajectoire d'une lettre. Les 22 lettres répertoriées par notre expériences sont proposées.
Procédure	• Etape 1 : Faire tracer aux sujets des lettres et enregistrer les trajectoires réalisées. • Etape 2 : Présenter les points lumineux se déplaçant le long de la trajectoire aux sujets en leur demandant de reconnaître la lettre. Pour chaque sujet, présenter ses propres trajectoires.
Résultats	Pour chaque lettre, on détermine le taux de reconnaissance. On détermine également la matrice de confusion des lettres.
Analyse	L'analyse des résultats correspond à comparer cette matrice de confusion avec la matrice de confusion prédite par le modèle.

5.2.3 Reconnaissance multi-scripteur

Lors de la phase d'apprentissage des distributions de probabilité relatives aux séquences de points de contrôle, quatre scripteurs ont participé à l'élaboration de notre base de données : *Estelle, Julienne, Jean-Louis* et *Christophe*. Nous nous intéressons maintenant à la question de reconnaissance de caractères ne connaissant pas le scripteur à l'origine des traces.

5.2.3.1 Question et inférence

La question correspond à déterminer la distribution de probabilité de la lettre associée à la trace fournie et ne connaissant pas le scripteur est :

Quelle est la lettre correspondant à la trace que je perçois ne sachant pas qui l'a écrite ?

$$P(L \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]). \tag{5.24}$$

Cette question est similaire à la question 5.16. Cependant, cette fois, nous n'ayons pas précisé le scripteur à l'origine de la trace $\{v_x^{0:M}, v_y^{0:M}\}$.

L'inférence bayésienne nous permet de répondre à cette question à partir de la décomposition de la distribution conjointe. Nous ne détaillerons pas l'inférence ici car elle est similaire à celle réalisée section 5.2.1. La seule différence vient du fait que cette fois nous avons une

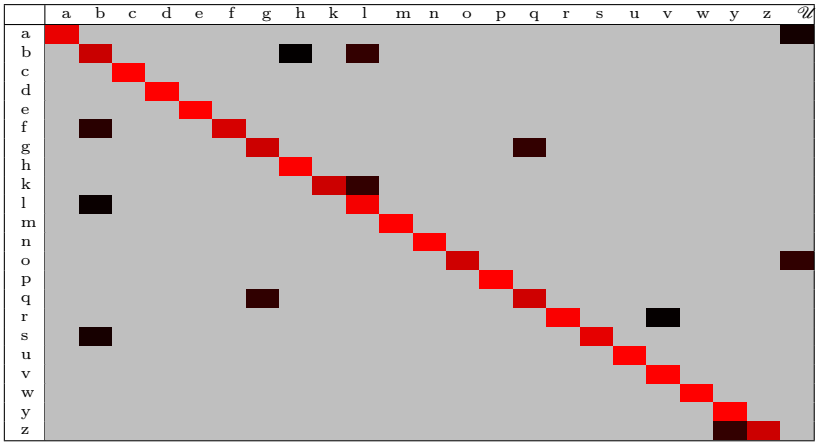
sommation sur la variable W. Une fois l'inférence calculée, nous obtenons :

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]) \propto \sum_W P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ W). \tag{5.25}$$

La tâche de lecture revient donc à comparer les points de contrôle de la trace lue avec les points de contrôle de chacune des lettres mémorisées et pour chaque scripteur connu.

5.2.3.2 Taux de reconnaissance

Si l'on pose cette question sur la même base de données que dans la section 5.2.2.2, nous obtenons la matrice de confusion suivante. Il est important de noter que bien que l'on ne précise pas l'identité du scripteur, les traces ont été produites par un scripteur connu du modèle.

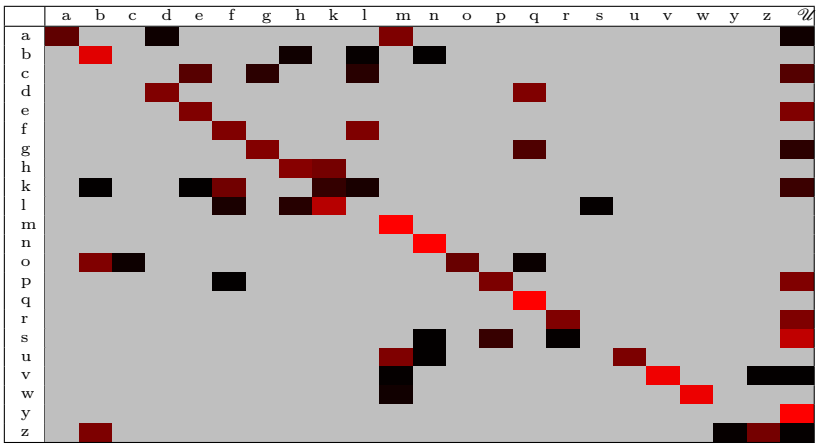


Afin de mieux représenter visuellement les confusions entre les données dans les matrices de confusions, nous avons supprimé le chiffre correspondant à la probabilité de reconnaître la lettre et gardé uniquement notre code couleur : lorsque la probabilité est égale à 0, nous représentons la case en gris et lorsqu'elle est égale à 1, nous la représentons en rouge. Pour les valeurs intermédiaires, nous la représentons avec un dégradé entre le rouge et le noir : plus la valeur est élevée, plus la case va tendre vers le rouge et inversement, plus la valeur est faible, plus la couleur va tendre vers le noir. Cette représentation nous permet de visualiser les confusions entre lettres en faisant ressortir les valeurs non nulles.

Nous observons que la reconnaissance des caractères est légèrement dégradée par rapport à la reconnaissance de caractères lorsque l'on connaît le scripteur : le taux de reconnaissance est de 92,72% alors qu'il était de 93,72% auparavant. Ce résultat est cohérent étant donné que l'on a supprimé l'information correspondant à l'identité du scripteur.

5.2.3.3 Scripteur inconnu

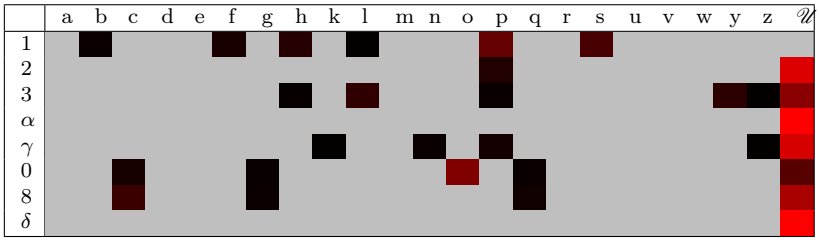
Dans le paragraphe précédent, nous avons posé la question de la reconnaissance des caractères sans préciser le scripteur mais sachant qu'il est connu dans la base de donnée. Cette fois nous allons poser la même question mais l'auteur des traces fournies en entrée n'est pas connu du modèle (aucun apprentissage sur ces traces n'a été réalisé et mémorisé). Nous observons dans le tableau suivant qu'il y a de nombreuses confusions entre les lettres.



Le taux de reconnaissance est de 49,68%. Il est faible étant donné que les lettres tracées par ce scripteur ne correspondent pas obligatoirement aux allographes réalisés par les autres scripteurs. Pour avoir un taux de reconnaissance correct lorsque le scripteur n'est pas connu, il faudrait avoir une base de donnée avec un nombre beaucoup plus important de scripteurs. Cela permettrait d'apprendre un plus grand nombre d'allographes et donc de mieux reconnaître les lettres tracées par des nouveaux scripteurs.

5.2.3.4 Lettres inconnues

Dans cette section on s'intéresse à valider le fait que si la trace présentée ne correspond pas à une lettre, le modèle va bien le détecter comme tel. Pour réaliser ce test, nous avons présenté au modèle des chiffres (1,2,3,8,0) et des lettres grecques (α, γ, δ) et lui avons demandé quelle était la lettre correspondante. Une trace est reconnue comme n'étant pas une lettre si le modèle la désigne comme la lettre « inconnue » ($\mathscr{U}$).



Nous observons dans la matrice de confusion que presque tous les caractères ont été classés dans la dernière colonne et donc ont bien été reconnus comme n'étant pas des lettres (65,60% des caractères ont été classés comme des lettres inconnues). Nous avons effectué le même test en précisant le scripteur lors de la question et nous obtenons un taux de classification « lettre inconnue » de 77,94%. Les principales confusions sont le 1 et le p (p écrit en cursif minuscule) ce qui n'est pas surprenant étant donné que ces deux caractères ont une grande similitude sur le début de la trace. Nous avons aussi des confusions entre le chiffre 0 et les caractères o, c, g et q. Tous ces caractères commencent de la même façon, par « un rond ».

5.3 Reconnaissance du scripteur

Dans cette section, nous nous intéressons non plus à reconnaître les lettres mais les scripteurs. Comme précédemment, nous détaillons tout d'abord la question posée au modèle avant d'analyser les résultats.

5.3.1 Question et inférence

La question relative à la reconnaissance du scripteur est :

Qui a écrit cette trace sachant que je connais la lettre correspondante ?

$$P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]\ [L = l]). \tag{5.26}$$

Cette question correspond à : *Qui a écrit la trace* $\{v_x^{0:M}, v_y^{0:M}\}$ *représentant la lettre* l *?* Comme précédemment, l'inférence bayésienne nous permet de répondre à cette question à partir de la décomposition de la distribution conjointe. Nous ne la détaillerons pas étant donné qu'elle est très proche de celle décrite précédemment. Le résultat de l'inférence est le suivant :

$$\begin{aligned} &P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]\ [L = l]) \\ &\quad \propto P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid W\ [L = l]). \end{aligned} \tag{5.27}$$

Nous observons que le résultat de l'inférence est similaire à celui de la question de lecture : poser la question de reconnaissance du scripteur revient donc à comparer les points de contrôle de la trace lue avec les points de contrôle des lettres mémorisées pour chaque scripteur.

5.3.2 Taux de reconnaissance

Pour chaque scripteur, nous posons cette question sur un ensemble de 110 traces où chaque lettre est présentée cinq fois. Nous rappelons que nous avons quatre scripteurs dans notre base de donnée : pour chaque trace, le modèle donne donc la probabilité qu'elle ait été tracée par un de ces quatre scripteurs. Nous obtenons les confusions suivantes :

	Estelle	Julienne	Jean-Louis	Christophe
Estelle	0.76	0.03	0.07	0.14
Julienne	0.02	0.80	0.07	0.11
Jean-Louis	0	0	1	0
Christophe	0.10	0.14	0.13	0.62

Comme précédemment, la matrice de confusion se lit en ligne : par exemple, pour la première ligne, nous avons la probabilité que le modèle ait reconnu les différents scripteurs sachant que le scripteur était *Estelle*. Nous observons que c'est avec une forte probabilité que le modèle reconnaît les différents scripteurs. Les principales confusions sont entre le scripteur *Christophe* et les autres scripteurs. Le taux de reconnaissance moyen des scripteurs est de 79,5%.

5.3.3 Lettre inconnue

Dans le paragraphe précédent, nous avons analysé les résultats pour la question suivante : *Qui a écrit la trace* $\{v_x^{0:M}, v_y^{0:M}\}$ *représentant la lettre l ?* . Nous allons maintenant comparer ces résultats avec ceux de la question :

Qui a écrit cette trace sachant que j'ignore à quelle lettre elle correspond ?

Dans cette question, nous ne précisons pas la lettre correspondant à la trace. Mathématiquement, cette question et son inférence sont :

$$\begin{aligned} P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]) \\ \propto \sum_L P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid W\ L). \end{aligned} \tag{5.28}$$

Nous observons que par rapport à la question précédente, nous avons une sommation sur la variable L. Nous nous intéressons donc à déterminer le scripteur sans pour autant chercher à déterminer la lettre correspondant à la trace. Cette question a été posée sur le même jeu de données que précédemment et nous obtenons la matrice de confusion suivante :

	Estelle	Julienne	Jean-Louis	Christophe
Estelle	0.78	0.06	0.05	0.11
Julienne	0.03	0.76	0.12	0.09
Jean-Louis	0	0	0.98	0.02
Christophe	0.12	0.15	0.13	0.60

Cette matrice de confusion est intéressante car bien que l'on ait fourni au modèle moins d'information que précédemment (la lettre correspondant à la trace n'est pas connue), le taux de reconnaissance est à peu près similaire. Il était de 79,5% auparavant et il est de 78% dans ce cas. Cette faible différence provient du fait que le taux de reconnaissance des lettres est proche de 90%. Cela revient donc au même de connaître la lettre ou de faire une sommation sur la variable L. Ce résultat suggère donc qu'il n'est pas nécessaire de reconnaître la lettre pour pouvoir déterminer le scripteur.

5.4 Conclusion

5.4.1 Résumé

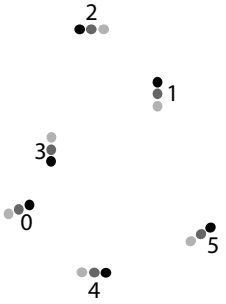
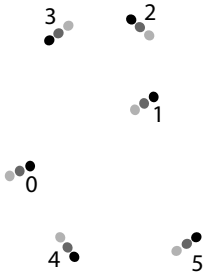
Dans ce chapitre, nous avons présenté le modèle de perception qui extrait une séquence de points de contrôle à partir d'une trace. Nous avons également défini des questions relatives aux tâches de lecture :

- Quelle est la lettre correspondant à cette trace connaissant le scripteur ?
  $P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1])$
- Quelle est la lettre correspondant à cette trace ne connaissant pas le scripteur ?
  $P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1])$
- Qui a écrit cette trace ? $P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1])$
- Qui a écrit cette lettre ? $P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]\ [L = l])$

Chacune de ces tâches correspond à une question probabiliste qui est résolue automatiquement par inférence bayésienne. Nous avons simulé informatiquement chacune de ces inférences pour obtenir le taux de reconnaissance de chaque caractère et analyser les confusions entre les lettres.

5.4.2 Suggestion d'un protocole expérimental

Une hypothèse fondamentale de ce travail est que la représentation interne des lettres correspond à des points de contrôle. Nous avons suggéré que les points de contrôle sont situés au niveau des points de tangence remarquable et au niveau des points de rebroussement. Nous suggérons un protocole expérimental pour étudier cette hypothèse.

<table>
<tr><td>Hypothèses</td><td>La représentation interne des lettres correspond à des points de contrôle. Les points de contrôle sont situés au niveau des points de tangence remarquable et au niveau des points de rebroussement.</td></tr>
<tr><td>Matériel</td><td>Points lumineux se déplaçant sur des portions de la trajectoire. Deux conditions seront testées :
<table>
<tr><td>C_0</td><td>C_1</td></tr>
<tr><td>Les points lumineux sont situés au niveau des points de contrôle.</td><td>Les points lumineux sont situés entre les points de contrôle.</td></tr>
<tr><td> 
</td><td> 
</td></tr>
</table></td></tr>
<tr><td>Procédure</td><td>Présenter les points lumineux se déplaçant sur des portions de trajectoire aux sujets en leur demandant de reconnaître la lettre. Les deux conditions seront présentées. Deux informations pourront être retenues : la lettre reconnue et le temps de réponse.</td></tr>
<tr><td>Résultats</td><td>Déterminer les taux de reconnaissance pour les deux conditions et le temps de réponse pour chacune des deux conditions.</td></tr>
<tr><td>Analyse</td><td>L'analyse des résultats correspond à la comparaison des taux de reconnaissance. Trois cas sont alors possibles :
• C_0 et C_1 sont équivalents : ce résultat suggère que les lettres ne sont pas représentées par des points particuliers de la trajectoire.
• C_0 est inférieur à C_1 : ce résultat invalide notre hypothèse. Les informations nécessaires à la reconnaissance de la lettre se situent entre les points de contrôle (points où l'accélération est nulle, par exemple).
• C_1 est inférieur à C_0 : ce résultat est en accord avec notre hypothèse.
La comparaison des temps de réponse pour les deux conditions peut être comparée de façon analogue.</td></tr>
</table>

Chapitre 6

Action

Dans ce chapitre, nous nous intéressons à la tâche cognitive de l'écriture et donc plus précisément à la production de traces représentant des lettres. Afin de déterminer le modèle de génération de trajectoires du modèle BAP, nous effectuons une revue des modèles issus de la théorie du contrôle. Cela nous amène tout aussi bien à présenter des modèles de contrôle en robotique (par exemple le modèle bang-bang) et des modèles de contrôle moteur humain (par exemple le modèle de secousses minimales). Dans la suite du chapitre, nous détaillerons donc le modèle de génération de trajectoire et les effecteurs que nous avons utilisés. Et enfin, nous analyserons les résultats de la tâche d'écriture.

6.1 Contrôle moteur dans la littérature

Les mouvements élémentaires du corps humain sont contraints principalement par la mécanique. Pour effectuer une tâche, il est possible de combiner de plusieurs façons des mouvements élémentaires. Cependant, bien que le répertoire des mouvements possibles soit très grand, on observe souvent un même mouvement résultant d'une même stratégie. En effet, le système nerveux central utilise une stratégie de planification qui impose des restrictions supplémentaires (Berthoz, 1998). Par exemple, dans une tâche point à point où il faut attraper des objets, la posture finale de la main est toujours la même quelque soit la trajectoire empruntée. Cette posture est définie en avance et est utilisée comme une variable de contrôle par le système nerveux central (Gréa *et al.*, 2000).

Dans cette section, nous allons définir la notion de trajectoire et étudier les différents modèles présents dans la littérature concernant le contrôle des mouvements lors de tâches de pointage.

6.1.1 Etudes expérimentales

Dans la suite de cette revue, nous nous intéresserons principalement aux trajectoires point à point. Les seuls paramètres spécifiés en entrée de ces modèles sont donc les positions initiales et finales des points ainsi que les vitesses initiales et finales. De nombreuses études expérimentales ont montré que la trajectoire décrite par la main (ou le doigt suivant l'étude) est quasiment une droite. De plus, le profil de vitesse correspond à une courbe en cloche. La figure 6.1 présente la trajectoire effectuée par la main dans une tâche point à point (vitesses initiales et finales nulles). On observe que la trajectoire est quasiment une droite et que le profil de

vitesse associé est celui d'une courbe en cloche.

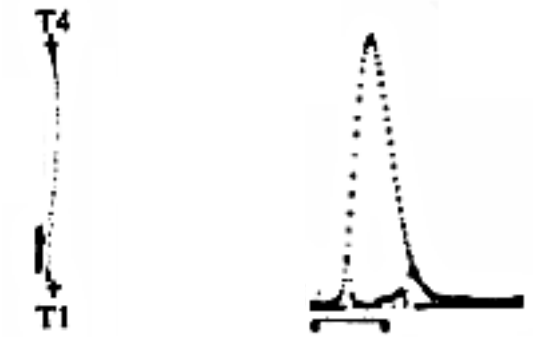


FIG. 6.1 – A gauche, la trajectoire décrite par la main entre les cibles T_1 et T_4. A droite, le profil de vitesse associé à la trajectoire (figure tirée de Abend *et al.* (1982)).

6.1.2 Modèles de génération de trajectoires

Il y a dans la littérature plusieurs types de modèles de génération de trajectoires pour les tâches de pointage. Par exemple, Feldman (1966); Flanagan *et al.* (1993) ont émis l'hypothèse que les mouvements des membres sont produits par une modification des conditions d'équilibre mécanique du système musculaire. Ils distinguent la variable contrôlée par le système moteur (longueur du muscle l) des variables cinématiques (vitesse, force, etc.) qui correspondent à des variables d'état. Cependant, ce modèle et donc ce type de contrôle (contrôle musculaire) fait l'objet d'importants débats (Stein, 1982). De plus, Atkeson et Hollerbach (1985) ont montré que si l'on positionne une masse sur le bras des sujets, la trajectoire reste inchangée, alors que ce n'est pas le cas avec les trajectoires générées par la théorie du point d'équilibre.

Il existe aussi de nombreux modèles utilisant une information visuelle pour engendrer le mouvement (Li *et al.*, 2005; Wolpert *et al.*, 1995; Wolpert et Ghahramani, 2000; Ghahramani *et al.*, 1997; Haruno *et al.*, 2001; Kawato, 1999; Bays et Wolpert, 2007). Cependant, Paillard a émis l'hypothèse, reprise par Sirat *et al.* (1988), que dans le cadre de l'écriture, le rôle de l'information visuelle est essentielle dans la phase de construction des programmes moteurs au cours de l'apprentissage de l'écriture et, ultérieurement, dans le contrôle de l'enchaînement des lettres composant le mot. Dans notre cas, nous nous intéressons uniquement au tracé de lettres isolées (sans levé de stylo) et en aucun cas à l'apprentissage de l'écriture, ni au tracé de mots.

Enfin, une autre série de modèles, que nous allons détailler par la suite, se base sur la notion de contrôle optimal.

6.1.2.1 Contrôle optimal

La « *théorie du contrôle* » concerne les propriétés des systèmes sur lesquels on peut agir au moyen d'une commande (ou contrôle) (Trélat, 2005). Le but est d'amener le système d'un état initial donné à un état final, en respectant éventuellement certains critères.

Considérons un système dynamique général où x est l'état du système dynamique et u la

variable de commande, tel que :

$$\dot{x}(t) = f\left(t, x(t), u(t)\right), \quad x(t_0) = x_0, \tag{6.1}$$

avec $u(t) \in \mathcal{U}$, où $\mathcal{U}$ est l'espace des commandes admissibles.

Un problème de contrôle optimal se décompose en deux parties : pour déterminer une trajectoire optimale joignant un ensemble initial à une cible, il faut d'abord savoir si cette cible est atteignable. C'est le *problème de contrôlabilité*. Ensuite, parmi toutes les trajectoires possibles, on cherche une trajectoire qui minimise une certaine fonction de *coût*. Une telle trajectoire, si elle existe, est dite *optimale* pour ce coût.

Contrôlabilité Un système dynamique est dit contrôlable si on peut l'amener (en temps fini) d'un état initial arbitraire vers un état final prescrit. Le système 6.1 est dit *contrôlable* en temps T depuis x_0 si

$$\mathbb{R}^n = Acc(x_0, T), \tag{6.2}$$

avec $Acc(x_0, T)$ l'ensemble des points accessibles au temps T en partant de x_0 au temps $t = 0$. Pour tous x_0, $x_1 \in \mathbb{R}^n$, il existe un contrôle u tel que la trajectoire associée relie x_0 à x_1 en temps T.

Contrôle optimal Une fois le problème de contrôlabilité résolu, on peut vouloir passer de l'état initial à l'état final en minimisant un certain critère : on parle alors d'un problème de *contrôle optimal*. Le critère à minimiser peut être par exemple le temps, l'énergie dépensée, la distance... Pour tout contrôle $u \in \mathcal{U}$, on définit le coût de la trajectoire associée c_u sur l'intervalle $[0, T]$

$$C(T, u) = \int_0^T f^0(t, x(t), u(t))dt + g(T, c_u(T)), \tag{6.3}$$

avec f^0 une fonction de classe C_1 de $I * V * U$ dans $\mathbb{R}^n$, g une fonction continue sur V, I un intervalle de $\mathbb{R}$, V ouvert de $\mathbb{R}^n$ et U un ouvert de $\mathbb{R}^n$

Principe du minimum de Pontryagin Pour trouver la commande optimale permettant d'amener un système dynamique d'un état à un autre en présence de contraintes portant sur l'état ou les commandes d'entrée, Pontryagin a énoncé les conditions nécessaires d'optimalité. L'utilisation du principe du minimum de Pontryagin[1] nous permet donc d'obtenir la commande optimale.

La commande $u \in \mathcal{U}$ doit être déterminée pour tout $t \in [0, T]$ tout en maximisant la fonctionnelle J, définie par :

$$J = \int_0^T L(t, x(t), u(t))dt, \tag{6.4}$$

[1]Ces résultats sont parfois connus sous le nom de principe du maximum de Pontryagin : le travail initial de Pontryagin était centré sur la maximisation d'une fonctionnelle de bénéfice plutôt que la minimisation d'une fonctionnelle de coût. La démonstration historique du principe du minimum est basée sur une maximisation et non sur sa minimisation. Dans ce cadre minimiser une fonction de coût plutôt qu'un bénéfice revient à maximiser son opposé.

avec le Lagrangien[2] $L(.)$ qui décrit la dépendance du coût vis à vis de la position $x(t)$ et du contrôle $u(t)$.

Par la suite, nous n'utiliserons pas la version générale du principe de maximum de Pontryagin car son énoncé et sa démonstration sont inutilement complexes pour notre cas. Nous nous plaçons donc dans le cas où il n'y a pas de contraintes sur le contrôle ($L(.)$ dépend uniquement de $x(t)$). Dans ce cas, on parle de maximum de Pontryagin faible.

Principe de maximum de Pontryagin faible Nous considérons la fonction de coût suivante :

$$J = \int_0^T L(x(t), x^{(1)}(t), ..., , x^{(n)}(t), t)dt, \tag{6.5}$$

où $L(.)$ décrit la dépendance du coût vis-à-vis de la position $x(t)$ de l'effecteur et de ses dérivées $x^{(n)}(t) = d^n x/dt^n$. La trajectoire optimale $x^*(t)$ est la trajectoire qui minimise J. Une condition pour que J soit minimale est que l'on ait :

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial x^{(1)}}\right) + ... + (-1)^n \frac{d^n}{dt^n}\left(\frac{\partial L}{\partial x^{(n)}}\right) = 0. \tag{6.6}$$

La trajectoire optimale peut donc être calculée en résolvant les équations d'Euler-Lagrange énoncées ci-dessus.

6.1.3 Génération de trajectoires à partir de la théorie de contrôle optimal

Par la suite, nous nous intéresserons principalement à la notion de *contrôle optimal*. D'un point de vue expérimental, cette théorie est plausible : Shadmehr et Krakauer (2008) ont montré que les ganglions de la base[3] semblent jouer un rôle dans le calcul d'un coût qui est ensuite minimisé par les aires du cortex moteurs pour déterminer les sorties motrices.

Nous partons du principe que la position finale est toujours atteignable, et donc que la notion de *contrôlabilité* est toujours vérifiée (nous considérons qu'il n'y a aucun obstacle sur la trajectoire et que la position initiale et la position finale sont des configurations plausibles). Dans la littérature, il existe de nombreux modèles de génération de trajectoires. Ces modèles définissent une grande variété de coût à minimiser : l'énergie (Nelson, 1983), le temps, les secousses, etc. C'est ce que nous allons voir par la suite. Nous avons classé ces différents modèles suivant l'espace dans lequel le coût est représenté (espace cartésien, articulaire). Bien que nous ayons fait l'hypothèse section 2.3 que dans notre modèle la génération de trajectoire se ferait dans l'espace cartésien, il nous a semblé intéressant de faire une revue des modèles plus complète.

6.1.4 Modèles dans l'espace cartésien

Plusieurs modèles de génération de trajectoires basés sur le principe de minimisation ont été définis dans l'espace cartésien. La plupart du temps, le Lagrangien correspondant est égal

[2]Le Lagrangien $L(.)$ d'un système dynamique est une fonction des variables dynamiques qui décrit de manière concise les équations du mouvement du système.

[3]Les ganglions de la base sont un ensemble sous-cortical de substance grise situés approximativement au milieu du cerveau humain.

au carré de la dérivée $n^{\text{ème}}$:

$$J = \frac{1}{2} \int_{t_1}^{t_2} \left(\frac{d^n x}{dt^n} \right)^2 + \left(\frac{d^n y}{dt^n} \right)^2 dt. \tag{6.7}$$

Pour $n = 2$, $n = 3$ et $n = 4$, les modèles qui minimisent cette fonction de coût se nomment respectivement *minumum-acceleration*, *minumum-jerk*, *minumum-snap*. Le Lagrangien du modèle *minimum-acceleration* est égal au carré de la dérivée seconde (Ben-Itzhak et Karniel, 2008). Pour le modèle *minimum-jerk* (Flash et Hogan, 1985), le coût est proportionnel aux secousses (carré des dérivées troisièmes). Le Lagrangien du *miminum-snap* est quant à lui proportionnel au carré de la dérivée quatrième.

Dans le cas où le Lagrangien est le carré d'une dérivée $n^{\text{ème}}$ ($L(t) = (x^{(n)}(t)^2)$, la solution de l'équation 6.6 est donnée par un polynôme par morceaux d'ordre $2n - 1$ au plus où les coefficients sont déterminés par les conditions aux limites (Harris et Harwood, 2005).

Démonstration Soit le coût défini par :

$$J = \int_0^T \left(x^{(n)}(t) \right)^2 dt. \tag{6.8}$$

La trajectoire optimale est calculée en résolvant les équations d'Euler-Lagrange :

$$(-1)^n \frac{d^n}{dt^n} \left(\frac{\partial \left(x^n(t) \right)^2}{\partial x^n} \right) = 0. \tag{6.9}$$

On calcule la dérivée partielle $n^{\text{ème}}$ selon x :

$$(-1)^n \frac{d^n 2x^n(t)}{dt^n} = 0, \tag{6.10}$$

puis la dérivée $n^{\text{ème}}$ selon t :

$$(-1)^n 2x^{2n}(t) = 0. \tag{6.11}$$

Cette équation implique que :

$$x^{2n}(t) = 0. \tag{6.12}$$

Quelque soit t appartenant à $[0, T]$, nous avons donc $x^{2n}(t) = 0$. Cette équation implique que la trajectoire est définie par un polynôme de degré inférieur ou égal à $2n - 1$.

Par la suite nous nous intéresserons au modèle *minimum-jerk*, c'est-à-dire que la fonction de coût est proportionnelle au carré de la norme euclidienne des secousses ($n = 3$) (Richardson et Flash (2002) ont montré que c'est pour la dérivée 3ème que les résultats de la modélisation correspondent le plus aux données expérimentales).

Puis, nous aborderons un autre exemple de la théorie de contrôle appliquée dans l'espace cartésien : les modèles de type Bang-Bang.

6.1.4.1 Minimum-Jerk point à point

Fonction de coût Flash et Hogan (1985) ont défini la fonction de coût suivante :

$$J = \frac{1}{2}\int_{t_1}^{t_2} \left(\frac{d^3x}{dt^3}\right)^2 + \left(\frac{d^3y}{dt^3}\right)^2 dt. \tag{6.13}$$

Résolution Nous sommes donc dans le cas où le Lagrangien est le carré d'une dérivée $n^{\text{ème}}$. La solution de l'équation 6.6 est donnée par un polynôme d'ordre $2n-1$ où les coefficients sont déterminés par les conditions aux limites.

Par linéarité de l'intégration et comme $\left(\frac{d^3x}{dt^3}\right)^2$ et $\left(\frac{d^3y}{dt^3}\right)^2$ sont positifs, nous résolvons indépendamment les deux dimensions. Le Lagrangien correspondant au modèle *minimum-jerk* est donc pour la dimension x :

$$L_x(t) = (x^{(3)}(t))^2. \tag{6.14}$$

En résolvant les équations d'Euler-Lagrange, la trajectoire optimale $x^*(t)$ est définie comme un polynôme d'ordre 5 :

$$x*(t) = a_5t^5 + a_4t^5 + a_3t^3 + a_2t^2 + a_1t + a_0. \tag{6.15}$$

En supposant que les vitesses et les accélérations initiales et finales sont nulles, Flash et Hogan (1985) ont défini les paramètres de ce polynôme :

$$x(t) = x_0 + (x_0 - x_f)(15\tau^4 - 6\tau^5 - 10\tau^3), \tag{6.16}$$

avec x_0 correspondant à la position initiale, x_f à la position finale et $\tau = \frac{t}{t_f}$. En appliquant cette résolution sur les deux dimensions, on obtient le système suivant :

$$\begin{aligned} x(t) &= x_0 + (x_0 - x_f)(15\tau^4 - 6\tau^5 - 10\tau^3) \\ y(t) &= y_0 + (y_0 - y_f)(15\tau^4 - 6\tau^5 - 10\tau^3), \end{aligned} \tag{6.17}$$

où x_0, y_0 sont les positions initiales et x_f, y_f les positions finales.

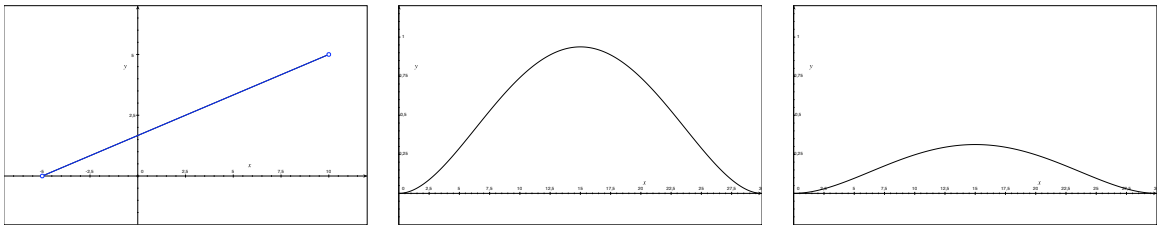


FIG. 6.2 – A gauche, la trajectoire générée par le modèle du *minimum-jerk*. Au centre, le profil de vitesse en x et à droite en y.

La figure 6.2 présente les résultats du modèle *minimum-jerk*. A gauche, on observe la trajectoire générée entre les points $(-5, 0)$ et $(10, 5)$. On observe que la trajectoire produite est une ligne droite. Au milieu et à droite on observe respectivement le profil de vitesse $\dot{x}$ et $\dot{y}$. Ces profils de vitesse sont de type courbe en cloche.

Résultats La figure 6.3 présente les résultats du modèle *minimum-jerk* comparé aux résultats expérimentaux. La tâche consistait à relier le point T_3 au point T_6. On observe (en haut) que pour relier deux points avec vitesse initiale et finale nulle, le modèle produit une ligne droite (trait continu) alors que les sujets ont tendances à tracer des légères courbes (trait en pointillé). On observe aussi les profils de vitesse (en bas à gauche) : les prédictions du modèle sont en accord avec les données expérimentales. Dans les deux cas, le profil de vitesse correspond à une courbe en cloche. En bas à droite sont présentés les courbes d'accélération en x et en y. Bien que la superposition des courbes ne soit pas parfaite, nous retrouvons quand même une certaine similitude entre les prédictions expérimentales et les sujets.

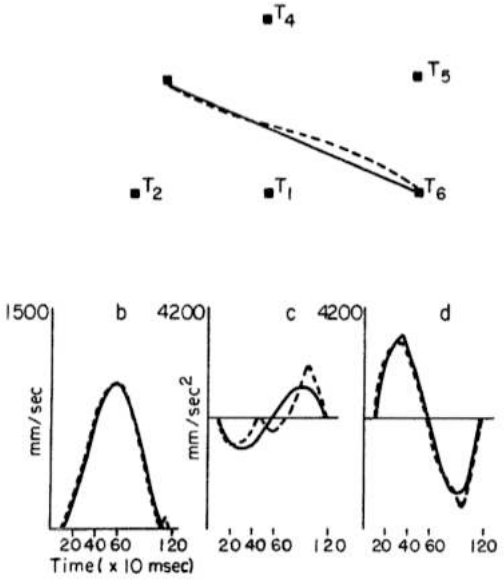


FIG. 6.3 – Courbes prédites par le modèle de minimum-jerk (trait continu) et mesurées expérimentales sur un sujet (en pointillé) des positions de la main (en haut), de la vitesse (en bas à gauche) et des accélérations en x et y (en bas à droite) (figure tirée de Flash et Hogan (1985)).

Todorov et Jordan (1998) ont défini une autre version de ce modèle (que nous ne détaillerons pas) : le modèle du minimum-jerk contraint (*constrained minimum-jerk model*). Ce modèle est une combinaison du modèle du minimum-jerk et de la loi de la puissance 2/3.

6.1.4.2 Minimum-Jerk pour trois points

Nous venons de voir que le modèle du minimum-jerk permettait très bien de reproduire des tâches point à point avec vitesses initiales et finales nulles, c'est-à-dire de produire des lignes droites. Maintenant, nous allons nous intéresser aux trajectoires courbes. Pour cela, Flash et Hogan ont rajouté en consigne de passer par un point intermédiaire :

$$\begin{aligned} x(t_1) &= x_1 \\ y(t_1) &= y_1. \end{aligned} \tag{6.18}$$

On note t_1 le temps de passage au point (x_1,y_1).

Fonction de coût La fonction de coût reste inchangée :

$$J = \frac{1}{2}\int_{t_1}^{t_2} \left(\frac{d^3x}{dt^3}\right)^2 + \left(\frac{d^3y}{dt^3}\right)^2 dt. \tag{6.19}$$

Résolution En apliquant le même principe que précédemment, et en effectuant la résolution par morceaux, la trajectoire optimale correspond à deux polynômes d'ordre 5 (un pour $t \in [0, t_1]$ et un pour $t \in [t_1, T]$.

Afin de déterminer les 6 paramètres du premier polynôme (pour $t \in [0, t_1]$), Flash et Hogan ont défini 6 équations :

$$x(0) = x_0, \quad x(t_1) = x_1, \quad \dot{x}(0) = 0, \quad \dot{x}(t_1^-) = \dot{x}(t_1^+), \quad \ddot{x}(t_1^-) = \ddot{x}(t_1^+), \quad \dddot{x}(t_1^-) = \dddot{x}(t_1^+). \tag{6.20}$$

Ils ont également défini 6 équations pour le second polynôme :

$$x(t_1) = x_1, \quad x(T) = x_f, \quad \dot{x}(T) = 0, \quad \dot{x}(t_1^-) = \dot{x}(t_1^+), \quad \ddot{x}(t_1^-) = \ddot{x}(t_1^+), \quad \dddot{x}(t_1^-) = \dddot{x}(t_1^+). \tag{6.21}$$

La trajectoire optimale obtenue pour $t \leq t_1$ est :

$$\begin{aligned} x^-(\tau) &= \frac{t_f^5}{720}(\pi_1(\tau_1^4(15\tau^4 - 30\tau^3) + \tau_1^3(80\tau^3 - 30\tau^4) \\ &\quad -60\tau^3\tau_1^2 + 30\tau^4\tau_1 - 6\tau^5) + c_1(15\tau^4 - 10\tau^3 - 6\tau^5)) + x_0, \end{aligned} \tag{6.22}$$

et pour $t > t_1$, l'expression est :

$$\begin{aligned} x^+(\tau) &= \frac{t_f^5}{720}(\pi_1(\tau_1^4(15\tau^4 - 30\tau^3 + 30\tau - 15) + \tau_1^3(-30\tau^4 + 80\tau^3 - 60\tau^2 + 10)) \\ &\quad +c_1(-6\tau^5 + 15\tau^4 - 10\tau^3 + 1)) + x_f. \end{aligned} \tag{6.23}$$

Ces équations dépendent des variables $\tau_1 = t_1/t_f$ et $\tau = t/t_f$. Des expressions similaires sont obtenues pour $y^-(\tau)$ et $y^+(\tau)$ avec des coefficients π_2, c_2 remplaçant π_1 et c_1. Ces quatre coefficients π_1, c_1, π_2 et c_2 dépendent des positions initiales et finales ainsi que de la position du point intermédiaire.

Résultats La figure 6.4 présente les résultats pour le modèle du minimum-jerk (à gauche) comparé une nouvelle fois aux données expérimentales d'un sujet (à droite). Dans cet exemple, le point intermédiaire est placé à égale distance entre le point initial et le point final. On observe (en haut) que la trajectoire prédite correspond assez bien à la trajectoire mesurée. On observe de plus que les profils de vitesse (2ème ligne), les profils d'accélération (3ème ligne), et les profils de courbure (en bas) en prédiction et en mesure sont similaires.

Comparaison avec le modèle *Minimum-energy* Partant de l'idée que le muscle ressemble à un ressort, Hollerbach (1977) cherche à déterminer la trajectoire qui consomme le moins d'énergie. Une étude (Veeraklaew *et al.*, 2005) a montré que le modèle *minimum-jerk* est comparable avec le modèle *minimum-energy* pour un système dynamique linéaire. C'est-à-dire que minimiser les secousses revient à minimiser l'énergie et inversement. Les auteurs proposent donc une alternative aux systèmes qui minimisent les secousses : ils proposent d'utiliser des modèles qui minimisent l'énergie sans se soucier de la fluidité du mouvement étant donné que cette dernière est liée à l'énergie.

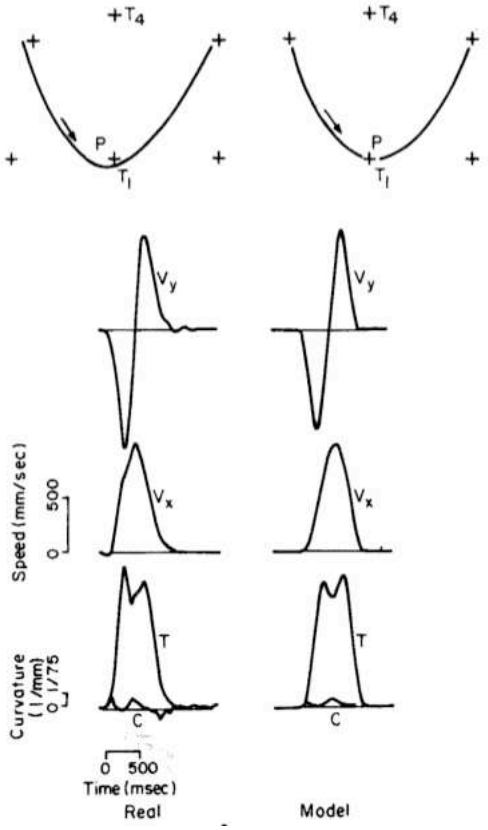


FIG. 6.4 – Courbes prédites par le modèle de minimum-jerk (trait continu) et mesurées expérimentales sur un sujet (en pointillé) des positions de la main (en haut), de la vitesse (en bas à gauche) et des accélérations en x et y (en bas à droite) (figure tirée de Flash et Hogan (1985)).

6.1.4.3 Minimum-Time ou contrôle en Bang-Bang

Jusqu'à présent, nous avons toujours spécifié le temps nécessaire pour effectuer la trajectoire (T). Nous le définissions comme un paramètre du modèle. Le problème de *minimum-time*, très similaire aux problèmes précédents, cherche à définir la trajectoire optimale en minimisant le temps et tout en respectant les limites aux contraintes. Afin de pouvoir résoudre ce problème avec le maximum de Pontryagin, il faut définir une contrainte supplémentaire sur la variable de contrôle (Bryson et Ho, 1975) :

$$|u(t)| \leq u_m, \tag{6.24}$$

avec u_m la « force » maximale que l'on peut appliquer au système. Le principe de maximum de Pontryagin indique que le contrôle optimal pour ce type de contrainte nécessite d'avoir :

$$|u(t)| = u_m, \tag{6.25}$$

où $u(t)$ peut alterner instantanément entre $-u_m$ et u_m. Pour cette raison, ce contrôle prend, dans la communauté robotique, le nom de contrôle en Bang-Bang (une partie de la trajectoire est sous commande maximum u_m et l'autre partie de la trajectoire est sous commande minimale $-u_m$).

Une nouvelle fois, nous nous intéressons aux tâches de pointage. Il a été montré que dans ce cas, il ne peut y avoir qu'une seule alternance entre $-u_m$ et u_m durant toute la trajectoire.

Résolution Lorsque l'on définit la variable de contrôle comme étant l'accélération fournie au modèle, et que cette accélération suit un profil de type bang-bang pour une tâche point à point, la composition de polynôme est la suivante :

$$\begin{aligned} x*(t) &= a_2 t^2 + a_1 t + a_0, \quad 0 \ \leq \ T/2 \\ x*(t) &= b_2 (t-T)^2 + b_1 (t-T) + b_0, \quad T/2 \ \leq \ T, \end{aligned} \tag{6.26}$$

dont les coefficients sont déterminés à partir des contraintes initiales, finales et de continuité en $t = T/2$:

$$\left\{ \begin{array}{l} x(t=0) = x_0, \quad \dot{x}(t=0) = 0 \\ x(t=T) = x_f, \quad \dot{x}(t=T) = 0 \\ x(t=T/2^-) = x(t=T/2^+) \\ \dot{x}(t=T/2^-) = \dot{x}(t=T/2^+) \end{array} \right. \Rightarrow \left\{ \begin{array}{l} a_2 = 2(x_f - x_0)/T^2 \\ a_1 = 0; a_0 = x_0 \\ b_2 = -2(x_f - x_0)/T^2 \\ b_1 = 0; b_0 = x_f. \end{array} \right.$$

Il est donc possible de calculer les vitesses et accélérations optimales au cours du temps et on peut en déduire l'expression optimale du temps T en fonction de la donnée des contraintes de vitesse et d'accélération :

$$T = max \left(\frac{2|x_f - x_0|}{\dot{x}_{max}}, 2\sqrt{\frac{|x_f - x_0|}{\ddot{x}_{max}}} \right). \tag{6.27}$$

Résultats La figure 6.5 présente la position au cours du temps pour une trajectoire point à point ainsi que les profils de vitesse et d'accélération. On observe que le profil d'accélération

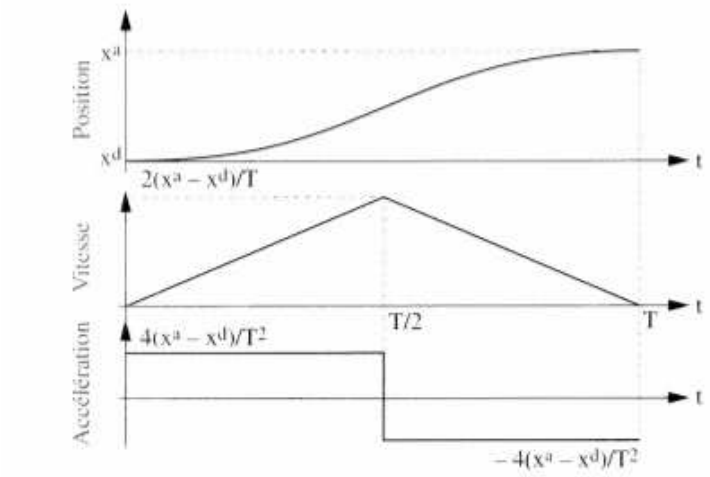


FIG. 6.5 – Mouvement point à point à profil d'accélération bang-bang (figure tirée de Khalil (2002)).

est bien de type bang-bang (avec deux plateaux et un changement abrupt à mi-parcours). Le

profil de vitesse correspond à un triangle avec une vitesse linéaire croissante puis décroissante. En haut de la figure, on observe la trajectoire en x au cours du temps. Si l'on effectue le même contrôle en y au cours du temps et que l'on trace la courbe y en fonction de x, on obtient une droite.

Cependant le contrôle de type bang-bang n'est pas forcément approprié pour reproduire le contrôle moteur humain. Tout d'abord, ce type de contrôle ne reproduit pas forcément les formes de trajectoires observées expérimentalement (par exemple, les profils de vitesse ne correspondent pas à des courbes en cloche). De plus, effectuer de telles alternances en accélération peut être très coûteux en énergie (Harris, 2004).

6.1.5 Modèles dans l'espace articulaire

Le modèle du minimum-jerk est invariant aux rotations et aux translations. En effet, quelque soit la position initiale et la position finale dans le plan de travail, la droite ou la courbe prédite sera identique. Or, des études expérimentales (Morasso, 1981; Osu *et al.*, 1997) ont montré que ce n'était pas le cas chez les sujets. La figure 6.6 présente les traces expérimentales obtenues lorsque les sujets devaient relier deux points.

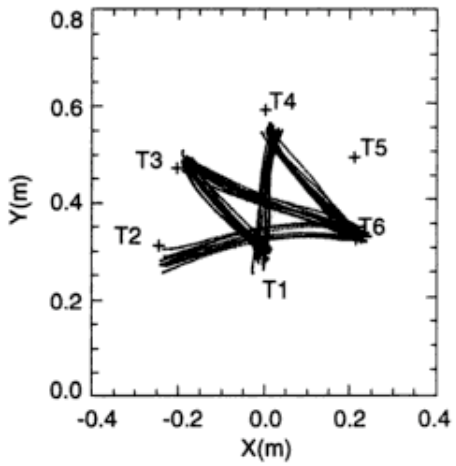


FIG. 6.6 – Trajectoires correspondantes à des tâches de pointage avec différentes orientations (figure tirée de Zelaznik (1996)).

On observe par exemple, qu'entre les points T_1 et T_4 (mouvement haut bas dans la feuille) les sujets ont eu tendance à tracer une droite alors qu'ils ont plutôt tracé des courbes entre les points T_2 et T_6 (mouvement gauche droite dans la feuille). C'est pourquoi, de nouveaux modèles basés sur une représentation articulaire ont émergé dans la littérature (Shadmehr et Mussa-Ivaldi, 1994). Nous allons étudier par la suite, le modèle *minimum-torque* (couple minimum) et le modèle *minimum-variance*, tous deux basés sur une représentation articulaire.

6.1.5.1 Minimum-torque

Le repère utilisé pour le modèle du minimum-jerk est un repère extracorporel. Le modèle que nous allons étudier maintenant (*minimum-torque* (Uno *et al.*, 1989; Nakano *et al.*, 1999))

se base sur des coordonnées corporelles. La trajectoire produite est définie en terme de coordonnées corporelles et plus précisement à partir des angles des articulations et de la longueur des muscles. Ce modèle génère donc une commande motrice qui est le couple fourni par les muscles.

Fonction de coût Le modèle du minimum-torque permet de tracer des trajectoires point à point. Le coût à minimiser est :

$$C_T = \frac{1}{2}\int_0^T \sum_{i=1}^{n} (\frac{dz_i}{dt})^2 dt, \tag{6.28}$$

avec z_i qui correspond au couple à l'articulation i.

Etant donné que le système musculo-squeletal est très difficile à décrire exactement, ils utilisent les équations ci-dessous d'un bras robotique à deux articulations illustré figure 6.7.

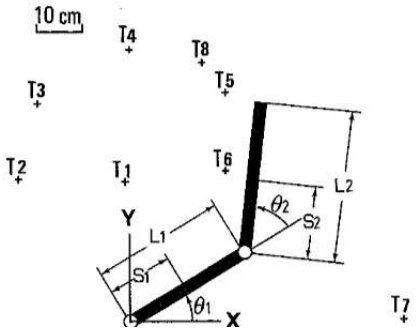


FIG. 6.7 – Un bras robotique à deux articulations qui bouge dans un plan horizontal. L'origine des coordonnées X et Y représente la localisation de la première articulation, c'est-à-dire l'épaule (figure tirée de Uno *et al.* (1989)).

$$\begin{aligned} z_1 &= (I_1 + I_2 + 2M_2L_1S_2cos\theta_2 + M_2(L_1)^2)\frac{d^2\theta_1}{dt^2} \\ &\quad +(I_2 + M_2L_1S_2cos\theta_2)\frac{d^2\theta_2}{dt^2} \\ &\quad -M_2L_1S_2(2\frac{d\theta_1}{dt} + \frac{d\theta_2}{dt})\frac{d\theta_2}{dt}sin\theta_2 + b_1\frac{d\theta_1}{dt} \qquad (6.29) \\ z_2 &= (I_2 + M_2L_1S_2cos\theta_2)\frac{d^2\theta_1}{dt^2} + I_2\frac{d^2\theta_2}{dt^2} \\ &\quad M_2L_1S_2(\frac{d\theta_1}{dt})^2 sin\theta_2 + b_2\frac{d\theta_2}{dt}. \qquad (6.30) \end{aligned}$$

M_i, L_i, S_i and I_i représentent respectivement la masse, la longueur, la distance du centre de la masse à l'articulation et le moment d'inertie. b_i et z_i représentent les coefficients de frottements et le couple à l'articulation i. Les angles des articulations sont définis selon la figure 6.7.

Résolution Le problème de trouver une unique trajectoire qui minimise C_T est un problème d'optimisation non linéaire. Par conséquent, il est impossible d'obtenir des expressions analytiques comme celles du minimum-jerk. Cependant, en utilisant un apprentissage itératif, il est possible de déterminer la trajectoire optimale.

Résultats La figure 6.8 présente les résultats de ce modèle. Les trajectoires sont pratiquement rectilignes et le profil de vitesse est bien une courbe en cloche. On observe aussi que la forme de la trajectoire dépend de la région dans laquelle la main a bougé. Ceci est dû au fait que ce modèle est formulé sur la dynamique du bras.

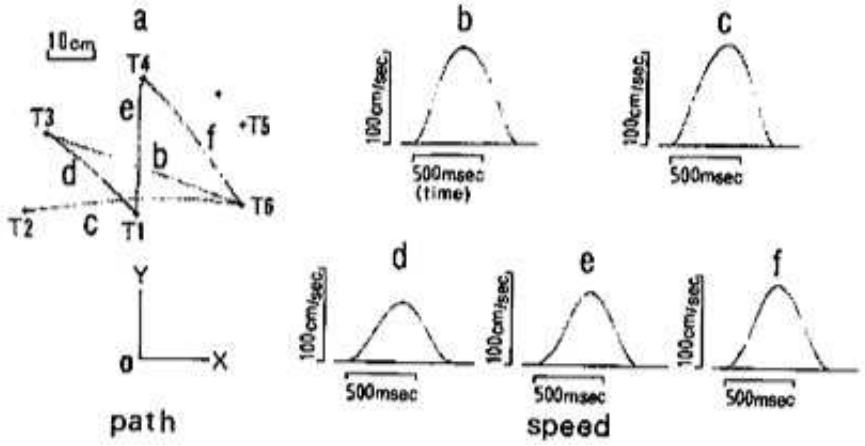


FIG. 6.8 – Mouvements libres entres deux points. A gauche, les 5 trajectoires prédites par le modèle minimum-torque et à droite les profils de vitesses correspondants (figure tirée de (Uno *et al.*, 1989)).

Ces résultats ont été comparés qualitativement à des données expérimentales réalisées par 7 sujets. On observe figure 6.9 les trajectoires générées. Bien que quelques sujets aient parfois généré des lignes droites, les chemins tracés sont en général convexes.

Le modèle *minimum-torque* reproduit avec succès les trajectoires générées sous plusieurs conditions :

- les mouvements libres (comme présentés figure 6.8),
- les mouvements avec des points de passages,
- les mouvements contraints avec des forces extérieures.

Chacun de ces mouvements est généré en minimisant la fonction de coût C_T. Un avantage à minimiser une telle fonction est que le système génère des couples minimaux et donc une trajectoire fluide. Un tel contrôle va donc réduire l'usure du système musculo-squelettique. De plus, la consommation d'énergie est relativement faible parce que les forces non nécessaires aux mouvements sont évitées.

6.1.5.2 Modèle de minimum de variance

Dans cette section, nous allons étudier un modèle qui n'est pas basé sur un coût relatif au mouvement comme les modèles précédents : le modèle *minimum-variance* introduit les notion de :

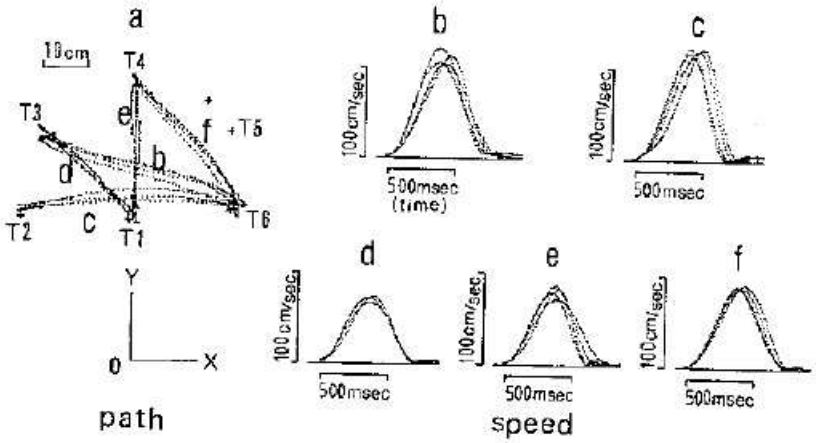


FIG. 6.9 – Mouvements libres entres deux points. A gauche, les 5 trajectoires générées par les sujets et à droite les profils de vitesses correspondants (figure tirée de (Uno *et al.*, 1989)).

– bruit dépendant du signal,
– variance de la position finale à minimiser.

Le bruit résultant de l'activation des neurones moteurs provoque des déviations dans les trajectoires par rapport au chemin désiré. Ces déviations sont accumulées au cours du temps, amenant à une variation de la position finale. Si le bruit était indépendant du mouvement, alors il suffirait d'effectuer le mouvement le plus rapidement possible. Or, ce type de bruit dépend du niveau moyen du signal et une solution est donc de ralentir le signal pour augmenter la précision ce qui revient à choisir un profil de vitesse pour le signal symétrique et fluide (profil de vitesse de type courbe en cloche). Il est donc possible de supposer que nous ne bougeons pas de façon fluide par chance ou nécessité mais que cela est dû au bruit des neurones moteurs (Sejnowski, 1998).

Le modèle de minimum de variance (Harris et Wolpert, 1998) minimise la variance de la position finale. Cette variance, cependant, dépend de l'amplitude de la commande motrice : l'optimisation est donc réalisée de façon dynamique.

Fonction de coût Le système moteur utilisé dans le modèle *minimum-torque* est un système du second ordre linéaire dont la dynamique est spécifiée par :

$$\ddot{\theta}(t) + a_1\dot{\theta}(t) + a_0\theta(t) = \tau(t). \tag{6.31}$$

Cette équation est basée sur la mécanique Newtonienne et a été utilisée pour modéliser, entre autres, les mouvements d'une articulation du bras (Hogan, 1984). a_0 et a_1 représentent respectivement les constantes d'élasticité et de frottements. θ correspond à la variable d'état de l'angle de l'articulation et $\tau(t)$ est le couple musculaire. Cette équation différentielle du second ordre peut être réduite à une équation du premier ordre en introduisant un vecteur de représentation des états :

$$x \equiv \begin{pmatrix} \theta \\ \dot{\theta} \end{pmatrix}. \tag{6.32}$$

L'équation 6.31 devient donc

$$\dot{x}(t) = Ax(t) + B\tau(t). \tag{6.33}$$

où les matrices A et B dont définies par :

$$A \equiv \begin{pmatrix} 0 & 1 \\ -a_0 & -a_1 \end{pmatrix},\ B \equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{6.34}$$

D'apres Harris et Wolpert (1998), un terme ξ relatif au bruit (dépendant de l'amplitude du mouvement) doit être ajouté à l'équation 6.35 :

$$\dot{x}(t) = Ax(t) + B[\tau(t) + \xi(t)]. \tag{6.35}$$

ξ suit une distribution gaussienne de moyenne zéro et de variance proportionnelle au carré du signal de contrôle τ :

$$E[\xi(t)] = 0, \qquad E[\xi(t)\xi(t')] = K\tau^2(t)\delta(t-t'), \tag{6.36}$$

où K est une constante proportionnelle à l'amplitude du bruit. La notation $E[.]$ correspond à l'espérance. La valeur du vecteur d'état et sa matrice de covariance sont définis par :

$$E[x(t)] = e^{At}x_0 + \int_0^t e^{A(t-t')}B\tau(t')dt', \tag{6.37}$$

$$Cov[x(t)] = K\int_0^t e^{A(t-t')}B(e^{A(t-t')}B)^T\tau^2(t')dt'. \tag{6.38}$$

Si l'on considère un mouvement sur un intervalle $0 \leq t \leq t_f$, la fonction de coût à minimiser est la variance de la somme des positions sur une certaine durée post-mouvement $[t_f, t_f + t_p]$. Nous considérons dans un premier cas que le coût est une fonction de la variance à la position finale t_f uniquement. La contrainte d'optimisation peut être résolue en introduisant un coefficient de Lagrange μ avec la fonction de coût décrite ci-dessous :

$$C_{MV} = K\int_0^{t_f} f(t;t_f)\tau^2(t)dt - \mu^T(E[x(t_f)] - x_f), \tag{6.39}$$

où $f(t;t_f)$ est un facteur de poids qui détermine l'influence du bruit sur la variation finale de la position.

$$f(t;t_f) \equiv [e^{A(t-t')}B(e^{A(t-t')}B)^T]. \tag{6.40}$$

Résolution Une solution analytique est obtenue pour le contrôle optimal :

$$\tau(t) = \frac{\mu^T e^{A(t_f-t)}B}{2Kf(t;t_f)}. \tag{6.41}$$

Le vecteur de Lagrange μ est une constante qui peut être déterminée par l'équation 6.39 et par la condition $E[x(t_f)] = x_f$.

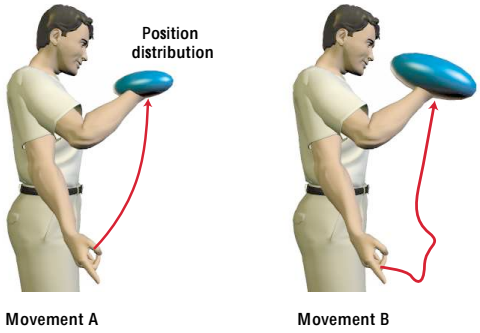


FIG. 6.10 – Les trajectoires sont représentées en rouge et la variabilité de la position finale par des ellipses bleues. La variabilité de la position finale de la trajectoire A est plus petite que celle de la trajectoire B. Cette trajectoire est donc moins coûteuse si l'on considère le modèle de minimum-variance (figure tirée de Wolpert et Ghahramani (2000)).

Résultats La figure 6.10 illustre ce modèle. Bien que les deux trajectoires permettent à la main de se positionner approximativement dans une position finale similaire, la distribution de la variabilité de la position finale (ellipses bleues) est différente. Ceci est due au bruit présent au niveau des commandes motrices. Nous observons que le mouvement A est moins coûteux que le mouvement B vu qu'il produit une position finale moins dispersée. Etant donné que le modèle de minimum-variance minimise la variabilité de la position finale, c'est la trajectoire A qui va être choisie par rapport à ce critère.

La figure 6.11 présente les résultats de génération de trajectoire point-à-point. A gauche, on observe les données expérimentales issues de l'expérience de Uno *et al.* (1989) que nous avons déjà présentés figure 6.9 et à droite les trajectoires prédites par le modèle. On observe que l'allure générale des trajectoires générées est similaire aux trajectoires expérimentales. De plus, la génération de la trajectoire dépend de la région dans laquelle le bras évolue.

D'autres versions de ce modèle ont été proposées tel que le modèle de Matsui minimisant un critère basé sur la fluidité de la trajectoire et le bruit dépendant du signal (Matsui et Wada, 2003). Un autre modèle appelé TOPS (*Task Optimization in the Presence of Signal-Dependent Noise*) (Miyamoto *et al.*, 2003, 2004) combine quant à lui un critère basé sur le bruit et sur la consommation d'énergie.

6.1.6 Confrontation des modèles

Dans cette section, nous allons comparer de manière qualitative les modèles de génération de trajectoires. Nous allons tout d'abord nous intéresser à comparer les deux modèles dynamiques entre eux. Une fois cette comparaison effectuée, nous décrirons brièvement un grand problème présent dans la littérature qui est de déterminer l'espace dans lequel est effectué le contrôle : espace de la tâche ou espace dynamique. C'est cette dernière question que nous traiterons avant de présenter le contrôle moteur choisi dans notre modèle.

6.1.6.1 Comparaison des modèles dynamiques

Nous avons vu figure 6.8 et figure 6.11 que les modèles *minimum-torque* et *minimum-variance* font des prédictions expérimentales similaires dans les trajectoires point à point.

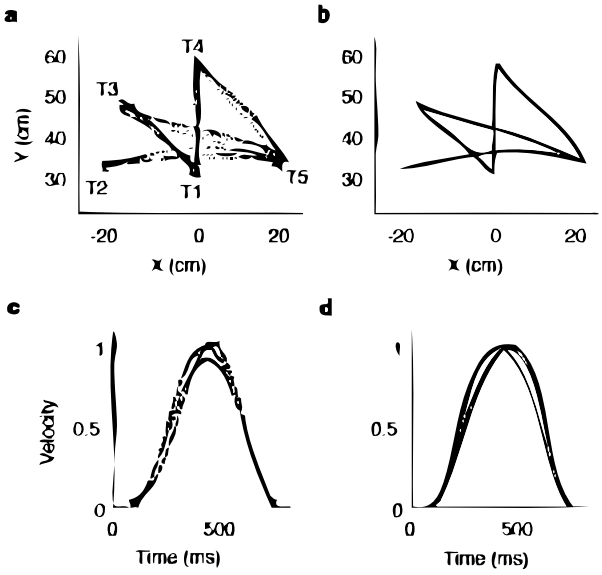


FIG. 6.11 – Comparaison des données expérimentales observées et prédites par le modèle de minimum de variance. A gauche, trajectoires et profils de vitesse observés (issues de Uno *et al.* (1989)), à droite, trajectoires et profils de vitesse prédits par le modèle de minimum-variance (figure tirée de (Harris et Wolpert, 1998)).

Nous retrouvons ces résultats figure 6.12 où sont présentés pour une trajectoire point à point, les chemins et profils de vitesse générés par les modèles *minimum-jerk*, *minimum-torque* et *minimum-variance* et implémentés sur un robot à deux degrés de liberté.

La similitude des deux trajectoires des modèles dynamiques (*minimum-torque* et *minimum-variance*) est d'autant plus surprenante que les deux modèles ont des stratégies d'optimisation basées sur des critères totalement différents : le modèle *minimum-torque* garantit la fluidité des couples musculaires durant le mouvement alors que le modèle *minimum-variance* est basé sur la précision de la position à la fin de la trajectoire.Tanaka *et al.* (2004) ont analysé les situations dans lesquelles ces deux modèles font des prédictions expérimentales qui divergent : la principale différence entre ces modèles vient du fait que le profil de vitesse prédit par le modèle *minimum-variance* peut être asymétrique lorsque l'on change la durée du mouvement ou le rapport entre les forces de frottements et élastiques. Par exemple, le modèle *minimum-variance* appliqué au mouvement oculaire permet d'obtenir un profil de vitesse asymétrique dû à la durée du mouvement. Ce résultat est plausible au vu des données expérimentales observées (Harris et Wolpert, 1998).

6.1.6.2 Modèles dynamiques ou modèles cartésiens

Nous avons déjà analysé la principale différence entre les modèles *minimum-jerk* et *minimum-torque*. Dans le premier cas et pour les trajectoires point à point, la trajectoire générée est toujours une ligne droite alors que dans le second cas, la trajectoire dépend de la position initiale et finale des points. Il en est de même entre les modèles *minimum-jerk* et *minimum-variance*. Cependant, dans certains cas, il n'est pas possible de distinguer les trajectoires issues d'une planification dans l'espace cartésien et d'une planification dans l'espace dynamique (Hol-

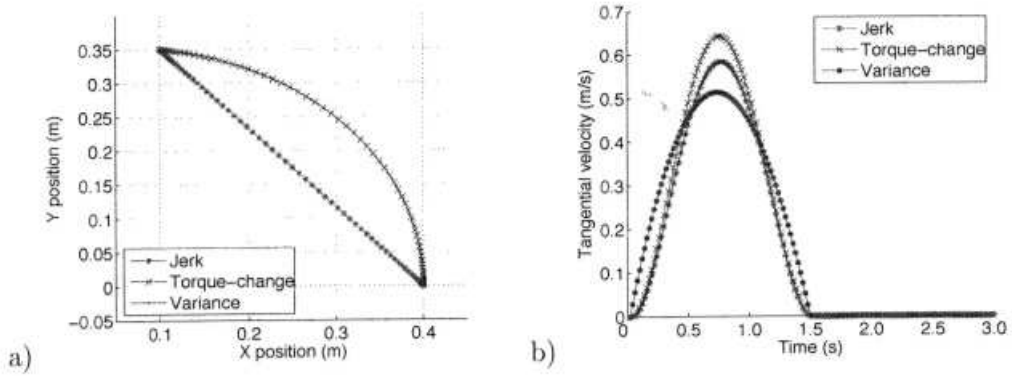


FIG. 6.12 – Comparaison des modèles *minimum-jerk*, *minimum-torque* et *minimum-variance*. A gauche, les trajectoires générées et à droite les profils de vitesse correspondants (figure tirée de Simmons et Demiris (2005)).

lerbach et Atkeson, 1987; Diard, 2009) (par exemple, lorsque la zone à atteindre est en limite de la zone atteignable).

De plus, les données expérimentales observées des tâches de pointage ne sont pas toutes cohérentes. Certaines prédisent une trajectoire rectiligne (Abend *et al.*, 1982; Vindras et Viviani, 1998). Ces données ont été observées lors de tâches contraintes en deux dimensions. D'autres études observent de légères courbes (Soechting et Lacquaniti, 1981). Cependant, cette fois la tâche n'était pas contrainte (trois dimensions).

Desmurget fait l'hypothèse que le système nerveux central est capable d'utiliser différentes stratégies de génération de trajectoires selon le contexte (Desmurget *et al.*, 1998). Par conséquent, toutes les recherches visant à définir un modèle global de génération de trajectoire sont vaines. Il a proposé de tester cette hypothèse et a comparé des trajectoires contraintes et des trajectoires non contraintes lors de tâches de pointage (Desmurget, 1997). Les résultats semblent indiquer que la planification de la trajectoire dépend de la nature de la tâche. Les mouvements non contraints seraient planifiés dans l'espace dynamique alors que les mouvements contraints dans un plan seraient planifiés dans l'espace de la tâche.

Il a de plus suggéré que lorsque le mouvement est contraint en deux dimensions, une représentation implicite (« suivre une ligne droite ») dans l'espace des tâches serait présente.

6.1.7 Contrôle optimal dans notre modèle

Nous faisons l'hypothèse que la génération de trajectoire se fait dans l'espace de la tâche, c'est-à-dire dans l'espace cartésien. Cette hypothèse est en accord avec les résultats de Desmurget (1997) stipulant que lorsque les mouvements sont contraints (ce qui est notre cas ici), la trajectoire est planifiée dans l'espace de la tâche.

De plus, chaque point de contrôle est défini par sa position et sa vitesse. Il nous est donc impossible d'appliquer un modèle type *minimum-jerk* étant donné que nous n'avons aucune information d'accélération. De plus, un modèle de type bang-bang (avec accélération discontinue) ne permet pas de générer des profils de vitesse de type courbe en cloche.

Etant donné que pour chaque segment de lettres (entre deux points de contrôle), nous

avons 4 paramètres (positions et vitesses initiales et finales). Il nous est seulement possible de définir un polynôme de degré 3. C'est pourquoi nous avons choisi d'appliquer le modèle *minimum-acceleration*.

6.1.7.1 Fonction de coût

La fonction de coût correspondante à ce modèle est :

$$J = \frac{1}{2}\int_{t_1}^{t_2} \left(\frac{d^2x}{dt^2}\right)^2 + \left(\frac{d^2y}{dt^2}\right)^2 dt. \tag{6.42}$$

Comme précédemment, la résolution se fait indépendamment sur chaque dimension. Le Lagrangien correspondant pour la dimension x est donc :

$$L(t) = \left(\frac{d^2x}{dt^2}\right)^2. \tag{6.43}$$

6.1.7.2 Résolution

En résolvant les équations d'Euler-Lagrange, nous obtenons que la solution générale est donnée par un polynôme d'ordre 3 :

$$\forall\ t\ \in\ [0,T] \quad x^*(t) = a_3t^3 + a_2t^2 + a_1t + a_0, \tag{6.44}$$

avec 4 paramètres a_3, a_2, a_1 et a_0 à définir. Les contraintes aux limites sont

$$x(0) = x_0, \quad x(T) = x_f, \quad \dot{x}(0) = v_0, \quad \dot{x}(T) = v_f. \tag{6.45}$$

En utilisant les contraintes aux limites, nous déterminons les paramètres du polynôme :

$$\begin{cases} a_3 = \frac{2}{T^3}(x_0 - x_f) + \frac{1}{T^2}(v_f + v_0) \\ a_2 = \frac{1}{T}(-v_f - 2v_0) + \frac{3}{T^2}(x_f - x_0) \\ a_1 = v_0 \\ a_0 = x_0. \end{cases} \tag{6.46}$$

6.1.7.3 Résultats

La figure 6.13 présente plusieurs trajectoires générées par le modèle minimum-accélération avec des vitesses initiales et finales différentes (la position du point initial (2,8) et du point final (10,4) est toujours la même). On observe que la trajectoire dépend des vitesses initiales et finales fournies au modèle. On note que dans le cas où les vitesses initiales et finales sont nulles, c'est-à-dire dans le cas d'une tâche point à point, la trajectoire générée est une ligne droite.

La figure 6.14 présente la trajectoire en fonction du temps, le profil de vitesse, d'accélération et de la dérivée de l'accélération selon la dimension x et en fonction du temps t. On observe que la trajectoire est fluide et que la vitesse est une courbe en cloche.

6.2 Définition du modèle

Dans cette section, nous allons définir les termes du modèle BAP relatifs au modèle moteur : nous allons définir les variables ainsi que les formes paramétriques et leurs paramètres. Dans un premier temps, nous nous concentrerons sur le modèle de génération de trajectoire, que nous définirons de façon probabiliste, puis nous aborderons le contrôle de l'effecteur.

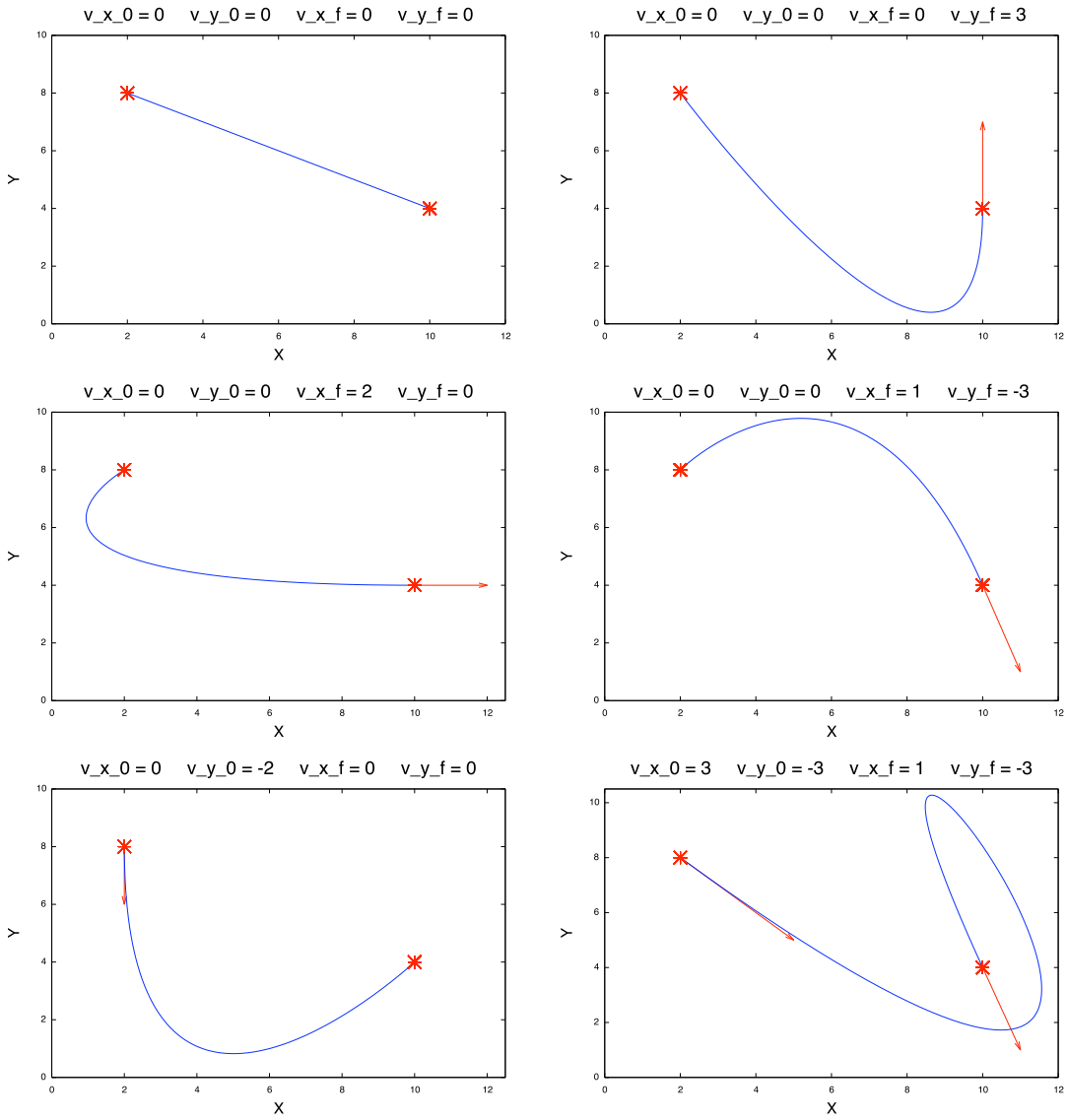


FIG. 6.13 – Trajectoires générées par le modèle minimum-accélération.

6.2.1 Modèle de génération de trajectoires

Le modèle de génération de trajectoire est le symétrique du modèle de perception : ce dernier extrait, à partir d'une trajectoire, un ensemble de points de contrôle alors que le modèle de génération de trajectoires détermine une trajectoire à partir d'un ensemble de points de contrôle.

6.2.1.1 Variables

Les variables correspondant à la trace produite générée (P) sont définies comme un ensemble de variables de positions $\{P_X^{0:T}, P_Y^{0:T}\}$ dans l'espace cartésien. T correspond au nombre maximal de points de la trace générée. Les variables P_x^t et P_y^t sont les positions en x et en y du $t^{\text{ième}}$ point de la trajectoire générée. Ce sont des variables entières à valeur dans $[0 : 40]$.

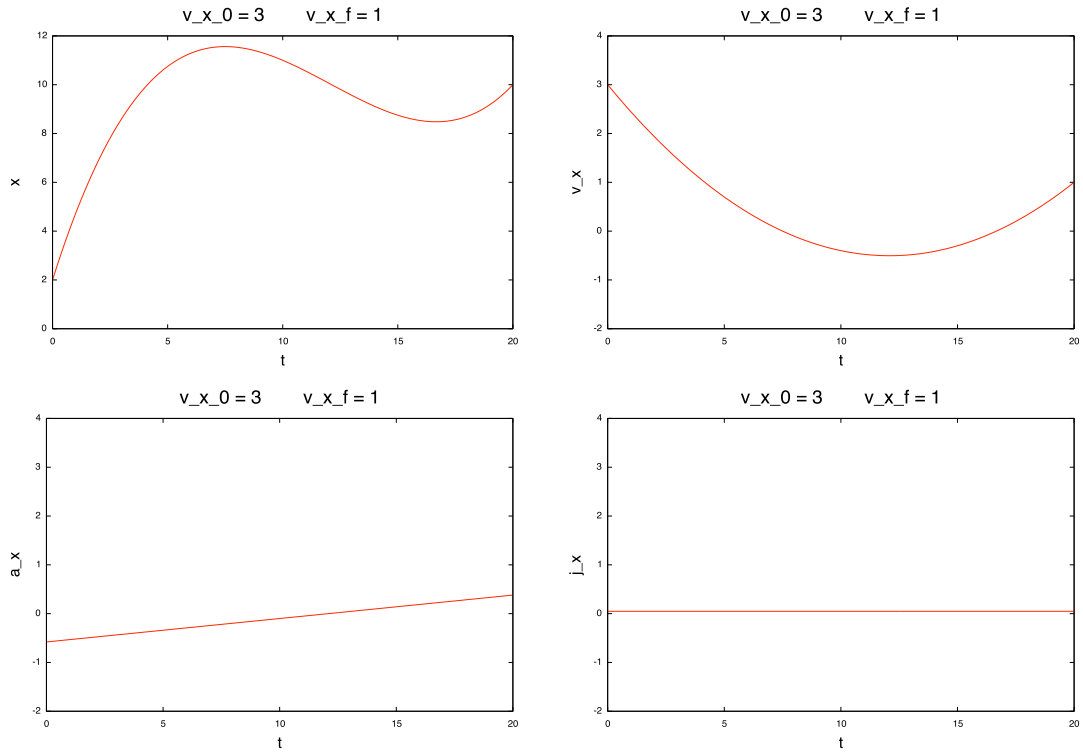


FIG. 6.14 – En haut à gauche, trajectoire en x générée par le modèle minimum-accélération. En haut à droite, profil de vitesse en x de la trajectoire générée. En bas à gauche, le profil d'accélération en x et en bas à droite, le profil du jerk en x.

6.2.1.2 Termes de la décomposition

Le terme correspondant au modèle de génération de trajectoire dans la distribution de probabilité conjointe est :

$$P(P \mid C_P). \tag{6.47}$$

Il exprime la relation entre la trajectoire et les points de contrôle. Avec la définition des variables de représentations internes et de positions, le terme s'exprime de façon équivalente en :

$$P(P_X^{0:T}\ P_Y^{0:T} \mid C_{Px}^{0:N}\ C_{Py}^{0:N}\ C_{P\dot{x}}^{0:N}\ C_{P\dot{y}}^{0:N}). \tag{6.48}$$

Nous faisons l'hypothèse que la trajectoire générée entre deux points de contrôle n et $n+1$ ne dépend que de la position et de la vitesse de ces deux points. Nous obtenons donc l'égalité suivante :

$$\begin{aligned} &P(P_X^{0:T}\ P_Y^{0:T} \mid C_{Px}^{0:N}\ C_{Py}^{0:N}\ C_{P\dot{x}}^{0:N}\ C_{P\dot{y}}^{0:N}) \\ &= \prod_{n=0}^{N-1} P(P_x^{Kn:K(n+1)}\ P_y^{Kn:K(n+1)} \mid C_{Px}^{n}\ C_{Py}^{n}\ C_{P\dot{x}}^{n}\ C_{P\dot{y}}^{n}\ C_{Px}^{n+1}\ C_{Py}^{n+1}\ C_{P\dot{x}}^{n+1}\ C_{P\dot{y}}^{n+1}), \end{aligned} \tag{6.49}$$

avec K le nombre de points générés entre deux points de contrôle ($KN \leq T$). Quelque soit la paire de points de contrôle, le nombre de points intermédiaires est toujours le même (K constant).

6.2.1.3 Formes paramétriques

Ce terme est associé à une loi Dirac et vaut 1 pour l'ensemble de valeur $\{p_x^{Kn:K(n+1)}, p_y^{Kn:K(n+1)}\}$ défini par une fonction déterministe g qui prend en entrée deux points de contrôle :

$$P\left(\begin{array}{c} P_X^{Kn:K(n+1)} \\ P_Y^{Kn:K(n+1)} \end{array} \middle| \begin{array}{cccc} C_{Px}^n & C_{Py}^n & C_{P\dot{x}}^n & C_{P\dot{y}}^n \\ C_{Px}^{n+1} & C_{Py}^{n+1} & C_{P\dot{x}}^{n+1} & C_{P\dot{y}}^{n+1} \end{array}\right) = \left\{\begin{array}{l} 1 \text{ si } \left\{ p_x^{Kn:K(n+1)}, p_y^{Kn:K(n+1)} \right\} = g(c_P^n, c_P^{n+1}) \\ 0 \text{ sinon.} \end{array}\right. \quad (6.50)$$

Cette fonction g est définie par l'algorithme de génération de trajectoire tel que nous l'avons défini section 6.1.7.

6.2.2 Modèle de l'effecteur

Dans un premier temps, nous allons utiliser un effecteur simulé à deux degrés de libertés comme le présente la figure 6.15. Il s'agit d'un bras dans un plan à deux articulations, une qui représente l'articulation de l'épaule (en rouge) et l'autre du coude (en vert).

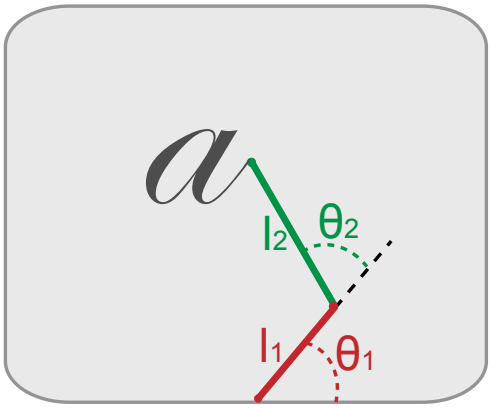


FIG. 6.15 – Modélisation du bras composé de deux articulations (l'épaule et le coude).

l_1 et l_2 correspondent à la longueur des deux segments du bras. En pratique, nous utilisons les mêmes données qu' Uno *et al.* (1989), c'est-à-dire $l_1 = 25cm$ et $l_2 = 35cm$. La variable θ_1 correspond à l'angle entre l'horizontale et le premier segment du bras et θ_2 à l'angle entre les deux segments. Le bras sera contrôlé en accélération.

6.2.2.1 Variables

La variable correspondant à l'effecteur E est un ensemble de variables :

$$E = \{\theta_1^{0:T},\ \theta_2^{0:T},\ \dot{\theta}_1^{0:T},\ \dot{\theta}_2^{0:T},\ \ddot{\theta}_1^{0:T},\ \ddot{\theta}_2^{0:T}\}. \quad (6.51)$$

Les variables θ_1 et θ_2 ont été définies ci-dessus. Leurs valeurs sont comprises entre -0.05 et 1.00 radians (cet espace est suffisant pour atteindre n'importe quel point de la grille 40×40 de l'espace cartésien). Les variables $\dot{\theta}_1$, $\dot{\theta}_2$ et $\ddot{\theta}_1$, $\ddot{\theta}_2$ correspondent aux vitesses articulaires et accélérations articulaires. Elles sont définies sur $[-0.5; 0.5]$ pour les vitesses et $[-0.5; 0.5]$ pour les accélérations.

6.2.2.2 Termes de la décomposition

Le terme de la décomposition correspondant à l'effecteur est :

$$P(\ E\ \mid\ P\ ). \tag{6.52}$$

Avec notre définition des variables, nous obtenons :

$$P(\theta_1^{0:T}\ \theta_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\mid P_X^{0:T}\ P_Y^{0:T}). \tag{6.53}$$

Nous faisons l'hypothèse que l'accélération articulaire est indépendante de la position du bout de l'effecteur si l'on suppose connu la vitesse articulaire. En appliquant cette hypothèse nous obtenons :

$$\begin{aligned} &P(\theta_1^{0:T}\ \theta_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\mid P_X^{0:T}\ P_Y^{0:T}) \\ &= \left(\begin{array}{l} P(\theta_1^{0:T}\ \theta_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\mid P_X^{0:T}\ P_Y^{0:T}) \\ P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}) \end{array}\right). \end{aligned} \tag{6.54}$$

Nous nommerons par la suite *modèle cinématique* le terme $P(\theta_1^{0:T}\ \theta_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\mid P_X^{0:T}\ P_Y^{0:T})$ et *contrôle de l'effecteur* le terme $P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T})$ (le contrôle de l'effecteur ne correspond pas à un vrai modèle dynamique car nous ne considérons ni masse, ni forces, ni matrices d'inertie).

Modèle cinématique Le modèle cinématique définit la conversion des coordonnées de l'espace cartésien dans l'espace articulaire. Nous faisons l'hypothèse que les couples positions et vitesses sont indépendants les uns des autres en fonction du temps. Nous obtenons donc :

$$\begin{aligned} &P(\theta_1^{0:T}\ \theta_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\mid P_X^{0:T}\ P_Y^{0:T}) \\ &\quad= \prod_{t=0}^{T} P(\theta_1^t\ \theta_2^t\ \dot{\theta}_1^t\ \dot{\theta}_2^t\mid P_X^t\ P_Y^t). \end{aligned} \tag{6.55}$$

Chaque terme est associé à un loi Dirac et vaut 1 pour l'ensemble de valeur $\{\theta_1^t\ \theta_2^t\ \dot{\theta}_1^t\ \dot{\theta}_2^t\}$ défini par une fonction déterministe h :

$$P\left(\theta_1^t\ \theta_2^t\ \dot{\theta}_1^t\ \dot{\theta}_2^t\ \mid\ P_X^t\ P_Y^t\right) = \left\{\begin{array}{l} 1 \text{ si } \{\ \theta_1^t,\ \theta_2^t,\ \dot{\theta}_1^t,\ \dot{\theta}_2^t\ \} = h(p_x^t,\ p_y^t) \\ 0 \text{ sinon.} \end{array}\right. \tag{6.56}$$

La fonction h est définie par les équations suivantes qui décrivent la transformation cinématique directe :

$$h(p_x,p_y) = \left(\begin{array}{l}\theta_1\\ \theta_2\end{array}\right) = \left(\begin{array}{l} \tan^{-1}\left(\frac{p_y}{p_x}\right) - \tan^{-1}\left(\frac{l_2\sin(\theta_2)}{l_1+(l_2\cos(\theta_2))}\right) \\ \cos^{-1}\left(\frac{p_x^2+p_y^2-l_1^2-l_2^2}{2l_1l_2}\right)\end{array}\right), \tag{6.57}$$

$$\ddot{\theta_1}^t = h_2(\dot{\theta}_1^t,\ \dot{\theta}_1^{t+1}) = \frac{\dot{\theta}_1^{t+1}-\dot{\theta}_1^t}{\Delta t}, \tag{6.58}$$

$$\ddot{\theta_2}^t = h_2(\dot{\theta}_2^t,\ \dot{\theta}_2^{t+1}) = \frac{\dot{\theta}_2^{t+1}-\dot{\theta}_2^t}{\Delta t}. \tag{6.59}$$

Contrôle de l'effecteur Le modèle de contrôle de l'effecteur permet à partir d'une séquence de vitesses de définir la séquence d'accélération correspondante. Nous faisons l'hypothèse que l'accélération au temps t ne dépend que des vitesses au temps t et $t+1$.

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}) \\ & = \prod_{t=0}^{T-1} P(\ddot{\theta}_1^t\ \ddot{\theta}_2^t \mid \dot{\theta}_1^t\ \dot{\theta}_2^t\ \dot{\theta}_1^{t+1}\ \dot{\theta}_2^{t+1}). \end{aligned} \tag{6.60}$$

De plus, nous faisons une hypothèse d'indépendance entre les deux dimensions :

$$\begin{aligned} & \prod_{t=0}^{T-1} P(\ddot{\theta}_1^t\ \ddot{\theta}_2^t \mid \dot{\theta}_1^t\ \dot{\theta}_2^t\ \dot{\theta}_1^{t+1}\ \dot{\theta}_2^{t+1}) \\ & = \prod_{t=0}^{T-1} \begin{pmatrix} P(\ddot{\theta}_1^t \mid \dot{\theta}_1^t\ \dot{\theta}_1^{t+1}) \\ P(\ddot{\theta}_2^t \mid \dot{\theta}_2^t\ \dot{\theta}_2^{t+1}) \end{pmatrix}. \end{aligned} \tag{6.61}$$

Le terme $P(\ddot{\theta}_1^t \mid \dot{\theta}_1^t\ \dot{\theta}_1^{t+1})$ est associé à une loi Dirac qui vaut 1 pour la valeur $\ddot{\theta}_1^t$ définie par une fonction déterministe j :

$$\ddot{\theta_1}^t = j(\dot{\theta}_1^t,\ \dot{\theta}_1^{t+1}) = \frac{\dot{\theta}_1^{t+1} - \dot{\theta}_1^t}{\Delta t}. \tag{6.62}$$

Il en est de même pour $\ddot{\theta}_2^t$.

6.3 Tâche d'écriture

6.3.1 Question et inférence

La question relative à l'écriture est la suivante :

Quelles sont les accélérations à appliquer au bras pour qu'il trace une lettre donnée selon le style d'écriture d'un scripteur ?

En termes probabilistes, cette question correspond à l'équation 6.63 :

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1]). \tag{6.63}$$

L'inférence bayésienne permet de répondre à cette question. En appliquant la règle de Bayes, nous obtenons :

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1]) \\ & = \frac{P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\ [L = l]\ [W = w]\ [\lambda_P = 1])}{P([L = l]\ [W = w]\ [\lambda_P = 1])}. \end{aligned} \tag{6.64}$$

Le dénominateur ne dépend pas des valeurs des variables que l'on recherche et n'est donc qu'un facteur de normalisation. Nous obtenons donc :

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1]) \\ & = \frac{1}{Z_1} P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\ [L = l]\ [W = w]\ [\lambda_P = 1]). \end{aligned} \tag{6.65}$$

Pour le numérateur, nous effectuons une sommation sur les variables manquantes (règle de marginalisation). Nous obtenons une nouvelle expression :

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]) \\ & \quad = \quad \frac{1}{Z_1}\sum_A P\left(\begin{array}{l} \ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T}\ V\ [L=l]\ [W=w]\ [\lambda_P=1] \\ C_{LV}\ C_V\ C_{LP}\ C_{LS}\ \lambda_v\ \lambda_s\ \lambda_l\ C_P\ C_S\ P\ S \end{array}\right) \\ & \quad avec \quad A = \{C_{LV}\ C_V\ C_{LP}\ C_{LS}\ \lambda_v\ \lambda_s\ \lambda_l\ C_P\ C_S\ P\ S\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T}\ V\}. \end{aligned} \tag{6.66}$$

Puis nous appliquons la décomposition que nous avons présentée équation 4.53. Le numérateur correspond donc maintenant à un produit de seize termes.

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]) \\ & = \frac{1}{Z_1}\sum_A \left(\begin{array}{l} P(C_{LV} \mid [L=l]\ [W=w])\ P(\lambda_V \mid C_{LV}\ C_V)\ P([L=l])\ P([W=w]) \\ P(C_V \mid V)\ P(V)\ P(\lambda_l \mid C_{LV}\ C_{LP})\ P(S \mid P)\ P(P)\ P(\lambda_S \mid C_{LS}\ C_S) \\ P(C_{LP} \mid [L=l]\ [W=w])\ P(C_{LS} \mid [L=l]\ [W=w])\ P([\lambda_P=1] \mid C_{LP}\ C_P) \\ P(P \mid C_P)\ P(C_S \mid S)\ P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T} \mid P) \end{array}\right). \end{aligned} \tag{6.67}$$

Nous avons vu section 4.1.3 qu'étant donné que λ_V, λ_S et λ_L ne sont pas définis, les modèles de perception et de simulation sont désactivés et peuvent être supprimés de l'équation. Nous obtenons donc :

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]) \\ & \quad = \quad \frac{1}{Z_1}\sum_B \left(\begin{array}{l} P([L=l])\ P([W=w])\ P(P) \\ P(C_{LP} \mid [L=l]\ [W=w])\ P([\lambda_P=1] \mid C_{LP}\ C_P) \\ P(P \mid C_P)\ P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T} \mid P) \end{array}\right) \\ & \quad avec \quad B = \{C_{LP}\ C_P\ P\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T}\ \}. \end{aligned} \tag{6.68}$$

Etant donné que le terme $P(P \mid C_P)$ est défini par une loi Dirac, il n'y a qu'une seule valeur de P telle que ce terme soit non nul. La sommation sur cette variable disparaît donc :

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]) \\ & \quad = \quad \frac{1}{Z_1}\sum_D \left(\begin{array}{l} P([L=l])\ P([W=w])\ P([P=g(C_P)]) \\ P(C_{LP} \mid [L=l]\ [W=w])\ P([\lambda_P=1] \mid C_{LP}\ C_P) \\ P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T} \mid [P=g(C_P)]) \end{array}\right) \\ & \quad avec \quad D = \{C_{LP}\ C_P\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T}\ \}. \end{aligned} \tag{6.69}$$

D'après la définition du terme $P([\lambda_P=1] \mid C_{LP}\ C_P)$, nous obtenons :

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]) \\ & \quad = \quad \frac{1}{Z_1}\sum_F \left(\begin{array}{l} P([L=l])\ P([W=w])\ P([P=g(C_{LP})]) \\ P(C_{LP} \mid [L=l]\ [W=w]) \\ P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T} \mid [P=g(C_{LP})]) \end{array}\right) \\ & \quad avec \quad F = \{C_{LP}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T}\ \}. \end{aligned} \tag{6.70}$$

Nous avons fait l'hypothèse que l'accélération articulaire était indépendante de la position de l'effecteur si l'on suppose la vitesse articulaire connue. En appliquant cette hypothèse nous obtenons :

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]) \\ & \quad = \quad \frac{1}{Z_1}\sum_F \left(\begin{array}{l} P([L=l])\ P([W=w])\ P([P=g(C_{LP})])\ P(C_{LP} \mid [L=l]\ [W=w]) \\ P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}) P(\dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T} \mid [P=g(C_{LP})]) \end{array}\right) \\ & \quad avec \quad F = \{C_{LP}\ \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T}\ \}. \end{aligned} \tag{6.71}$$

Etant donné que les termes $P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T})$ et $P(\dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T} \mid [P = f(C_{LP})])$ suivent des lois Dirac, il n'y a donc qu'une seule trajectoire possible pour un ensemble de points de contrôle.

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1]) \\ = \ & \frac{1}{Z_1} \sum_{C_{LP}} \left(\begin{array}{l} P([L = l])\ P([W = w])\ P([P = g(C_{LP})])\ P(C_{LP} \mid [L = l]\ [W = w]) \\ P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [\dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T} = h(P) = h(g(C_{LP}))]) \end{array} \right). \end{aligned}$$

Les termes $P([L = l])$, $P([W = w])$, $P([P = f(C_{LP})])$ suivent des lois uniformes. Leurs valeurs numériques sont donc toujours les mêmes, quelque soient les valeurs de w, l et c_{lp}. Ils peuvent donc être intégrés dans la constante de normalisation qui devient Z_2 :

$$\begin{aligned} & P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1]) \\ = \ & \frac{1}{Z_2} \sum_{C_{LP}} \left(\begin{array}{l} P(C_{LP} \mid [L = l]\ [W = w]) \\ P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [\dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T} = h(P) = h(g(C_{LP}))]) \end{array} \right). \end{aligned} \tag{6.72}$$

On remarque que l'inférence de la tâche d'écriture est égale au produit de deux termes. Techniquement, nous avons choisi de faire un tirage sur C_{LP}. Le premier terme correspond donc au tirage des points de contrôle selon la lettre et le scripteur et le second correspond à déterminer les accélérations du bras selon les points de contrôle générés.

Nous observons donc que la question d'écriture ne dépend que du modèle de production (génération de trajectoire et description de l'effecteur) et du modèle interne de représentation des lettres. En aucun cas, le modèle de simulation interne ni le modèle de vision n'interviennent dans cette question.

6.3.2 Résultats

Dans cette section, nous allons décrire les résultats de la question d'écriture en analysant les traces produites par le modèle : nous allons tout d'abord expliquer en détail les résultats pour une question donnée. Puis, nous poserons plusieurs fois cette même question afin d'étudier les variabilités de production. Enfin, nous comparerons les résultats pour plusieurs scripteurs. Nous utilisons la même base de données que lors de la perception des lettres : 3520 traces réalisées par 4 scripteurs.

6.3.3 Analyse d'une trace

Nous analysons les résultats pour la question suivante :

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Estelle]\ [\lambda_P = 1]). \tag{6.73}$$

Cette question correspond à : quelles sont les accélérations à appliquer pour faire tracer la lettre « *a* » par le scripteur « *Estelle* »

6.3.3.1 Analyse des profils d'accélération, de vitesse et de position

La figure 6.16 présente les résultats pour la question $P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Estelle]\ [\lambda_P = 1])$.

Afin de pouvoir analyser ces résultats, nous avons intégré les accélérations articulaires ($\ddot{\theta}_1$ et $\ddot{\theta}_2$) au cours du temps pour obtenir les profils de vitesse articulaire ($\dot{\theta}_1$ et $\dot{\theta}_2$). La figure 6.17 présente le profil de vitesse articulaire $\dot{\theta}_1$ à gauche et le profil de vitesse articulaire $\dot{\theta}_2$ à droite au cours du temps. Nous avons ajouté, à titre indicatif sur ces figures, les vitesses articulaires des

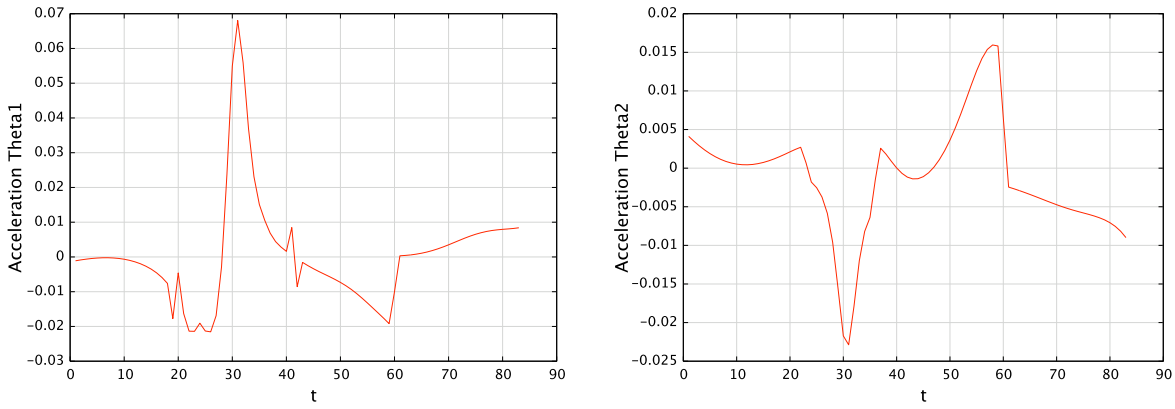


FIG. 6.16 – Accélérations articulaires $\ddot{\theta}_1$ (à gauche) et $\ddot{\theta}_2$ (à droite) obtenues en réponse à la question $P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Estelle]\ [\lambda_P = 1])$.

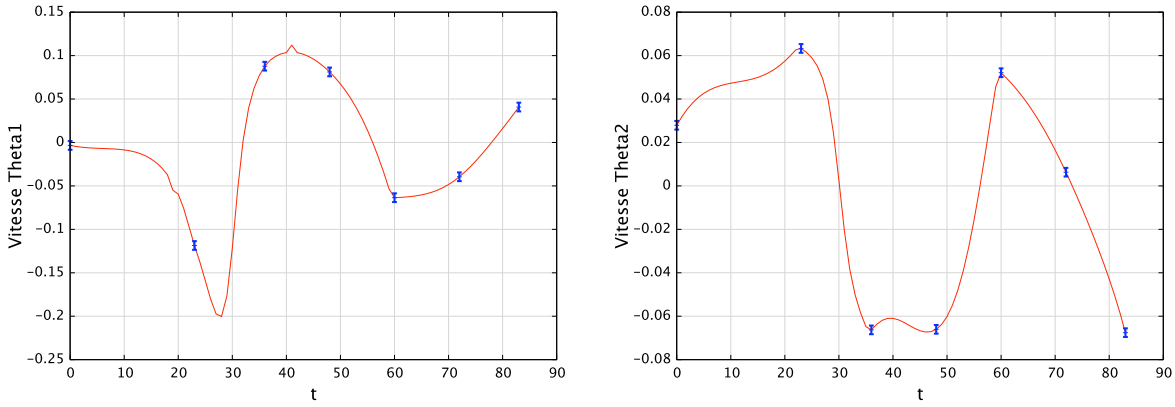


FIG. 6.17 – En rouge, vitesse articulaire $\dot{\theta}_1$ (à gauche) et $\dot{\theta}_2$ (à droite) au cours du temps (t). En bleu, vitesse articulaire des points de contrôle.

points de contrôle en bleu. En effet, les points de contrôle sont fournis dans l'espace cartésien et non dans l'espace articulaire mais il nous a semblé intéressant de les faire apparaître sur ces schémas après conversion. Nous observons en effet que les points de contrôle qui se trouvent à des positions remarquables dans l'espace cartésien ne le sont plus dans l'espace articulaire. Nous observons qu'à chaque point de contrôle, la vitesse de consigne est bien respectée étant donné que la courbe rouge et les points bleu se superposent exactement.

Nous avons une nouvelle fois intégré ces résultats, pour obtenir les profils de positions articulaires θ_1 et θ_2.

Nous observons figure 6.18 que les positions au cours du temps sont fluides et qu'elles respectent bien les positions des points de contrôle.

6.3.3.2 Application sur le bras simulé

Pour pouvoir visualiser la trace obtenue, nous appliquons ces données (accélérations en θ_1 et en θ_2) sur le bras simulé.

Les résultats obtenus sont présentés figure 6.19. La trace est représentée en bleue et les

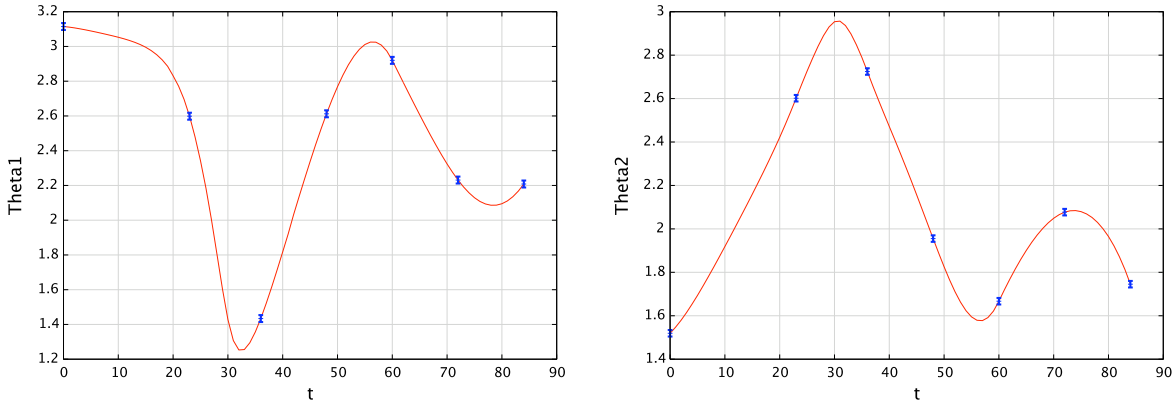


FIG. 6.18 – En rouge, θ_1 à gauche et θ_2 à droite au cours du temps. En bleu, les positions des points de contrôle.

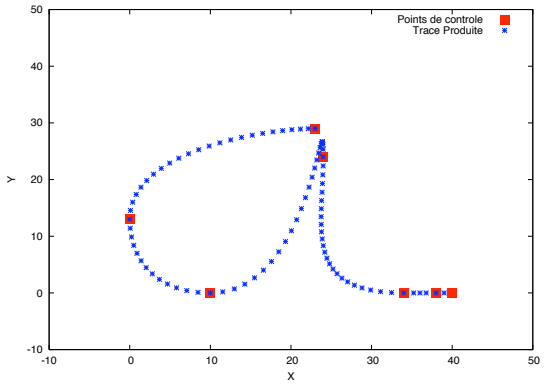


FIG. 6.19 – Les points de contrôle sont présentés en rouge et la trace obtenue en bleu.

points de contrôle en rouge. On observe que la trace obtenue est lisible et correspond bien à la lettre « a ».

Nous avons posé plusieurs fois la question en variant la lettre demandée. La figure 6.20 présente un ensemble de traces correspondant aux lettres « y », « r », « o », « m ». Nous observons que les traces sont lisibles.

Une des caractéristiques de presque tout système moteur humain tient dans les variations d'une performance à l'autre. En ce qui concerne l'écriture, il semble qu'une lettre donnée puisse prendre un nombre de formes quasiment infini (Sirat *et al.*, 1988). Certaines variations peuvent être considérées comme intentionnelles. Celles qui correspondent au style du scripteur (nous les nommerons variabilités inter-scripteurs). Les autres variabilités peuvent être considérées comme des erreurs du point de vue du lecteur (variabilité intra-scripteur). Dans la suite de ce chapitre, nous allons étudier des deux types de variations : intra-scripteur et inter-scripteurs.

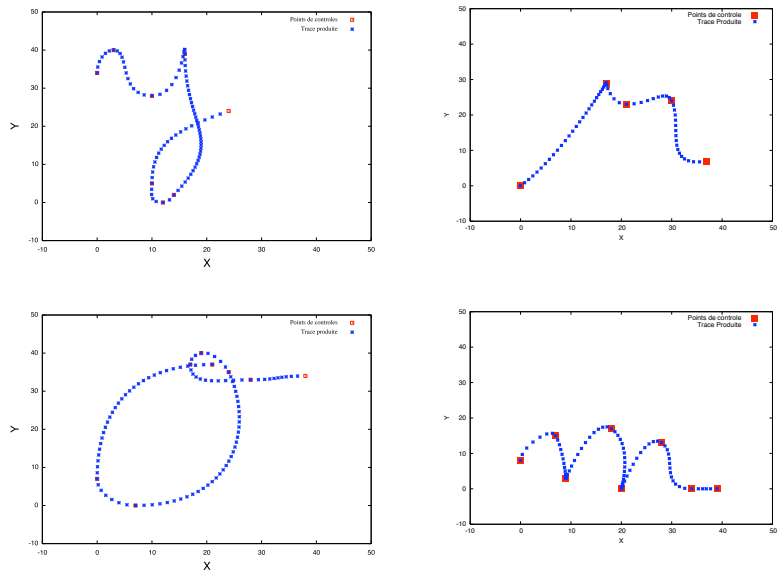


FIG. 6.20 – Traces obtenues en réponse à la question $P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = Estelle]\ [\lambda_P = 1])$ avec $l = y$ ou r ou o ou m.

6.3.4 Variabilité intra-scripteur

Si l'on pose plusieurs fois la même question d'écriture au modèle, on observe que les résultats ne sont pas totalement identiques (figure 6.21). En effet, chacun des points de contrôle est tiré selon la loi de probabilité que nous avons défini section 3.2.2.1. A chaque nouvelle lettre, les points de contrôle ne sont donc pas tirés exactement au même endroit et les traces résultantes sont donc légèrement différentes. Dans la figure 6.21, nous observons une certaine variabilité entre les lettres. Ces résultats sont en accord avec les observations expérimentales : lorsque nous écrivons, nous ne réalisons jamais exactement la même trace. On observe que les traces obtenues par le modèle représentent plus ou moins bien la lettre a. Avec notre représentation des lettres, nous ne pouvons garantir que la trace soit identifiable comme une lettre. En effet, étant donné que nous avons une indépendance sur les quatre composantes x, y, $\dot{x}$, $\dot{y}$, lors du tirage des points de contrôle nous ne pouvons pas être sûr de leurs cohérences. C'est pourquoi, il arrive que, dans certains cas, les lettres ne soient pas reconnaissables.

Si l'on considère l'écriture comme un moyen de communication on est amené à évaluer sa qualité en fonction de l'exactitude des informations transmises. Plusieurs études ont traité de la variabilité intra-scripteur chez l'humain : elles portaient principalement sur la variabilité de taille entre des lettres successives (Wann, 1988), sur la variabilité des profils cinématiques lorsque les lettres sont tracées à plusieurs échelles (Djioua et Plamondon, 2009) et sur la taille des traits verticaux et horizontaux dans les lettres (Teulings et Schomaker, 1993). D'autres études portent sur la variabilité des traces en vue d'améliorer les systèmes de reconnaissance d'écriture (Nosary *et al.*, 1999). Il serait intéressant de compléter ces travaux en proposant une étude portant sur la variabilité spatiale et cinématique des lettres dans le cas le plus simple : des caractères isolés tous produits dans les mêmes conditions expérimentales (même échelle, même orientation, etc.).

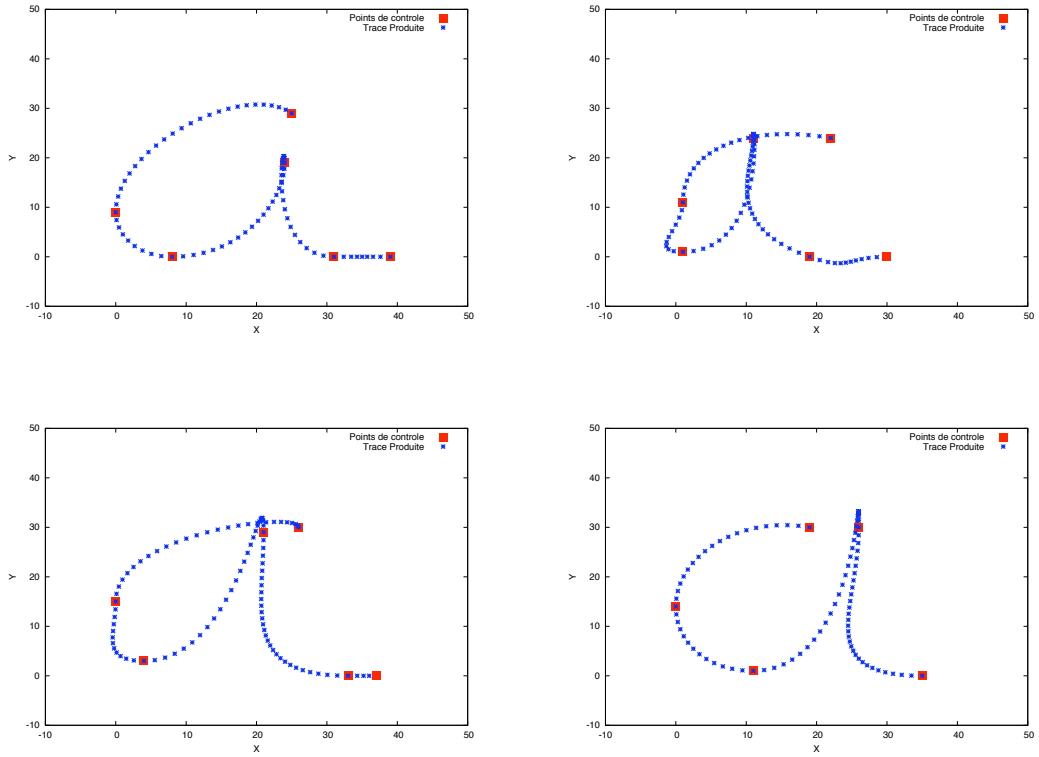


FIG. 6.21 – Traces obtenues en réponse à la question $P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Estelle]\ [\lambda_P = 1])$.

6.3.5 Variabilité inter-scripteurs

Souvent, nous modifions notre écriture afin de la rendre plus rapide et nous utilisons des formes s'écartant des formes standardisées. Deux scripteurs peuvent donc diverger vers des formes différentes : c'est une des sources de la variabilité inter-scripteurs. Le modèle BAP nous permet de simuler la question d'écriture avec différents scripteurs. Nous avons donc testé cette tâche cognitive afin de vérifier que le modèle BAP permet bien de capturer et restituer cette variabilité inter-scripteurs. En effet, il est possible de poser la même question d'écriture que précédemment mais pour un autre scripteur (quelles sont les accélérations à appliquer au bras pour tracer la lettre « *a* » avec le scripteur « *Christophe* ») :

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Christophe]\ [\lambda_P = 1]). \tag{6.74}$$

Tout d'abord, nous présentons figure 6.22 un ensemble de traces réalisées par les scripteurs *Estelle* (en haut) et *Christophe* (en bas).

Nous observons que les « a » du scripteur *Estelle* sont plutôt arrondis alors que les « a » du scripteur *Christophe* sont plutôt allongés.

Ces traces ont servi à l'identification des paramètres des formes paramétriques du modèle (4 traces parmi les 35 utilisées). Nous les avons présentées afin de pouvoir comparer le style d'écriture des scripteurs à celui des traces générées par le modèle. La figure 6.23 présente les traces obtenues par le modèle lors de la tâche d'écriture en simulant le scripteur « *Estelle* » (à

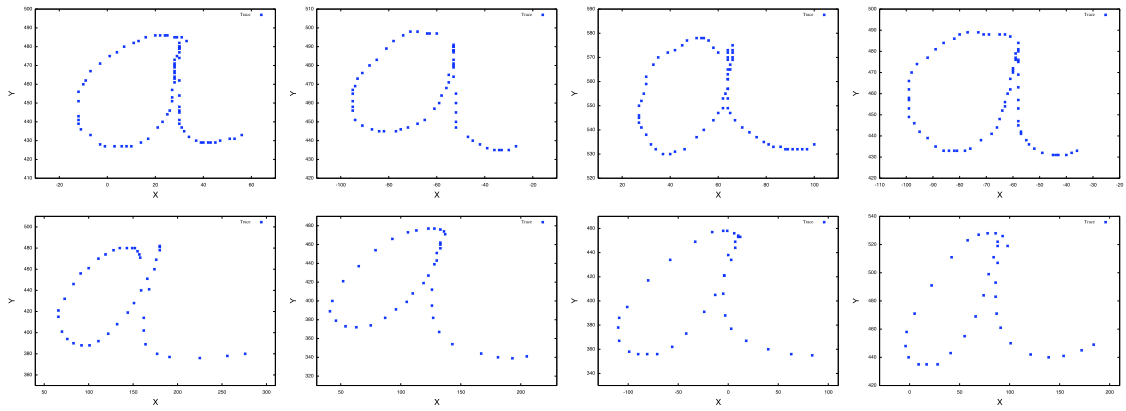


FIG. 6.22 – En haut : traces réalisées par le scripteur *Estelle* et en bas par le scripteur *Christophe*. Ces traces sont extraites de la base de données.

gauche) et en simulant le scripteur « *Christophe* » (à droite). La trace de gauche correspond donc à la question :

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Estelle]\ [\lambda_P = 1]), \tag{6.75}$$

et le trace de droite à la question :

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Christophe]\ [\lambda_P = 1]). \tag{6.76}$$

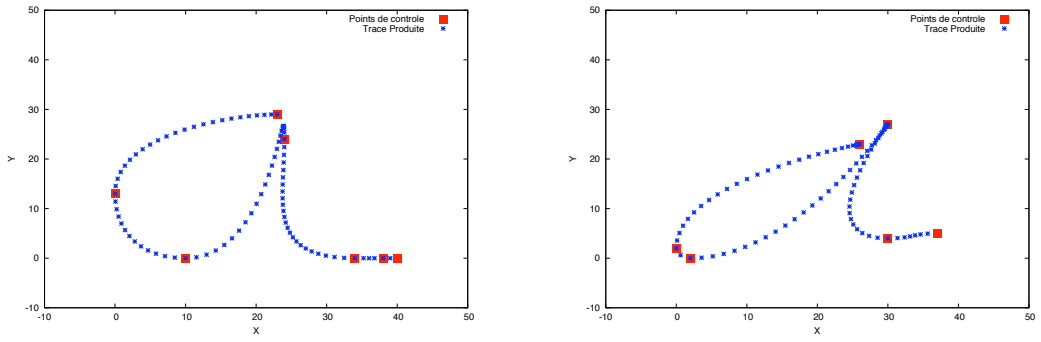


FIG. 6.23 – A gauche : trace obtenue en réponse à la question $P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Estelle]\ [\lambda_P = 1])$. A droite : trace obtenue en réponse à la question $P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = a]\ [W = Christophe]\ [\lambda_P = 1])$

On observe que dans les deux cas, les lettres sont lisibles. On retrouve que le « *a* » de gauche a une forme plutôt arrondie ce qui correspond au style d'écriture du scripteur « *Estelle* » alors qu'à droite, l'écriture est plutôt allongée (style d'écriture du scripteur « *Christophe* »).

Nous venons d'observer que le modèle est capable de reproduire une variabilité inter-scripteurs avec la lettre « a ».

La figure 6.24 présente les résultats pour les lettres « *n* », « *v* », « *e* » et « *w* ».

Nous observons que les traces diffèrent selon les scripteurs simulés par le modèle. Cette expérience permet de valider le fait que le modèle est capable de reproduire la variabilité inter-scripteurs.

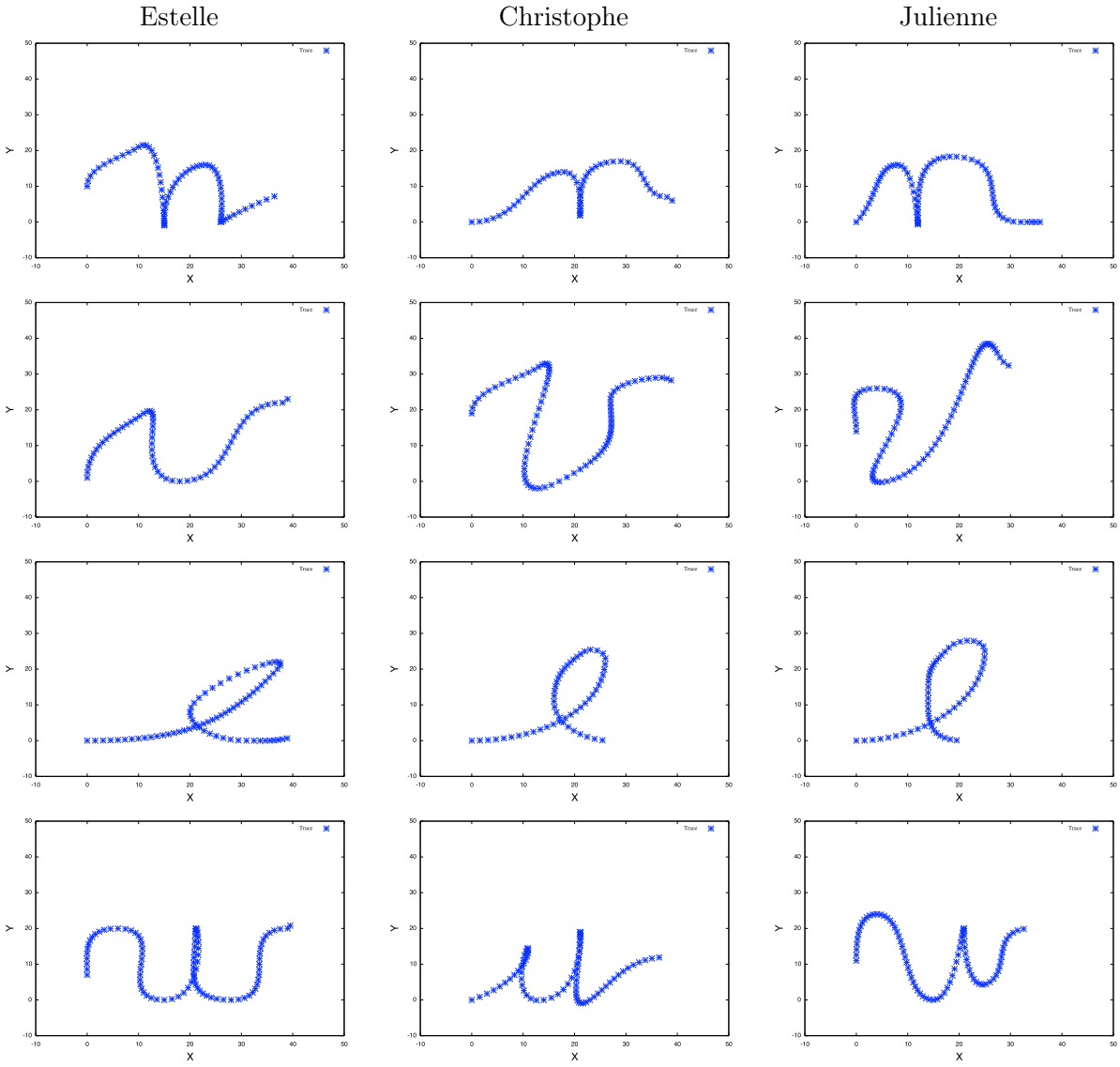


FIG. 6.24 – A gauche, les traces produites par le modèle simulant le scripteur « *Estelle* », au centre les lettres produites par le modèle simulant le scripteur « *Christophe* » et à droite les lettres produites par le modèle simulant le scripteur « *Julienne* »

6.4 Equivalence motrice

Cette section a pour but d'illustrer l'hypothèse d'équivalence motrice : cette hypothèse suggère qu'il est possible d'écrire avec n'importe quelle partie de son corps de façon instantanée.

Un des objectifs de cette thèse est d'apporter des éléments de réponse à la question suivante : **Quels sont les mécanismes qui permettent de réaliser un même mouvement avec différents effecteurs ?** Afin de répondre à cette question, et donc d'assurer l'équivalence motrice dans le modèle BAP, nous avons fait une hypothèse : la représentation interne des mouvements est indépendante de l'effecteur généralement utilisé pour effectuer le mouvement. Dans le cadre de l'écriture, nous avons défini la représentation interne des lettres dans

l'espace de la feuille (espace cartésien). L'espace de la représentation interne choisi est donc indépendant des effecteurs. Dans cette section, nous allons simuler le modèle BAP avec deux effecteurs et montrer qu'on retrouve bien le phénomène d'équivalence motrice dans les traces réalisées. Tout d'abord, nous avons réalisé un bras à deux articulations sur lequel nous avons implanté la tâche d'écriture. Puis, nous avons utilisé un second effecteur : un robot holonome. Pour chaque effecteur, nous avons posé un ensemble de questions relatives à la tâche d'écriture et vérifié la variabilité intra-scripteur et inter-scripteurs.

6.4.1 Bras à deux articulations

Le bras a deux articulations que nous avons réalisé a pratiquement les mêmes caractéristiques que celui que nous avons utilisé en simulation. La seule différence vient du fait que le bras simulé était contrôlé en accélérations alors que le bras réel est contrôlé en vitesses La figure 6.25 présente une photo du bras robotique utilisé.

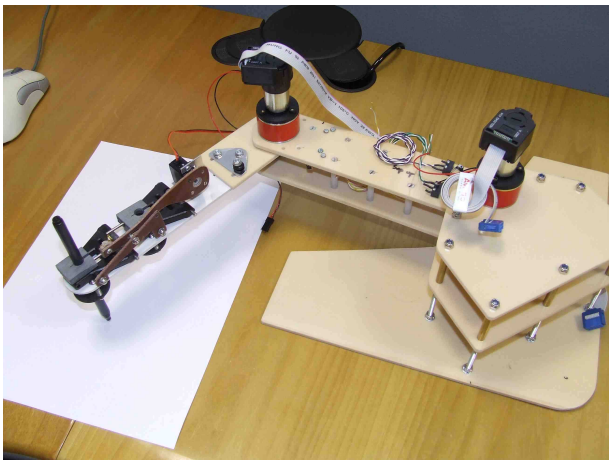


FIG. 6.25 – Bras robotique utilisé par le modèle BAP.

La question posée au modèle est similaire aux précédentes (cette fois, on détermine la distribution de probabilité des vitesses articulaires) :

Quelles sont les vitesses à appliquer au bras robotique pour qu'il trace la lettre l comme le scripteur w ?

$$P(\dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]). \tag{6.77}$$

Nous avons posé cette question sur un ensemble de scripteurs et de lettres donnés et nous avons comparé les traces à celles obtenues en simulation. La figure 6.26 présente la lettre « a » obtenue par le modèle en utilisant le bras robotique (à droite) et la compare à la lettre « a » obtenue en simulation (à gauche).

Nous remarquons que les traces sont similaires : leurs formes sont identiques. Cependant, nous observons que dans le cas du bras robotique, les traces sont légèrement fluctuantes. En effet, cela vient du contrôle de l'effecteur et d'une plus forte discrétisation des vitesses. Nous analyserons plus en détail les traces réalisées par le bras robotique après avoir présenté le dernier effecteur : le robot holonome.

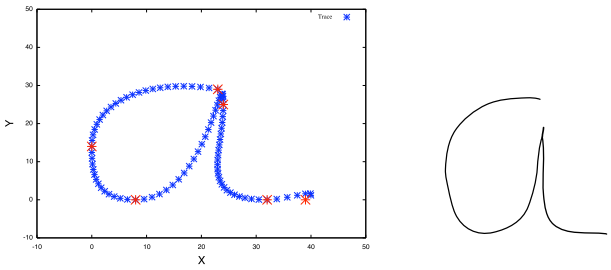


FIG. 6.26 – A gauche, une trace écrite par le bras simulé. A droite, la même trace écrite par le bras robotique.

6.4.2 Robot holonome

Nous avons modélisé un second effecteur : un robot holonome[4]. Bien que cet effecteur ne puisse correspondre à aucun des effecteurs humain, il permet d'illustrer une nouvelle fois le principe d'équivalence motrice. En effet, la fonction première du robot n'est pas d'écrire et aucun processus d'apprentissage lié à l'écriture n'a été réalisé. En utilisant ce robot n'ayant jamais écrit auparavant, nous montrons que le modèle BAP est capable d'utiliser n'importe quel type d'effecteur de façon instantanée. Avant de détailler la modélisation du robot, nous allons présenter ses principales caractéristiques.

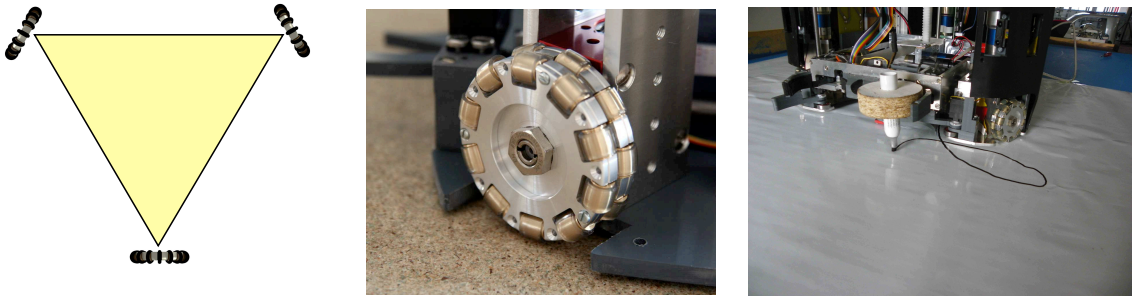


FIG. 6.27 – A gauche, schéma de la base roulante du robot holonome : le robot est composé de trois roues omnidirectionnelles (photo du centre). A droite, le stylo fixé sur le robot.

La figure 6.27, à gauche, présente la base roulante du robot holonome qui est composée de trois roues. Ces roues, comme l'illustre la photo de droite de cette même figure, permettent d'avoir un mouvement omnidirectionnel : les roulettes au niveau de la surface de la roue permettent d'avoir un mouvement parallèle à l'axe de la roue. Le stylo est accroché sur une des faces du robot holonome.

Dans un premier temps, nous présentons le modèle de ce nouvel effecteur (modèle cinématique et de contrôle) avant d'illustrer l'équivalence motrice avec des résultats expérimentaux.

6.4.2.1 Modèle de l'effecteur « robot holonome »

Variables La variable E correspondant à cet effecteur est un ensemble de variables :

$$E = \{V_1^{0:T}, V_2^{0:T}, V_3^{0:T}, \omega_1^{0:T}, \omega_2^{0:T}, \omega_3^{0:T}\}. \tag{6.78}$$

[4]Un robot holonome ou robot omni-directionnel est un robot qui possède 3 degrés de liberté sur un plan.

Ces variables correspondent donc aux vitesses linéaires dans le repère de la roue $(V_1^{0:T}, V_2^{0:T}, V_3^{0:T})$ et aux vitesses angulaires des roues $(\omega_1^{0:T}, \omega_2^{0:T}, \omega_3^{0:T})$.

Termes de la décomposition Le terme de la décomposition correspondant à l'effecteur est :

$$P(\ E\ \mid\ P\ ). \tag{6.79}$$

Avec notre définition des variables, nous obtenons :

$$P(V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T}\ \omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T} \mid P_X^{0:T}\ P_Y^{0:T}). \tag{6.80}$$

Nous avons une indépendance conditionnelle entre les variables $\omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T}$ et les variables $P_X^{0:T}\ P_Y^{0:T}$. Cette indépendance est due à la définition mathématique des différents termes. En appliquant cette indépendance nous obtenons :

$$\begin{aligned} &P(V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T}\ \omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T} \mid P_X^{0:T}\ P_Y^{0:T}) \\ &= \begin{pmatrix} P(V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T} \mid P_X^{0:T}\ P_Y^{0:T}) \\ P(\omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T} \mid V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T}) \end{pmatrix}. \end{aligned} \tag{6.81}$$

Nous nommerons *modèle cinématique* le terme $P(V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T} \mid P_X^{0:T}\ P_Y^{0:T})$ et modèle de *contrôle de l'effecteur* le terme $P(\omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T} \mid V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T})$.

Modèle cinématique Nous faisons l'hypothèse que la vitesse au temps t des roues est indépendante des vitesses précédentes connaissant la vitesse du bout de l'effecteur au temps t. Nous faisons également une hypothèse d'indépendance entre chacune des dimensions du robot. Nous obtenons :

$$\begin{aligned} &P(V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T} \mid P_X^{0:T}\ P_Y^{0:T}) \\ &= \prod_{t=0}^{T} \begin{pmatrix} P(V_1^t \mid P_X^t\ P_Y^t) \\ P(V_2^t \mid P_X^t\ P_Y^t) \\ P(V_3^t \mid P_X^t\ P_Y^t) \end{pmatrix}. \end{aligned} \tag{6.82}$$

Dans le cas d'un robot holonome, il faut distinguer la vitesse contrôlée et la vitesse non contrôlée au niveau de chacune des roues. La figure 6.28 présente le vecteur vitesse réel de la roue ainsi que les vitesses contrôlée et non contrôlée. La vitesse réelle correspond à la somme

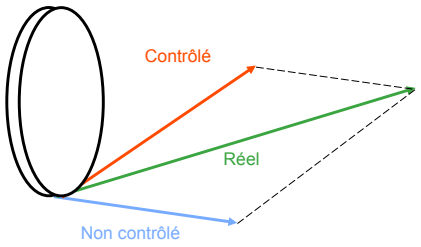


FIG. 6.28 – Vitesse réelle de la roue composée de la vitesse contrôlée et non contrôlée.

des vitesses contrôlées et non contrôlées. Le modèle cinématique s'intéresse donc à calculer la vitesse contrôlée de chacune des roues en fonction de la vitesse du bout de l'effecteur.

V_1, V_2 et V_3 correspondent à la vitesse contrôlée de chacune des roues comme le présente la figure 6.29. Sur cette figure, nous observons la vitesse réelle du robot (flèche violette), la vitesse réelle de chacune des roues (flèches vertes), la vitesse non contrôlée de chacune des roues (flèches bleues) et la vitesse contrôlée de chacune des roues (flèches rouges).

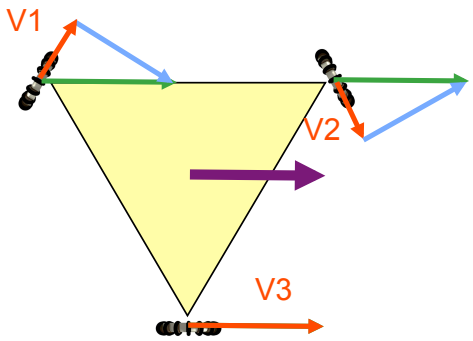


FIG. 6.29 – En violet, vitesse réelle du robot lors d'une translation. En vert, vitesse réelle de chacune des roues, en rouge vitesse contrôlée et en bleu vitesse non contrôlée.

Afin de déterminer la vitesse contrôlée de la roue ($V_i = (x_i, y_i)$) en fonction de la vitesse réelle souhaitée ($\{p_{\dot{x}}, p_{\dot{y}}\}$) nous avons défini les notations suivantes qui sont présentées figure 6.30 :

- α est l'angle entre la vitesse réelle de la roue (en vert) et la vitesse contrôlée (en rouge),
- β est l'angle entre le centre du robot et la roue,
- d est la distance entre le centre du robot et la roue.

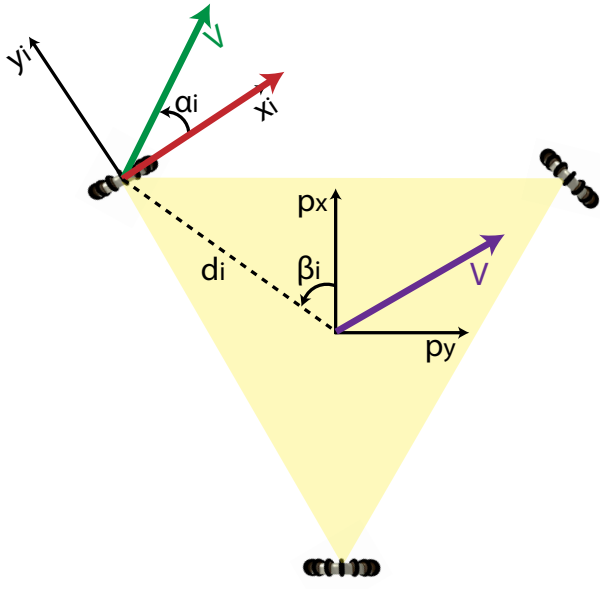


FIG. 6.30 – Repère global du robot et repère centré sur une roue.

Dans le cas d'une translation, nous avons la relation suivante entre la vitesse contrôlée et la vitesse réelle :

$$\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = M_i \begin{pmatrix} p_{\dot{x}} \\ p_{\dot{y}} \\ 1 \end{pmatrix}. \tag{6.83}$$

avec M_i la matrice suivante :

$$M_i = \begin{pmatrix} \cos\alpha_i & \sin\alpha_i & 0 \\ -\sin\alpha_i & \cos\alpha_i & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{6.84}$$

Nous notons que dans le cas d'une translation β_i et d_i n'interviennent pas.

Le terme $P(V_1 \mid P_{\dot{x}}\ P_{\dot{y}})$ est associé à une loi Dirac qui vaut 1 pour la valeur v_1 donnée par une fonction déterministe f_1 qui prend en entrée le vecteur vitesse du bout de l'effecteur :

$$P\left([V_1 = v_1] \mid [P_{\dot{X}} = p_{\dot{x}}]\ [P_{\dot{Y}} = P_{\dot{y}}]\right) = \begin{cases} 1 \text{ si } \{\ v_1\ \} = f_1(p_{\dot{x}}, p_{\dot{y}}) \\ 0 \text{ sinon.} \end{cases} \tag{6.85}$$

La fonction f_1 est définie de la façon suivante :

$$f_1(p_{\dot{x}}, p_{\dot{y}}) = p_{\dot{x}} \cos\alpha_i + p_{\dot{y}} \sin\alpha_i. \tag{6.86}$$

Les termes $P(V_2 \mid P_{\dot{x}}\ P_{\dot{y}})$ et $P(V_3 \mid P_{\dot{x}}\ P_{\dot{y}})$ sont définis de façon similaire.

Sous-modèle 2 Le modèle de contrôle de l'effecteur permet à partir d'une séquence de vitesses dans le repère de la roue de définir la séquence des vitesses des roues correspondantes :

$$P(\omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T} \mid V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T}). \tag{6.87}$$

Nous faisons des hypothèses d'indépendance, entre les dimensions, et temporellement :

$$\begin{aligned} &P(\omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T} \mid V_1^{0:T}\ V_2^{0:T}\ V_3^{0:T}) \\ &\qquad = \prod_{t=0}^{T} \begin{pmatrix} P(\omega_1^t \mid V_1^t) \\ P(\omega_2^t \mid V_2^t) \\ P(\omega_3^t \mid V_3^t) \end{pmatrix}. \end{aligned} \tag{6.88}$$

Chacun de ces termes est associé à une loi Dirac qui vaut 1 pour la valeur ω_i définie par une fonction déterministe f_2 :

$$\omega_i = f_2(v_i) = \frac{1}{r_i} v_i, \tag{6.89}$$

avec r_i le rayon de la roue.

6.4.3 Résultats

Cette section présente les résultats de la tâche d'écriture réalisée avec les effecteurs « bras robotique » et « robot holonome ». Les questions posées au modèle sont donc les suivantes :

– pour le bas robotique : **Quelles sont les vitesses à appliquer au bras robotique pour qu'il trace la lettre l comme le scripteur w ?**

$$P(\dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]). \tag{6.90}$$

– pour le robot holonome : **Quelles sont les vitesses à appliquer sur chacune des roues pour que le robot trace la lettre l comme le scripteur w ?**

$$P(\omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T} \mid [L=l]\ [W=w]\ [\lambda_P=1]) \tag{6.91}$$

Dans la suite, nous allons comparer en détail une trace réalisée par le robot avec les traces réalisées par le bras. Puis, nous analyserons la variabilité multi-scripteur.

Production d'une lettre Nous nous intéressons tout d'abord à la production de la lettre « a ».

La figure 6.31 présente une trace produite en simulation (à gauche), produite par le bras robotique (au centre) et par le robot holonome (à droite). Nous observons que les deux effecteurs robotiques sont capables de tracer la lettre « a » aussi bien que l'effecteur simulé.

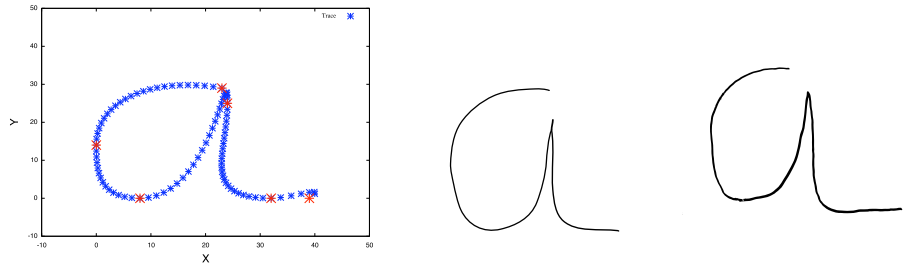


FIG. 6.31 – A gauche, la trace a été réalisée en simulation, au centre sur le bras robotique et à droite par le robot holonome.

Avant de comparer les traces produites par les trois effecteurs, nous présentons figure 6.32 un ensemble de lettres tracées par le robot holonome.

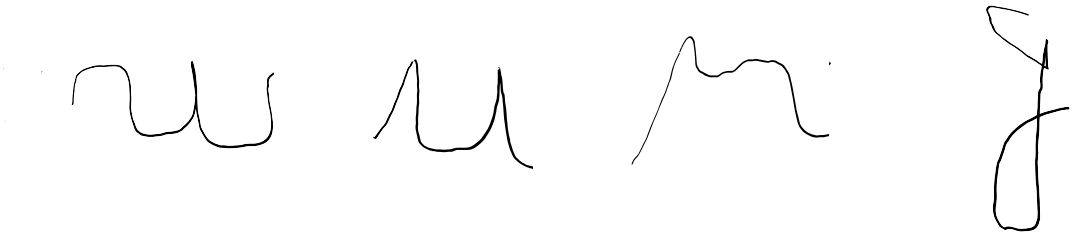


FIG. 6.32 – Traces réalisées par le robot holonome

Nous observons que par endroits (par exemple dans la boucle de la lettre « g »), la trajectoire est légèrement fluctuante. Cela provient des roues du robot qui ne sont pas totalement rondes (roues holonomes avec des rouleaux). A chaque changement de rouleau, le robot tremble légèrement et la trajectoire obtenue est donc moins fluide qu'en simulation.

Variabilité inter-scripteurs Nous avons testé la variabilité inter-scripteurs du modèle BAP. Cette fois, nous avons utilisé le robot holonome et le bras robotique pour tracer les lettres simulant le style d'écriture des différents scripteurs. La figure 6.33 présente la lettre « a » tracée par le robot holonome en simulant l'écriture du scripteur *Estelle* à gauche et du scripteur *Christophe* à droite.

FIG. 6.33 – Trace écrite par le robot holonome reproduisant le sytle d'écriture du scripteur *Estelle* à gauche et du scripteur *Christophe* à droite.

Le robot holonome a également tracé les lettres « e » et « n » suivant le style des scripteurs *Estelle*, *Christophe* et *Julienne*. Nous avons comparé ces résultats aux traces effectuées par l'effecteur bras.

Nous observons que quelque soit l'effecteur utilisé le modèle a tracé des lettres qui sont lisibles. De plus, ces lettres correspondent au style d'écriture de chaque scripteur. Nous pouvons donc conclure que les hypothèses que nous avons réalisées dans le chapitre 2 (représentation des lettres indépendante des effecteurs) permettent d'obtenir l'équivalence motrice.

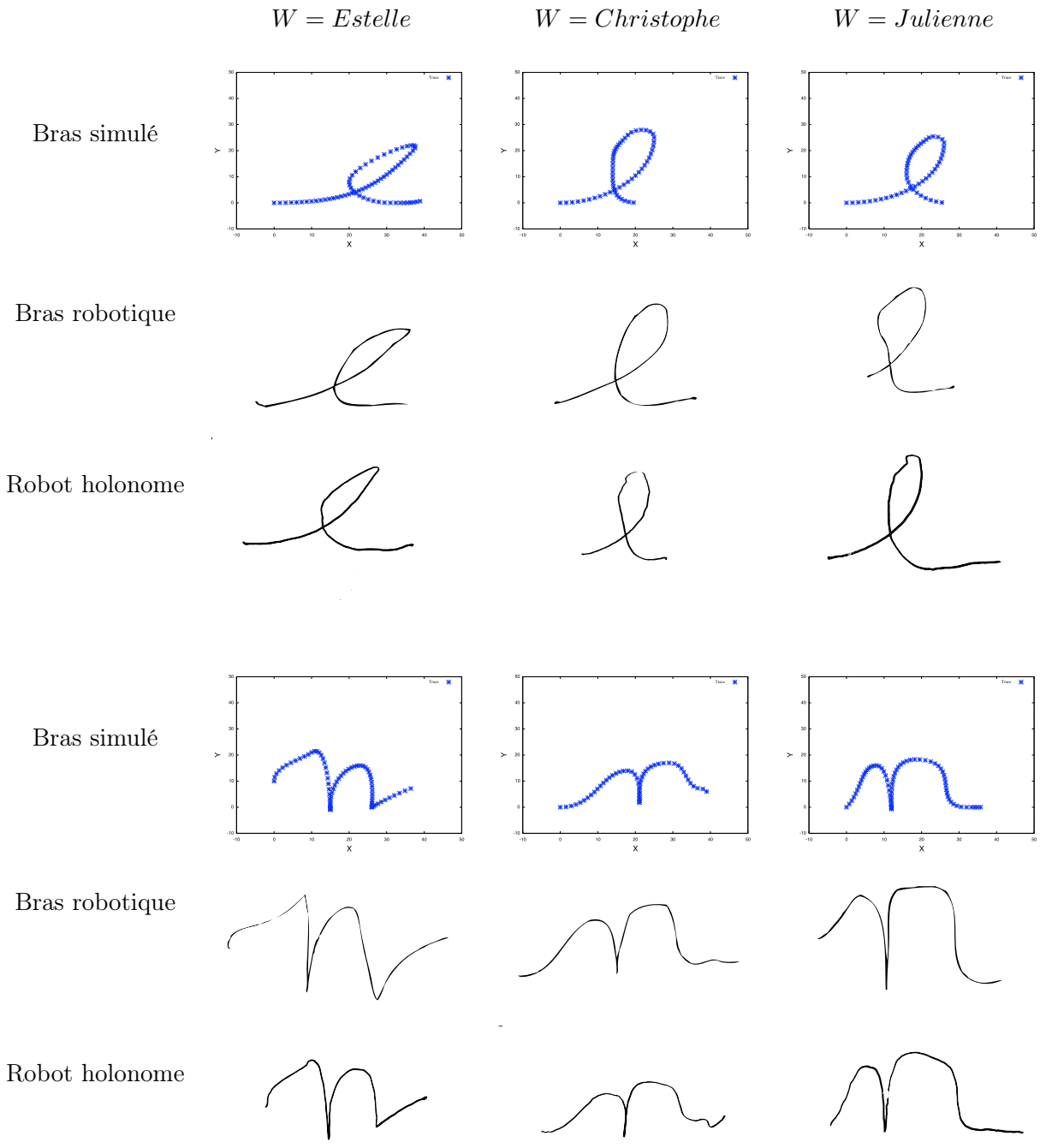


FIG. 6.34 – A gauche, les traces produites par le modèle simulant le scripteur « *Estelle* », au centre les lettres produites par le modèle simulant le scripteur « *Christophe* » et à droite les lettres produites par le modèle simulant le scripteur « *Julienne* ». Pour chaque lettre, la ligne du haut correspond aux traces produites pas le bras simulé, la signe du centre par le bras robotique et la ligne du bas par le robot holonome.

6.5 Résumé

Dans ce chapitre, nous avons tout d'abord effectué une revue des modèles du contrôle moteur. Cela nous a permis de déterminer le modèle de génération de trajectoire du modèle BAP. Nous avons choisi d'utiliser un modèle qui minimise l'accélération. Nous avons également modélisé trois effecteurs : un bras simulé, un bras robotique et un robot holonome. Nous avons présenté figure 6.35 les différents modèles et variables nécessaires au contrôle de l'effecteur « bras » et de l'effecteur « robot holonome ».

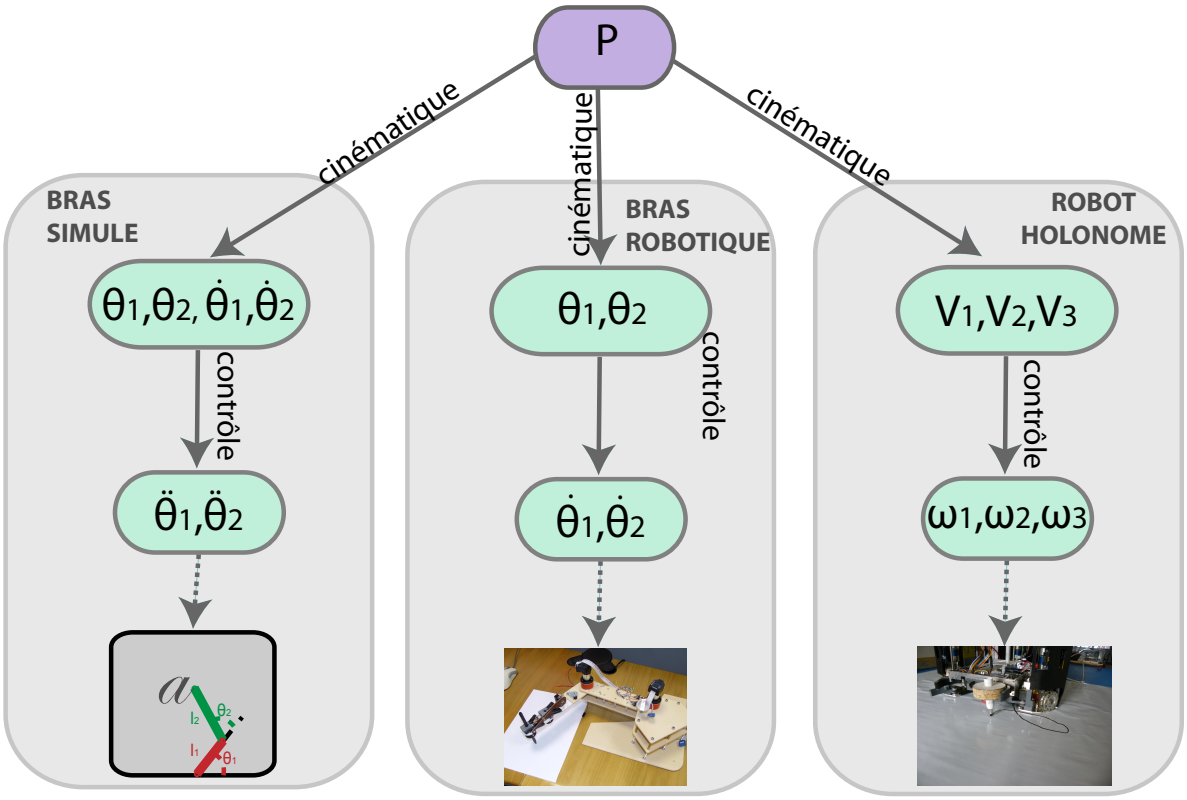


FIG. 6.35 – Modélisation des trois effecteurs du modèle BAP. A gauche : le bras simulé. Au centre : le bras robotique. A droite : le robot holonome.

Nous avons dans les deux cas un modèle cinématique qui détermine la vitesse dans le repère de l'effecteur à partir de la vitesse dans le repère cartésien. Dans le cas du bras, nous avons vu précédemment que la vitesse dans le repère de l'effecteur correspondait aux vitesses articulaires $(\dot{\theta}_1, \dot{\theta}_2)$. Pour le robot holonome, la vitesse cartésienne sera traduite en vitesse dans l'espace de chacune des roues (V_1, V_2, V_3).

Les modèles des effecteurs sont également composés d'un modèle de contrôle de l'effecteur. L'effecteur « bras » est contrôlé en accélération articulaire $(\ddot{\theta}_1, \ddot{\theta}_2)$ et l'effecteur « robot » est contrôlé par la vitesse angulaire de chacune des roues $(\omega_1, \omega_2, \omega_3)$.

Le modèle de production étant complètement défini, nous avons pu résoudre les tâches cognitives liées à l'écriture. Nous avons vu que le modèle BAP est capable de tracer des lettres et de reproduire les variabilités intra-scripteur et inter-scripteurs. Enfin, nous avons

montré qu'en utilisant une représentation interne des lettres indépendante de l'effecteur, le modèle BAP reproduit l'effet d'équivalence motrice. Afin de pouvoir valider ce résultat, il serait intéressant de réaliser une expérience de psychologie cognitive. Comme les propositions précédentes de protocoles expérimentaux, nous rappelons qu'il s'agit uniquement d'une piste de recherche et non de la définition complète d'un protocole.

Hypothèses	Les trois effecteurs du modèle BAP reproduisent la variabilité intra-scripteur et l'équivalence motrice
Matériel	Traces produites par le modèle BAP. Ces traces simulent le style d'écriture de plusieurs scripteurs et ont été tracées avec les 3 effecteurs.
Procédure	Présenter au sujet un trace de référence (une lettre, un scripteur et un effecteur donnés) ainsi qu'un ensemble de traces. Ces traces, représentant la même lettre, ont été générées par le modèle BAP en utilisant un autre effecteur que celui de la trace de référence. Elles simulent le style d'écriture de plusieurs scripteurs. La consigne correspond à demander au sujet de retrouver parmi toutes les traces présentées celle qui correspond au même scripteur que la trace de référence.
Résultats	Pour chaque trace choisie par le sujet, on compare le scripteur avec le scripteur de la trace de référence. S'il s'agit du même scripteur, on valide la réponse du sujet et on incrémente donc le score de l'effecteur ayant généré l'ensemble des traces. Pour chaque effecteur, on peut donc déterminer son taux de reconnaissance.
Analyse	L'analyse des résultats correspond à analyser les taux de reconnaissance. Pour un effecteur donné, si son taux de reconnaissance est élevé, cela indique que l'effecteur reproduit le style d'écriture des différents scripteurs. A l'inverse, si le taux est faible, cela indique que l'effecteur ne permet pas de reproduire l'observation d'équivalence motrice.

Chapitre 7

Etude du lien perception-action

Ce chapitre concerne l'étude de l'interaction entre la perception et l'action. Nous nous intéressons à un premier exemple de tâche cognitive où la perception et l'action sont sollicités : la copie. Puis, nous présentons le modèle de simulation motrice qui permet de simuler le geste d'écriture lors de la reconnaissance de lettres. Nous analysons sur des exemples le rôle de la simulation motrice lors de la lecture. Enfin, nous reproduisons une expérience de psychophysique où la simulation motrice aide à la prédiction des lettres à venir.

7.1 Copie

Dans cette section, nous allons présenter les résultats liés à l'interaction entre la perception et l'action sur le thème de la copie. Nous avons distingué deux tâches de copie : la *copie de traces* qui prend une trace en entrée et la recopie telle quelle, et la *copie de lettres* qui prend une trace en entrée, l'analyse et la recopie suivant les connaissances du scripteur. Par la suite, nous allons présenter les questions, les inférences et les résultats relatifs à ces deux tâches.

7.1.1 Copie de traces

7.1.1.1 Question et Inférence

La question relative à la tâche de copie de traces est :

Quelles sont les accélérations à appliquer au bras pour recopier la trace donnée ?

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_l = 1]\ [\lambda_p = 1]). \tag{7.1}$$

L'inférence bayésienne permet de répondre à cette question à partir de la décomposition de la distribution conjointe. Nous présentons uniquement le résultat de l'inférence et pas le détail du calcul car il est similaire à ceux présentés précédemment. Nous allons analyser les principales simplifications et le résultat. En appliquant la règle de Bayes, la règle de normalisation et la décomposition de la distribution conjointe, nous obtenons :

$$\begin{aligned} &P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_l = 1]\ [\lambda_p = 1]) \\ &\quad \propto \left(\begin{array}{l} P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T})\ P(\theta_1^{0:T}\ \theta_2^{0:T} \mid P)\ P(P \mid C_P) \\ P([\lambda_P = 1] \mid C_P\ C_{LP})\ P([\lambda_L = 1] \mid C_{LV}\ C_{LP})\ P([\lambda_V = 1] \mid C_{LV}\ C_V) \\ P(C_V \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]) \end{array} \right). \end{aligned} \tag{7.2}$$

Nous notons que tous les termes du résultat de l'inférence ont été associés à des lois Dirac. Nous pouvons donc effectuer la simplification suivante :

$$\begin{aligned} &P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_l = 1]\ [\lambda_p = 1]) \\ &\quad \propto \left(\begin{array}{l} P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T})\ P(\theta_1^{0:T}\ \theta_2^{0:T} \mid P)\ P(P \mid C_P) \\ P([\lambda_P = 1] \mid C_P\ C_{LP})\ P([\lambda_L = 1] \mid C_{LV}\ C_{LP}) \\ P([\lambda_V = 1] \mid C_{LV}\ [C_V = f(v_x^{0:M}, v_y^{0:M})]) \end{array} \right). \end{aligned} \tag{7.3}$$

avec f la fonction déterministe définie section 5.1.3. Etant donné que $[\lambda_V = 1]$ impose que les séquences de points de contrôle soient identiques dans les branches perceptives et motrices du modèle, nous obtenons :

$$\begin{aligned} &P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_l = 1]\ [\lambda_p = 1]) \\ &\quad \propto \left(\begin{array}{l} P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T})\ P(\theta_1^{0:T}\ \theta_2^{0:T} \mid P)\ P(P \mid C_P) \\ P([\lambda_P = 1] \mid C_P\ C_{LP})\ P([\lambda_L = 1] \mid C_{LP}\ [C_{LV} = f(v_x^{0:M}, v_y^{0:M})]) \end{array} \right). \end{aligned} \tag{7.4}$$

Nous effectuons de même avec $[\lambda_L = 1]$ et $[\lambda_P = 1]$ pour obtenir :

$$\begin{aligned} &P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_l = 1]\ [\lambda_p = 1]) \\ &\quad \propto P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T})\ P(\theta_1^{0:T}\ \theta_2^{0:T} \mid P)\ P(P \mid [C_P = f(v_x^{0:M}, v_y^{0:M})]). \end{aligned} \tag{7.5}$$

Nous observons que seuls les modèles de perception et d'action sont impliqués dans cette question (la figure 7.1 présente les variables et relations impliquées dans cette question). De manière schématique, cette inférence peut être décrite de la façon suivante : des points de contrôle sont déterminés à partir de la trace perçue, puis ils sont utilisés pour déterminer une nouvelle trajectoire ($P(P \mid [C_P = f(v_x^{0:M}, v_y^{0:M})])$). Cette trajectoire est convertie dans l'espace articulaire ($P(\theta_1^{0:T}\ \theta_2^{0:T} \mid P)$) et enfin, l'effecteur est asservi sur cette trajectoire ($P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T})$).

Nous observons que les notions de lettres et de scripteurs n'apparaissent nulle part dans ce résultat : à aucun moment le modèle n'analyse la trace perçue pour déterminer la lettre. En effet, la variable λ_L a court-circuité le modèle des lettres et des scripteurs. La *copie de traces* permet donc de recopier n'importe quel type de traces (des lettres, des chiffres, etc.).

7.1.1.2 Résultats

La figure 7.2 présente une trace fournie en entrée au modèle. En rouge, nous avons représenté les points de contrôle de cette trace.

La trace obtenue en résultat de la question de copie de traces est présentée figure 7.3.

Nous observons que les deux traces sont très similaires. Les positions des points de contrôle de la trace perçue et de la trace recopiée sont identiques. Les différences sont situées entre les points de contrôle, ce qui est normal étant donné que la production se base uniquement sur la position et vitesse aux points de contrôle. La trajectoire entre les points de contrôle a été générée par notre modèle de génération de trajectoires. Il est important de noter que le modèle n'a en aucun cas cherché à reconnaître la lettre correspondant à cette trace.

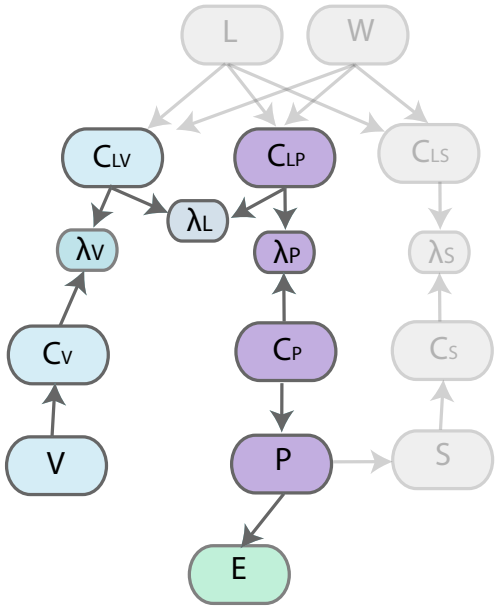


FIG. 7.1 – Architecture du modèle où on a grisé les parties non utilisées lors la question de la copie de traces.

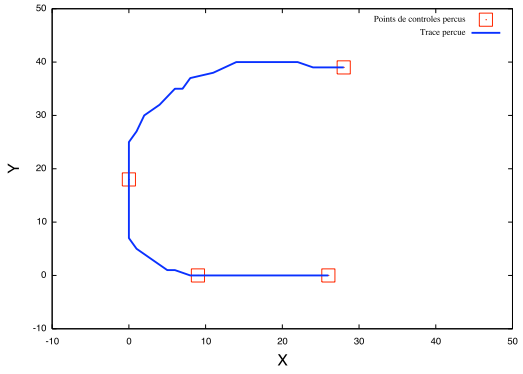


FIG. 7.2 – Trace perçue (en bleu) qui est fournie en entrée au modèle pour la question de copie de traces. Les carrés rouges représentent la position des points de contrôle.

La figure 7.4 présente plusieurs traces résultantes de la copie de traces. Pour chaque cas, nous avons représenté en bleu la trace d'origine, en rose la trace résultat et en rouge les points de contrôle. Nous observons que quelque soit la trace fournie en entrée (traces correspondantes aux symboles : m, 1, G, α, 3 et e), le modèle est capable de les recopier.

7.1.2 Copie de lettres

La tâche que nous avons nommée *copie de lettres* correspond à la recopie de lettres en fonction des connaissances relatives à chaque scripteur (connaissances apprises dans le modèle

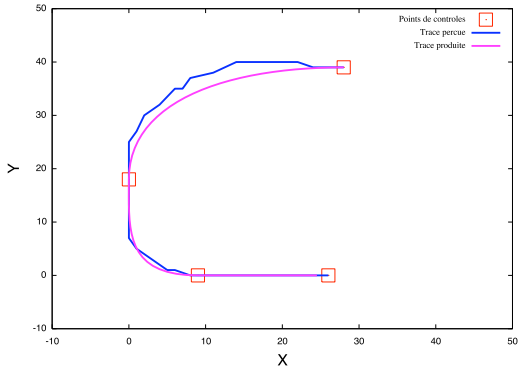


FIG. 7.3 – Trace perçue (en bleu) et trace recopiée (en rose).

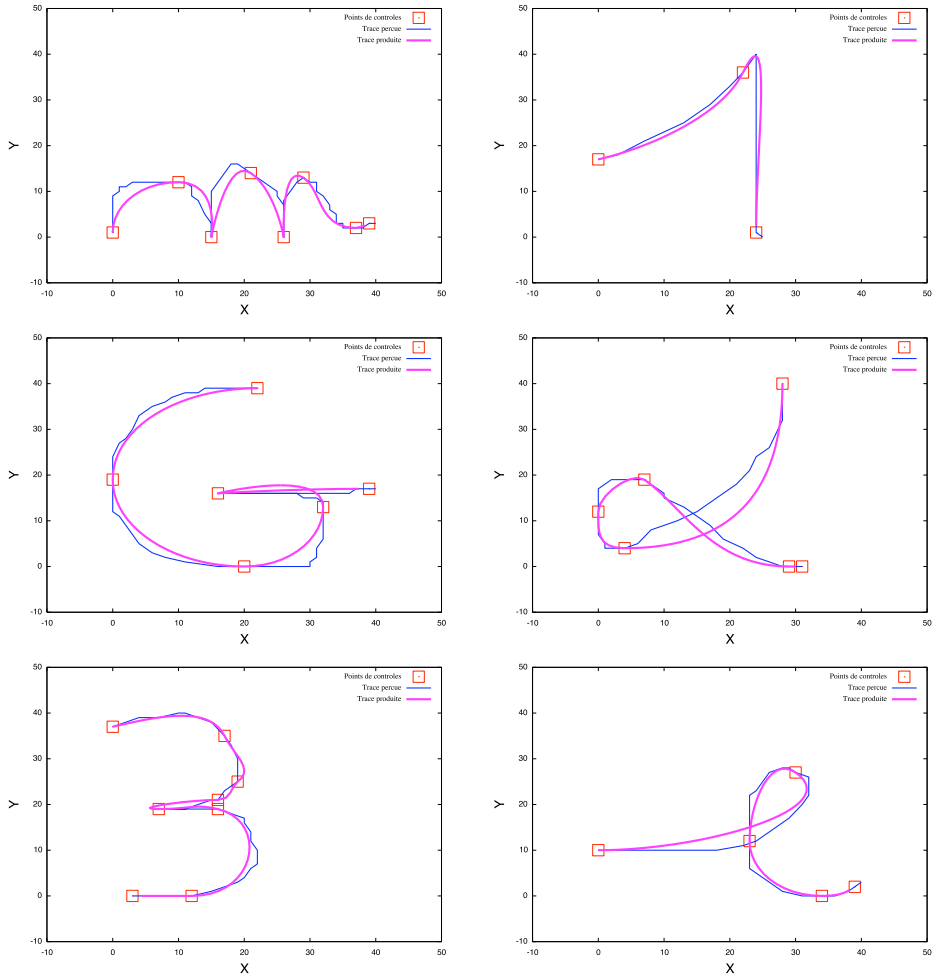


FIG. 7.4 – Traces perçues (en bleu) et traces recopiées (en rose). Plusieurs symboles ont été recopiés : m, 1, G, α, 3 e.

de représentation interne des lettres). Intuitivement, cela correspond à percevoir une lettre, la reconnaître et la tracer. Dans la suite de ce travail, nous allons détailler la question correspondante et son inférence avant de donner les résultats et de les comparer à ceux de la copie de traces.

7.1.2.1 Variables probabilistes

Pour cette tâche, nous modifions légèrement le modèle. En effet, nous avons besoin de spécifier un scripteur pour la reconnaissance des lettres et un autre pour leur production. Nous avons donc deux nouvelles variables scripteurs : W_V et W_P. Elles sont définies de façon identiques à W. La relation entre ces variables et la représentation interne des lettres est présentée figure 7.5.

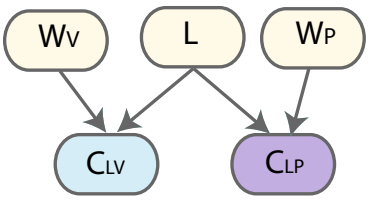


FIG. 7.5 – Architecture du modèle où on a représenté les deux variables scripteurs.

7.1.2.2 Question et Inférence

La question relative à la copie de lettres est :

Quelles sont les accélérations à appliquer au bras pour qu'il recopie la lettre ?

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_p = 1]\ [W_P = w_p]\ [W_V = w_v]). \tag{7.6}$$

En d'autres termes, cette question correspond à : quelles sont les accélérations à appliquer au bras pour que le scripteur w_p recopie la lettre $(v_x^{0:M}, v_y^{0:M})$ précédemment écrite par le scripteur w_v ?

D'un point de vue mathématique, la différence avec la question précédente vient du fait que cette fois la valeur de la variable de cohérence λ_L n'est pas spécifiée et que les valeurs des scripteurs W_V et W_P sont spécifiées. Le résultat de l'inférence va nous indiquer les implications de cette différence. Comme précédemment, nous ne détaillons pas l'inférence mais analysons uniquement le résultat :

$$\begin{aligned} &P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_p = 1]\ [W_P = w_p]\ [W_V = w_v]) \\ &\quad \propto \left(\begin{array}{l} P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid \dot{\theta}_1^{0:T}\ \dot{\theta}_2^{0:T}\ \theta_1^{0:T}\ \theta_2^{0:T})\ P(\theta_1^{0:T}\ \theta_2^{0:T} \mid P)\ P(P \mid C_{LP}) \\ \sum_L P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W_V = w_v])\ P(C_{LP} \mid L\ [W_P = w_p]) \end{array} \right). \end{aligned} \tag{7.7}$$

Nous observons que le résultat de l'inférence comprend une sommation sur la variable L. De manière schématique, la question de copie de lettres peut être décrite de la façon suivante : des points de contrôle sont déterminés à partir de la trace perçue, ce qui permet de déterminer une distribution de probabilité sur la variable L. Cette distribution de probabilité permet

ensuite de déterminer une nouvelle séquence de points de contrôle qui va servir à générer une trajectoire. Enfin, le modèle de l'effecteur contrôle l'effecteur sur la trajectoire au préalable convertie dans l'espace articulaire. La figure 7.6 présente l'architecture du modèle avec les variables et les actions impliquées dans cette question. La lettre n'est donc pas à proprement

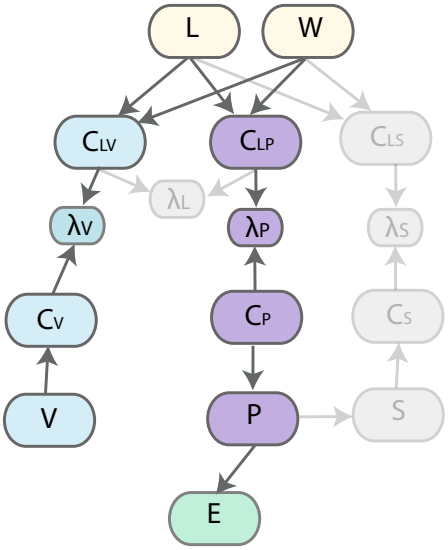


FIG. 7.6 – Architecture du modèle où les parties non utilisées lors de la question de la copie de lettres ont été grisées.

dite reconnue par le modèle : seule une distribution de probabilité sur la lettre est déterminée.

7.1.2.3 Résultats

Nous présentons tout d'abord le cas particulier où les scripteurs W_V et W_P sont identiques. La trace fournie en entrée du modèle est présentée figure 7.7 (à gauche). Elle a été tracée par le scripteur *Estelle* ($W_V = Estelle$). En rouge, nous avons représenté les points de contrôle de la trace. Cette lettre a été recopiée par le scripteur *Estelle* (figure 7.7 (à droite)). La trace perçue a été recopiée selon le style d'écriture du scripteur *Estelle*. On observe que les deux lettres sont semblables. Les points de contrôle de la trace produite (en rose) ont été déterminés en fonction des points de contrôle de la trace perçue (en bleu) et de la représentation interne des lettres précédemment apprise. Ils sont très proches de ceux de la trace perçue mais ne sont pas confondus.

La figure 7.8 présente un autre exemple de copie de lettres. La lettre présentée au modèle est un *e* (en bleu sur les figures). A gauche, la trace a été recopiée par le scripteur *Estelle* ($W_P = Estelle$), au centre par le scripteur *Jean-Louis* ($W_P = Jean\text{-}Louis$) et à droite par le scripteur *Christophe* ($W_P = Christophe$). On observe que pour chaque scripteur, le modèle a « reconnu » la lettre *e* et qu'il a généré trois nouveaux *e* personnalisés. On observe donc que la lettre produite dépend du scripteur W_P spécifié au modèle.

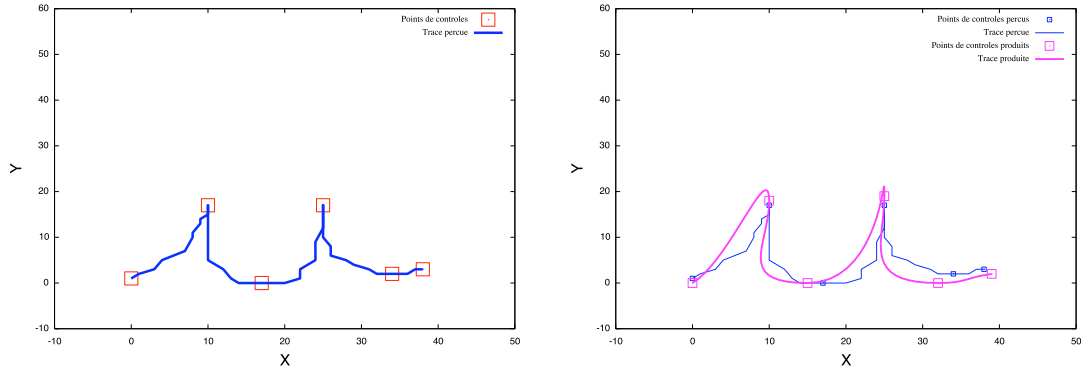


FIG. 7.7 – Trace perçue (en bleu) qui est fournie en entrée au modèle pour la question de copie de lettres. A gauche, les carrés rouges représentent la position des points de contrôle de la trace perçue et à droite, les carrés roses ceux de la trace produite. La trace a été initialement tracée par le scripteur *Estelle* (trace de gauche) et elle a été recopiée pas le même scripteur (trace de droite).

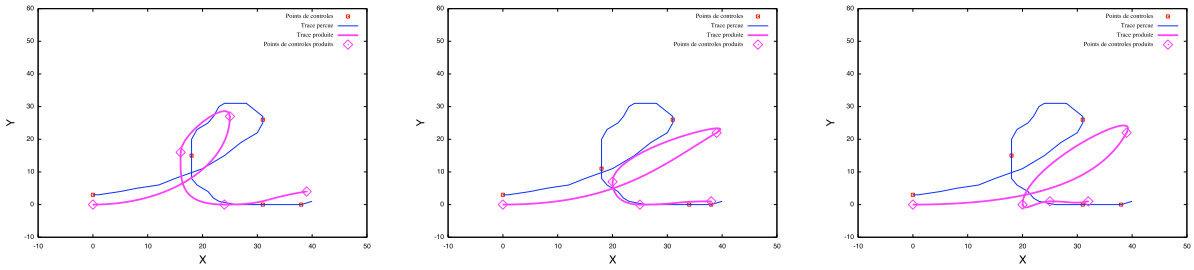


FIG. 7.8 – Trace perçue (en bleu) et trace recopiée (en rose). A gauche, la trace a été recopiée par le scripteur *Estelle*, au centre par le scripteur *Jean-Louis* et à droite par le scripteur *Christophe*.

7.1.3 Comparaison des deux types de copie

Nous avons observé que les scripteurs *Estelle* et *Christophe* ne réalisent pas la lettre w suivant le même allographe. La figure 7.9 présente l'allographe de la lettre w du scripteur *Estelle* à gauche et du scripteur *Christophe* à droite.

Nous avons posé la question suivante au modèle :

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_p = 1]\ [W_p = Estelle]\ [W_v = Christophe]). \quad (7.8)$$

Cette question correspond à déterminer la distribution de probabilité sur les accélérations à appliquer au bras pour que le scripteur *Estelle* recopie la lettre décrite par les points $\{v_x^{0:M}, v_y^{0:M}\}$ qui a été produite par le scripteur *Christophe*.

Nous avons appliqué cette question à la trace présentée en bleu figure 7.10 et la trace résultante est présentée en rose, figure de gauche. Nous avons comparé ce résultat à celui de la copie de traces. Le résultat de la copie de traces est présenté figure 7.10, à droite.

Dans le cas de la copie de lettres, nous observons que le scripteur n'est pas influencé par l'allographe proposé : la trace produite correspond au w qu'il a l'habitude de tracer. Dans le

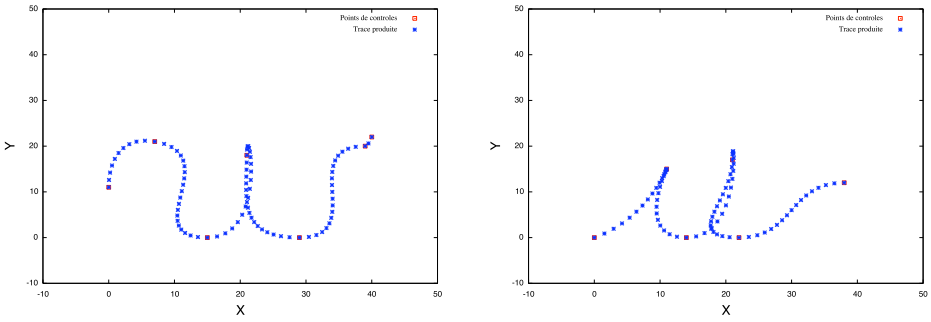


FIG. 7.9 – Prototypes de la lettre *w* par le scripteur *Estelle* à gauche et par le scripteur *Christophe* à droite.

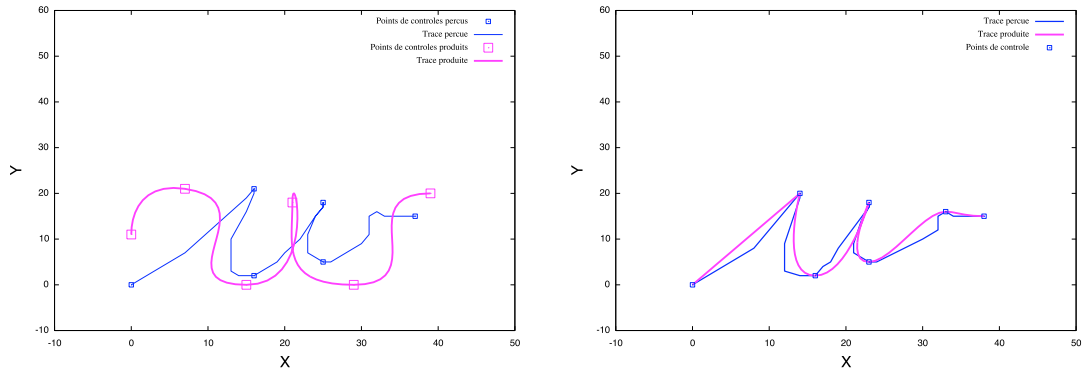


FIG. 7.10 – Copie de lettre : la lettre à recopier est proposée en bleu, elle a été écrite par le scripteur *Christophe*. La trace produite (copie de la trace bleue) est présentée en rose, elle a été produite par le scripteur *Estelle* à gauche et par le scripteur *Christophe* à droite.

cas de la copie de traces, la trace résultante correspond à l'allographe proposé : il n'y a eu aucune influence des connaissances apprises relatives aux scripteurs (connaissances apprises dans le modèle de représentation interne des lettres). Le résultat n'a pas été influencé par les habitudes du scripteur.

7.1.4 Résumé

Dans notre modèle, nous avons distingué les tâches cognitives de copie de traces et de copie de lettres. Pour une même trace fournie en entrée, on observe deux résultats différents : la copie de traces génère une trace très similaire à la trace perçue alors que la copie de lettres génère une lettre correspondant à la trace perçue.

De plus, la copie de traces est indépendante du scripteur : quelque soit le scripteur spécifié la trace produite sera identique, ce qui n'est pas le cas de la copie de lettres : les traces produites sont dépendantes de chaque scripteur. Enfin, la copie de traces permet de recopier n'importe quelle trace alors que la copie de lettres ne permet de recopier que des traces correspondantes à des lettres apprises au préalable.

Ces deux types de copie nous ont permis de réaliser des prédictions expérimentales : lorsque

l'on recopie une trace, on ne s'intéresse en aucun cas à reconnaître le symbole correspondant à la trace, alors que lorsque l'on recopie une lettre, on détermine la lettre (plus précisément la distribution de probabilité de la lettre) puis elle est recopiée suivant la représentation interne de la lettre correspondant au scripteur spécifié.

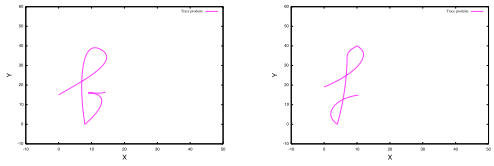
En d' autres termes, recopier une lettre, c'est reconnaître la lettre perçue (plus précisément inclure dans l'inférence la distribution de probabilité sur les lettres) puis la reproduire. En revanche, recopier une trace c'est directement reproduire la trace perçue, sans se référer aux formes des lettres connues.

De plus, dans cette section, nous avons illustré l'outil de court-circuit que nous avions défini section 4.1.3 : en activant à 1 la variable λ_L, nous avons observé que l'inférence ne dépendait plus de la lettre et du scripteur. Cette partie du modèle a donc été court-circuité.

7.1.5 Suggestion de protocole expérimental : Etude de l'influence de la représentation interne sur la copie de lettres

Nous l'avons vu, notre modèle prédit, pour ces deux tâches différentes, des propriétés différentes des traces résultantes. Cela suggère une question. Si on demande à un sujet de « recopier un texte » sa production sera-t-elle similaire à notre copie de traces, ou à notre copie de lettres ? En d'autres termes, si je vous demande de recopier une trace, faites-vous référence à vos connaissances et vos prototypes, ou bien copiez-vous la trace ?

Il serait intéressant de tester cette prédiction expérimentale chez l'humain pour déterminer l'influence de la trace perçue et de la représentation interne sur la trace produite. Pour cela, nous suggérons une expérience.

Hypothèses	Lors de la copie de lettres, la représentation interne des lettres joue un rôle majeur. La perception des lettres sert uniquement à déterminer la lettre à recopier mais n'influence pas la copie.
Matériel	L'hypothèse pourra être testée sur des lettres telles que le w ou le f où plusieurs allographes sont clairement identifiables. Par exemple, pour la lettre f, nous avons les deux allographes suivants. 
Procédure Etape 1	Le sujet doit tracer un certain nombre de fois la lettre f afin que l'on puisse identifier l'allographe lui correspondant.
Procédure Etape 2	On montre à des sujets une trace correspondante à la lettre f avec en consigne : recopier la lettre présentée. On effectue cette tâche plusieurs fois en présentant alternativement les deux allographes.
Analyse	Si les lettres tracées par le sujet correspondent à l'allographe lui correspondant, alors il est possible de suggérer que la représentation interne joue un rôle majeur : le sujet trace les lettres selon son habitude. A l'inverse, si les lettres tracées correspondent aux allographes présentés, alors il est possible de suggérer que la représentation interne n'intervient pas et que le sujet recopie la trace et non la lettre.

7.2 Définition du modèle de simulation motrice

Dans cette section, nous allons définir le modèle de simulation motrice (figure 7.11) : nous définissons les variables, les termes de la décomposition et leurs formes paramétriques. Une fois le modèle de simulation motrice décrit, notre modèle BAP de la boucle sensori-motrice de la lecture et de l'écriture sera complètement défini.

7.2.1 Variables probabilistes

Les variables correspondant à la simulation motrice sont :

- S : cette variable correspond à la trajectoire simulée. Elle correspond à un ensemble de variables de positions $\{S_X^{0:T}, S_Y^{0:T}\}$ dans l'espace cartésien (espace de la feuille).
- C_{LS} : cette variable correspond à la représentation interne simulée. Elle correspond également à un ensemble de variables de positions $\{C_{LS_x}^{0:N}, C_{LS_y}^{0:N}\}$ et de vitesses $\{C_{LS\dot{x}}^{0:N}, C_{LS\dot{y}}^{0:N}\}$.

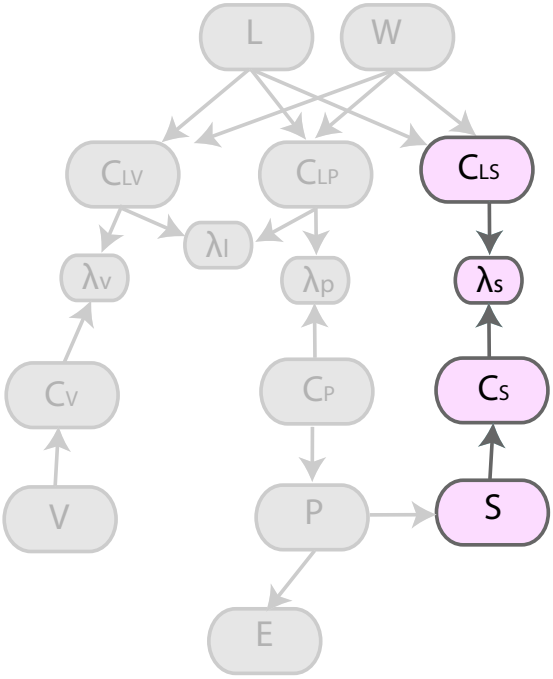


FIG. 7.11 – Architecture du modèle où l'on a grisé les variables non impliquées dans le sous-modèle de simulation motrice.

– C_S : cette variable est définie de la même manière que C_{LS}.

7.2.2 Termes de la décomposition

Les termes de la décomposition correspondant au modèle de simulation motrice dans la distribution de probabilité conjointe sont :

$$P(S \mid P)\ P(C_S \mid S)\ P(C_{LS} \mid L\ W)\ P(\lambda_S \mid C_{LS}\ C_S). \tag{7.9}$$

Nous simplifions chaque terme par des hypothèses d'indépendances conditionnelles. Nous définissons également les formes paramétriques associées ainsi que leurs paramètres.

- $P(S \mid P)$: ce terme exprime la relation entre la trajectoire écrite simulée (S) et la trajectoire écrite générée (P). Avec la définition des variables fournies ci-dessus, il s'exprime de façon équivalente en :

$$P(S \mid P) \quad = \quad P(S_X^{0:T}\ S_Y^{0:T} \mid P_X^{0:T}\ P_Y^{0:T}). \tag{7.10}$$

  Ce terme est associé à une loi Dirac centrée sur l'ensemble de valeurs $\{s_x^{0:T}, s_y^{0:T}\}$ et définie par une fonction déterministe f_0 qui prend en entrée un ensemble de valeurs (correspondant à une trajectoire) $\{p_x^{0:T}, p_y^{0:T}\}$:

$$P\left(\ S_X^{0:T}\ S_Y^{0:T}\ \mid\ P_X^{0:T}\ P_Y^{0:T}\ \right) = \left\{ \begin{array}{l} 1 \text{ si } \{\ s_x^{0:T}, s_y^{0:T}\ \} = f_0(p_x^{0:T}, p_y^{0:T}) \\ 0 \text{ sinon.} \end{array} \right. \tag{7.11}$$

  avec f_0 la fonction identitée.

- $P(\ C_S\ \mid\ S\ )$: ce terme correspond à l'extraction des points de contrôle simulés à partir de la trajectoire écrite simulée. Dans le paragraphe précédent, nous avons vu que C_S correspondait à l'ensemble des quatre variables $\{C_{Sx}^{0:N}, C_{Sy}^{0:N}, C_{S\dot{x}}^{0:N}, C_{S\dot{y}}^{0:N}\}$. Le terme de la décomposition correspondant à l'extraction des points de contrôle simulés est :

$$P(C_{Sx}^{0:N}\ C_{Sy}^{0:N}\ C_{S\dot{x}}^{0:N}\ C_{S\dot{y}}^{0:N}\ \mid\ S_x^{0:T}\ S_y^{0:T}). \tag{7.12}$$

  Il est associé à une loi Dirac qui vaut 1 pour l'ensemble de valeur $\{c_{sx}^{0:N}, c_{sy}^{0:N}, c_{s\dot{x}}^{0:N}, c_{s\dot{y}}^{0:N}\}$ donnée par une fonction déterministe f qui prend en entrée l'ensemble des points d'une trace.

$$P\left(\begin{array}{c|c} [C_{Sx}^{0:N}=c_{sx}^{0:N}]\ [C_{Sy}^{0:N}=c_{sy}^{0:N}] & [S_x^{0:T}=s_x^{0:T}] \\ [C_{S\dot{x}}^{0:N}=c_{s\dot{x}}^{0:N}]\ [C_{S\dot{y}}^{0:N}=c_{s\dot{y}}^{0:N}] & [S_y^{0:T}=s_y^{0:T}] \end{array}\right) = \left\{\begin{array}{l} 1 \text{ si } \left\{\begin{array}{l} c_{sx}^{0:N}, c_{sy}^{0:N} \\ c_{s\dot{x}}^{0:N}, c_{s\dot{y}}^{0:N} \end{array}\right\} = f(s_x^{0:T}, s_y^{0:T}) \\ 0 \text{ sinon.} \end{array}\right. \tag{7.13}$$

  Cette fonction f est définie par un l'algorithme d'extraction des points de contrôle qui est basé sur la description des points de contrôle que nous avons effectuée section 2.3.3. Il s'agit du même algorithme d'extraction que pour le modèle de perception. En d'autres termes, $P(C_S \mid S)$ est une duplication du modèle $P(C_V \mid V)$.
- $P(\ C_{LS}\ \mid\ L\ \ W\ )$: ce terme est défini de façon identique à $P(\ C_{LV}\ \mid\ L\ \ W\ )$ (définition section 3.2.2.1). Il exprime la relation entre la séquence des points de contrôle, la lettre et le scripteur.
- $P(\ \lambda_S\ \mid\ C_{LS}\ \ C_S\ )$: la définition de ce terme a été présentée section 4.1. Il permet d'activer ou de désactiver le modèle de simulation motrice.

7.3 Lecture avec simulation motrice

Nous avons vu, lors de l'étude bibliographique (section 2.4.2), que pour reconnaître une lettre l'action d'écriture est simulée. Lorsqu'on observe des traces manuscrites, on n'observe pas directement le mouvement effectué par le scripteur pour produire ces traces. Nous simulons mentalement l'action que le scripteur a effectué pour obtenir cette trace. Cette section est consacrée à l'étude de la simulation motrice pour la reconnaissance des lettres.

7.3.1 Question et inférence

La question relative à lecture avec simulation motrice est :

Quelle est la lettre correspondant à la trace que je perçois ?

$$P(L \mid [V_X^{0:M}=v_x^{0:M}]\ [V_Y^{0:M}=v_y^{0:M}]\ [W=w]\ [\lambda_V=1]\ [\lambda_L=1]\ [\lambda_P=1]\ [\lambda_S=1]). \tag{7.14}$$

L'inférence bayésienne permet de répondre à cette question à partir de la décomposition de la distribution de probabilité conjointe présentée équation 4.53 :

$$\begin{aligned} &P(L \mid [V_X^{0:M}=v_x^{0:M}]\ [V_Y^{0:M}=v_y^{0:M}]\ [W=w]\ [\lambda_V=1]\ [\lambda_L=1]\ [\lambda_P=1]\ [\lambda_S=1]) \\ &\propto \left(\begin{array}{l} P([C_{LV}=f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W=w]) \\ P([C_{LS}=h(p_x^{0:M}, p_y^{0:M})] \mid L\ [W=w]))P(p_x^{0:M}\ p_y^{0:M} \mid [C_{LP}=C_{LV}]) \end{array}\right) \\ &\propto \left(\begin{array}{l} P([C_{LV}=f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W=w]) \\ P([C_{LS}=h(g(C_{LV}))] \mid L\ [W=w])) \end{array}\right). \end{aligned} \tag{7.15}$$

Nous observons que la lecture avec simulation motrice correspond au produit de deux termes de perception : le terme $P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])$ compare les points de contrôle perçus avec ceux mémorisés (il s'agit du modèle de perception) et le terme $P([C_{LS} = h(g(C_{LV}))] \mid L\ [W = w]))$ compare les points de contrôle issus de la trace simulée avec ceux mémorisés (il s'agit du modèle de perception simulée).

7.3.2 Analyse d'une trace

Nous avons posé la question de lecture avec simulation motrice au modèle avec les valeurs de $\{v_x^{0:M}, v_y^{0:M}\}$ présentées figure 7.12.

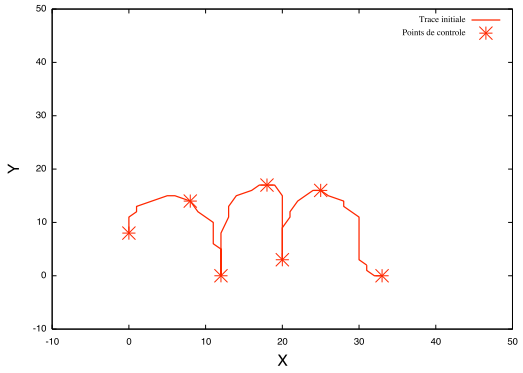


FIG. 7.12 – Trace fournie au modèle pour la reconnaissance avec simulation motrice.

D'un point de vue schématique, la reconnaissance avec simulation motrice correspond à extraire les points de contrôle de la trace perçue (croix sur la figure 7.12). Ces points de contrôle servent à générer une nouvelle trajectoire que nous avons présentée figure 7.13. De

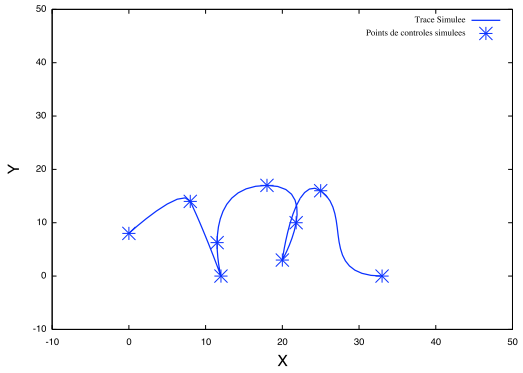


FIG. 7.13 – Trace simulée générée par le modèle pour la reconnaissance avec simulation motrice.

nouveaux points de contrôle sont extraits de la trace générée. Et enfin, la lettre est déterminée à partir des points de contrôle de la trace perçue (en rouge sur la figure 7.12) et de la trace générée (en bleu sur la figure 7.13). Pour notre exemple, la lettre m est reconnue avec une probabilité de 1. On observe que la génération de trajectoire a modifié le nombre et la situation des points de contrôle par rapport à la trace perçue.

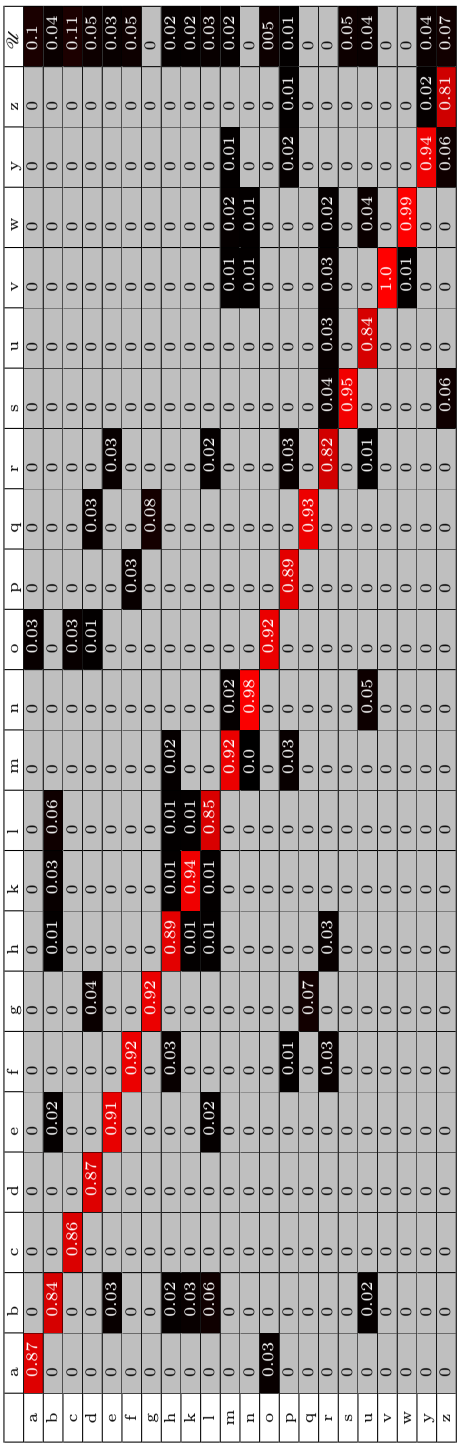


FIG. 7.14 – Matrice de confusion de perception avec simulation motrice.

7.3.3 Taux de reconnaissance

Si l'on se place dans les mêmes conditions que dans la section 5.2.2.2, c'est-à-dire que l'on précise le scripteur à l'origine des traces lorsque l'on pose la question au modèle, on obtient

la matrice de confusion présentée figure 7.14.

Le taux de reconnaissance obtenu est de $90,22\%$. Nous observons qu'il est donc similaire à celui de la reconnaissance sans simulation motrice (93,36%). Nous observons que des lettres correctement reconnues sans simulation motrice ont été mal classifiées avec la simulation motrice (dégradation de la reconnaissance) et inversement (amélioration de la reconnaissance). Par exemple, la lettre b était reconnue avec une probabilité de 0.72 d'être un b sans simulation motrice et est reconnue avec une probabilité de 0.84 d'être un b avec simulation motrice. Dans ce cas, la simulation motrice a donc amélioré la reconnaissance de la lettre. Inversement, la lettre o était reconnue avec une probabilité de 1 comme étant la lettre o sans simulation motrice et n'est reconnue qu'avec une probabilité de 0.92 avec simulation motrice. Dans le suite de cette section, nous allons analyser un cas où la simulation motrice améliore les résultats (nous ne présenterons pas le cas de dégradation des résultats car il fonctionne sur le même principe que l'amélioration des résultats).

Analyse des résultats

La trace présentée figure 7.15 correspond à la lettre d. Elle a été reconnue sans simulation motrice comme étant une lettre a avec une probabilité de 0.87 ou comme étant un d avec une probabilité de 0.13. En effet, ce type d'allographe correspond tout particulièrement à celui de la lettre a.

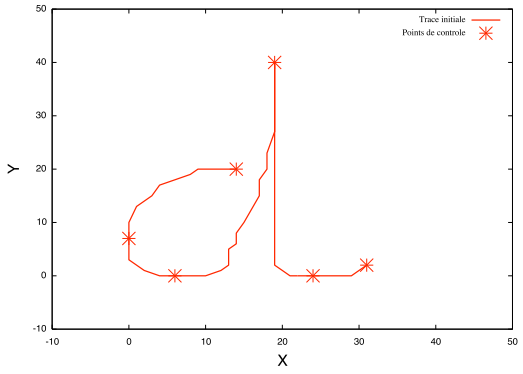


FIG. 7.15 – Trace correspondant à la lettre d reconnue par le modèle comme une lettre a.

A partir des points de contrôle de cette trace (croix rouges sur la figure 7.15), une nouvelle trace a été générée. Cette trace est présentée figure 7.16. Les points de contrôle issus de cette trace sont présentés par les croix bleues de la figure 7.16. Nous observons qu'il y a un point supplémentaire dans la partie descendante de la lettre d. L'allographe de cette trace n'est donc pas le même que celui de la trace présentée figure 7.15. Or, ce nouvel allographe généré est connu par le modèle, il correspond à la lettre d.

La trace simulée peut correspondre à un allographe différent de la trace à reconnaître. Cela peut amener à améliorer la reconnaissance (comme c'est le cas sur cet exemple) ou à la détériorer.

7.3.4 Comparaison avec et sans simulation motrice

Nous avons comparé les résultats de la reconnaissance de caractères avec et sans simulation motrice. Nous avons testé plusieurs cas : scripteur unique, scripteurs multiples, sans connaître

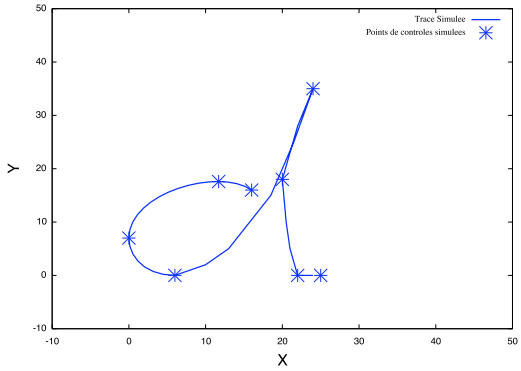


FIG. 7.16 – Trajectoire générée à partir des points de contrôle de la figure 7.15.

le scripteur, etc. Dans le tableau suivant, le paramètre A correspond au scripteur à l'origine des traces, c'est-à-dire à la personne qui a écrit la lettre à reconnaître. Nous distinguons le paramètre W_L du paramètre W_S. W_L correspond à l'information sur le scripteur fourni au modèle correspondant au scripteur à l'origine de la trace. En d'autres termes W_L est « est-ce que le modèle sait qui a écrit la lettre ? ». Le paramètre W_S correspond à l'information sur le scripteur qui effectue la simulation motrice. Par exemple, le scripteur *Estelle* reconnaît une lettre produite par le scripteur *Julienne* ($W_L = Julienne$), tout en effectuant la simulation interne sur ses connaissances ($W_S = Estelle$). La figure 7.17 présente l'architecture du modèle comprenant les deux variables scripteur.

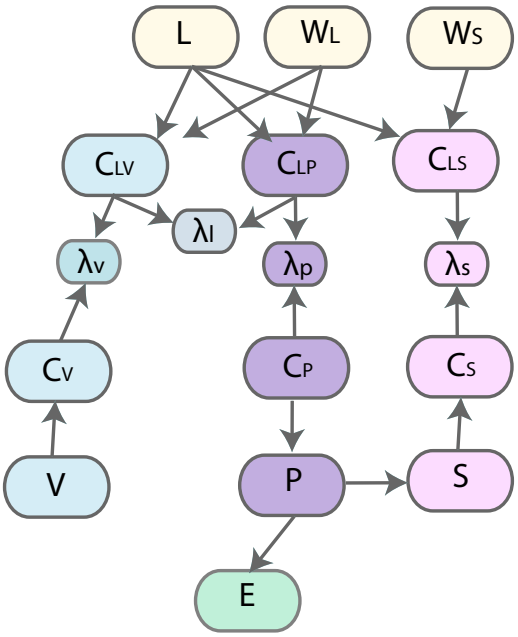


FIG. 7.17 – Architecture du modèle comprenant les deux variables scripteurs W_L et W_S.

La question correspondant à la tâche de lecture avec simulation motrice est (avec des

scripteurs différents spécifiés) :

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W_L = w_l]\ [W_S = w_s]\ [\lambda_V = 1]\ [\lambda_L = 1]\ [\lambda_P = 1]\ [\lambda_S = 1]) \quad (7.16)$$

Dans le tableau suivant, nous notons ? lorsque le scripteur n'est pas spécifié.

A	W_L	W_S	Sans Simulation Motrice	Avec Simulation Motrice
Estelle	Estelle	Estelle	93,72	90,22
Estelle	Estelle	?		92,27
Estelle	?	Estelle	92,72	91,78
Estelle	?	Jean-Louis		84,31
Estelle	?	?		93,45
?	?	?	49,68	53,59

Nous observons que, quelque soit le cas, les taux de reconnaissance avec et sans simulation motrice sont similaires. Nous notons un score supérieur pour la reconnaissance avec simulation motrice lorsque le scripteur à l'origine des traces n'est pas connu dans le modèle et n'appartient pas à la base de données (53,59% au lieu de 49,69% sans simulation motrice).

7.3.5 Reconnaissance du scripteur

Il est également possible de reconnaître le scripteur à l'aide de la simulation motrice. La question correspondant à cette tâche est :

Qui a écrit cette trace sachant que je connais la lettre correspondante ?

$$P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [L = l]\ [\lambda_V = 1]\ [\lambda_L = 1]\ [\lambda_P = 1]\ [\lambda_S = 1]). \quad (7.17)$$

L'inférence est similaire à celle de la reconnaissance de caractères :

$$\begin{aligned} &P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [L = l]\ [\lambda_V = 1]\ [\lambda_L = 1]\ [\lambda_P = 1]\ [\lambda_S = 1]) \\ &\propto \begin{pmatrix} P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid [L = l]\ W) \\ P([C_{LS} = h(g(C_{LV}))] \mid [L = l]\ W)) \end{pmatrix}. \end{aligned} \quad (7.18)$$

La question est posée sur la même base de données que dans la section 5.3.2. Les tableaux suivants correspondent aux matrices de confusion lors de la reconnaissance de scripteurs sans simulation motrice (à gauche) et avec simulation motrice (à droite).

Sans Simulation Motrice

	Estelle	Julienne	Jean-Louis	Christophe
Estelle	0.76	0.03	0.07	0.14
Julienne	0.02	0.80	0.07	0.11
Jean-Louis	0	0	1	0
Christophe	0.10	0.14	0.13	0.62

Avec Simulation Motrice

	Estelle	Julienne	Jean-Louis	Christophe
Estelle	0.84	0.07	0.04	0.04
Julienne	0.02	0.97	0	0.01
Jean-Louis	0.24	0.14	0.51	0.11
Christophe	0.05	0.01	0	0.93

Les taux de reconnaissance sont similaires : 79,5% sans simulation motrice et 81,25% avec simulation motrice. Nous observons que pour trois scripteurs, la simulation motrice améliore la reconnaissance. Il n'y a que pour le scripteur *Jean-Louis* que la reconnaissance est fortement dégradée. Etant donné que ce scripteur a un style d'écriture italique, les points de contrôle des traces perçues sont souvent des points d'annulation de vitesse à la fois en x et en y. Les lettres sont donc représentées par un nombre très faible de points de contrôle. Lors de la simulation motrice, des points de contrôle intermédiaires sont souvent générés et les traces simulées ont donc un nombre de points de contrôle plus important. C'est pourquoi la reconnaissance du scripteur Jean-Louis est plus faible avec simulation motrice que sans.

7.3.6 Lettres tronquées

Dans cette section, nous nous intéressons aux informations apportées par la simulation motrice lors de la reconnaissance de caractères. Nous avons observé précédemment, que la simulation motrice n'améliorait pas les performances lors de la reconnaissance de caractères. Nous proposons maintenant d'analyser un cas où la simulation motrice aide à la reconnaissance des caractères.

Nous nous intéressons au cas où les traces sont partiellement effacées, ce que nous appelons une lettre tronquée. Par exemple, la figure 7.18 présente une trace g partiellement effacée.

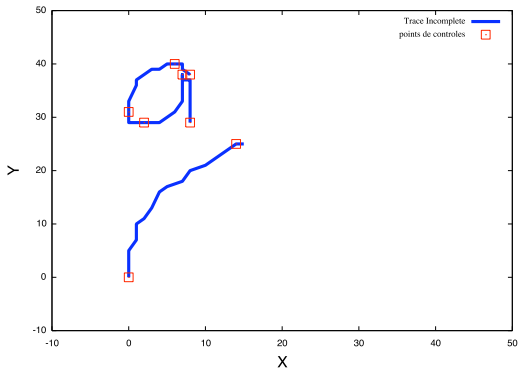


FIG. 7.18 – Lettre tronquée.

Nous posons les deux questions suivantes au modèle :

- $P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = Estelle]\ [\lambda_V = 1])$,
- $P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = Estelle]\ [\lambda_V = 1]\ [\lambda_L = 1]\ [\lambda_P = 1]\ [\lambda_S = 1])$,

avec $v_x^{0:M}, v_y^{0:M}$ correspondant à la trace présentée figure 7.18. La première question correspond à la reconnaissance de la lettre sans simulation motrice et la seconde avec.

La distribution de probabilité résultante aux deux questions est présentée dans le tableau ci-dessous.

	a	b	c	d	e	f	g	h	k	l	m	n	o	p	q	r	s	u	v	w	y	z	$\mathscr{U}$
Sans	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Avec	0	0	0	0	0	0	0.9	0	0	0	0	0	0	0	0.1	0	0	0	0	0	0	0	0

Nous observons que le modèle a reconnu avec certitude cette trace comme étant une lettre q lorsque la simulation motrice n'est pas active et comme un g lorsqu'elle l'est.

Analyse de la trace simulée Lors de la question de la lecture avec simulation motrice, le modèle simule l'écriture de la trace à partir des points de contrôle. Il détermine donc la trajectoire à l'origine de la trace afin de mieux reconnaître la lettre. La trajectoire simulée est présentée figure 7.19. Nous observons que la trajectoire simulée correspond à la lettre g.

Autres exemples La figure 7.20 présente un second exemple de reconnaissance de traces tronquées. A gauche, nous observons la trace fournie au modèle : une lettre o dont une partie a été effacée. Lors de la reconnaissance sans simulation motrice, la trace a été reconnue avec une probabilité de 1 comme étant une lettre a. Avec la simulation motrice, la trace a été reconnue avec certitude (probabilité de 1) comme étant une lettre o.

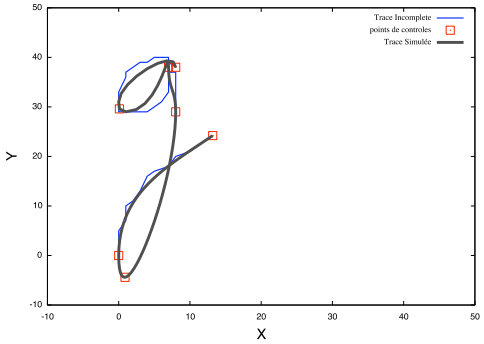


FIG. 7.19 – Simulation motrice de la lettre tronquée.

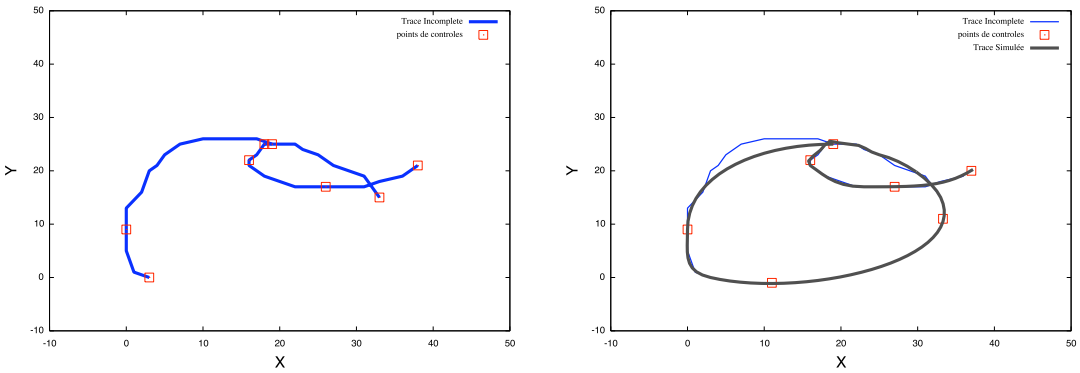


FIG. 7.20 – A gauche, trace représentant une lettre *o* dont une partie a été effacée. A droite, trajectoire simulée lors de la lecture avec simulation motrice.

La figure 7.21 présente une dernier exemple de lecture de traces tronquées avec simulation motrice. Lorsque la simulation motrice n'est pas activée, la trace est reconnue comme étant une lettre *p* (avec une probabilité de 1). Avec la simulation motrice, une trajectoire est générée (figure 7.21, à droite) et la trace est reconnue comme étant une lettre *u* (probabilité de 1).

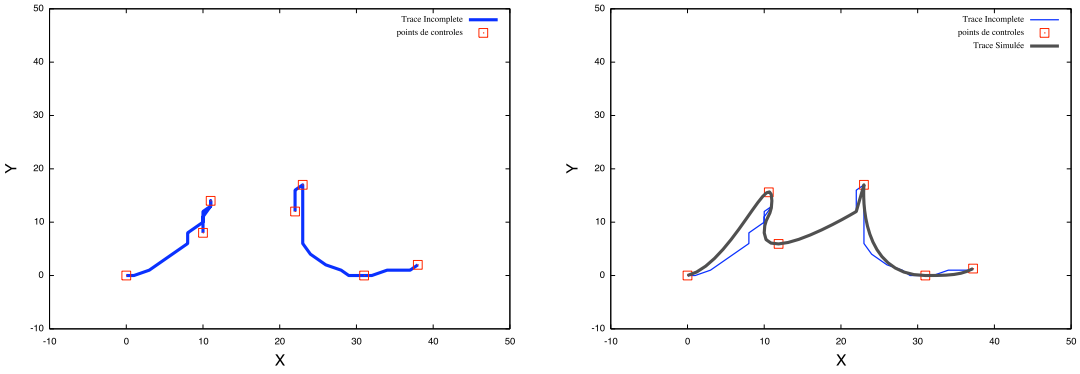
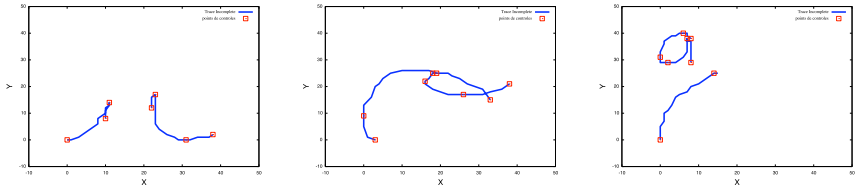
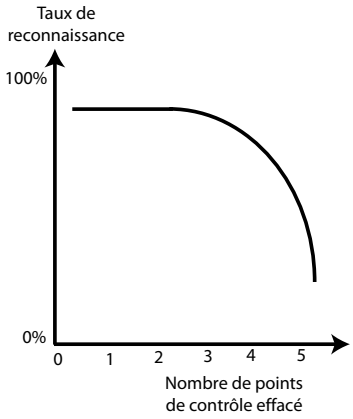
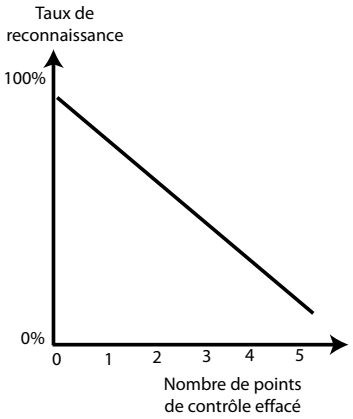


FIG. 7.21 – A gauche, trace représentant une lettre *u* dont une partie a été effacée. A droite, trajectoire simulée lors de la lecture avec simulation motrice.

Suggestion de protocole expérimental Nous venons de voir que pour le modèle, la reconnaissance de traces tronquées est améliorée avec la simulation motrice du geste. Or, nous n'avons aucune comparaison de résultats avec l'humain. Il est possible de se poser la question suivante : *sommes-nous capable de reconnaître des lettres tronquées et ce jusqu'à quel niveau d'effacement ?* Afin d'apporter des éléments de réponse à cette question, nous suggérons deux protocoles expérimentaux. Le premier met en relation le niveau de bruit de la lettre présentée et l'apport d'information dû à la simulation motrice. Le second, quant à lui, fixe le niveau de bruit et compare la lecture lorsque le processus moteur est biaisé ou non par une tâche d'interférence.

<table>
<tr><td>Hypothèses</td><td>Lors de la lecture, l'action d'écriture est simulé.</td></tr>
<tr><td>Matériel</td><td>Traces représentant des lettres dont une partie a été effacée aléatoirement.
 
Plusieurs niveaux pourront être déterminés suivant le nombre de points de contrôle effacés.</td></tr>
<tr><td>Procédure</td><td>Présenter aux sujets des points lumineux se déplaçant le long de la trajectoire et demander aux sujets de déterminer la lettre correspondante.</td></tr>
<tr><td>Résultats</td><td>Etablir les taux de reconnaissance pour chacune des lettres ainsi que les matrices de confusion pour chaque niveau.</td></tr>
<tr><td>Analyse</td><td>Si la courbe (taux de reconnaissance en fonction du nombre de points de contrôle effacé) comporte un palier, comme présenté à droite, nous pouvons valider l'hypothèse que lors de la lecture, l'action d'écriture est simulée. En effet, nous pouvons faire l'hypothèse que l'écriture simulée permet de déterminer la trajectoire à l'origine de cette trace et donc de déterminer les points de contrôle manquants. C'est pourquoi la courbe du taux de reconnaissance ne chutera pas si l'on enlève des points de contrôle (un nombre raisonnable). La simulation motrice comblera ce manque d'information.
 

A l'inverse, si la courbe obtenue a le même profil que celui présenté dans la figure à droite, l'hypothèse de l'écriture simulée ne pourra pas être validée. Cette courbe suggère que le taux de reconnaissance est directement proportionnel au nombre de points de contrôle effacés. Aucune information simulée n'est donc ajoutée lors de la reconnaissance.
 
</td></tr>
</table>

Hypothèses	Lors de la lecture, l'action d'écriture est simulé.
Matériel	Traces représentant des lettres dont une partie a été effacée aléatoirement.
Procédure	Présenter aux sujets des points lumineux se déplaçant le long de la trajectoire et demander aux sujets de déterminer la lettre correspondante. Pendant qu'ils effectuent cette tâche, demander au sujet d'effecteur une tâche motrice pour tuer le processus moteur. Signer leur nom ou manipuler un objet par exemple pourra permettre d'inhiber d'éventuels processus moteurs par une tâche d'interférence. Deux cas seront testés : avec et sans interférence.
Résultats	Etablir les taux de reconnaissance pour chacune des lettres et comparer la reconnaissance (dans un même niveau de difficulté) avec et sans tâche d'interférence.
Analyse	Si les taux de reconnaissance (avec ou sans interférence) sont similaires, cela suggère que le processus moteur n'est pas impliqué dans la tâche de perception. A l'inverse, si les taux de reconnaissance avec interférence sont très inférieurs aux taux de reconnaissance sans interférence, il est possible de dire qu'étant donné que le processus moteur était en train de réaliser une autre tâche, il n'est pas intervenu lors de la reconnaissance. Cela suggère que la simulation motrice du geste aide à la reconnaissance.

7.4 Reproduction d'expérience

Kandel *et al.* (2000a) se sont intéressés aux informations cinématiques contenues dans le tracé des lettres qui permettent de prédire la lettre à venir et plus précisément à quel moment cette prédiction a lieu. Des recherches ont montré que nous exploitons l'information cinématique contenue dans la partie descendante de la lettre l pour prédire la lettre suivante. Dans la suite de ce travail, nous allons présenter l'expérience en détail et, pas à pas, nous donnerons la modélisation correspondante.

7.4.1 Analyse et modélisation

Afin de mieux présenter l'expérience et la modélisation correspondante, nous avons utilisé un code couleur. Lorsque nous présentons l'expérience, nous utilisons la couleur grise alors que le modèle est présenté en noir. De plus, les questions de lecture sans simulation motrice sont présentées en vert, et celles avec simulation motrice en rouge.

7.4.1.1 Modélisation des lettres

Expérience L'expérience d'anticipation perceptive a consisté à présenter sur un écran d'ordinateur un « l » appartenant aux couples « ll » ou « le ».

Variables Afin de reproduire cette expérience, nous avons défini des couples de lettres (« ll » et « le ») comme étant des lettres du modèle. Le modèle connaît donc uniquement deux lettres : la lettre « ll » et la lettre « le ». La variable L est donc redéfinie comme ci-dessous :

$$L = \{ll, le\}. \tag{7.19}$$

Décomposition de la distribution conjointe La décomposition de la distribution de probabilité conjointe reste inchangée.

Termes de la décomposition Etant donné que nous avons changé la définition de la variable L, nous devons effectuer une nouvelle phase d'apprentissage des paramètres des formes paramétriques des termes suivants :

$$P(C_{LV} \mid L\ W)\ \ P(C_{LP} \mid L\ W)\ \ P(C_{LS} \mid L\ W). \tag{7.20}$$

Comme précédemment, ces trois termes sont défini de façon identiques. Les mêmes hypothèses d'indépendance sont appliquées. Chacun des termes peut donc être décomposé de la façon suivante :

$$P(C_{LV} \mid L\ W) = \begin{pmatrix} P(C^0_{LVx} \mid L\ W)\ P(C^0_{LVy} \mid L\ W) \\ P(C^0_{LV\dot{x}} \mid L\ W)\ P(C^0_{LV\dot{y}} \mid L\ W) \end{pmatrix} \prod_{n=1}^{N} \begin{pmatrix} P(C^n_{LVx} \mid C^{n-1}_{LVx}\ L\ W)\ P(C^n_{LVy} \mid C^{n-1}_{LVy}\ L\ W) \\ P(C^n_{LV\dot{x}} \mid C^{n-1}_{LV\dot{x}}\ L\ W)\ P(C^n_{LV\dot{y}} \mid C^{n-1}_{LV\dot{y}}\ L\ W) \end{pmatrix} \tag{7.21}$$

La forme paramétrique de ces termes reste inchangée : la distribution de probabilité de chacun des termes du produit est une loi de succession de Laplace.

Identification des paramètres Il nous reste maintenant à déterminer les paramètres des lois de successions de Laplace. Ces paramètres sont définis lors d'une phase d'apprentissage de notre modèle. L'apprentissage est réalisé par un unique scripteur ($W = Estelle$). L'apprentissage est similaire à celui réalisé section 3.2.2.2. Il est réalisé sur un total de 40 traces (20 fois chaque paire de lettre).

7.4.1.2 Prédiction des lettres

Expérience Durant l'expérience, un « l » est progressivement tracé sur un écran et le sujet doit prédire à quel couple de lettre ce « l » appartient.

Modélisation Afin de réaliser cette prédiction, la question suivante est posée au modèle : **A quel couple de lettres, la trace $\{v_x^{0:M}, v_y^{0:M}\}$ correspondant à un « l » appartient-elle ?**

Mathématiquement, cette question correspond à :

$$Q_1 \ : \ P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]). \tag{7.22}$$

Etant donné que nous souhaitons analyser l'influence de la simulation motrice lors de la reconnaissance de lettres, nous posons également la question suivante :

$$Q_2 \ : \ P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]\ [\lambda_P = 1]\ [\lambda_L = 1]\ [\lambda_S = 1]). \quad (7.23)$$

Nous rappelons qu'ajouter λ_P, λ_L et λ_S en partie droite permet d'activer la simulation motrice.

Afin de simplifier l'analyse de ces deux questions, nous utilisons le code couleur suivant : vert lorsque nous parlons de reconnaissance sans simulation motrice (Q_1) et rouge avec (Q_2). L'inférence bayésienne permet de répondre aux questions Q_1 et Q_2. Elle est la même que pour les questions de lecture. Nous avons donc pour Q_1 :

$$\begin{aligned} &P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]) \\ &\quad \propto P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]). \end{aligned} \quad (7.24)$$

et pour Q_2 :

$$\begin{aligned} &P\left(L \ \middle| \ \begin{array}{c} [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w] \\ [\lambda_V = 1]\ [\lambda_L = 1]\ [\lambda_P = 1]\ [\lambda_S = 1] \end{array}\right) \\ &\quad \propto \left(\begin{array}{c} P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LS} = h(g(C_{LV}))] \mid L\ [W = w])) \end{array}\right). \end{aligned} \quad (7.25)$$

Ces questions sont équivalentes à déterminer la lettre en fonction des points de contrôle. En effet, étant donné que le terme $P(C_V \mid V)$ suit une distribution de probabilité de type Dirac, il n'y a donc, pour une trace, qu'une seule séquence de points de contrôle telle que la distribution de probabilité soit non nulle. La questions Q_1 et son inférence sont donc équivalentes à :

$$\begin{aligned} &P(L \mid [C_{LV} = f(v_x^{0:M}, v_y^{0:M})]\ [W = w]) \\ &\quad \propto P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]). \end{aligned} \quad (7.26)$$

Il en est de même pour Q_2 :

$$\begin{aligned} &P\left(L \ \middle| \ [C_{LV} = f(v_x^{0:M}, v_y^{0:M})]\ [C_{LS} = h(g(C_{LV}))]\ [W = w]\right) \\ &\quad \propto \left(\begin{array}{c} P([C_{LV} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LS} = h(g(C_{LV}))] \mid L\ [W = w])) \end{array}\right). \end{aligned} \quad (7.27)$$

7.4.1.3 Information cinématique

Expérience Nous avons vu précédemment que durant l'expérience des « l » sont progressivement tracés sur un écran. Les traces correspondant au « l » ont été au préalable normalisées et modifiées pour qu'aucune information géométrique ne permette de distinguer les différents « l ». Seule l'information cinématique permet de discriminer entre le « l » appartenant au couple « ll » ou « le ».

Modélisation Avec la définition des variables de points de contrôle (C_{LV} et C_{LS}) et les hypothèses d'indépendance entre les quatre dimensions X, Y, $\dot{X}$ et $\dot{Y}$, les questions sont équivalentes à :

$$\begin{aligned} &P\left(L \ \middle| \ \begin{array}{cc} [C_{LVX} = f(v_x^{0:M}, v_y^{0:M})] & [C_{LVY} = f(v_x^{0:M}, v_y^{0:M})] \\ [C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] & [C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})]\ [W = w] \end{array}\right) \\ &\quad \propto \left(\begin{array}{cc} P([C_{LVX} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) & P([C_{LVY} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) & P([C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \end{array}\right), \end{aligned} \quad (7.28)$$

et

$$P\left(L \;\middle|\; \begin{array}{l} [C_{LVX} = f(v_x^{0:M}, v_y^{0:M})]\ [C_{LVY} = f(v_x^{0:M}, v_y^{0:M})] \\ [C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \\ [C_{LSX} = h(g(C_{LVX}))]\ [C_{LSY} = h(g(C_{LVY}))] \\ [C_{LS\dot{X}} = h(g(C_{LV\dot{X}}))]\ [C_{LS\dot{Y}} = h(g(C_{LV\dot{Y}}))]\ [W = w] \end{array}\right)$$
$$\propto \begin{pmatrix} P([C_{LVX} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LVY} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LSX} = h(g(C_{LV}))] \mid L\ [W = w]))\ P([C_{LSY} = h(g(C_{LV}))] \mid L\ [W = w])) \\ P([C_{LS\dot{X}} = h(g(C_{LV}))] \mid L\ [W = w]))\ P([C_{LS\dot{Y}} = h(g(C_{LV}))] \mid L\ [W = w])) \end{pmatrix}. \quad (7.29)$$

Etant donné que seule l'information cinématique nous intéresse, nous éliminons les informations de position de la question. Les variables connues correspondent donc uniquement aux vitesses des points de contrôle. Les questions correspondent alors à déterminer à quel couple de lettres appartiennent les vitesses des points de contrôle. Pour Q_1, nous avons :

$$P\left(L \mid [C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})]\ [W = w]\right) \quad (7.30)$$
$$\propto \sum_{C_{LVX}\ C_{LVY}} \begin{pmatrix} P([C_{LVX} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LVY} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \end{pmatrix},$$

et pour Q_2 :

$$P\left(L \;\middle|\; \begin{array}{l} [C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \\ [C_{LS\dot{X}} = h(g(C_{LV}))]\ [C_{LS\dot{Y}} = h(g(C_{LV}))]\ [W = w] \end{array}\right) \quad (7.31)$$
$$\propto \begin{pmatrix} \sum\limits_{\substack{C_{LVX}\ C_{LVY} \\ C_{LSX}\ C_{LSY}}} & \begin{array}{l} P([C_{LVX} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LVY} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LSX} = h(g(C_{LVX}))] \mid L\ [W = w]))\ P([C_{LSY} = h(g(C_{LVY}))] \mid L\ [W = w])) \\ P([C_{LS\dot{X}} = h(g(C_{LV}))] \mid L\ [W = w]))\ P([C_{LS\dot{Y}} = h(g(C_{LV\dot{Y}}))] \mid L\ [W = w])) \end{array} \end{pmatrix}.$$

En appliquant la règle de normalisation, nous obtenons :

$$P\left(L \mid [C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})]\ [W = w]\right)$$
$$\propto P([C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]), \quad (7.32)$$

et

$$P\left(L \;\middle|\; \begin{array}{l} [C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \\ [C_{LS\dot{X}} = h(g(C_{LV}))]\ [C_{LS\dot{Y}} = h(g(C_{LV}))]\ [W = w] \end{array}\right)$$
$$\propto \begin{pmatrix} P([C_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C_{LS\dot{X}} = h(g(C_{LV}))] \mid L\ [W = w]))\ P([C_{LS\dot{Y}} = h(g(C_{LV}))] \mid L\ [W = w])) \end{pmatrix}. \quad (7.33)$$

7.4.1.4 Découpage des traces

Expérience Afin de déterminer à quel moment la prédiction de la lettre à venir à lieu, chaque « l » est coupé en plusieurs morceaux à partir du début de la partie descendante et jusqu'à la fin de la trace. La figure 7.22 présente le découpage réalisé. Le dernier morceau de chaque stimulus se termine au point le plus bas de la descente afin de ne fournir aucune information sur la lettre à venir. Dans cette expérience, les « l » ont été coupés en huit morceaux (G_1 à G_8). Chaque morceau a exactement la même longueur : le périmètre de la partie descendante a été divisé en huit secteurs identiques.

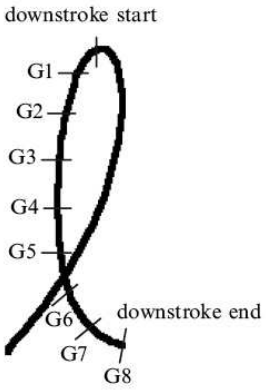


FIG. 7.22 – La lettre « l » est coupée en huit secteurs (pour les points G_1 à G_8)(figure tirée de Kandel *et al.* (2000a)).

Modélisation Nous souhaitons également modéliser ce découpage de la lettre. Cependant, étant donné que notre reconnaissance de la lettre se base uniquement sur les points de contrôle, nous ne pouvons pas reproduire à un degré aussi précis le découpage. Chaque « l » sera donc découpé en trois morceaux : G_0, G_{34} et G_8 comme le présente la figure 7.23. G_0

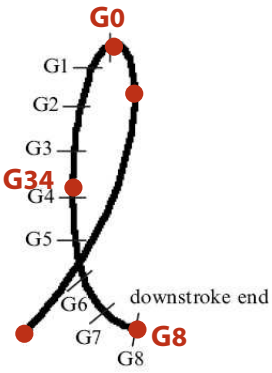


FIG. 7.23 – La lettre « l » est coupée à trois niveaux (G_0, G_{34} et G_8). (figure tirée de Kandel *et al.* (2000a))

correspond au troisième point de contrôle de la trace. Il est donc situé au point le plus haut du « l ». Le quatrième point de contrôle que nous nommerons G_{34} se situe dans la partie descendante du « l ». Il se situe entre les points G_3 et G_4 du découpage de la lettre dans l'expérience. Enfin, le cinquième point de contrôle se situe au point le plus bas de la trajectoire. Il correspond donc au point G_8. Ce découpage revient à tronquer les traces à trois niveaux différents.

Pour modéliser le premier stimulus (G_0), nous posons la question suivante au modèle (pour Q_1) :

$$P\left(L \;\middle|\; [C_{LV\dot{X}}^{0:2} = f(v_x^{0:M}, v_y^{0:M})]\ [C_{LV\dot{Y}}^{0:2} = f(v_x^{0:M}, v_y^{0:M})]\ [W = w]\right)$$
$$\propto P([C_{LV\dot{X}}^{0:2} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C_{LV\dot{Y}}^{0:2} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]). \quad (7.34)$$

Les variables connues correspondent aux vitesses des trois premiers points de contrôle

$(0:2)$. Etant donné que G_0 correspond au troisième point de contrôle, nous ne fournissons pas les valeurs des autres points de contrôle au modèle. L'inférence de cette question (en appliquant la règle de normalisation) est proportionnelle aux vitesses des trois premiers points de contrôle. Pour Q_2, nous faisons de manière identique :

$$P\left(L \;\middle|\; \begin{array}{l}[C^{0:2}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C^{0:2}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \\ [C^{0:2}_{LS\dot{X}} = h(g(C_{LV}))]\ [C^{0:2}_{LS\dot{Y}} = h(g(C_{LV}))]\ [W = w]\end{array}\right) \propto \left(\begin{array}{l}P([C^{0:2}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C^{0:2}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C^{0:2}_{LS\dot{X}} = h(g(C_{LV}))] \mid L\ [W = w]))\ P([C^{0:2}_{LS\dot{Y}} = h(g(C_{LV}))] \mid L\ [W = w]))\end{array}\right). \quad (7.35)$$

Les deux autres stimuli sont modélisés de façon identique. Le stimulus G_{34} correspond au quatrième point de contrôle, nous avons donc :

$$P\left(L \;\middle|\; [C^{0:3}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C^{0:3}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})]\ [W = w]\right) \propto P([C^{0:3}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C^{0:3}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]), \quad (7.36)$$

et

$$P\left(L \;\middle|\; \begin{array}{l}[C^{0:3}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C^{0:3}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \\ [C^{0:3}_{LS\dot{X}} = h(g(C_{LV}))]\ [C^{0:3}_{LS\dot{Y}} = h(g(C_{LV}))]\ [W = w]\end{array}\right) \propto \left(\begin{array}{l}P([C^{0:3}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C^{0:3}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C^{0:3}_{LS\dot{X}} = h(g(C_{LV}))] \mid L\ [W = w]))\ P([C^{0:3}_{LS\dot{Y}} = h(g(C_{LV}))] \mid L\ [W = w]))\end{array}\right). \quad (7.37)$$

Le stimulus G_8 correspond au cinquième point de contrôle, ce qui nous donne :

$$P\left(L \;\middle|\; [C^{0:4}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C^{0:4}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})]\ [W = w]\right) \propto P([C^{0:4}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C^{0:4}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]), \quad (7.38)$$

et

$$P\left(L \;\middle|\; \begin{array}{l}[C^{0:4}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C^{0:4}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \\ [C^{0:4}_{LS\dot{X}} = h(g(C_{LV}))]\ [C^{0:4}_{LS\dot{Y}} = h(g(C_{LV}))]\ [W = w]\end{array}\right) \propto \left(\begin{array}{l}P([C^{0:4}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w])\ P([C^{0:4}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \mid L\ [W = w]) \\ P([C^{0:4}_{LS\dot{X}} = h(g(C_{LV}))] \mid L\ [W = w]))\ P([C^{0:4}_{LS\dot{Y}} = h(g(C_{LV}))] \mid L\ [W = w]))\end{array}\right). \quad (7.39)$$

7.4.1.5 Seuil de probabilité

Expérience Une fois la trace présentée, les sujets devaient prédire à quel couple le « l » présenté appartenait en appuyant sur un bouton. Plus précisément, il devait appuyer sur le bouton LE s'ils étaient *sûrs* ou *presque sûrs* que le « l » appartenait au couple « le ». Sinon, ils devaient appuyer sur l'autre bouton.

Modélisation Afin de modéliser les termes *sûr* et *presque sûr*, nous avons défini un paramètre α correspondant à la probabilité minimale pour que la réponse soit validée. Quelque soit le stimulus présenté, nous analysons le terme suivant :

$$P\left([L = LE] \;\middle|\; [C^{0:i}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C^{0:i}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})]\ [W = w]\right) \geq \alpha. \quad (7.40)$$

Plusieurs cas sont alors possibles :

- Le stimulus présenté appartient au couple LE :
  - l'inégalité 7.40 est vraie : la réponse du modèle est considérée comme juste.
  - l'inégalité 7.40 est fausse : la réponse du modèle est considérée comme fausse.
- Le stimulus présenté appartient au couple LL :
  - l'inégalité 7.40 est vraie : la réponse du modèle est considérée comme fausse.

– l'inégalité 7.40 est fausse : la réponse du modèle est considérée comme juste.

Le même procédure est appliquée pour Q_2 :

$$P\left([L = LE] \;\middle|\; \begin{array}{l}[C^{0:i}_{LV\dot{X}} = f(v_x^{0:M}, v_y^{0:M})]\ [C^{0:i}_{LV\dot{Y}} = f(v_x^{0:M}, v_y^{0:M})] \\ [C^{0:i}_{LS\dot{X}} = h(g(C_{LV}))]\ [C^{0:i}_{LS\dot{Y}} = h(g(C_{LV}))]\ [W = w]\end{array}\right) \geq \alpha. \tag{7.41}$$

7.4.2 Résultats

Dans cette section, nous présentons les résultats de l'expérience de psychophysique que nous comparons aux résultats du modèle.

7.4.2.1 Résultats de l'expérience

Le pourcentage de réponses correctes *le* en fonction des huit coupures de la partie descendante (G_1 à G_8) est présenté figure 7.24.

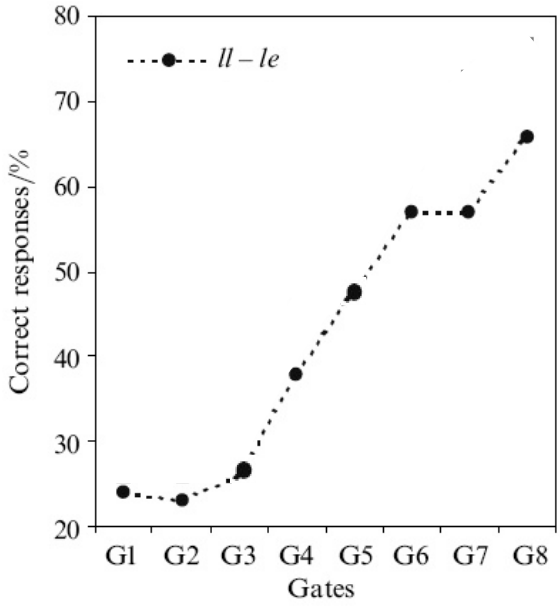


FIG. 7.24 – Taux de reconnaissance en fonction de la quantité d'information montrée aux sujets (figure extraite de Kandel *et al.* (2000a)).

Nous observons que le pourcentage de réponses correctes augmente significativement avec le nombre de morceaux présenté aux sujets. Par exemple, lorsque la trajectoire est présentée jusqu'au niveau G_6 (ce qui correspond à 75% de la trajectoire de la partie descendante du « l »), le taux de reconnaissance est de 57%. Nous observons que la performance est stable durant les trois premiers morceaux et augmente significativement à partir de G_4. Cela indique que le taux de prédiction correcte n'est pas directement proportionnel à la quantité d'information montrée aux sujets. Mais plutôt, que l'information cinématique nécessaire à la détection se trouve entre G_4 et G_6. Ce résultat révèle qu'avec 75% de la trajectoire les sujets peuvent identifier avec certitude la lettre à venir (taux de réponse correcte égal à 57%). La présentation de la trajectoire entière n'est donc pas nécessaire pour une prédiction correcte.

7.4.2.2 Analyse des résultats

Grâce à ces questions, nous pouvons simuler le modèle dans les mêmes conditions que l'expérience de Kandel *et al.* (2000a). Cette simulation nous permet d'évaluer les résultats du modèle et donc de déterminer les taux de reconnaissance. Pour chaque condition, nous posons la question sur 5 traces (différentes des traces ayant servies à l'apprentissage). De plus, nous fixons la valeur de α à 0.7 : la probabilité de reconnaître la lettre doit donc être supérieure à 0.7 pour être validée.

Sans simulation motrice Nous allons, dans un premier temps, nous intéresser aux résultats de la question Q_1, c'est-à-dire à la prédiction de la lettre à venir sans utiliser la simulation motrice. La figure 7.25 présente le pourcentage de réponses correctes en fonction du nombre de points de contrôle présenté. En rouge, nous avons reproduit la courbe présentée figure 7.24.

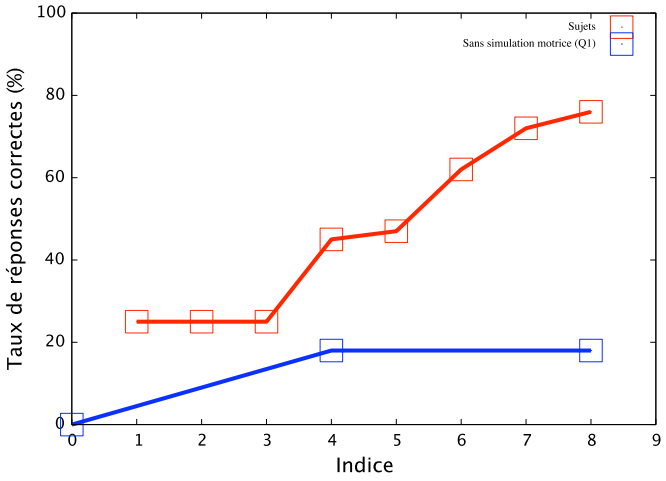


FIG. 7.25 – Pourcentage de réponses correctes en fonction du nombre de points de contrôle. En rouge, pourcentage des sujets et en bleu pourcentage de la question Q_1.

Elle correspond donc au pourcentage de réponses correctes des sujets. En bleu, nous avons tracé celui du modèle. Nous observons que les pourcentages de réponses correctes obtenus par le modèle sont différents des ceux obtenus par les sujets. Les informations cinématiques des points de contrôle sont donc insuffisantes pour prédire la lettre à venir.

Avec simulation motrice Nous nous intéressons maintenant au pourcentage de réponses correctes de la question Q_2 qui correspond à la prédiction de la lettre à venir en utilisant la simulation motrice. La figure 7.26 présente le pourcentage de réponses correctes en fonction du nombre de points de contrôle présentés.

Nous avons ajouté, en vert, le pourcentage de réponses correctes pour la question Q_2. Nous observons que l'allure de la courbe verte est similaire à celui de la courbe rouge : le modèle avec simulation motrice permet de reproduire correctement l'expérience.

Analyse de l'information apportée par la simulation motrice Seule l'information cinématique entre les deux couples de lettre « ll » et « le » permet de prédire la lettre à venir. Or, la différence entre les distributions de probabilité $P(C_{VY} \mid [L = LE])$ et $P(C_{VY} \mid [L = LL])$

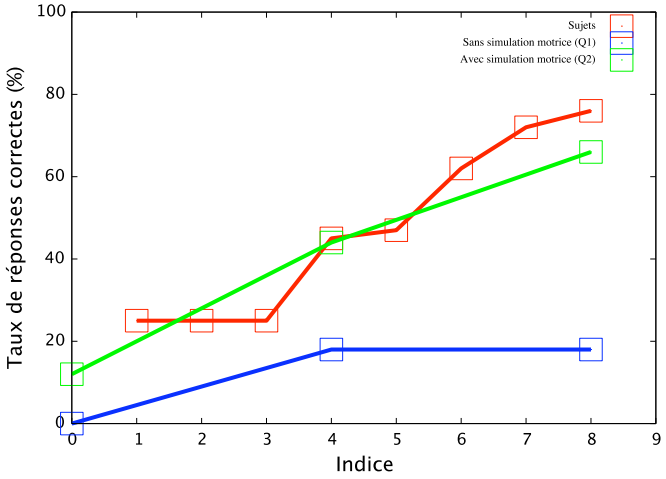


FIG. 7.26 – Pourcentage de réponses correctes en fonction du nombre de points de contrôle. En rouge, pourcentage des sujets, en bleu pourcentage de la question Q_1 et en vert, pourcentage de la question Q_2.

est faible. C'est pourquoi, la distribution de probabilité de $P([L = LE] \mid C_{VY})$ est souvent proche de 0.5.

Nous observons que l'inférence de la question Q_2 correspond au produit de quatre termes alors qu'il correspond au produit de 2 termes dans le cas de l'inférence de Q_1. Etant donné que les distributions de probabilité sont multipliées, plus le nombre de termes est important, plus le « signal » est augmenté. C'est ce qu'illustre la figure 7.27.

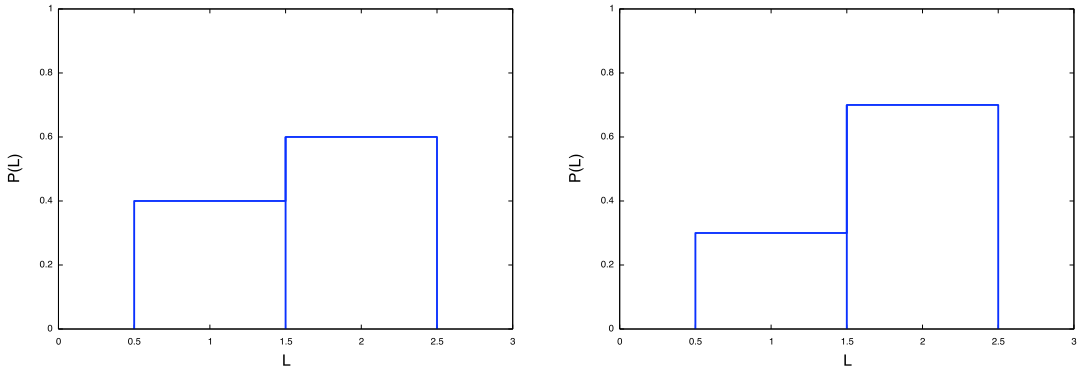


FIG. 7.27 – A gauche, distribution de probabilité sur la variable L. Nous l'avons multiplié par elle même pour obtenir la distribution de probabilité de droite.

A gauche, la distribution de probabilité est proche de l'uniforme ($0,6$ et $0,4$). Nous avons multiplié cette distribution de probabilité par elle même pour obtenir celle de droite. Nous observons qu'elle est plus « piquée ». C'est le même phénomène qui se produit lors de la reconnaissance. La simulation motrice permet donc d'apporter plus d'informations qui permettent de différencier les deux cas « ll » et « le ». Les taux de reconnaissance sont donc améliorés.

Dans la section suivante, nous verrons qu'il y a un lien étroit entre le seuil α et les taux de reconnaissance. Cependant, nous pouvons d'ores et déjà suggérer qu'il y a une valeur minimale

du seuil α pour laquelle les résultats ont un sens. En effet, la plupart des distributions de probabilité sur L sont comprises entre 0.4 et 0.6 sans simulation motrice. Si l'on avait donc pris un seuil à 0.6, tous les taux de reconnaissance seraient proches de 50%, et ce quelque soit le nombre de points de contrôle présenté. On aurait donc un profil de résultat plat, avec ou sans simulation motrice. Il faut donc étudier des valeurs de α supérieur à 0.6 pour obtenir des profils en pente.

7.4.2.3 Variations du paramètre α

Dans cette première analyse des résultats, la valeur du paramètre α était fixée à 0.7. Cependant, nous n'avons aucune idée de la valeur de ce paramètre étant donné qu'il traduit les termes *sûr* et *presque sûr*. Nous avons donc de nouveau simulé le modèle avec cette fois $\alpha = 0.8$.

La figure 7.28 présente les résultats du modèle comparés aux résultats des sujets.

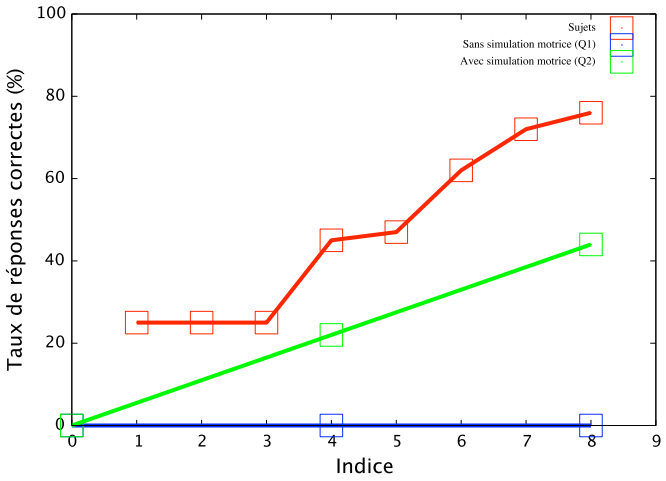


FIG. 7.28 – Pourcentage de réponses correctes en fonction du nombre de points de contrôle. En rouge, pourcentage des sujets, en bleu pourcentage de la question Q_1 et en vert, pourcentage de la question Q_2.

Une nouvelle fois, les pourcentages de réponses correctes des sujets sont présentés en rouge (courbe inchangée par rapport à la figure 7.26), ceux de la question Q_1 en bleu et ceux de la question Q_2 en vert. Nous observons que les courbes vertes et bleues ont été décalées vers le bas par rapport à la figure 7.26, mais que leurs allures n'ont pas changé. Avec le paramètre α égal à 0.8, nous avons une nouvelle fois les résultats du modèle utilisant la simulation motrice plus proche des résultats des sujets que le modèle ne l'utilisant pas.

7.4.2.4 Conclusion

Dans cette section, nous avons présenté une expérience de psychophysique. Nous avons simulé les conditions expérimentales de l'expérience dans le modèle BAP. Nous avons pu tester deux cas particuliers de la tâche de lecture : avec et sans simulation motrice. Pour chaque cas, nous avons recréé les courbes de pourcentage de réponses correctes. Ainsi, nous avons pu juger, qualitativement, de l'adéquation du modèle aux données dans chacun des cas.

Nous avons déterminé que la tâche de lecture avec simulation motrice obtenait des résultats plus proche de ceux de l'expérience que la tâche sans simulation motrice.

7.5 Résumé

Dans ce chapitre, nous avons présenté l'interaction de la perception et de l'action dans le cadre de la lecture et de l'écriture. Nous avons tout d'abord présenté une tâche cognitive sollicitant à la fois la perception et l'action : la copie. Nous avons ensuite présenté notre modèle de simulation motrice de l'écriture. Nous avons constaté que dans le cadre de la lecture « normale », la boucle simulée n'apportait aucune information supplémentaire pour la reconnaissance de lettres. Nous avons ensuite testé cette boucle sur un autre exemple : la reconnaissance de traces partiellement effacées. Nous avons analysé l'utilité de la simulation motrice dans ce cas. Enfin, nous avons reproduit une expérience de psychophysique ce qui nous amène une nouvelle fois à conclure que simuler le geste d'écriture aide à la reconnaissance des lettres.

Chapitre 8

Conclusion

Dans ce manuscrit, nous nous sommes intéressés à la modélisation d'un système sensori-moteur. Nous avons étudié :

– l'influence du système moteur lors de la perception,

– les mécanismes qui permettent de réaliser un même mouvement avec différents effecteurs.

Ces études ont été réalisées dans le cadre particulier de la lecture et de l'écriture. Notre principale contribution est donc la réalisation d'un système sensori-moteur de la lecture et de l'écriture : le modèle BAP. Dans ce chapitre, nous rappelons ses hypothèses et leurs traductions mathématiques. Nous reviendrons également sur la résolution des différentes tâches cognitives du modèle avant de nous intéresser, plus généralement, aux perspectives ouvertes par ce travail.

8.1 Hypothèses

Pour définir l'architecture du modèle BAP, nous nous sommes appuyés sur une étude bibliographique articulée autour de quatre questions principales.

Y a-t-il une représentation unique des lettres pour la perception et l'action ou s'agit-il de représentations distinctes ?

Cette première question concerne l'unicité des représentations internes sensorielles et motrices. Nous nous sommes intéressés aux différentes théories issues du domaine de la parole telles que la théorie motrice de la perception et la théorie pour le contrôle de l'action. Nous avons fait l'hypothèse que l'espace interne dans lequel est codée la perception est le même que celui dans lequel est codée la production. Nous avons donc une représentation unique qui est le pivot entre la perception et la production des lettres.

Quel est l'espace dans lequel les représentations internes sont définies ?

Nous nous sommes intéressés à l'espace dans lequel est définie la représentation interne des lettres. Nous avons tout d'abord étudié les observations expérimentales résultantes de l'équivalence motrice. Cet effet suggère que la représentation interne des lettres est indépendante de l'effecteur généralement utilisé pour effectuer le mouvement. Dans le cadre de l'écriture, cela suggère que la représentation interne est indépendante du bras et de la main utilisés pour écrire. Nous avons donc fait l'hypothèse que la représentation interne est définie dans le même espace que celui de la tâche, c'est-à-dire l'espace cartésien.

Comment sont définies les représentations interne des lettres ?

Dans le modèle BAP, les lettres sont représentées par des séquences de points particuliers de la trajectoire : ce sont les points où la vitesse en x ou en y s'annule. Nous avons choisi ces points de contrôle car ils sont pertinents perceptivement (ils sont géométriquement identifiables) mais aussi vis-à-vis de la production. En effet, à partir de la séquence de points de contrôle, le modèle génère une trajectoire. Dans le modèle BAP, la trajectoire est déterminée de façon à minimiser l'accélération (planification de trajectoire dans l'espace cartésien).

Une simulation de l'action influence t-elle la perception ?

Nous nous sommes intéressés au rôle de l'action lors de la perception. La théorie de la simulation suggère que lorsque nous percevons des gestes, nous simulons l'action qui a permis de réaliser ce geste. Dans le cadre de la perception des lettres, l'action d'écriture est simulée pour mieux reconnaître les lettres. Nous avons modélisé cette boucle dans le modèle BAP : le processus de reconnaissance des lettres correspond à extraire une séquence de points de contrôle de la trace lue et à simuler le geste à l'origine de cette trace. La simulation permet d'extraire une nouvelle séquence de points de contrôle et donc des informations supplémentaires pour la reconnaissance.

8.2 Modélisation

L'ensemble des hypothèses que nous venons de résumer nous a permis de définir l'architecture du modèle BAP. Cette architecture est rappelée figure 8.1.

Pour décrire notre modèle, nous avons choisi le formalisme de la Programmation Bayésienne. Le modèle BAP rassemble au sein de ce formalisme les hypothèses décrites ci-dessus. Dans ce formalisme, l'information est définie sous forme d'une distribution de probabilité. La distribution de probabilité conjointe et la décomposition choisie sont :

$$P\left(\begin{array}{cccccccc} C_{LV} & C_{LP} & C_{LS} & C_V & \lambda_p & \lambda_v & \lambda_s & \lambda_l \\ E & C_P & C_S & V & P & L & W & S \end{array}\right) \tag{8.1}$$

$$= \left(\begin{array}{l} P(\,C_{LV} \mid L \;\; W\,)\; P(\,C_{LP} \mid L \;\; W\,) \\ P(\,L\,)\; P(\,W\,)\; P(\,C_{LS} \mid L \;\; W\,) \\ P(\,\lambda_V \mid C_{LV} \;\; C_V\,)\; P(\,\lambda_P \mid C_{LP} \;\; C_P\,) \\ P(\,C_V \mid V\,)\; P(\,P \mid C_P\,)\; P(\,C_S \mid S\,) \\ P(\,\lambda_l \mid C_{LV} \;\; C_{LP}\,)\; P(\,S \mid P\,)\; P(\,P\,) \\ P(\,E \mid P\,)\; P(\,\lambda_S \mid C_{LS} \;\; C_S\,) \end{array}\right). \tag{8.2}$$

8.3 Résultats

Pour définir mathématiquement le modèle BAP, nous avons fait un ensemble d'hypothèses, spécifiés des connaissances et défini des termes mathématiques. Durant la phase d'utilisation du modèle, nous n'avons plus modifié le modèle BAP. La programmation bayésienne permet d'automatiser l'inférence. En effet, connaissant la distribution de probabilité conjointe, le

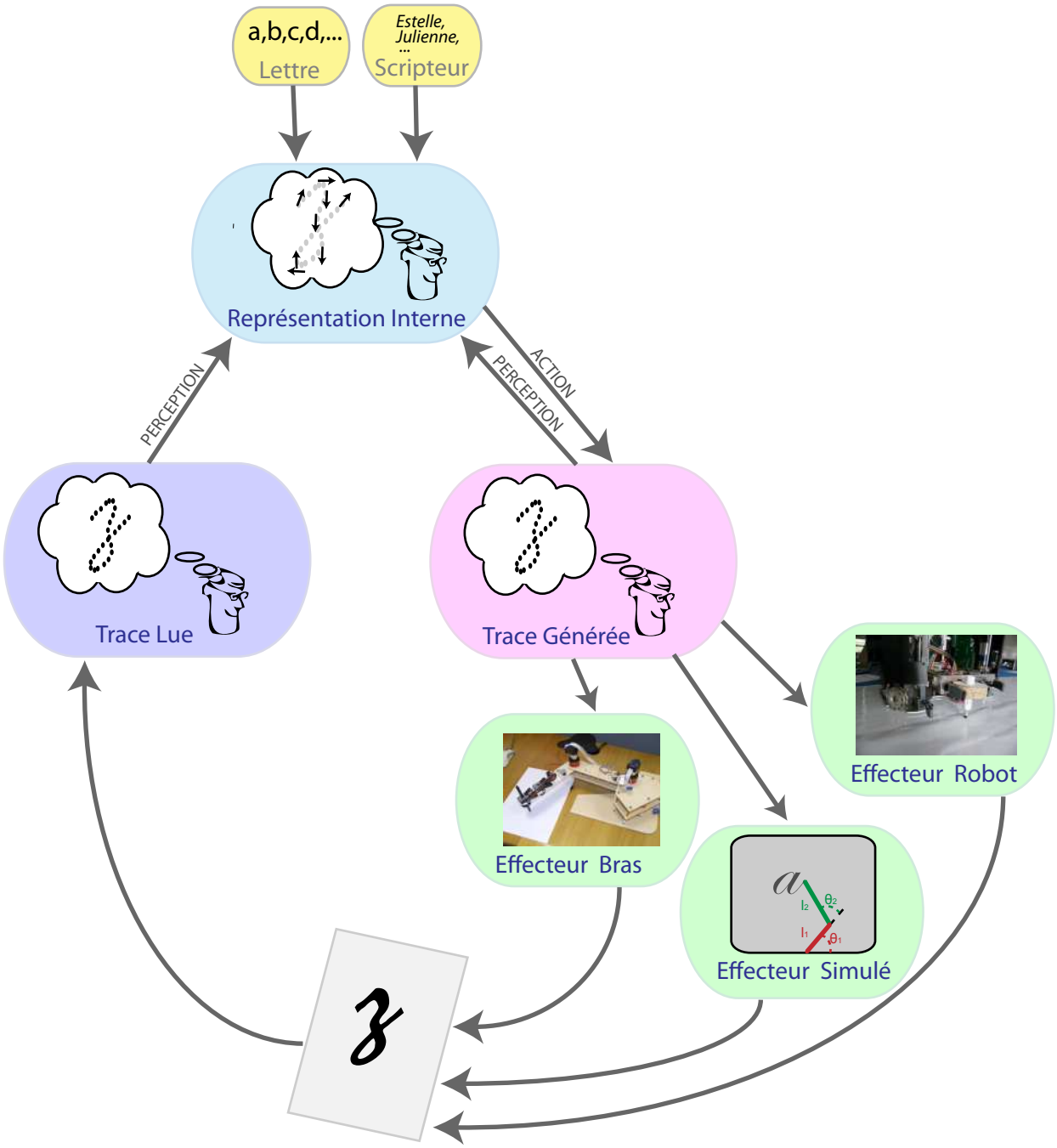


FIG. 8.1 – Architecture du modèle BAP : la représentation interne des lettres est le pivot entre la perception et la production. Les lettres sont représentées par une séquence de points de contrôle. Le processus de perception correspond à extraire les points de contrôle de la trace. Inversement, le processus d'action correspond à déterminer la trajectoire à partir des points de contrôle. Les trois effecteurs (bras simulé, bras robotique et robot holonome) permettent de réaliser la lettre à partir de la trajectoire déterminée par le processus d'action.

modèle est capable de résoudre plusieurs tâches cognitives. Chacune de ces tâches correspond à une *question probabiliste* :

- **Lecture** : nous avons simulé la tâche de lecture dans différentes conditions :

  Quelle est la lettre correspondant à la trace que je perçois et qui a été écrite par un scripteur que je connais ?

  – sans simulation motrice :

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]), \tag{8.3}$$

  – avec simulation motrice :

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]\ \ [\lambda_P = 1]\ [\lambda_S = 1]), \tag{8.4}$$

  Quelle est la lettre correspondant à la trace que je perçois ne sachant pas qui l'a écrite ?

  – sans simulation motrice :

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]), \tag{8.5}$$

  – avec simulation motrice :

$$P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]\ \ [\lambda_P = 1]\ [\lambda_S = 1]). \tag{8.6}$$

  Pour chacune de ces questions, nous avons déterminé les matrices de confusion des lettres ainsi que les taux de reconnaissance (entre 50% et 93% selon les cas). Nous rappelons que dans ce travail nous n'avons pas pour but de développer un système de reconnaissance d'écriture industriel. Nous ne cherchons donc pas à obtenir des taux de reconnaissance de 100% mais des taux le plus proche possible de ceux observés chez l'humain. Or, nous n'avons pas trouvé de telles expériences réalisées chez l'humain dans la littérature (reconnaissance de lettres à partir de l'information cinématique). Nous avons donc dû nous contenter de suggérer de nouveaux protocoles expérimentaux.
  De plus, la simulation de la boucle perception action du modèle BAP nous a permis de définir deux questions de lecture. Une question basée uniquement sur la perception de la trace et une basée sur la perception de la trace et sur la simulation du geste à l'origine de cette trace. Pour chacune de ces questions nous avons simulé le modèle et comparé les taux de reconnaissance. Nous avons vu que dans le cas de la reconnaissance de lettres, la simulation motrice n'améliore pas les performances. Cependant lorsque les traces sont incomplètes, la simulation motrice permet de déterminer la trajectoire correspondant à la trace et donc d'améliorer la reconnaissance. Une nouvelle fois nous avons suggéré un protocole expérimental pour valider les résultats du modèle lors de la reconnaissance de lettres tronquées.

- **Reconnaissance du scripteur** : cette tâche a également été simulée dans plusieurs conditions :

  Qui a écrit cette trace sachant que je sais de quelle lettre il s'agit ?

  – sans simulation motrice :

$$P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [L = l]\ [\lambda_V = 1]), \tag{8.7}$$

– avec simulation motrice :

$$P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [L = l]\ [\lambda_V = 1]\ \ [\lambda_P = 1]\ [\lambda_S = 1]), \quad (8.8)$$

Qui a écrit cette trace sachant que j'ignore à quelle lettre elle correspond ?

– sans simulation motrice :

$$P(W \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [\lambda_V = 1]), \quad (8.9)$$

Cette tâche correspond à reconnaître le scripteur à l'origine de la trace. Nous avons simulé cette question et obtenu les matrices de confusion entre scripteurs. Nous avons également déterminé les taux de reconnaissance dans les deux cas suivants : en connaissant ou non la lettre correspondant à la trace. Les taux de reconnaissance varient entre 78% et 81%.
Grâce à la simulation, nous avons effectué une constatation : connaître la lettre correspondant à la trace n'aide pas le modèle à mieux discriminer entre les scripteurs. Cela suggère qu'il n'est pas nécessaire de reconnaître la lettre pour pouvoir déterminer le scripteur.

- **Ecriture** : La tâche d'écriture correspond à la question suivante :
Quels sont les commandes à fournir à l'effecteur pour qu'il trace une lettre donnée ?

Nous avons vu que le modèle BAP est capable de tracer des lettres. Nous avons mis en avant dans les résultats du modèle les variabilités intra-scripteurs et inter-scripteurs obtenues. Pour un scripteur donné et une lettre donnée, nous observons une certaine variabilité entre chaque trace. Ce résultat est cohérent avec l'écriture humaine : lorsque nous écrivons, nous ne réalisons jamais exactement la même trace. Les résultats du modèle sont également cohérents avec l'écriture humaine du point de vue de la variabilité inter-scripteurs. Le modèle est capable de reproduire le style d'écriture de plusieurs scripteurs. Comme pour la perception, le but n'est pas d'avoir un système parfait et donc des traces parfaites. Le but est d'obtenir un système qui reproduit au mieux l'écriture humaine.
Cette tâche a été réalisée avec trois effecteurs : un bras simulé à deux articulations, un bras robotique également à deux articulations et un robot holonome.

– Bras à deux articulations simulé ou robotique :
Quelles sont les vitesses articulaires à appliquer au bras pour qu'il trace une lettre donnée ?

$$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1]). \quad (8.10)$$

Nous avons utilisé un effecteur à deux degrés de libertés : un bras avec l'articulation du coude et de l'épaule. Nous avons tracé des lettres avec cet effecteur à la fois en simulation et avec un robot réel. Nous avons vu que le modèle est capable de reproduire les variabilités intra-scripteur et inter-scripteurs avec cet effecteur.

– Robot holonome :
Quelles sont les vitesses à appliquer à chacune des roues pour que le robot trace une lettre donnée ?

$$P(\omega_1^{0:T}\ \omega_2^{0:T}\ \omega_3^{0:T} \mid [L = l]\ [W = w]\ [\lambda_P = 1]). \quad (8.11)$$

Nous avons également modélisé un robot holonome comme effecteur. Bien qu'il soit très loin d'être comparable à un effecteur humain, il est capable de tracer des lettres. On retrouve les mêmes variabilités intra-scripteur et inter-scripteurs qu'avec l'effecteur bras. Cette modélisation nous a permis d'illustrer l'équivalence motrice. Nous avons pu valider que les mécanismes que nous avons définis (représentation interne indépendante de l'effecteur) nous permettent de réaliser un même mouvement avec différents effecteurs.

- **Copie** : nous avons simulé la tâche de copie dans deux cas : la copie de traces et la copie de lettres.
  - copie de traces : cette tâche correspond à percevoir une trace et à la recopier sans en interpréter le sens. La question correspondant à cette tâche est :

    Quels sont les ordres à fournir à l'effecteur bras pour qu'il recopie la trace ?

    $$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_l = 1]\ [\lambda_p = 1]). \quad (8.12)$$

    Nous avons vu que le modèle est capable de recopier n'importe quelle trace (trace correspondant à une lettre ou non).
  - copie de lettres : cette tâche correspond à percevoir une trace, à reconnaître la lettre correspondante et à la recopier. La question correspondant à cette tâche est :

    Quels sont les ordres à fournir à l'effecteur bras pour qu'il recopie la lettre ?

    $$P(\ddot{\theta}_1^{0:T}\ \ddot{\theta}_2^{0:T} \mid [V_x^{0:M} = v_x^{0:M}]\ [V_y^{0:M} = v_y^{0:M}]\ [\lambda_v = 1]\ [\lambda_p = 1]). \quad (8.13)$$

    A l'inverse de la copie de traces, la copie de lettres ne peut recopier que des traces correspondant à des lettres. Des points de contrôle sont issus de la trace perçue. Ils servent à déterminer la lettre correspondant mais aussi à déterminer de nouveaux points de contrôle. Ce sont ces nouveaux points de contrôle qui servent à générer la nouvelle trace. Cette dernière correspond à la lettre perçue et est écrite dans le style du scripteur qui copie. Elle n'est donc pas identique à l'originale.
    La copie de traces et la copie de lettres sont donc deux tâches cognitives distinctes pour le modèle. Nous avons suggéré un protocole expérimental afin de déterminer dans quel cas la copie de traces ou de lettres est effectuée chez l'humain. En effet, nous n'avons pu effectuer aucune comparaison avec des données expérimentales existantes.

- **Prédiction des lettres** : nous avons simulé, avec le modèle, l'expérience de psychologie cognitive de prédiction de lettres de Kandel *et al.* (2000a). Cette simulation nous a permis d'évaluer les résultats du modèle dans les mêmes conditions que l'expérience. Tout d'abord, le modèle est capable de reproduire les données expérimentales étant donné que les taux de reconnaissance sont similaires aux taux obtenus par les sujets. De plus, nous avons pu comparer les résultats du modèle dans le cas où simulation motrice était activée ou non. Nous avons remarqué que la simulation motrice obtenait des résultats qui correspondaient mieux aux données expérimentales.

Pour résumé, cette thèse a illustré deux points. Tout d'abord, nous avons défini un système permettant d'assurer l'équivalence motrice. Nous l'avons modélisé et testé avec succès.

Deuxièmement, nous avons étudié l'influence du système moteur lors de la perception. Nous avons défini des hypothèses que nous avons également modélisées et testées. La modélisation nous a permis de simuler le modèle dans différentes conditions et de conclure que le système moteur peut jouer un rôle important lors de la perception.

8.4 Discussion et perspectives

Une première piste de recherche concerne les suggestions de protocoles expérimentaux que nous avons réalisés tout au long de ce manuscrit. Réaliser ces expériences et en confronter les résultats à les prédictions issues du modèle permettrait d'étudier son périmètre de validité et, éventuellement, proposer des améliorations.

Dans cette section, nous souhaitons discuter d'autres pistes de recherches futures : nous allons passer en revue les différents choix que nous avons effectué tout au long de la construction du modèle BAP. Nous allons discuter des implications de ces choix et des solutions alternatives envisageables. Afin de structurer cette discussion, nous allons suivre la structure du modèle comme le présente la figure 8.2 : nous commencerons donc par la « trace lue » pour terminer avec la « modélisation des mots ».

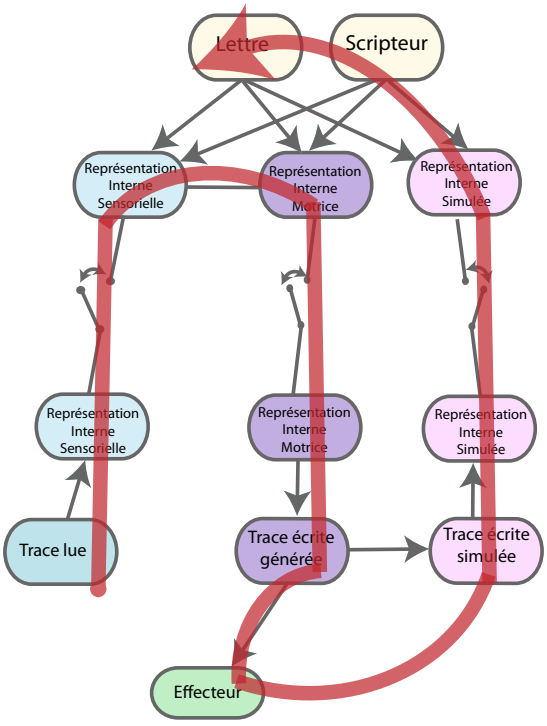
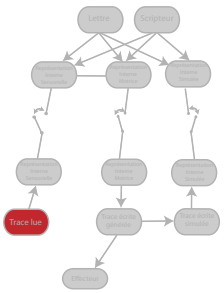


FIG. 8.2 – Ordre dans lequel nous présentons la discussion.

Nous commencerons donc par la « trace lue » pour terminer avec la « modélisation des mots ».

8.4.1 Espace visuel



La figure 8.3 présente le modèle sensoriel : la trace sur la feuille est perçue par l'œil. Le modèle de l'œil génère une trace correspondant à la trace sur la feuille : la « la trace lue ».

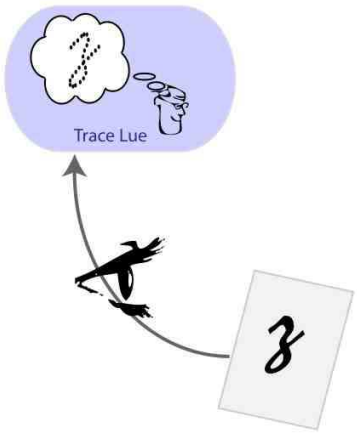


FIG. 8.3 – Modèle sensoriel : la trace sur la feuille est perçue par l'œil pour obtenir la « trace lue ».

L'espace dans lequel est représentée la « trace lue » correspond à l'espace visuel. Dans le modèle BAP, nous avons choisi un espace cartésien pour la rétine afin de simplifier la projection entre l'espace de la feuille et l'espace de la rétine (dans le modèle BAP, il s'agit donc du modèle identité).

Nous pouvons nous demander quelles sont les conséquences d'une telle simplification. En effet, il est connu que les traitements visuels suivent des transformations géométriques complexes : par exemple, certaines structures neurophysiologiques comme le colliculus supérieur présentent des transformations rétinotopiques log-polaires (Girard et Berthoz, 2005) (figure 8.4).

Un ensemble de questions se posent alors :

- **Comment à partir d'une trace représentée dans un espace non cartésien, peut-on déterminer la séquence des points de contrôle ?**
- **Est-ce que la représentation interne des lettres choisie (points où la vitesse s'annule en x ou en y) est toujours aussi pertinente du point de vue de la perception ?**

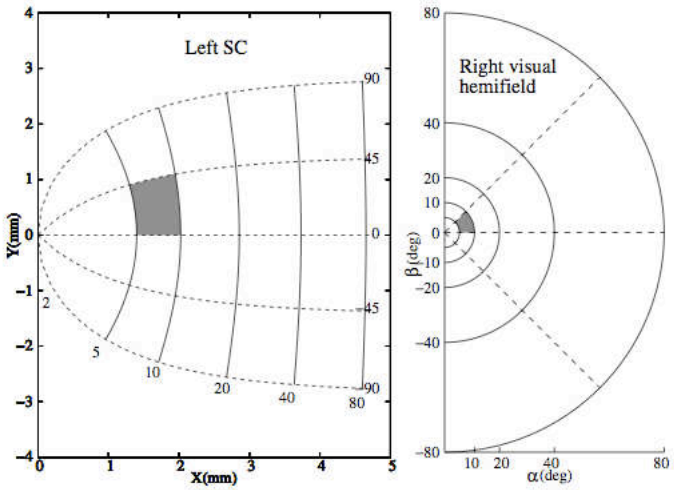
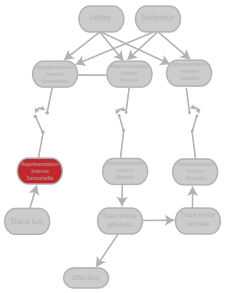


FIG. 8.4 – La position angulaire de la cible visuelle (à droite) est projetée dans l'espace du colliculus supérieur (à gauche), en utilisant une projection logarithmique. La surface grisée représente la même partie du champ visuel dans les deux représentations (figure extraite de Colas *et al.* (2009)).

8.4.2 Représentation interne des lettres



Dans le modèle BAP, la représentation interne des lettres correspond à une séquence de points de contrôle. Les points de contrôle sont définis aux points où au moins une des dérivées en x ou en y s'annule. Nous avons choisi cette représentation car elle pertinente du point de vue de la perception (géométriquement identifiable) et de la production (pertinent du point de vue cinématique). Nous nous sommes interrogés sur la problématique suivante : quelles seraient les conséquences sur le modèle BAP si la représentation interne des lettres était différente ?

Afin de donner des éléments de réponse à cette question, nous allons prendre un exemple où la représentation interne est différente et nous allons essayer de comparer les résultats. Par exemple, nous choisissons de représenter les lettres par une séquence de points où l'accélération est nulle en x ou en y. Nous nous posons donc la question suivante :

Quelles sont les différences pour la perception et l'action entre une représentation interne où les points sont déterminés aux points de vitesse nulle (comme pour le modèle BAP) ou aux points d'accélération nulle ?

La figure 8.5 présente en rouge la position des points de contrôle où l'accélération est nulle et en vert la position des points où la vitesse est nulle sur un exemple de trace extraite de notre base de données.

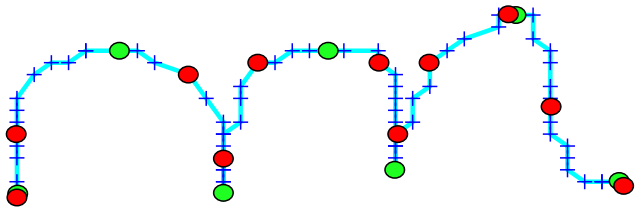


FIG. 8.5 – En vert, les points de contrôle correspondant à une annulation de la vitesse, et en rouge correspondant à une annulation en accélération.

Nous observons que la position des points rouges n'est pas aussi remarquable que dans le modèle BAP. En effet, il serait difficile de les déterminer géométriquement.

Pour la production, nous avons testé l'algorithme de minimum-accélération avec cette représentation des lettres et nous avons comparé les résultats à celui du modèle BAP.

La figure 8.6 compare les trajectoires générées à partir des points de contrôle d'accélération nulle et de vitesse nulle. La figure de gauche présente la trace d'origine. On a extrait des points de contrôle de cette trace : points de contrôle où l'accélération est nulle pour la figure du centre et points de contrôle où la vitesse est nulle pour la figure de droite. A partir de ces points de contrôle, l'algorithme de minimum-accélération a été appliqué dans les deux cas.

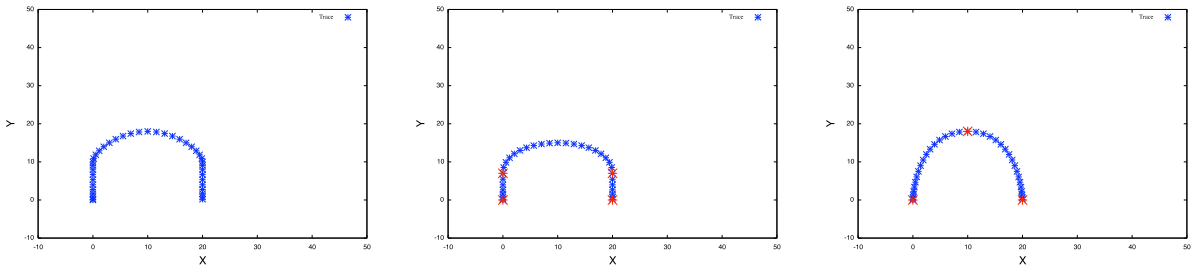
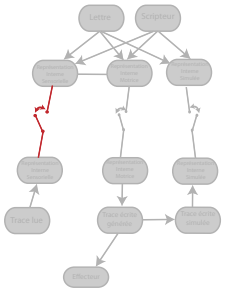


FIG. 8.6 – A gauche, la trace dont on a extrait les points de contrôle. Au centre, la trace générée à partir des points d'accélération nulle. A droite, la trace générée à partir des points de vitesse nulle.

Nous observons une différence entre ces deux trajectoires. La trajectoire du centre est « plate » alors que la trajectoire de droite est « pointue ». Celle du centre est donc plus proche de la trace d'origine que celle de droite. Il semble donc intéressant d'étudier plus en détail les passages à zéro de l'accélération dans la trace. Il faudrait étudier plus en détail les profils de vitesse et d'accélération correspondants et les comparer à ceux générée par l'humain pour pouvoir effectivement conclure.

Quelle est donc la meilleure stratégie pour positionner les points de contrôle ?

8.4.3 Rôle de la variable λ



Dans le modèle BAP, nous avons défini le terme suivant :

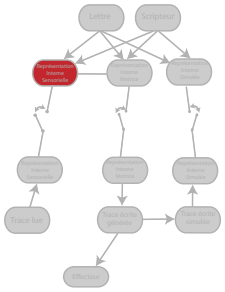
$$P([\lambda_V = 1] \mid C_{LV}\ C_V) \tag{8.14}$$

Etant donné qu'il suit une loi de probabilité de type Dirac et qu'il vaut 1 si C_{LV} est égal à C_V, ce terme permet d'activer ou de désactiver des sous-modèles (en l'occurrence ici, le modèle de perception). La première question que l'on peut se poser est :

Quelle est la signification du terme $P([\lambda_V = 0] \mid C_{LV}\ C_V)$?

De plus, afin de modéliser la notion de distance entre les séquences des points de contrôle C_{LV} et C_V, nous proposons deux solutions : la première consiste à utiliser des distributions de probabilité gaussiennes à la place des distributions de probabilité de type Dirac. La seconde solution consiste à définir la variable λ_V comme étant un réel. On pourrait alors dans $P(\lambda_V \mid C_{LV}\ C_V)$ représenter la notion de distance entre les séquences des points de contrôle C_{LV} et C_V. Le terme $P(C_{LV} \mid C_V\ [\lambda_V = \alpha])$ correspondrait alors à déterminer la séquence des points de contrôle C_{LV} en fonction de C_V et de la quantité de bruit (α) que l'on veut ajouter. Par exemple, poser la question avec $\alpha = 1$ revient à choisir une distance nulle entre les séquence de points (c'est le cas qu'on applique dans le modèle BAP) et $\alpha = 0$ correspondrait à choisir une séquence de points de contrôle (C_{LV}) la plus éloignée possible de celle perçue (C_V). Cette définition du terme $P([\lambda_V = \alpha] \mid C_{LV}\ C_V)$ permettrait alors de bruiter la représentation interne acquise avant de la comparer à la représentation interne mémorisée.

8.4.4 Allographe



Dans le modèle BAP, nous n'avons pas défini de façon explicite les différents allographes possibles d'une lettre (variantes de forme d'une même lettre). Ils sont encodés implicitement à travers les différents modes du terme $P(C_L \mid L\ W)$. Nous aurions pu définir explicitement cette notion. Cependant, plusieurs questions se posent :

Comment encoder explicitement la notion d'allographe dans le modèle ? Quelles sont les conséquences sur la perception de définir explicitement les allographes ? Est-ce-que la reconnaissance des lettres ou des scripteurs serait améliorée ?

Afin de donner des éléments de réponse et des pistes de recherche, nous proposons un exemple de modélisation : le terme $P(C_L \mid L\, W)$ pourrait être remplacé par les termes suivants (choix de décomposition possible mais pas unique) :

$$P(C_L \mid A) \;\; P(A \mid L\, W) \tag{8.15}$$

avec A correspondant aux différents allographes possibles d'une lettre. Le terme $P(C_L \mid A)$ correspondrait à la représentation d'une séquence de points de contrôle pour un allographe donné. Cette relation serait donc indépendante de la lettre et du scripteur. Quant au terme $P(A \mid L\, W)$, il mettrait en relation les allographes avec la notion de lettre et de scripteur. Il permettrait de définir pour chaque scripteur les allographes qu'il a l'habitude de tracer.

Cette notion nous aurait peut être permis de mieux classer les lettres lors de la reconnaissance de caractères et notamment les allographes non connus du modèle. De plus, cela nous aurait permis de rajouter des connaissances supplémentaires dans le modèle.

8.4.5 Influence de la base de données

Afin d'identifier les paramètres des formes paramétriques, nous avons utilisé un base de données de traces. Cette base de données est composée de traces fournies par 4 scripteurs, chaque scripteur ayant réalisé 40 traces pour chaque lettre. Etant donné que notre modèle est très riche en hypothèses, une petite base de données pour l'apprentissage est suffisante. Cependant, il est possible de s'interroger sur la pauvreté de cette base de données (seulement 3520 traces).

- **Quelle est la validité du modèle et surtout des résultats ?**
- **Est-ce-que la reconnaissance des lettres serait plus performante ou moins performante avec d'avantage de scripteurs et de traces apprises ?**
- **Est-ce-que le taux de reconnaissance du scripteur est proportionnel au nombre de scripteurs dans la base de données ?**

Bien qu'il aurait été intéressant de répondre à ces questions, cette thèse n'a pas pour but de développer un système de type industriel. Notre but était de modéliser une boucle sensori-motrice et d'illustrer son utilisation sur des exemples. Et, cette illustration a pu être réalisée malgré le peu de données.

8.4.6 Apprentissage du contrôle moteur

Dans le modèle BAP, nous avons utilisé deux types d'effecteurs : un effecteur simulé, et des effecteurs robotiques. Nous avons vu que les traces produites par les systèmes mécaniques, que ce soit par le bras ou par le robot holonome, diffèrent de la trace simulée. En effet, les traces produites par les effecteurs robotiques sont toujours bruitées par rapport aux traces

produites par le système simulé. Pour rendre le système simulé plus réel, on aurait pu modéliser l'apprentissage du contrôle moteur.

Comment modéliser cet apprentissage ? Quelles seraient les conséquences d'un tel apprentissage sur la réalisation des traces ?

Dans la littérature, l'apprentissage du contrôle moteur a été étudié par Wolpert *et al.* (2005); Ghahramani et Wolpert (1997). Dans BAP, nous avons modélisé un système ayant un contrôle moteur parfait : pour une séquence de points de contrôle donnée la trajectoire réalisée correspond toujours à la trajectoire planifiée. Une piste intéressante de recherche serait de modéliser l'apprentissage du contrôle moteur au sein du modèle BAP. En effet, plusieurs modèles de la littérature proposent des modèles de l'apprentissage du contrôle moteur (Barto *et al.*, 1999; Haruno *et al.*, 2001; Körding et Wolpert, 2004a,b). Or, ces modèles sont toujours proposés dans des tâches de pointage. Il serait donc intéressant de les intégrer dans un modèle complet de la boucle perception-action de la lecture et de l'écriture.

8.4.7 Sources de variabilité

Le modèle BAP permet de simuler la tâche d'écriture. Nous avons analysé les traces produites par le modèle et notamment observé une variabilité intra-scripteur : chaque lettre tracée diffère de la précédente. Nous souhaitons revenir sur ce point car nous avons observé deux sources de variabilité intra-scripteur : une variabilité provenant de la position des points de contrôle et une variabilité provenant du contrôle de l'effecteur.

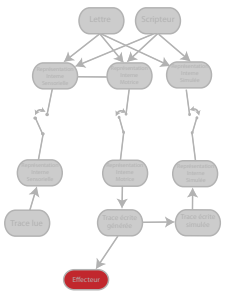
La première source de variabilité vient du fait que l'on effectue des tirages sur la position des points de contrôle selon une distribution de probabilité. A chaque nouvelle trace, des nouveaux points sont déterminés et la trace résultante est donc différente des précédentes.

La seconde source de variabilité vient du contrôle de l'effecteur (dans le cas du bras robotique ou du robot holonome). La variabilité se situe sur toute la trajectoire. Cette dernière source de variabilité a souvent été modélisée (Wolpert et Ghahramani, 2000; Ghahramani *et al.*, 1997; Kawato, 1999; Wolpert, 2007) dans la littérature.

Le modèle BAP propose donc une modélisaton basée sur deux sources de variabilité. Une question se pose alors :

Est ce que la variabilité des traces chez l'humain provient uniquement de la variabilité du contrôle de l'effecteur, de la représentation interne, ou des deux ?

8.4.8 Geste simulé



Nous avons vu que la simulation du geste d'écriture pouvait aider lors de la tâche de lecture. Dans le modèle BAP, nous avons traduit la théorie la simulation par la boucle rouge présentée figure 8.7.

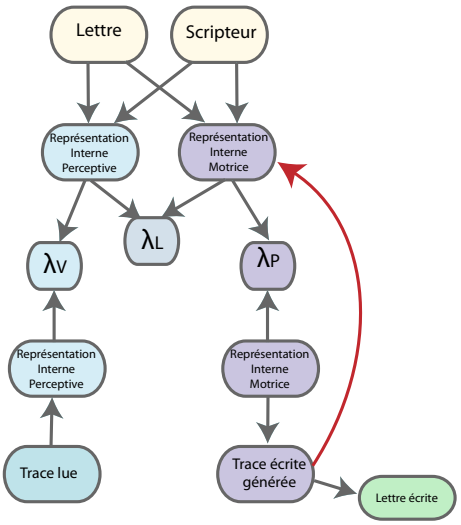


FIG. 8.7 – Boucle de perception simulée.

Or, cette théorie stipule que *le geste* est simulé. Il faudrait donc que la boucle comprenne la traduction de la trajectoire dans l'espace de l'effecteur comme le présente la figure 8.8. En effet, dans le modèle BAP, nous nous sommes contenté de simuler la trajectoire et non le geste. Il faut donc simuler l'effecteur et la trace produite par cet effecteur simulé.

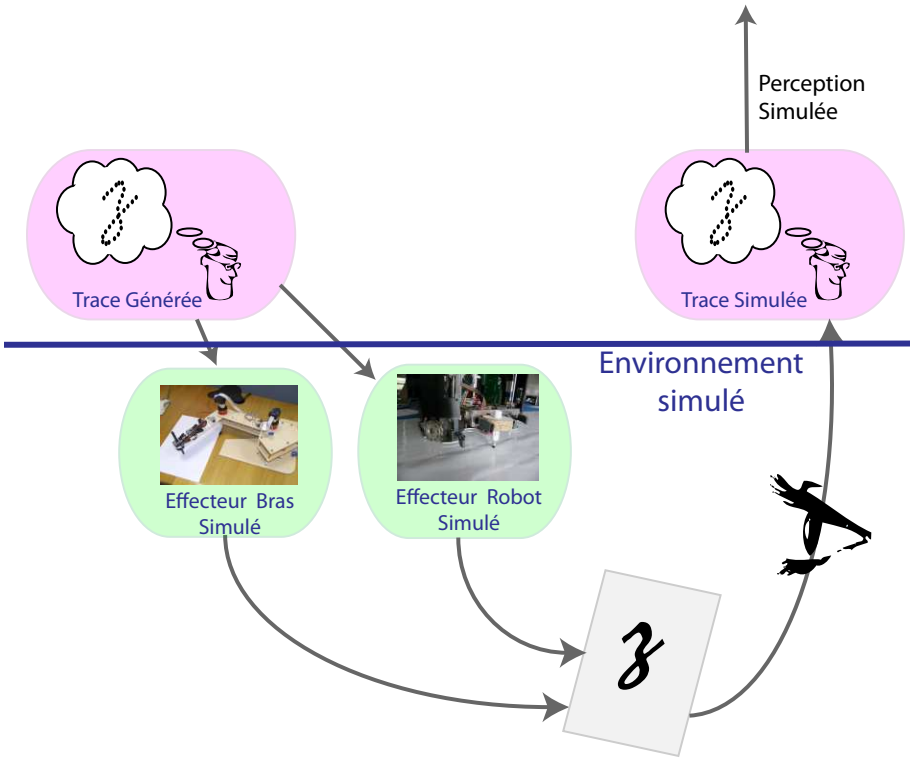


FIG. 8.8 – Boucle de perception simulée incluant la simulation du geste.

Sur la figure 8.8 nous observons la trajectoire générée. Cette trace est utilisée pour déterminer la trajectoire simulée sur la feuille (en simulant l'effecteur). Puis, l'œil est simulé afin de déterminer la trajectoire simulée.

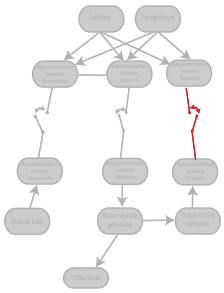
Or, dans le cas où les modèles internes des effecteurs et de la perception simulée sont des modèles identités, il s'avère que la boucle incluant le geste simulé est indistinguable de la boucle du modèle BAP : traduire la trajectoire simulée en commandes motrices puis de nouveaux en trajectoire simulée est identique à appliquer le modèle identité. Cependant, dans le cas où le modèle interne des effecteurs ou dans le cas où le modèle interne sensoriel ne sont pas des modèles identités, les boucles ne sont plus identiques.

La question de lecture avec simulation motrice correspond donc tout d'abord à déterminer la représentation interne à partir de la trace lue. Puis, la représentation interne est utilisée pour déterminer une trajectoire générée. Cette trajectoire est convertie dans l'espace de l'effecteur. Ensuite, l'effecteur simulé détermine une trace simulée qui est perçue par l'œil simulé. Et, de nouveau, une représentation interne est extraite de la trace simulée perçue par l'œil. Cette représentation interne ainsi que celle issue du modèle de perception sont utilisées pour déterminer la lettre correspondant à la trace. Dans le modèle BAP, nous avons modélisé l'œil par le modèle identité (l'espace de la rétine correspond à l'espace cartésien). La trajectoire simulée sur la feuille est donc la même que la trajectoire perçue par l'œil simulée.

Plusieurs questions se posent :

- **Quelles sont les implications de la modélisation du geste dans le perception avec simulation motrice ?**
- **Est-ce que modéliser le geste avec différents effecteurs donne des résultats différents ?**

8.4.9 Lecture avec ou sans simulation motrice



Nous avons identifié deux tâches cognitives relatives à la lecture :

- une tâche qui implique uniquement le système de perception et de représentation des lettres,
- une tâche qui simule le geste d'écriture pendant la lecture et qui fait donc intervenir les systèmes de perception, de perception simulée, de représentation des lettres et d'action.

Mathématiquement, nous avons modélisé ces deux tâches par deux questions distinctes au modèle BAP :

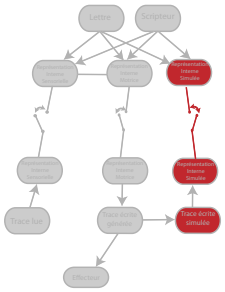
- $P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1])$,
- $P(L \mid [V_X^{0:M} = v_x^{0:M}]\ [V_Y^{0:M} = v_y^{0:M}]\ [W = w]\ [\lambda_V = 1]\ \ [\lambda_P = 1]\ [\lambda_S = 1])$.

Est-ce que l'on effectue toujours une simulation du geste lors de la perception ? Ou, est ce que le geste n'est simulé que dans les cas de reconnaissance difficile ?

Le modèle BAP ne dit rien que l'activation de l'une ou l'autre de ces questions. Nous pouvons nous interroger sur les mécanismes qui permettraient de passer de l'une à l'autre.

S'agirait-il d'un mécanisme explicite ? Ou bien, puisque l'une est une partie de l'autre, est-ce que la simulation est toujours activée, mais que le calcul ne va pas forcément jusqu'au bout à chaque fois (selon la difficulté de la tâche) ?

8.4.10 Dédoublement du modèle de perception

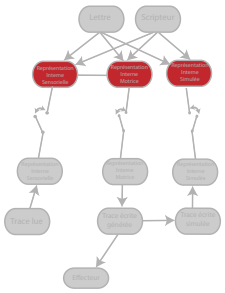


Dans le modèle BAP, nous avons dédoublé le sous-modèle de perception. Nous avons distingué le modèle de perception qui extrait une représentation interne à partir d'une « trace lue », du modèle de perception simulée qui extrait également une représentation interne à partir d'une « trace simulée ». Etant donné que la « trace lue » et la « trace simulée » sont toutes les deux définies dans l'espace cartésien, le mécanisme d'extraction des deux sous-modèles est identique.

Pour des raisons de modélisation, nous avons dû, dans le modèle BAP, définir deux sous-modèles distincts. Or, ce dédoublement n'indique pas que nous faisons l'hypothèse de deux traitements distincts réalisés par deux sous parties distinctes du cerveau. Il nous semblerait plus parcimonieux d'imaginer que la perception et la perception simulée sont réalisées par la même aire du cerveau mais à des moments distincts.

Serait-il possible que des expériences d'imagerie cérébrale étudient la boucle temporelle, en particulier le décours temporel entre la perception, l'action simulée et la perception simulée, prédit par le modèle BAP ?

8.4.11 Apprentissage



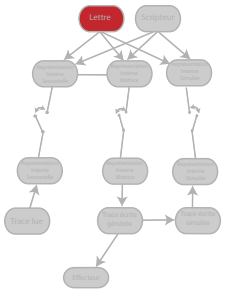
Dans cette thèse, nous avons présenté un modèle de la boucle perception-action dans le cadre de la lecture et de l'écriture. Nulle part, nous n'avons abordé la notion de processus d'apprentissage humain. C'est ce que nous proposons de faire dans cette piste de recherche. Tout d'abord, nous distinguons l'apprentissage du contrôle moteur de l'apprentissage de la représentation interne (Krakauer *et al.*, 1999, 2006).

Une piste de recherche concerne l'apprentissage de la représentation interne des lettres. Des études ont montré qu'un apprentissage haptique améliore la production (Feygen *et al.*, 2002; Henmi et Yoshikawa, 1998; Kahn *et al.*, 2006; Morris *et al.*, 2007; Teo *et al.*, 2002)

mais aussi la perception (Bara *et al.*, 2007). Il serait donc intéressant de modéliser dans BAP l'apprentissage haptique afin de renforcer la connaissance sur la représentation interne des lettres (plus particulièrement sur la position et la vitesse des points de contrôle). Dans le modèle BAP, nous avons utilisé une représentation interne unique pour la perception et la production. Cependant deux objets correspondent à cette représentation interne. Une piste consisterait à avoir une représentation interne des lettres apprise à partir d'un processus de perception visuelle et une autre apprise à partir d'un processus de proprioception. Une question pourrait alors porter sur la fusion de ces deux représentations internes, ou alors sur leurs co-constructions.

Comment co-construire deux représentations des lettres dont une est basée sur la vision et l'autre sur la proprioception ? Est-ce que l'apprentissage sur une modalité renforce les connaissances sur l'autre modalité ? Est-ce qu'il s'agit de deux objets appris indépendamment l'un de l'autre qu'on fusionne pour obtenir un troisième objet co-construit ? Est-ce qu'une défaillance d'une des deux modalités (par exemple, dyslexie ou défaillance du contrôle moteur) détériore la co-construction ? Et si oui, quels sont les impacts, par exemple, d'une défaillance en perception lors de la production ? Ou, inversement, est-ce-qu'une défaillance du contrôle moteur entraîne des moins bonnes performances à la lecture ?

8.4.12 Modélisation des mots



Dans cette thèse, nous nous sommes uniquement intéressés à la reconnaissance et à la production de caractères cursifs isolés. Pouvoir tracer et reconnaître un ensemble de lettres, et plus précisément des mots, nous semble une piste de recherche intéressante.

Comment modéliser les mots dans le modèle BAP ?

Les mots pourrait être modélisés dans BAP par une approche hiérarchique (Diard, 2003). Il serait intéressant d'analyser le mécanisme humain de la lecture pour la perception des mots et la production des ligatures pour leurs productions. En effet, l'identification des lettres et des mots est un processus actif de décodage dans lequel le cerveau ajoute de l'information au signal visuel. Nous ne reconnaissons pas une lettre seule sans bénéficier immédiatement du contexte dans lequel elle est présentée. Ce contexte (mot ou fragment de mot) donne accès à des niveaux supplémentaires de codage (graphème, syllabe, etc.) dont les informations facilitent la perception (Dehaene, 2007). Cette information contextuelle pourrait être modélisée dans le modèle BAP, ce qui permettrait d'étudier le mélange de processus descendants (« top-down ») et montant (« bottom-up ») dans la tâche cognitive de la lecture.

8.4.13 Application à d'autres domaines

Notre principale contribution concerne donc la modélisation de la boucle perception-action de la lecture et de l'écriture. Nous souhaitons, enfin aborder des questions concernant la pertinence de cette boucle vis-à-vis d'autres domaines, afin d'étudier le domaine de validité du modèle BAP.

Par exemple, est-il possible de modéliser la parole avec une représentation unique pivot entre la perception et la production ? Quel serait l'espace de représentation interne des vocables ? Comment seraient-ils codés ? Quel serait le sens des points de contrôle dans le domaine articulatoire ?

Glossaire

- **Allographes :** Variantes de formes d'une même lettre.
- **Cursif :** Type de lettre manuscrite issue d'une écriture liée.
- **Ecriture en ligne :** Ecriture connue par la trajectoire du stylo.
- **Ecriture hors-ligne :** Ecriture issue de la numérisation d'une image.
- **Modèle d'action :** Modèle qui génère une trajectoire à partir de la représentation interne.
- **Modèle de perception :** Modèle qui extrait la représentation interne à partir d'une trajectoire
- **Modèle du geste moteur :** Modèle qui détermine les commandes motrices de l'effecteur.
- **Modèle sensoriel :** Modèle qui détermine une trajectoire dans l'espace du capteur.
- **Mono-scripteur :** Un seul scripteur identifié.
- **Multi-scripteur :** Plusieurs scripteurs identifiés.
- **Omni-scripteur :** Scripteur quelconque.
- **Représentation interne des lettres :** Information mémorisée correspondant au codage des lettres.
- **Trace (en ligne) :** Séquence de points décrivant une trajectoire.
- **Trace générée :** Séquence de points décrivant la trajectoire planifiée (la trajectoire n'est pas réellement exécutée).
- **Trace lue :** Séquence de points décrivant la trajectoire perçue.

Bibliographie

W. Abend, E. Bizzi et P. Morasso : Human arm trajectory formation. *Brain*, 105(2):331–348, 1982.

S. V. Adamovich, P. S. Archambault, M. Ghafouri, M. F. Levin, H. Poizner et A. Feldman : Hand trajectory invariance in reaching movements involving the trunk. *Experimental Brain Research*, 138:288–303, 2001.

E. Anquetil et G. Lorette : Reconnaissance en ligne de lettres manuscrites cursives reconnaissance en ligne de lettres manuscrites cursives par chaînes de markov cachées. *Traitement du signal*, 12(6):561–566, 1995.

C. G. A. Atkeson et J. M. Hollerbach : Kinematic features of unrestrained vertical arm movements. *Journal of Neuroscience*, 5:2318–2330, 1985.

F. Bara, E. Gentaz et P. Colé : The visuo-haptic and haptic exploration increases the decoding level of children coming from low socio-ecnonomic status families. *British Journal of Developmental Psychology*, 2007.

A. Barto, A. Sitkoff et J. Houk : A cerebellar model of timing and prediction in the control of reaching. *Neural Computation*, 11:565–594, 1999.

P. M. Bays et D. Wolpert : Computational principles of sensorimotor control that minimize uncertainty and variability. *The Journal of Physiology*, 578(2):387–396, 2007.

S. Ben-Itzhak et A. Karniel : Minimum acceleration criterion with constraints implies bang-bang control as an underlying principle for optimal trajectories of arm reaching movements. *Neural Computation*, 20:779–812, 2008.

M. Berkenblit et A. Feldman : Some problems of motor control. *Journal of Motor Behavior*, 20:369 –373, 1988.

N. Bernstein : *The Co-ordination and Regulation of Movements*. Oxford, 1967.

A. Berthoz : *Le sens du mouvement*. Odile Jacob, 1998.

P. Bessière : Probabilistic methodology and techniques for artefact conception and development. Rap. tech., Institut national de recherche en informatique et en automatique, 2003.

P. Bessière, E. Dedieu, O. Lebeltel, E. Mazer et K. Meknacha : Interprétation vs description i : Proposition pour une théorie probabiliste des systèmes cognitifs sensorimoteurs. *Intellectica*, 26-27:257–311, 1998a.

P. Bessière, E. Dedieu, O. Lebeltel, E. Mazer et K. Meknacha : Interprétation vs description ii : Fondements mathématique. *Intellectica*, 26-27:313–336, 1998b.

S. B. Bortolami, P. Pigeon, P. Dizio et J. R. Lackner : Kinetic analysis of arm reaching movements during voluntary and passive rotation of the torso. *Experimental Brain Research*, 187:509–523, 2008.

A. E. Bryson et Y.-C. Ho : *Applied optimal control : optimization, estimation, and control*. Washington : Hemisphere Pub. Corp. New York : distributed by Halsted Press, 1975.

D. Bullock et S. Grossberg : Neural dynamics of planned arm movements : Emergent invariants and speed-accuracy properties during trajectory formation. *Psychological Review*, 95(1):49–90, 1988.

D. Bullock, S. Grossberg et C. Mannes : A neural network model for cursive script production. *Biological Cybernetics*, 70:15–28, 1993.

B. Calvo-Merino, D. Glaser, J. Grèzes, R. Passingham et P. Haggard : Action observation and acquired motor skills : An fmri study with expert dancers. *Cerebral Cortex*, 2004.

U. Castiello et G. E. Stelmach : Generalized representation of handwriting : evidence of effector independence. *Acta psychologica*, 82:53–68, 1993.

F. Colas : *Perception des objets en mouvement. Composition bayésienne du flux optique et du mouvement de l'observateur*. Thèse de doctorat, Institut National Polytechnique de Grenoble, 2006.

F. Colas, F. Flacher, T. Tanner, P. Bessière et B. Girard : Bayesian models of eye movement selection with retinotopic maps. *Biological Cybernetics*, 100(3):203–214, 2009.

J.-P. Crettez et G. Lorette : Reconnaissance de l'écriture manuscrite. *Techniques de l'Ingénieur, traité Informatique*, 1998.

S. Dehaene : *Les neurones de la lecture*. Odile Jacob, 2007.

M. Delalandre : *Analyse de documets graphiques : une approche par reconstruction d'objets*. Thèse de doctorat, Université de Rouen, 2005.

M. Desmurget : Constrained and unconstrained movements involve different control strategies. *Journal of Neurophysiology*, 77:1644–1650, 1997.

M. Desmurget, D. Pélisson, Y. Rossetti et C. Prablanc : From eye to hand : Planning goal-directed movements. *Neuroscience and Biobehavioral Reviews*, 22(6):761 – 788, 1998.

J. Diard : *La carte bayésienne un modèle probabiliste hiérarchique pour la navigation en robotique mobile*. Thèse de doctorat, Institut National Polytechnique de Grenoble, 2003.

J. Diard : Bayesian model comparison and distinguishability. *Proceedings of the International on Cognitive Modeling (ICCM 09)*, 2009.

M. Djioua et R. Plamondon : Studying the variability of handwriting patterns using the kinematic theory. *Human Movement Science*, 2009.

S. Edelman et T. Flash : A model of handwriting. *Biological Cybernetics*, 57:25–36, 1987.

M. Eden : Handwriting and pattern recognition. *IRE Trans. Information Theory*, 8:160–166, 1962.

A. Feldman : The origin and use of positional frames of reference in motor control. *Biophysics*, 11:565–578, 1966.

D. Feygen, M. Keehner et F. Tendick : Haptic guidance : Experimental evaluation of a haptic training method for a perceptual motor skill. *10th IEEE. Haptics Symposium*, p. 40–47, 2002.

J. R. Flanagan, A. G. Feldman et D. J. Ostry : Control of trajectory modifications in target-directed reaching. *Journal of Motor Behavior*, 25(3):140–152, 1993.

T. Flash et N. Hogan : The coordination of arm movements : An experimentally confirmed mathematical model. *Journal of Neuroscience*, 5(7):1688–1703, 1985.

J. A. Fodor : *The modularity of mind.* Cambridge, 1983.

B. Galantucci, C. A. Fowler et M. T. Turvey : The motor theory of speech perception reviewed. *Psychonomic Bulletin and Review*, 13(3):361–377, 2006.

Z. Ghahramani et D. Wolpert : Modular decomposition in visuomotor learning. *Nature*, 386:392–395, 1997.

Z. Ghahramani, D. Wolpert et M. Jordan : Computational models of sensorimotor integration. *Self-organization, computational maps and motor control*, p. 117–147, 1997.

B. Girard et A. Berthoz : From brainstem to cortex : computational models of the saccade generation circuitry. *Progress in Neurobiology*, 77(4):215–251, 2005.

H. Gréa, M. Desmurget et C. Prablanc : Postural invariance in three-dimensional reaching and grasping movements. *Experimental Brain Research*, 134:155–162, 2000.

C. Harris : Exploring smoothness and discontinuities in human motor behavior with fourier analysis. *Mathematical Biosciences*, 188:99–116, 2004.

C. Harris et M. R. Harwood : Boundary conditions in human movement i : Constraining constraints. *Proceedings : Biomedical Engineering*, 2005.

C. Harris et D. Wolpert : Signal-dependent noise determines motor planning. *Nature*, 394:780–784, 1998.

M. Haruno, D. Wolpert et M. Kawato : Mosaic model for sensorimotor learning and control. *Neural Computation*, 13:2201–2220, 2001.

K. Henmi et T. Yoshikawa : Virtual lesson and its application to virtual calligraphy system. *Proceedings of the 1998 IEEE International Conference on Robotics and Automation*, 1998.

N. Hogan : Adaptive control of mechanical impedance by coactivation of antagonist muscles. *IEEE Transactions on Automatic Control*, 29(8):681–690, 1984.

J. M. Hollerbach : The minimum energy movement for a sptring muscle model. *A.I. Memo*, 424, 1977.

J. M. Hollerbach : An oscillation theory of handriting. *Biological Cybernetics*, 39:139–240, 1981.

J. M. Hollerbach et C. G. Atkeson : Deducing planning variables from experimental arm trajectories : Pitfalls and possibilities. *Biological Cybernetics*, 56:279–292, 1987.

J. M. Hollerbach et T. Flash : Dynamic interactions between limb segments during planar arm movement. *Biological Cybernetics*, 44:67 – 77, 1982.

M. A. Huckvale : Illustrating speech : Analogies between speaking and writing, in speech, hearing and language. *Work in Progress, Phonetics and Linguistics*, 6, 1992.

C. Ildei-Bidet : *Motricité et perception visuelle des mouvements humains :expériences chez l'adulte, l'enfant et le patient parkinsonien.* Thèse de doctorat, Université Pierre Mendès France, 2006.

S. Impedovo, L. Ottaviano et S. Occhinegro : Optical character recognition - a survey. *International Journal on Pattern Recognition and Artificial Intelligence*, 5(1-2):1–24, 1991.

M. Jeannerod : Neural simulation of action : A unifying mechanism for motor cognition. *NeuroImage*, 14:103–109, 2001.

L. E. Kahn, P. S. Lum, W. Z. Rymer et D. J. Reinkensmeyer : Robot-assisted movement training for the sroke-impaired arm :does it matter what the robot does ? *Journal of Rehabilitation Research and Development*, 43(5):619–630, 2006.

S. Kandel, J.-P. Orliaguet et L.-J. Boë : Detecting anticipatory events in handwriting movements. *Perception*, 29:953–964, 2000a.

S. Kandel, J.-P. Orliaguet et P. Viviani : Perceptual anticipation in handwriting : The role of implicit motor competence. *Perception and Psychophysics*, 62:706–716, 2000b.

M. Kawato : Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9:718–727, 1999.

S. Keele, P. Jennings, S. Jones, D. Caulton et A. Cohen : On the modularity of sequence representation. *Journal of Motor Behavior*, 27:17–30, 1995.

D. Kerzel et H. Bekkering : Motor activation from visible speech : Evidence from stimulus response compatibility. *Journal of Experimental Psychology : Human Perception and Performance*, 26:634–647, 2000.

W. Khalil : *Commande des robots manipulateurs.* Lavoisier, 2002.

G. Knoblich et R. Flach : Action identity : Evidence from self-recognition, prediction, and coordination. *Consciousness and Cognition*, 12:620–632, 2006.

G. Knoblich et W. Prinz : Recognition of self-generated actions from kinematic displays of drawing. *Journal of Experimental Psychology : Human Perception and Performance*, 27:456–465, 2001.

G. Knoblich, E. Seigerschmidt, R. Flach et W. Prinz : Autorship effects in the prediction of handwriting strokes : Evidence for action simulation during action perception. *The quarterly journal of experimental psychology*, 55A(3):1027–1046, 2002.

C. M. C. Koike : *Bayesian Approach to Action Selection and Attention Focusing.* Thèse de doctorat, Institut National Polytechnique de Grenoble, 2005.

K. Körding et D. Wolpert : Bayesian integration in sensorimotor learning. *Nature*, 427:244–247, 2004a.

K. Körding et D. Wolpert : The loss function of sensorimotor learning. *PNAS*, 2004b.

J. W. Krakauer, M.-F. Ghilardi et C. Ghez : Independent learning of internal models for kinematic and dynamic control of reaching. *Nature Neuroscience*, 2(11), 1999.

J. W. Krakauer, P. Mazzoni, A. Ghazizadeh, R. Ravindran et R. Shadmehr : Generalization of motor learning depends on the history of prior action. *PLoS Biology*, 4 (10):1798–1808, 2006.

R. Le Hy : *Programmation et apprentissage bayésien de comportements pour des personnages synthétiques. Application aux personnages de jeux vidéos.* Thèse de doctorat, Institut National Polytechnique de Grenoble, 2007.

O. Lebeltel : *Programmation Bayésienne des Robots.* Thèse de doctorat, Institut National Polytechnique de Grenoble, 1999.

W. Li, E. Todorov et X. Pan : Hierarchical feedback and learning for multi-joint arm movement control. *Engineering in Medecine and Bilology*, 2005.

A. M. Liberman : Some results of research on speech perception. *Journal of the Acoustical Society of America*, 29:117–123, 1957.

A. M. Liberman et I. G. Mattingly : The motor theory of speech perception revised. *Cognition*, 21:1–36, 1985.

B. Lindblom : On the notion of possible speech sound. *Journal of Phonetics*, 18:135–152, 1990.

M. Longcamp : *Etude comportementale et neurofonctionnelle des interactions perceptivo-motrices dans la perception visuelle de lettres. Notre maniere d'écrire influence-t-elle notre manière de lire ?* Thèse de doctorat, Université de la Mediterranée -Aix Marseille II, 2003.

G. Lorette : Handwriting recognition or reading ? what is the situation at the drawn of the 3rd millenium ? *International Journal on Document Analysis and Recognition*, 2:2–12, 1999.

M. Mahjoub et N. Ellouze : Non stationary hidden markov model- application to on-line arabic character recognition. *Computational engineering in systems applications*, p. 111–116, 1998.

A. Matsui et Y. Wada : Examination of human trajectory planning criterion on signal dependent noise. *SICE Annual Conference in Fukui*, 2003.

D. Méary, E. Kitromilides, K. Mazens, C. Graff et E. Gentaz : Four-day-old human neonates look longer at non-biological motions of a single. *PlosOne*, 2007.

P. Mermelstein et M. Eden : Experiments on computer recognition of connected hanwritten words. *Information Control*, 7:255–270, 1964.

R. G. J. Meulenbroek, D. A. Rosenbaum, A. J. W. Thomassen, L. D. Loukopoulos et J. Vaughan : Adaptation of a reaching model to handwriting : How different effectors can produce the same written output, and other results. *Psychological Research*, 59:64–74, 1996.

H. Miyamoto, E. Nakano, D. Wolpert et M. Kawato : Tops (task optimization in the presence of signal-dependent noise) model. *Systems and Computers in Japan*, 35(11), 2004.

H. Miyamoto, D. M. Wolpert et M. Kawato : *Computing the optimal trajectory of arm movement : the TOPS (Task Optimization in the Presence of Signal-Dependent Noise) model*, p. 395–415. Physica-Verlag GmbH, 2003.

P. Morasso : Spatial control of arm movements. *Experimental Brain Research*, 42:223–227, 1981.

P. Morasso et I. Mussa : A trajectory formation and handwriting : a computational model. *Biological Cybernetics*, 45:131–142, 1982.

D. Morris, H. Z. Tan, F. Barbagli, T. Chang et K. Salisbury : Haptic feedback enhances force skill learning. *roceedings of the 2007 World Haptics Conference (WHC07) : The Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, p. 21–26, 2007.

E. Nakano, H. Imamizu, R. Osu, Y. Uno, H. Gomi, T. Yoshioka et M. Kawato : Quantitative examinations of internal representations for arm trajectory planning : minimum commanded torque change model. *Journal of Neurophysiology*, 1999.

W. Nelson : Physical principles for economies of skilled movements. *Biological Cybernetics*, 46:135 – 147, 1983.

A. Nosary, L. Heutte, T. Paquet et Y. Lecourtier : Defining writer's invariants to adapt the recognition task. *5th International Conference on Document Analysis and Recognition*, p. 765–768, 1999.

R. Osu, Y. Uno, Y. Koike et M. Kawato : Possible explanations for trajectory curvature in multijoint arm movements. *Journal of Experimental Psychology : Human Perception and Performance*, 23:890–913, 1997.

R. Plamondon : What does differential geometry tell us about handwriting generation. *Proceedings of the 3rd international Symposium on Handwriting and Computer Applications*, p. 11–13, 1987.

R. Plamondon et F. J. Maarse : An evaluation of motor models of handwriting. *IEEE transactions on systems, man, and cybernetics*, 19(5):1060–1072, 1989.

R. Plamondon, C. Suen et M. Simmer : Computer recognition and human production of handwriting, 1989.

C. Pradalier : *Navigation intentionnelle d'un robot mobile.* Thèse de doctorat, Institut National Polytechnique de Grenoble, 2004.

C. Pradalier, F. Colas et P. Bessière : Expressing bayesian fusion as a product of distributions : Applications in robotics. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2003.

J. Proust : Pour une théorie "motrice" de la simulation. *Psychologie Française*, 45(4):295–306, 2000.

M. H. Raibert : Motor control and learning by the state-space model. Rap. tech. A1-TR-439, Cambridge, MA : Artificial Intelligence Laboratory, 1977.

V. Raos, M. N. Evangeliou et H. E. Savaki : Mental simulation of action in the service of action perception. *The Journal of Neuroscience*, 27(46):12675–12683, November 2007.

V. Reid, J. Belsky et M. Johnson : Infant perception of human action : Toward a developmental cognitive neuroscience of individual differences. *Cognition Brain Behavior*, 2005.

M. J. Richardson et T. Flash : Comparing smooth arm movements with the two-thirds power law and the related segmented-control hypothesis. *The Journal of Neuroscience*, 22 (18):8201–8211, 2002.

G. Rizzolatti et M. Arbib : Language within our grasp. *Trends in Neurosciences*, 21:188–194, 1998.

I. Sallagoïty : *Dynamique de coordination spontanée de l'écriture.* Thèse de doctorat, Universite Toulouse III - Paul Sabatier, 2004.

C. Savariaux, P. Perrier et J.-P. Orliaguet : Compensation strategies for the perturbation of the rounded vowel [u] using a lip tube : A study of the control space in speech production. *Journal of the Acoustical Society of America*, 98:2428 – 2442, 1995.

L. Schomaker, H.-L. Teulings, E. Helsper et G. Abbink : Adaptive recognition of online, cursive handwriting. *Proceedings of the Sixth International Conference on Handwriting and Drawing*, 4-7:19–21, 1993.

J.-L. Schwartz : *Percevoir : monde et langage*, chap. Une théorie de la perception pour le contrôle de l'action, p. 261–270. Mardaga, 2001.

J.-L. Schwartz, C. Abry, L. Boë et M. Cathiard : *Phonetics, phonology, and cognition*, chap. Phonology in a theory of perception-for-action-control. Phonetics, Phonology and Cognition, 2002.

T. Sejnowski : Making smooth moves. *Nature*, 34, 1998.

J. Serkhane : *Un bébé androïde vocalisant :Etude et modélisation des mécanismes d'exploration vocale et d'imitation orofaciale dans le développement de la parole.* Thèse de doctorat, Institut National Polytechnique de Grenoble, 2005.

G. Serratrice et M. Habib : *L'écriture et le cerveau.* Masson, 1993.

R. Shadmehr et J. W. Krakauer : A computational neuroanatomy for motor control. *Experimental Brain Research*, 2008.

R. Shadmehr et F. A. Mussa-Ivaldi : Adaptive representation of dynamics during learning of a motor task. *The Journal of Neuroscience*, 14(5):3208–3224, 1994.

M. Shiffrar et J. J. Freyd : Apparent motion of the human body. *Psychological science*, 1(4):257–264, 1990.

G. Simmons et Y. Demiris : Optimal robot arm control using the minimum variance model. *Journal of Robotic Systems*, 2005.

C. Sirat, J. Irigoin et E. Poulle : *L'écriture : le cerveau, l'oeil et la main.* Brepols - Turnhout, actes du colloque international du cnrs. paris, collège de france édn, 1988.

J. Soechting et F. Lacquaniti : Invariant characteristics of a pointing movement in man. *The Journal of Neuroscience*, 1(7):710–720, 1981.

R. Stein : What muscle variable(s) does the nervous system control in linm movements. *The behavioral and brain sciences*, 5:535–577, 1982.

C. Y. Suen, C. Nadal, R. Legault, T. A. Mai et L. Lam : Computer recognition of unconstrained handwritten numerals. *Proceedings of the IEEE*, 80(7):1162–1180, 1992.

H. Tanaka, M. Tai et N. Qian : Different predictions by the minimum variance and minimum torque-change models on the skewness of movement velocity profiles. *Neural Computation*, 16:2021–2040, 2004.

C. C. Tappert, C. Y. Suen et T. Wakahara : The state of the art in on-line handwriting recognition. *IEEE Transactions on pattern analysis and machine intelligence*, 12(8):787–808, 1990.

C. Teo, E. Burdet et H. Lim : A robotic teacher of chinese handwriting. *Proceedings of the 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 2002.

H.-L. Teulings : Invariant handwriting features useful in cursive-script recognition. *In in Fundamentals in Handwriting Recognition, S. Impedovo ed*, p. 179–198. Springer-Verlag, 1993.

H.-L. Teulings et L. Schomaker : Invariant properties between stroke features in handwriting. *Acta Psychologica*, 82, 1993.

E. Todorov et M. Jordan : Smoothness maximization along a predefined path accurately predicts the speed profiles of complex arm movements. *Journal of Neurophysiology*, 80:696–714, 1998.

K. Tokunaga et Y. Wada : A model for movement pattern acquisition process. *SICE Annual Conference in Sapporo*, 2004.

E. Trélat : *Contrôle optimal : théorie & applications.* Paris : Vuibert, imprimerie, 2005.

Y. Uno, M. Kawato et R. Suzuki : Formation and control of optimal trajectory in human multijoint arm movement - minimum torque-change model. *Biological Cybernetics*, 61:89–101, 1989.

G. Vallortigara, L. Regolin et F. Marconato : Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biology*, 2005.

T. Veeraklaew, P. Piromsopa, K. Chirungsarpsook et C. Pattaravarangkur : A study on the comparison between minimum jerk and minimum energy of dynamic systems. *IEEE Computer Society*, 2005.

P. Vindras et P. Viviani : Frames of reference and control parameters in visuomanual pointing. *Journal of Experimental Psychology : Human Perception and Performance*, 24 (2):569–591, 1998.

P. Viviani et N. Stucchi : Motor-perceptual interactions. *Tutorials in Motor Behavior II*, p. 229–248, 1992.

V. Vuori : *Adaptive Mathods for On-Line Recognition of Isolated Handwritten Characters.* Thèse de doctorat, HelsinkiUniversity of Technology, 2002.

Y. Wada et M. Kawato : A via-point time optimization algorithm for complex sequential trajectory formation. *Neural Networks*, 17:353–364, 2004.

J. Wann : *The control of fine motor trajectories.* Thèse de doctorat, Cambridge, 1988.

A. M. Wing : Motor control : Mechanisms of motor equivalence in handwriting. *Current Biology*, 10(6):245–248, 2000.

D. Wolpert et Z. Ghahramani : Computational principles of movement neuroscience. *Nature*, 3:1212–1217, 2000.

D. Wolpert, Z. Ghahramani et J. Flanagan : Perspectives and problems in motor learning. *Trends in Cognitive Sciences*, 5(11):487–494, 2005.

D. Wolpert, Z. Ghahramani et M. Jordan : An internal model for sensorimotor integration. *Science*, 269:1880–1882, 1995.

D. M. Wolpert : Probabilistic models in human sensorimotor control. *Human Movement Science*, 26:511–524, 2007.

C. E. Wright : *Attention and Performance XIII : Motor Representation and Control*, chap. Generalized motor programs : Reexamining claims of effector independence in writing. Jeannerod, 1990.

A. E. Yacoubi : *Modélisation markovienne de l'écriture manuscrite, application à la reconnaissance des adresses portales.* Thèse de doctorat, Université de Rennes 1, 1996.

H. Zelaznik : *Advances in Motor Learning and Control.* Human Kinetics, 1996.

Abstract.

How can the motor system influence perception ? How can the same movements be performed with different effectors ?

In order to examine these questions, which are relevant in cognitive sciences, we studied the sensorimotor system involved in tasks of reading and writing isolated handwritten cursive letters. Our contribution is a formal model of this sensorimotor loop, that takes into account experimental findings as well as state-of-the-art theories.

Our main hypothesis is the existence of an internal encoding of letters that is common to both perception and action. This encoding is assumed to be independent of the space of the effectors' joints. It acts as a pivot between perception and action models, and is made of sequences of geometrically salient control point. These are extracted from trajectories of handwritten letters and allow to recover part of the kinematic information.

The formalism we used in this work is called Bayesian Programming. It allows to represent and manipulate incomplete and uncertain information using probabilities ; models are defined by specifying joint probability distributions.

Given the joint probability distribution, the model can solve several cognitive tasks. Each of these corresponds to a probabilistic question that is automatically solved by Bayesian inference : reading, reading with an internal simulation of the writing gesture, trace copy and letter copy, writer recognition, writing using different effectors.

We have implemented and simulated each of these probabilistic inferences. Our results reproduce observations from psychophysics experiments. Our model also suggests new experimental protocols, which we defined.

Keywords : bayesian modelling, perception-action loop, motor simulation, handwriting, motor control.

Résumé.

Comment notre système moteur peut-il influencer notre perception ? Quels sont les mécanismes qui permettent de réaliser un même mouvement avec différents effecteurs ?

Afin d'apporter des éléments de réponse à ces deux questions issues des sciences cognitives, nous avons étudié un système sensori-moteur : celui mis en oeuvre pour les tâches de lecture et écriture de lettres cursives isolées. Notre contribution est un modèle formel de cette boucle sensori-motrice, qui tient compte des observations et théories issues de la littérature.

L'hypothèse centrale est qu'il existe un codage interne des lettres commun à la perception et à l'action, et, de plus, indépendant de l'effecteur. Ce codage, pivot entre la perception et l'action, consiste en des séquences de points de contrôle, géométriquement remarquables, identifiables à partir de la trace des lettres, et permettant de récupérer une partie de l'information cinématique.

Le formalisme utilisé est celui de la *programmation bayésienne*. La définition d'un programme bayésien se fait par l'écriture d'une distribution de probabilité conjointe.

Connaissant cette probabilité conjointe, le modèle est capable de résoudre plusieurs tâches cognitives. Chacune de ces tâches correspond à une *question probabiliste* qui est résolue automatiquement par inférence bayésienne : la lecture, la lecture en simulant le geste d'écriture, la copie de traces, la reconnaissance du scripteur, l'écriture avec différents effecteurs.

Nous avons simulé informatiquement chacune de ces inférences et comparé leurs résultats aux observations issues des expériences de psychophysique. Le modèle permet d'effectuer des prédictions et ainsi de suggérer de nouveaux protocoles expérimentaux.

Mots clés : modélisation bayésienne, boucle perception-action, simulation motrice, écriture, contrôle moteur.