



HAL
open science

Détection d'hypovigilance chez le conducteur par fusion d'informations physiologiques et vidéo

Antoine Picot

► **To cite this version:**

Antoine Picot. Détection d'hypovigilance chez le conducteur par fusion d'informations physiologiques et vidéo. Traitement du signal et de l'image [eess.SP]. Institut National Polytechnique de Grenoble - INPG, 2009. Français. NNT : . tel-00440959

HAL Id: tel-00440959

<https://theses.hal.science/tel-00440959>

Submitted on 14 Dec 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.

Résumé

Titre : Détection de l'hypovigilance chez le conducteur par fusion d'informations physiologiques et vidéo.

Résumé : L'hypovigilance correspond à la transition entre la veille et le sommeil durant laquelle l'organisme voit ses facultés d'observation et d'analyse fortement réduites. Elle est responsable de nombreux accidents sur la route. Le but de cette thèse est de proposer un système de détection de cette phase accidentogène à partir de l'analyse conjointe de l'activité cérébrale (électroencéphalogramme ou EEG) et d'une vidéo du conducteur. Dans un premier temps, une méthode non-supervisée de détection d'hypovigilance fonctionnant à l'aide d'un unique canal EEG a été proposée. Cette méthode, qui met en œuvre différentes techniques de traitement du signal et de diagnostic, obtient de bonnes performances sur un ensemble de conducteurs, sans qu'il soit nécessaire de régler de paramètres. Dans un deuxième temps, nous nous sommes intéressés à la caractérisation des signes visuels de l'hypovigilance par une analyse vidéo des clignements. Une comparaison entre l'approche vidéo développée et l'approche traditionnelle par électro-oculogramme (EOG) a permis d'étudier dans quelle mesure la vidéo peut remplacer l'EOG pour la caractérisation des clignements. Elle a également permis de souligner la nécessité d'utiliser d'une caméra rapide (pouvant aller jusqu'à 200fps) pour caractériser les clignements. Un algorithme de détection d'hypovigilance à partir de la caractérisation vidéo des clignements a ainsi été développé. Pour finir, un algorithme de détection d'hypovigilance fusionnant, à l'aide de logique floue, les informations obtenues par les approches physiologique et vidéo est présenté. Toutes ces méthodes ont été testées et validées sur une base de données conséquente de conduite en état d'hypovigilance, la base de données ayant été expertisée par un spécialiste.

Mots clés : Traitement de signaux biomédicaux (EEG, EOG) - Diagnostic par des méthodes sans modèle - Traitement d'images (Analyse vidéo, algorithmes de détection de contour, détection de caractéristiques, Caméra rapide) - Fusion de données - Logique floue

Abstract

Title : Driver drowsiness detection using both physiological and video information.

Abstract : Drowsiness is the transition between the awake state and sleep where one's abilities to observe and analyse are strongly reduced. So, drowsiness is responsible for a huge number of road accidents. A drowsiness detection system is presented in this PhD. This system uses both driver's brain activity (through electroencephalogram or EEG) and driver's video analysis to detect drowsiness. A non-supervised method using a single EEG channel is first proposed. Several techniques of diagnostic and signal processing are used in this method. The method obtains good results on a large number of different drivers without tuning any parameters. Drowsiness visual indicators have also been studied thanks to a video analysis. The proposed video analysis has been compared to the traditional electro-oculogram (EOG) approach to study the relevance of the video analysis. This comparison highlights the need of a high frame rate camera (which frame rate can reach 200fps) for blinks characterisation. This study leads to drowsiness detection system based on blinks video characterisation. Both EEG and video methods are at last been merged using fuzzy logic to obtain a 2-levels drowsiness detection system. All these methods have been tested and validated on a consistent database of twenty different drivers which was evaluated by a specialist doctor.

Mots clés : Biomedical signal processing (EEG, EOG) - Diagnostic without model - Image processing (video analysis, edge detection, face and eye detection, high frame rate camera) - data fusion - fuzzy logic

Remerciements

Avant d'entrer dans le vif du sujet, je souhaite remercier toutes les personnes qui ont rendu ce travail possible.

Je tiens tout d'abord à remercier chaleureusement mes directrices de thèse, Mmes. Alice Caplier et Sylvie Charbonnier, pour avoir accepté d'encadrer ce travail. Leur confiance, leur aide ainsi que les libertés scientifiques qu'elles m'ont offertes tout au long de cette thèse m'ont permis d'effectuer mes recherches dans des conditions idéales.

Je tiens également à exprimer un grand merci à Mr. Thierry Pébayle et au laboratoire CEPA à Strasbourg qui m'ont fourni la base de données de conduite en état d'hypovigilance sur laquelle j'ai pu tester l'ensemble de mes travaux. Sans eux, ce travail n'aurait jamais été possible. Leur accueil et leur aide ont été très précieux pour moi.

Un grand merci à Mme. Sylviane Gentil pour avoir accepté de présider ce jury et pour avoir suivi mes travaux au sein de l'équipe SA-IGA tout au long de ces trois ans.

Je tiens à remercier Mr. Frédéric Chausse pour avoir accepté de rapporter et d'évaluer ce travail de recherche.

Je remercie également Mr. Jean-Christophe Popieul pour son accueil chaleureux au LAMIH à Valenciennes ainsi que pour avoir accepté de rapporter et d'évaluer ce travail de recherche.

Merci à Mr. Alain Muzet pour le temps qu'il a pu m'accorder afin de discuter du sujet. Je suis très sensible à sa participation au jury en tant qu'examineur.

Un grand merci également à Mmes. Suzanne Lesecq et Nadine Martin ainsi qu'à toute l'équipe SA-IGA au sein de laquelle j'ai évolué durant ces trois ans. Merci pour leur aide et leur conseil. Un grand merci également à Mr. Julien Huillery et Fabien Milloz, anciens doctorants, pour leur participation à mes travaux.

Merci à Mr. Laurent Bonnaud pour toute l'aide qu'il m'a apportée ; celle-ci m'a permis de résoudre de nombreux problèmes concernant l'acquisition vidéo.

Un grand merci à Mmes. Marie-Thérèse Descotes-Genon, Patricia Reynier et Virginie Messina ainsi qu'à toute l'équipe administrative, technique et informatique du Gipsa-Lab et du Département Automatique en particulier sans qui il serait difficile de travailler.

J'aimerais aussi exprimer mes remerciements à toutes les personnes du Gipsa-Lab qui, de près ou de loin, ont participé au bon déroulement de cette thèse.

Je remercie particulièrement toutes les personnes qui ont eu la patience et le courage de participer à mes expériences parfois très (voir trop...) expérimentales.

Un merci très spécial à André pour avoir éclaboussé tous les jours notre bureau d'un peu de soleil brésilien.

Merci à Sylvain et à toute l'équipe du GAG (association des doctorants) pour tous les bons moments et toutes les grosses activités organisées au sein du laboratoire.

Merci à toute l'équipe organisatrice de la Journée Des Doctorants 2008 pour les bons moments scientifiques et extra-scientifiques.

"Trugarez" à Tom B. pour ses bons conseils et l'intérêt qu'il a porté à mes travaux depuis sa bretonnie adoptive.

Une pensée particulière pour Tom R. et Xavier, compagnons de galère. Courage les amis, la fin est proche !

Un grand merci à tous les camarades de la musique, de la voile et du comptoir qui ont fait paraître ces trois années bien plus courtes qu'elles ne l'ont été en réalité.

Merci à mon père pour son aide et son soutien au cours de toutes ces années. Merci de m'avoir aidé lors des différents choix que j'ai eu à réaliser et d'avoir respecté mes décisions.

Merci à Catherine pour toutes les corrections de dernière minute.

Un énorme merci à toute ma famille pour leur confiance et leur soutien depuis toujours. J'espère que ce travail vous rendra fier.

Enfin, je tiens à remercier tout particulièrement Mathilde pour son soutien, sa patience et son amour. J'espère pouvoir faire la même chose pour toi.

Bonne lecture,
Antoine

Table des matières

Résumé	i
Abstract	iii
Remerciements	v
Table des matières	vii
Table des figures	x
Liste des tableaux	xiv
Liste des algorithmes	xvii
Introduction	1
1 Caractérisation de l'hypovigilance par analyse EEG	9
1.1 État de l'art	9
1.1.1 L'électroencéphalographie	10
1.1.2 Manifestations de l'hypovigilance dans l'EEG	12
1.1.3 Les systèmes de détection	14
1.2 Méthode de détection	18
1.2.1 Pré-traitements	19
1.2.2 Comparaison de moyennes	22
1.2.3 La détection d'artéfacts	25
1.2.4 Intérêt de la méthode	28
1.3 Résultats et discussion	29
1.3.1 Base de données	30
1.3.2 Techniques de validation	31
1.3.3 Résultats à partir de la puissance relative alpha	35
1.3.4 La détection d'artéfacts	36
1.3.5 Résultats à partir d'autres indices	38
1.3.6 Résultats de la fusion des informations α et β	40
1.3.7 Comparaison à la littérature	45
1.4 Conclusion du chapitre	46

2	Caractérisation vidéo des clignements	49
2.1	État de l'art	50
2.1.1	Hypovigilance et EOG	50
2.1.2	Les systèmes de détection d'hypovigilance par analyse des clignements	55
2.1.3	Problématique	57
2.2	Caractérisation des clignements par analyse vidéo	58
2.2.1	Filtrages rétiniens	59
2.2.2	Extraction de zones d'intérêt	61
2.2.3	Caractérisation des clignements	67
2.3	Comparaison entre EOG et vidéo	73
2.3.1	Base de données	73
2.3.2	Résultats et discussion	74
2.4	Conclusion du chapitre	81
3	Détection d'hypovigilance par analyse vidéo	83
3.1	Revue des techniques existantes	83
3.2	Analyse mono-variable	87
3.2.1	Préambule	87
3.2.2	Techniques décisionnelles	88
3.2.3	Sélection des variables les plus pertinentes	91
3.2.4	Choix de la fenêtre de traitement	100
3.2.5	Synthèse	102
3.3	Méthode proposée de détection d'hypovigilance	103
3.3.1	Description de la méthode	104
3.3.2	Intérêt de la méthode	107
3.4	Résultats et discussion	107
3.4.1	Résultats obtenus par notre méthode	108
3.4.2	Comparaison à la littérature	111
3.4.3	Approche par les arbres de décision	113
3.5	Conclusion du chapitre	115
4	Vers une fusion des approches physiologiques et vidéo	117
4.1	État de l'art	117
4.1.1	La fusion de données	118
4.1.2	Revue des systèmes existants pour la détection d'hypovigilance	121
4.2	Notre méthode	124
4.2.1	Les détections EEG et video	125
4.2.2	Le système de fusion	127
4.2.3	Synchronisation temporelle	129
4.2.4	Intérêt de la méthode	131
4.3	Résultats et discussions	132
4.3.1	La base de données	132
4.3.2	Résultats obtenus	132
4.3.3	Comparaison avec la littérature	138
4.3.4	Approche par les arbres de décisions	140

4.3.5	Vers une validation des résultats	141
4.4	Conclusion	143
Conclusions et perspectives		147
A Algorithme de détection des clignements à partir de l'EOG		151
A.1	Le modèle de clignement utilisé	151
A.2	L'algorithme de détection	152
A.3	Calcul des caractéristiques des clignements	154
B Introduction à la logique floue		157
B.1	Les sous-ensembles flous	157
B.2	Opérations sur les sous-ensembles flous	158
B.2.1	Opérateur NON	159
B.2.2	Opérateur ET	160
B.2.3	Opérateur OU	160
B.2.4	Compromis logiques	161
C Introduction à l'induction d'arbres de décision		163
C.1	Présentation	163
C.2	L'apprentissage par arbre de décision	164
C.3	L'algorithme C4.5	165
C.3.1	Algorithme de base	166
C.3.2	Élagage	167
C.3.3	Cost Sensitive C4.5	167
Bibliographie		168
Publications		179

Table des figures

1.1	Signal EEG décomposé selon ses différentes bandes de fréquence (source : wikipédia)	11
1.2	Système 10-20 ([Malmivuo95])	12
1.3	Méthode de détection d'hypovigilance à partir de l'EEG	18
1.4	Transformée de Fourier avec fenêtre glissante et méthode de Welch	21
1.5	Illustration des fenêtres utilisées pour la MCT	23
1.6	Exemple d'artéfact de haute amplitude sur l'EEG (a) et son spectrogramme (b)	26
1.7	Traitement du signal depuis l'EEG brut jusqu'à la détection d'hypovigilance	29
1.8	Le simulateur dynamique PAVCAS	30
1.9	Comparaison entre la décision experte (a,b) et la décision de notre système (c)	33
1.10	Résultats obtenus sur les différents canaux EEG	35
1.11	Exemple de détection d'artéfacts de haute amplitude	37
1.12	Nombre d'artéfacts détectés en fonction du seuil λ_{art}	38
1.13	Courbes ROC de la détection d'hypovigilance lorsque la détection d'artéfact est utilisée	39
1.14	Courbes ROC obtenues en utilisant différents indices pour la détection d'hypovigilance	40
1.15	Probabilité expérimentale d'être "fatigué" en fonction du seuil λ	42
1.16	Fonction d'appartenance à la classe fatiguée	43
1.17	Performances obtenues en utilisant α_{rel} , β_{rel} et $(\alpha_{rel}, \beta_{rel})$ fusionnés par logique floue	43
2.1	Changements du potentiel EOG liés aux mouvements des yeux	51
2.2	Placement des électrodes pour l'enregistrement EOG	51
2.3	Exemple EOG d'un clignement, sa dérivée et les paramètres considérés	52
2.4	Définition de la durée à 50%	53
2.5	Enregistrement EOG de la baisse de vigilance ([Svensson04])	55
2.6	Méthode de détection des clignements	59
2.7	Fonction de transfert du filtre OPL	60
2.8	Exemple de filtrage OPL	61
2.9	Exemple de filtrage IPL lors d'une fermeture des yeux	61
2.10	Exemples de détection de visages	62
2.11	Illustration des différentes boîtes englobantes pour la détection des yeux	63

2.12	Méthode de détection de l'œil	64
2.13	Exemples de détections de la position de l'œil	65
2.14	Comparaison de détection d'un œil fermé	67
2.15	Exemple des signaux E_{OPL} (a) et E_{IPL} (b) lors d'un clignement	68
2.16	Méthode de détection des clignements	70
2.17	Exemple de mauvaise détection de l'œil sur les signaux E_{OPL} et E_{IPL}	71
2.18	Paramètres extraits des signaux E_{OPL} (a) et E_{IPL} (b) lors d'un clignement	72
2.19	Extraits de la base de données	73
2.20	Exemple de détection EOG (a) et vidéo (b) de deux clignements rapides successifs	75
2.21	Exemple des signaux EOG et E_{OPL} (a) et leur dérivées (b) lors d'un clignement	79
3.1	Représentation schématique du système expert flou proposé par Damousis et al. ([Damousis09])	86
3.2	Technique décisionnelle mono-variable	90
3.3	Résultats obtenus pour la durée à 50%	92
3.4	Résultats obtenus pour le PERCLOS80	93
3.5	Résultats obtenus pour la fréquence	94
3.6	Résultats obtenus pour le rapport amplitude-vélocité	95
3.7	Résultats obtenus pour l'énergie de fermeture de l'œil	96
3.8	Résultats optimaux obtenus pour chaque indice	98
3.9	Résultats obtenus avec et sans considérer les courts clignements pour les différentes variables	100
3.10	Résultats obtenus pour les différentes longueurs de la fenêtre d'étude	101
3.11	Représentation des résultats optimaux obtenus avec une détection par seuil fixe	103
3.12	Méthode de détection d'hypovigilance à partir des données oculaires	104
3.13	Fonction d'appartenance en fonction du seuil s_F de la variable considérée	105
3.14	Résultats par fusion des différentes variables comparés à ceux obtenus variable par variable	108
3.15	Résultats obtenus par tirage aléatoire des seuils de détections	111
3.16	Arbre de décision d'hypovigilance	114
4.1	Architecture générale du système de fusion	125
4.2	Architecture complète du système de fusion	128
4.3	Comparaisons des résultats obtenus pour la détection vidéo validée par la détection EEG et la détection vidéo seule	135
4.4	Détections d'hypovigilance pour le sujet n° 12	137
4.5	Détections d'hypovigilance pour le sujet n° 8	137
4.6	Arbre de décision d'hypovigilance	141
4.7	Exemple de sujet durant un enregistrement	142
A.1	Modèle de clignement et sa dérivée	152
A.2	Algorithme de détection des clignements	153
A.3	Définition des caractéristiques d'un clignement	154

B.1	Caractérisation de la température de l'eau en logique binaire (a) et en logique floue (b)	158
B.2	Caractérisation d'une température $T=17,5$ ° C en logique floue	159
B.3	Fonction d'appartenance au sous-ensemble A (a) et son complémentaire \bar{A} (b)	159
B.4	Fonctions d'appartenance aux sous-ensembles A et B (a) et leur intersection (b)	160
B.5	Fonctions d'appartenance aux sous-ensembles A et B (a) et leur union (b)	161

Liste des tableaux

1	Résumé des critères d'évaluation d'hypovigilance proposés par A.Muzet . . .	5
1.1	Critères OSS proposés par Muzet et al.([Muzet03])	14
1.2	Critères OSS restreints au contenu EEG	32
1.3	Table de contingence	32
1.4	Table de contingence pour le test de McNemar	34
1.5	Table de contingence pour la comparaison entre la détection par fusion avec celle utilisant MCT_{alpha} ($\lambda_{MCT} = 3$)	44
1.6	Synthèse de la comparaison de nos résultats à ceux de la littérature concernant la détection de l'hypovigilance à partir des signes cérébraux	46
2.1	Performances de l'algorithme de détection des yeux sur la base BioID (1521 images)	66
2.2	Nombre de clignements	74
2.3	Résultats de la détection vidéo de clignements	75
2.4	Durée moyenne des clignements manqués	76
2.5	Résultats de la détection vidéo de clignements avec localisation manuelle de l'œil	76
2.6	Coefficients de corrélation pour les différents paramètres	78
2.7	Coefficients de corrélation entre l'EOG et la vidéo	80
3.1	Critères OSS restreints aux indices visuels	89
3.2	Résultats optimaux obtenus pour la durée à 50%	93
3.3	Résultats optimaux obtenus pour le PERCLOS80	94
3.4	Résultats optimaux obtenus pour la fréquence	95
3.5	Résultats optimaux obtenus pour le rapport amplitude-vélocité	96
3.6	Résultats optimaux obtenus pour l'énergie de fermeture de l'œil	97
3.7	Synthèse des résultats obtenus avec une détection par seuil fixe	103
3.8	Table de contingence pour la comparaison entre la détection par fusion floue avec celle utilisant uniquement le PERCLOS80	109
3.9	Table de contingence pour la comparaison entre la détection par fusion floue avec celle par fusion booléenne	110
3.10	Synthèse de la comparaison de nos résultats à ceux de la littérature concernant la détection d'hypovigilance à partir d'indices visuels	113
3.11	Données d'entrée utilisées pour la construction de l'arbre de décision	114
4.1	Critères OSS restreints aux niveaux 1 et 2 d'hypovigilance	124

4.2	Résultats obtenus par notre système de détection	133
4.3	Résultats obtenus par fusion “aveugle”	134
4.4	Table de contingence pour la comparaison entre notre algorithme avec un algorithme de fusion “aveugle”	134
4.5	Table de contingence pour la comparaison de la fusion des détection vidéo et EEG à la détection vidéo seule	136
4.6	Synthèse de la comparaison de nos résultats à ceux de la littérature concernant la détection de l’hypovigilance par fusion d’informations	139
4.7	Données d’entrée utilisées pour la construction de l’arbre de décision	140

Liste des algorithmes

1	Détection d'hypovigilance par fusion floue	107
2	Détection d'hypovigilance par fusion EEG et vidéo	130
3	Apprentissage par arbre de décision	164

Introduction

Contexte général

Le problème de la sécurité routière est un problème important dont l'enjeu est la diminution du nombre d'accidents, et par là même du nombre de morts et de blessés, sur les routes. D'après le rapport annuel de l'ONISR (Observatoire National Interministériel de Sécurité Routière), plus de 76000 accidents ont eu lieu en 2008 sur les routes de France, provoquant la mort de plus de 4400 personnes et environ 97000 blessés ([ONISR09]). De nombreux efforts ont permis de réduire le nombre d'accidents liés à la vitesse ou à l'alcool, comme en atteste la multiplication des campagnes de prévention et de contrôles routiers.

En revanche, la prévention concernant la vigilance et la sensation de fatigue du conducteur reste quasi-inexistante. La baisse du niveau de vigilance, appelée *hypovigilance*, engendrée par la transition involontaire de l'état de veille vers le sommeil est pourtant responsable du tiers des accidents mortels sur autoroute d'après un rapport d'accidentologie de la Sécurité Routière ([SR08]). Ce problème est d'autant plus sérieux avec l'augmentation du travail de nuit et des cadences de travail dans une société du "24 heures sur 24" où les gens favorisent souvent gain et plaisir plutôt que sommeil. L'augmentation du nombre d'accidents liés à l'hypovigilance ces dernières années (un tiers des accidents en 2007 contre un sixième en 1997) va dans ce sens. Il apparaît donc primordial de parvenir à détecter et à prévenir cet état accidentogène du conducteur afin d'améliorer sa sécurité.

Définition de l'hypovigilance

Avant toutes choses, il semble important de donner une définition aussi précise que possible du phénomène qui nous intéresse. Si l'on regarde la définition du terme hypovi-

gilance dans un dictionnaire, on peut lire :

Hypovigilance : état intermédiaire entre la veille et le sommeil durant lequel l'organisme a ses facultés d'observation et d'analyse très réduites.

(Source : Reverso, dictionnaire en ligne)

Effectivement, il s'agit bien ici de la transition entre les deux états fondamentaux que sont l'éveil et le sommeil. Cette transition est involontaire et accompagnée d'une baisse de vigilance, c'est-à-dire de la capacité à porter son attention sur une tâche. Il faut faire cependant attention à différencier l'hypovigilance de la fatigue. Le terme de fatigue correspond à un épuisement du corps après une activité soutenue trop longue. Elle existe sous deux formes : physique et mentale. Ces deux formes ne sont d'ailleurs pas forcément liées. La fatigue nécessite du repos, ce qui n'est pas forcément synonyme de sommeil. Elle engendre cependant une baisse de la capacité à maintenir son attention et à se concentrer car comme le corps a besoin de repos, il y a moins de ressources disponibles. L'hypovigilance n'est donc qu'une conséquence de la fatigue. On préférera donc parler de niveau d'éveil ou, en anglais, de "drowsiness". Cependant, pour des modalités de rédaction, nous utiliserons dans la suite le terme "fatigué" en opposition à celui de "éveillé" de manière à nommer l'état d'un conducteur souffrant d'hypovigilance.

Les conséquences de l'hypovigilance au niveau physiologique sont nombreuses. Une étude du ministère américain des transports (*FHWA* pour *Federal Highway Administration*) recense ces différentes conséquences ([Wylie96]) :

- Baisse de l'attention,
- Augmentation du temps de traitement de l'information et de prise de décision,
- Augmentation du temps de réaction aux évènements critiques,
- Réponses contrôles plus variables et moins efficaces,
- Baisse de la motivation et des performances,
- Baisse de l'activité psychophysique,
- Augmentation des sentiments de somnolence et de fatigue,
- Baisse de la vivacité.

À la vue de ces différentes conséquences, il apparaît clairement pourquoi la conduite dans cet état représente un danger, pour le conducteur et les autres usagers, ainsi que l'enjeu de développer un système capable de prévenir l'hypovigilance. De plus, de nom-

breux mécanismes tels que les cycles circadiens et homéostatiques (cycles physiologiques régissant l'éveil et le sommeil), l'inertie du sommeil (temps nécessaire à chaque individu pour récupérer l'intégralité de ses capacités cognitives au réveil), les micro-sommeils (épisode de sommeil de quelques secondes) ainsi que l'environnement du conducteur (moment de la journée, conditions atmosphériques, durée de conduite, etc) peuvent favoriser l'hypovigilance, ce qui ne fait que complexifier ce phénomène.

État des recherches

Depuis une trentaine d'années, des recherches ont été menées afin de mieux comprendre cet état à risque et de mettre en œuvre des systèmes permettant sa détection automatique. Deux types d'approches ont été proposés afin d'étudier ce phénomène. La première approche, la plus ancienne, est orientée sur le comportement du véhicule. Son but est de détecter les comportements anormaux du véhicule sur la route engendrés par l'hypovigilance. Pour cela, différents indices tels que la position du véhicule sur la route, les mouvements du volant ou la force de pression sur les différentes pédales ont été étudiés afin de pouvoir prévenir les conduites à risque liées à l'hypovigilance. Cette approche reste cependant très dépendante de la route, du trafic routier et des habitudes de conduite du conducteur.

La seconde approche est orientée sur le comportement du conducteur. Dans ce cas, le but est de détecter les signes physiologiques et comportementaux de l'hypovigilance chez le conducteur. Des études sur les signes physiologiques au niveau de l'activité cérébrale et oculaire ont été menées afin de mieux comprendre et caractériser ce problème d'hypovigilance. C'est d'ailleurs à partir de ces signaux physiologiques que les médecins travaillent afin de quantifier le niveau de vigilance chez un patient. Plus récemment, l'utilisation de caméras embarquées dans le véhicule a permis de développer des systèmes de détection basés sur les indices visuels de la baisse de vigilance tels que les clignements des yeux, les bâillements ou la direction du regard du conducteur. Cependant, aucun système ne s'est révélé suffisamment efficace à l'heure actuelle pour être commercialisé et l'approche physiologique reste aujourd'hui la plus efficace pour caractériser l'hypovigilance.

Au cours de ces différents travaux, plusieurs problèmes ont été soulevés, le principal étant la difficulté de proposer une définition précise de l'hypovigilance et de ses caractéristiques. De plus, l'hypovigilance se manifeste de manière différente chez les différents

individus, ce qui augmentent la difficulté à définir un référentiel commun pour la détection de ce phénomène. Cette dernière constatation ajoutée au fait qu'aucune approche ne se soit révélée suffisamment pertinente a conduit à essayer de fusionner différentes techniques afin de d'améliorer les systèmes de détection d'hypovigilance.

Problématique

Nous proposons ici une approche résolument orientée sur l'étude du conducteur afin de détecter l'hypovigilance à partir d'informations physiologiques (liées à l'activité cérébrale) et vidéo (liées à l'activité oculaire). Le but de ce travail consiste à développer un système de détection automatique de l'hypovigilance chez le conducteur à partir des données électroencéphalographiques (décrivant l'activité cérébrale) et vidéo du conducteur. Cette approche est motivée par le fait que les médecins spécialistes de l'hypovigilance travaillent essentiellement à partir des données cérébrales et visuelles pour détecter l'hypovigilance. De plus, la complémentarité des activités cérébrale et oculaire semble indiquer que l'apport des informations cérébrales permettrait d'améliorer la fiabilité des approches par caméra embarquées (n'utilisant donc que les indices visuels) utilisées pour la détection automatique de la baisse de vigilance.

En effet, l'hypovigilance résulte de nombreux facteurs complexes qui interagissent entre eux. De plus, si ces facteurs favorisent la baisse de vigilance, tout le monde ne réagit pas de la même manière et il existe de nombreuses différences inter-individuelles à ce niveau ([Karrer04]). La difficulté à définir l'hypovigilance est liée au fait qu'il s'agit d'une transition entre deux états et non d'un état à part entière. En effet, nous avons présenté ici les différents facteurs ainsi que les effets de l'hypovigilance mais il est difficile d'en établir les caractéristiques précises. En fait il n'existe pas d'alphabet commun de l'hypovigilance qui permette de définir à coup sûr si une personne est victime d'hypovigilance ou non, comme les règles de Rechtschaffen & Kales ([Rechtschaffen68]) pour définir les différentes phases du sommeil par exemple. Les travaux de Dinges ([Dinges95]) ont cependant montré que les indicateurs physiologiques, notamment les activités cérébrale et oculaire, sont les meilleurs indicateurs de cet état de transition entre éveil et sommeil. En effet, Renner & Mehring ([Renner97]) ont suggéré que l'étude de l'activité cérébrale permet de repérer la défaillance dans le traitement de l'information et que l'étude de l'activité oculaire permet de repérer l'insuffisance des capacités de perception. Nous commencerons donc par nous focaliser sur l'activité électroencéphalographique pour construire notre système de détec-

tion d'hypovigilance. Des échelles ont ainsi été proposées par des médecins spécialisés afin d'évaluer le niveau d'éveil d'un patient mais aucune n'a pour l'instant été standardisée. On peut citer parmi elles l'approche développée par l'institut Karolinska appelée *Karolinska Drowsiness Scale* ou *KDS* ([Gillberg96]) ou celle développée par Muzet et al. ([Muzet03]) appelée *Objective Sleepiness Scale* ou *OSS* résumée dans le tableau 1. Les notions d'activités α et θ correspondent à des bandes de fréquences de l'activité cérébrale qui seront définies dans le chapitre suivant.

État	Activité cérébrale	Activité oculaire
Éveil	Normale	Normale
Hypovigilance	Apparition d'activités α et/ou θ	Ralentissement des clignements

TAB. 1 – Résumé des critères d'évaluation d'hypovigilance proposés par A.Muzet

Nous reviendrons en détail sur ces échelles dans les chapitres suivants. Ces échelles ont néanmoins toutes pour point commun qu'elles permettent l'évaluation du niveau d'éveil à partir de deux indices : l'activité cérébrale et l'activité oculaire. Dans ce cadre, il nous semble judicieux d'étudier le problème de l'hypovigilance au volant sous ces deux aspects complémentaires.

Plan de la thèse

Ce document présente le travail effectué afin d'élaborer un système de détection automatique de l'hypovigilance chez le conducteur par fusion d'informations physiologiques et d'indices visuels. Nous commençons par présenter une approche physiologique du problème puis nous présentons une approche vidéo. Enfin, une fusion des deux approches est proposée.

Le problème délicat et complexe de la définition de l'hypovigilance à partir des signaux électroencéphalographiques est introduit dans le chapitre 1. Nous nous intéressons dans ce chapitre aux signes de l'hypovigilance présents dans l'électroencéphalogramme (*EEG*). Après une présentation de ces différents signes, un algorithme de détection d'hypovigilance à partir de l'EEG est proposé. Ce système est ensuite validé et discuté à partir d'une base de données de conduite en état d'hypovigilance.

Dans le deuxième chapitre, nous nous intéressons aux signes visuels de l'hypovigilance que sont les clignements. Pour des raisons d'ergonomie, nous comparons deux approches permettant la caractérisation de ces clignements : l'analyse par électrooculogramme (*EOG*) et l'analyse par vidéo. Nous nous interrogeons quant à la possibilité de remplacer l'analyse des signes oculaires traditionnellement faite par étude d'EOG par l'étude d'une vidéo des yeux du conducteur. Un système de caractérisation des clignements des yeux par analyse d'une vidéo haute fréquence (jusqu'à 200 images par seconde) est présenté. Les résultats obtenus par cette méthode sont comparés à ceux obtenus à partir de la caractérisation EOG des clignements afin de mettre en évidence les différents indices pouvant être estimés avec la même précision par caméra rapide et par EOG.

Le chapitre 3 propose un système de détection d'hypovigilance à partir des informations vidéo. Les différents indices mis en évidence dans le chapitre 2 sont d'abord étudiés afin d'évaluer leur pertinence pour caractériser l'hypovigilance. Les indices pertinents sont alors fusionnés afin d'obtenir un détecteur fiable d'hypovigilance. Ce système de détection est ensuite évalué sur une base de données conséquente.

Enfin, dans un quatrième chapitre, nous analysons comment fusionner les approches physiologique et vidéo afin d'obtenir un système de détection automatique d'hypovigilance plus fiable. Nous achevons cette réflexion par une conclusion générale sur le travail réalisé ainsi que par une présentation des différentes perspectives envisagées.

*“ J’ai peur de m’endormir au volant.
– Ah... Tu sais ce que tu fais ?
T’attaches tes cheveux au plafond.”*

Extrait du film Eldorado de Bouli Lanners (2008)

Chapitre 1

Caractérisation et détection de l'hypovigilance par analyse de l'activité cérébrale

Ce chapitre a pour objet d'étudier dans quelle mesure l'activité cérébrale est révélatrice de notre niveau de vigilance. Il en résulte l'élaboration d'un algorithme de détection d'hypovigilance à partir de l'analyse de l'électroencéphalogramme (EEG). Après une présentation de la caractérisation de l'hypovigilance du point de vue de l'activité cérébrale ainsi qu'un tour d'horizon des techniques de détection existantes basées sur l'EEG, un algorithme de détection d'hypovigilance fonctionnant en ligne à partir du spectre EEG filtré médian d'un seul canal EEG est proposé. Cet algorithme est basé sur un test statistique de comparaison de moyenne et a l'avantage de posséder un seuil de détection indépendant du conducteur. Une technique de détection d'artéfacts utilisant un test de variance est également développée afin d'étudier en temps réel la fiabilité du signal EEG. Notre méthode est ensuite évaluée sur une base de données conséquente et comparée aux autres méthodes de la littérature.

1.1 État de l'art

Après une présentation de l'électroencéphalographie qui permet la mesure de l'activité cérébrale, nous présentons les différentes manifestations de l'hypovigilance au niveau de

cette activité. Une revue des différents systèmes de détection d'hypovigilance existants est ensuite présentée.

1.1.1 L'électroencéphalographie

L'*électroencéphalographie* (*EEG*) est la mesure de l'activité électrique du cerveau par des électrodes placées sur le cuir chevelu, souvent représentée sous la forme d'un tracé appelé électroencéphalogramme ([Blinowska06]).

Typiquement, l'EEG est décrit en termes d'activités rythmiques et transitoires. L'activité rythmique de l'EEG est divisée en bandes de fréquence. Ces différentes bandes de fréquence sont avant tout une affaire de nomenclature et ont été définies en fonction de la signification biologique des différents rythmes. Il se peut cependant que certaines bandes changent de quelques Hertz d'un auteur à l'autre. L'activité EEG est principalement étudiée dans la bande [1-30]Hz car en dessous et au dessus de cette bande, l'EEG a tendance à être très bruité (lorsqu'il est enregistré dans des conditions cliniques standards).

Voici une brève présentation des différentes bandes de fréquence présentes dans l'EEG :

Delta L'activité δ correspond à la bande de fréquence [0,5-4]Hz. Cette activité est généralement composée d'ondes de haute amplitude qui apparaissent dans la région frontale chez les adultes et pariétale chez les enfants. Cette activité est principalement présente chez les adultes lors du sommeil profond et chez les bébés.

Theta L'activité θ correspond à la bande de fréquence [4-8]Hz. Elle se manifeste principalement chez les enfants, lors de la somnolence chez les adolescents et adultes et en cas d'inactivité.

Alpha L'activité α correspond à la bande de fréquence [8-12]Hz. Elle apparaît principalement au niveau des régions pariétales de la tête mais principalement du côté dominant et également dans les zones centrales lorsqu'on se repose. Cette activité est caractéristique de la relaxation. Elle est également très présente lorsqu'on ferme les yeux.

Beta L'activité β correspond à la bande de fréquence [12-30]Hz. Elle est présente des deux côtés du cerveau de manière symétrique. Elle est constituée d'ondes de faible amplitude et apparaît plus clairement dans la zone frontale. Cette activité est caractéristique de l'éveil, de la réflexion active ainsi que de la concentration.

Gamma L'activité γ correspond à la bande de fréquence au dessus de 30Hz. Elle apparaît

lors de certaines tâches cognitives ou fonctions motrices et est difficile à enregistrer dans des conditions cliniques standards.

La figure 1.1 présente une seconde de signal EEG puis toutes les composantes de ce signal, correspondant aux différentes bandes de fréquence.

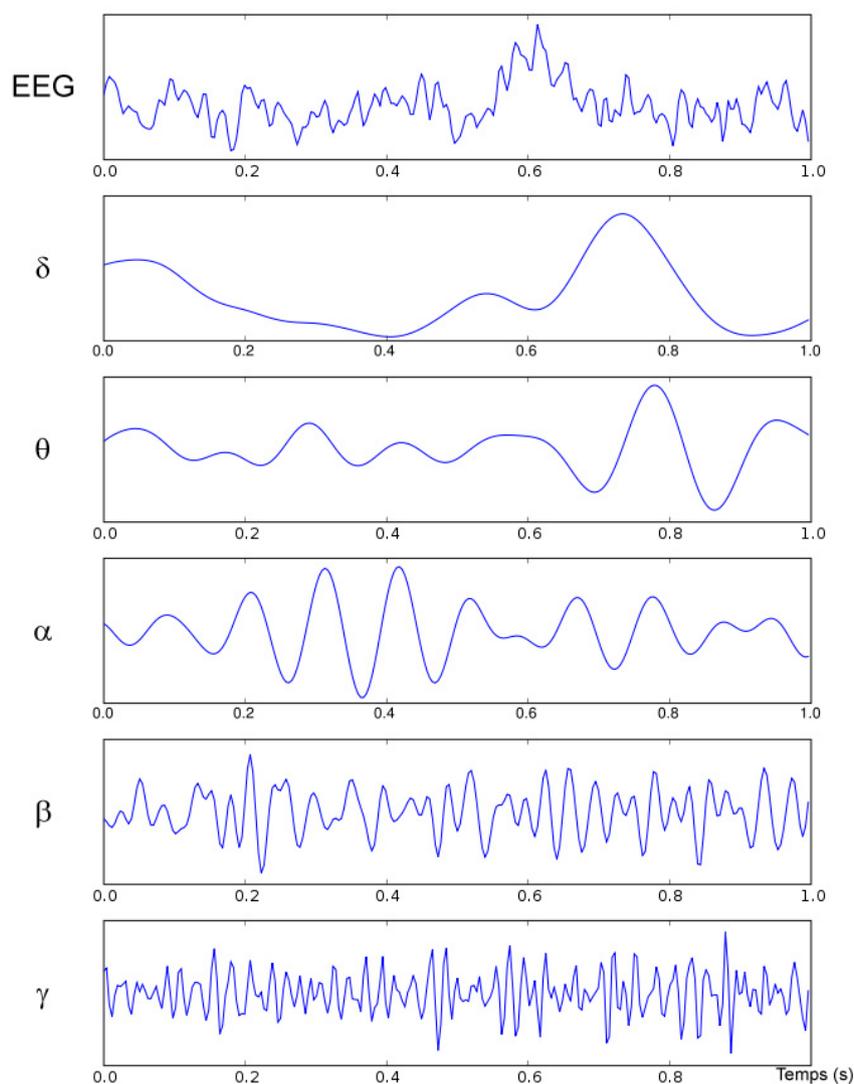


FIG. 1.1 – Signal EEG décomposé selon ses différentes bandes de fréquence (source : wikipédia)

Traditionnellement, l'EEG est enregistré selon le système 10-20 présenté figure 1.2. Ce système tient son nom du fait que les électrodes adjacentes sont placées à 10% ou 20% de la distance totale entre l'avant et l'arrière et entre la droite et la gauche du crâne, comme indiqué sur la figure. La valeur des tensions relevées va d'environ $10\mu V$ à $100\mu V$. L'échantillonnage a souvent lieu autour de 250Hz.

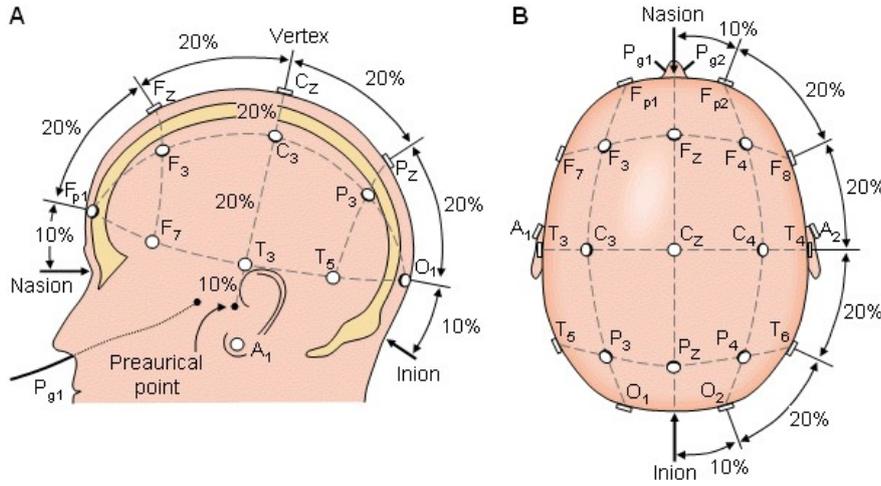


FIG. 1.2 – Système 10-20 ([Malmivuo95])

Le nombre de canaux utilisés pour un enregistrement peut monter jusqu'à 256. La multiplication du nombre de canaux d'enregistrement permet d'améliorer la précision du calcul de l'activité de chaque bande de fréquence ainsi que leur localisation. En revanche, elle diminue le confort du patient. De plus, dans le cas d'un patient sain ne présentant ni souffrance hémisphérique, ni dépression d'amplitude correspondant à la perte de neurones, on constate une symétrie de rythme et d'amplitude de l'activité cérébrale en ce qui concerne les activités de fond α , β et θ ([Vion08]). On peut donc se contenter selon l'application d'étudier un seul hémisphère cérébral sur les deux. De même, le nombre de canaux étudiés dépend de l'application.

1.1.2 Manifestations de l'hypovigilance dans l'EEG

Au regard de la signification physiologique des différentes bandes de fréquence EEG, il apparaît que les bandes susceptibles d'être caractéristiques de l'hypovigilance sont les bandes θ dont l'activité correspond à la somnolence et à l'inactivité ainsi que la bande α qui correspond au repos et à la fermeture des yeux. En effet, plusieurs études ([Gillberg96, Cajochen00, Otmani05, Papadelis06]) ont mis en avant que la baisse de vigilance est caractérisée par une hausse de l'activité électrique dans les bandes α et θ . Les récentes études de Kaida et al. ([Kaida07]) ainsi que celles de Huang et al. ([Huang09]) ont révélé que la bande α est particulièrement caractéristique de la baisse de vigilance.

De plus, on constate que la bande β est caractéristique de l'éveil et de la concentra-

tion. La baisse de vigilance se caractérise également par une baisse de l'activité dans la bande β . Cette baisse d'activité a été mise en évidence par plusieurs études ([Makeig00, Oken06, Peiris06]). Cependant la baisse de l'activité β en cas d'hypovigilance doit être nuancée suite aux travaux de Smit et al. ([Smit04]). En effet, leur étude a montré que la baisse de vigilance peut s'accompagner d'une hausse de l'activité β due à une hausse de concentration pour tenter de compenser l'apparition d'hypovigilance.

Au niveau spatial, l'hypovigilance apparaît principalement dans la zone pariétale ([Bittner01, Lin05a, Huang09]) et dans la zone centrale ([Papadelis06, Oken06]). La zone pariétale correspond aux électrodes P_3 , P_4 et P_z de la figure 1.2 et la zone centrale aux électrodes C_3 , C_4 et C_z . Cela semble cohérent avec le fait que l'activité α apparaît principalement au niveau des régions pariétales de la tête.

Plusieurs échelles ont été proposées afin d'évaluer le niveau d'éveil mais aucune n'a été standardisée. Deux échelles d'évaluation ont particulièrement retenu notre attention. La première est la méthode développée par l'institut Karolinska appelée *Karolinska Drowsiness Scale* ou *KDS* ([Gillberg96]). Cette méthode propose de quantifier le niveau d'éveil à partir des relevés EEG et EOG. La méthode de notation est inspirée des règles de notation des phases de sommeil développées par Rechtschaffen & Kales (Rechtschaffen68). Les données sont analysées par segment de 20s eux mêmes divisés en sous-segments de 2s. Chaque fois qu'un signe d'hypovigilance est trouvé sur un des sous-segments le score attribué au segment est incrémenté de 10. Ainsi le score maximum est 100 et le sujet sera considéré comme victime d'hypovigilance si le score du segment est supérieur à 50. La seconde échelle qui nous a intéressée est celle développée par Muzet et al. ([Muzet03]). Il s'agit d'une échelle d'évaluation objective de la vigilance (*OSS* pour *Objective Sleepiness Scale*). Cette OSS est séparée en cinq niveaux de vigilance allant de 0 (complètement éveillé) à 4 (hypovigilance très avancée). Les décisions sont prises toutes les 20s en fonction de la longueur des bouffées d'activité α et θ présentes dans l'EEG ainsi que de la vitesse des mouvements des yeux et des clignements. Ces critères sont présentés dans le tableau 1.1.

Les deux échelles présentées servent à des médecins experts pour évaluer visuellement le niveau de vigilance à partir des différentes données EEG et oculaires d'une personne. Tout d'abord, on constate que ces deux échelles confirment les propos de Renner & Mehring ([Renner97]) sur le fait d'utiliser à la fois l'activité cérébrale et les clignements pour quantifier le niveau de vigilance. En revanche, une étude de Peiris et al. ([Peiris05]) a montré que deux experts différents ne donnent pas forcément le même diagnostic sur un

Objective Sleepiness Score	Contenu EEG	Clignements et mouvements d'yeux
0	Activité β continue en fond, pas d'activité α et θ	et clignements et mouvements d'yeux normaux
1	Apparition d'ondes α et/ou θ pendant au moins 5s cumulées	et clignements et mouvements d'yeux normaux
2	Apparition d'ondes α et/ou θ pendant au moins 5s cumulées ou Apparition d'ondes α et/ou θ pendant 5 à 10s cumulées	et clignements et mouvements d'yeux lents et clignements et mouvements d'yeux normaux
3	Apparition d'ondes α et/ou θ pendant 5 à 10s cumulées ou Apparition d'ondes α et/ou θ pendant plus de 10s cumulées	et clignements et mouvements d'yeux lents et clignements et mouvements d'yeux normaux
4	Apparition d'ondes α et/ou θ pendant plus de 10s cumulées	et clignements et mouvements d'yeux lents

TAB. 1.1 – Critères OSS proposés par Muzet et al.([Muzet03])

même relevé. Cela tend à prouver que le problème de la caractérisation de l'hypovigilance reste un sujet délicat et qu'il n'existe pas d'échelle universelle.

Ceci dit, l'utilisation de l'EEG par les médecins experts afin de quantifier le niveau de vigilance des sujets nous encourage à proposer un algorithme qui se base sur l'EEG pour détecter les moments d'hypovigilance. Le but de cet algorithme sera la détection, dans la mesure du possible, des différents critères EEG caractéristiques de l'hypovigilance, notamment la hausse de l'activité α .

1.1.3 Les systèmes de détection

La majorité des systèmes de détection d'hypovigilance à partir de l'EEG s'appuie sur l'analyse du contenu fréquentiel des signaux EEG. En effet, comme on a pu le voir dans la section précédente, l'hypovigilance se traduit principalement par une hausse de l'activité dans les bandes α et θ . Les systèmes de détection cherchent donc à analyser le contenu fréquentiel afin de détecter ces hausses.

On distingue deux méthodes pour analyser le contenu fréquentiel d'un signal de fré-

quence variable : la transformée de Fourier à court terme (*STFT* pour *Short Time Fourier Transform*) et la transformée en ondelettes (*WT* pour *Wavelet Transform*). Le principe de la STFT est de calculer la transformée de Fourier, i.e. calculer le contenu fréquentiel du signal, sur une fenêtre glissante, c'est-à-dire une partie du signal, le but étant d'avoir une représentation temps-fréquence. Ce principe est expliqué plus en détails dans la section 1.2.1. La WT quant à elle est une analyse temps-fréquence réalisée à partir d'une fonction appelée ondelette de base que l'on peut contracter ou dilater de manière à spécifier les caractéristiques du signal que l'on souhaite détecter. Cette technique a été développée par Mallat ([Mallat89]). Une comparaison entre l'utilisation de ces deux techniques dans le cadre de l'étude EEG a été réalisée par Kiyimik et al. ([Kiyimik05]). Il ressort de ce travail que la STFT est plus adaptée dans le cadre d'applications temps réel et que la WT est plus adaptée dans le cadre d'applications qui nécessitent une grande résolution. Dans la pratique, on constate que la majorité des travaux a été réalisée en utilisant la STFT. Les récents travaux de Sharabaty et al. ([Sharabaty08]) proposent d'utiliser la transformée de Hilbert Huang afin de localiser les ondes α et θ dans le signal EEG de manière précise. Cette technique n'a cependant pas encore été testée sur des signaux réels d'hypovigilance.

On distingue deux types de méthode pour analyser le contenu fréquentiel de l'EEG : les techniques avec apprentissage et sans apprentissage. Parmi les techniques supervisées à base d'apprentissage, on peut citer l'utilisation de réseaux de neurones. Le réseau de neurones est un algorithme d'apprentissage composé d'une structure de neurones connectés associant des entrées à une ou plusieurs sorties. Lorsqu'il est entraîné avec des données d'hypovigilance, le réseau de neurones va permettre l'apprentissage d'un modèle type "boîte noire" capable de détecter la baisse de vigilance. La principale différence entre toutes les approches utilisant des réseaux de neurones réside dans la nature des entrées utilisées ainsi que le type de réseau utilisé. Jung et al. ([Jung97]) ont suggéré d'appliquer une PCA sur un spectre EEG à trois canaux afin d'entraîner un réseau de neurones multicouches sur le spectre de la composante principale. Wilson & Bracewell ([Wilson00]) ont eux estimés le spectre EEG à la fois par STFT et par WT pour entraîner leur réseau. Leur technique n'a cependant pas été testée sur des signaux réels. Ben Khalifa et al. ([BenKhalifa04]) se sont servis d'un seul canal EEG pariéto-occipital (P4-O2) afin d'instruire deux réseaux différents. Un réseau de type SOM (Self-Organizing Map) est utilisé pour la détection de l'hypovigilance alors qu'un autre réseau de type LVQ (Learning Vector Quantization) est utilisé pour la détection des artéfacts. Subasi ([Subasi05]) a proposé quant à lui d'appliquer un réseau de neurones sur les bandes α , θ , β et δ d'un spectre EEG obtenu par WT afin de détecter les états suivants : "éveillé", "fatigué" et "endormi".

D'autres techniques de classification ont également été utilisées telles que les modèles markoviens cachés (*HMM* pour *Hidden Markov Model*). Le HMM est un modèle statistique dans lequel le système modélisé est supposé être un processus Markovien de paramètres inconnus. Un processus Markovien est un processus qui estime l'état futur à partir des seules informations de l'état présent et non celles passées. Il s'agit encore d'une technique nécessitant un apprentissage. Cette technique a été utilisée par Huang et al. ([Huang96]) pour détecter des formes (*pattern*) particulières correspondant aux différents niveaux d'éveil. Le HMM est appliqué sur un vecteur à huit dimensions où chaque dimension représente une bande particulière du spectre EEG.

Ces différentes techniques donnent de bons résultats, notamment les réseaux de neurones, mais nécessitent un apprentissage sur un grand nombre de données afin d'élaborer le modèle de détection d'hypovigilance. En pratique, cela n'est pas toujours réalisable.

D'autres techniques ne nécessitant pas d'apprentissage existent également. On peut citer les modèles régressifs. Le but est d'approcher le signal par un modèle mathématique qui est mis à jour à l'aide des valeurs passées du signal. Les états d'hypovigilance sont alors détectés lorsque le signal s'éloigne trop du modèle. Cette technique n'est pas récente et a été utilisée par Matousek & Petersén en 1983 ([Matousek83]). Ils ont proposé d'utiliser de multiples modèles régressifs sur 22 indices extraits du spectre EEG. Lin et al. ([Lin05a]) ont proposé d'appliquer un modèle auto-régressif d'ordre 50 sur le spectre d'un enregistrement EEG à deux canaux. Le spectre est divisé en 10 bandes de fréquence entre 1 et 40Hz. Les canaux sont traités par analyse en composantes principales (*PCA* pour *Principal Component Analysis*), ce qui permet de récupérer un signal qui représente le mieux la corrélation entre les deux canaux. Les indices ainsi récupérés servent à entraîner un modèle auto-régressif qui estime les performances de conduite du conducteur, c'est-à-dire sa capacité à maintenir son véhicule au milieu de la route. Cette technique a été légèrement améliorée par l'application d'une analyse en composantes indépendantes (*ICA* pour *Independent Component Analysis*) sur un enregistrement EEG à 33 canaux ([Lin05b]). L'ICA sert à séparer les différentes sources indépendantes d'une multitude de signaux. Les deux composantes principales résultantes de l'ICA serviront alors à l'entraînement du modèle auto-régressif. Les techniques de PCA et d'ICA sont très pratiques pour sortir l'information principale lors d'un enregistrement EEG multi-sources. L'ICA est également très utilisée pour retirer les artéfacts d'un enregistrement EEG ([Vigario97, Jung00]) car c'est une technique de séparation de sources. Plus récemment, Trejo et al. ([Trejo05]) ont proposé d'appliquer un modèle de régression par les moindres carrés partiels (Kernel Partial Least-Squares) sur un spectre EEG à 30 canaux. Cette étude réalisée pour la NASA

a donné d'excellents résultats. Les différents résultats sont présentés dans la section 1.3.5. Des modèles de mélange ont également été proposés par Rosipal et al. ([Rosipal07]). Un modèle de mélange considère qu'un échantillon de données suit non pas une loi de probabilité usuelle, mais une loi dont la fonction de densité est une densité mélange, c'est-à-dire une combinaison linéaire de plusieurs fonctions de densité. Rosipal et al. ont appliqué ce type de modèle sur le contenu spectral de segments EEG de 4s transférés sous la forme compacte du vecteur des coefficients du modèle auto-régressif associé. Ce système donne une indication d'hypovigilance à deux niveaux.

Ces modèles dépendent cependant du conducteur. En effet, les modèles régressifs présentés nécessitent un entraînement sur le conducteur afin d'estimer l'hypovigilance sur ce même conducteur. Ainsi, cette technique trouve sa limite dans les fortes différences inter-individuelles qui existent entre conducteurs ([Karrer04]). En effet, un modèle estimé sur un conducteur ne fonctionnera pas (ou mal) sur un autre conducteur. Chaque nouveau conducteur nécessite une phase d'entraînement qui est assez longue et demande beaucoup de données.

Enfin, il existe quelques méthodes ne nécessitant pas d'apprentissage et s'adaptant sans réglage aux différents conducteurs. On peut noter le travail de Bittner et al. ([Bittner01]) qui ont étudié différents indices EEG tels que l'énergie dans les bandes α , β et θ ainsi que des combinaisons des différents indices, notamment le rapport $\frac{\alpha+\theta}{\beta}$ qui avait été suggéré par De Waard & Brookhuis ([DeWaard91]). En effet, comme l'hypovigilance se traduit par une hausse de l'activité dans les bandes α et θ ainsi qu'une baisse de l'activité cognitive dans la bande β , l'indice $\frac{\alpha+\theta}{\beta}$ est censé augmenter avec l'hypovigilance. Bittner et al. ont comparé l'évolution de ces indices avec la baisse de vigilance mais n'ont trouvé qu'une faible corrélation. Enfin, assez récemment, Pal et al. ([Pal08]) ont suggéré d'étudier la déviation de puissance dans les bandes α et θ à l'aide de la distance de Mahalanobis. L'état de référence est pris lors la première minute de conduite en supposant que le conducteur est parfaitement éveillé lorsqu'il commence à conduire. La corrélation entre la combinaison linéaire de leurs deux indices et la dégradation des performances de conduite est satisfaisante.

Dans la littérature, il existe peu de méthodes fonctionnant sans apprentissage et indépendantes du conducteur. En effet, la plupart des méthodes présentées nécessitent soit un apprentissage sur un grand nombre de données afin d'élaborer le système de détection soit un entraînement spécifique sur chaque conducteur. Ces deux points sont clairement des inconvénients dans le but d'obtenir une méthode pouvant s'appliquer en ligne sur tous les

conducteurs. La méthode que nous présentons dans la section suivante s'inscrit clairement dans cette optique, celle d'être utilisable directement sans entraînement et en ligne.

1.2 Méthode de détection

Le but de cette partie est de proposer un algorithme de détection d'hypovigilance qui puisse fonctionner en ligne sans réglage préalable à partir de données expertisées sur le conducteur. L'étude de la littérature nous a permis de mettre en avant que les signes de l'hypovigilance apparaissent principalement dans la zone pariétale du cerveau, de manière symétrique. Nous proposons donc d'utiliser un seul canal pariétal pour notre système de détection afin de réduire au minimum le nombre de canaux EEG et ainsi améliorer le confort du conducteur. De plus, l'hypovigilance étant principalement caractérisée par l'évolution des activités α , θ et β , nous proposons de nous concentrer sur l'évaluation de l'activité dans ces différentes bandes de fréquence afin de concevoir notre algorithme. Une vue d'ensemble de notre méthode est présentée figure 1.3.

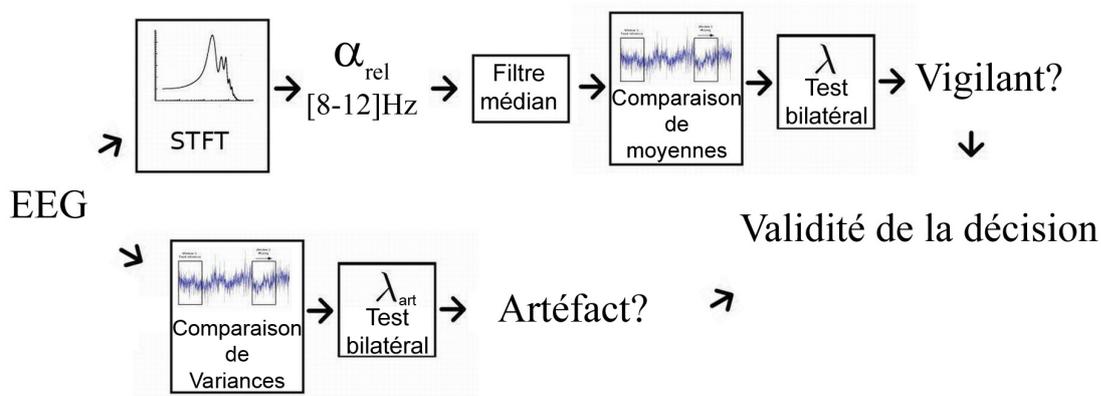


FIG. 1.3 – Méthode de détection d'hypovigilance à partir de l'EEG

Tout d'abord, on établit le spectre de puissance EEG à l'aide d'une transformée de Fourier à court terme (STFT) qui semble plus adaptée à une application temps réel. Ce spectre permet alors de calculer la puissance relative dans les différentes bandes EEG. On applique alors un filtrage médian sur la puissance relative des différentes bandes à l'aide d'une fenêtre glissante, de manière à lisser le signal et rejeter les valeurs anormales qui auraient pu être produites par les nombreuses perturbations du signal EEG. Enfin,

une comparaison de moyennes (MCT) est réalisée afin de comparer l'énergie à un niveau de référence, appris au début de l'enregistrement alors que le conducteur est supposé éveillé. Comme la MCT est normalisée, un seuil commun de détection peut alors être proposé, de manière à considérer un niveau acceptable de fausses alarmes. Ce seuil sera validé par expérimentation. En parallèle, un test de comparaison de variances (VCT) est réalisé sur les données EEG brutes afin de détecter les artéfacts de haute amplitude. Cette information sur la présence d'artéfacts donne alors une indication quant à la fiabilité de la décision sur le niveau de vigilance du conducteur.

L'ensemble des différents traitements est détaillé ici. Cet algorithme a été présenté dans [Picot08] et [Picot09].

1.2.1 Pré-traitements

Spectre de puissance EEG

Le spectre du signal EEG est calculé par une transformée de Fourier à court terme (*STFT*). Le but de la STFT est d'analyser le signal par fractions supposées stationnaires à l'aide d'une fenêtre glissante. On applique ainsi la transformée de Fourier sur chaque fraction de signal de manière à avoir une représentation à la fois temporelle et fréquentielle. Dans le cas d'un signal continu, la formule de la STFT est exprimée par l'équation (1.1). Dans cette équation $x(t)$ représente le signal et $g(t - \tau)$ la fenêtre de longueur τ par laquelle on regarde le signal.

$$S_{\tau}(f) = \int_{-\infty}^{+\infty} x(t)g(t - \tau)e^{-j2\pi ft} dt \quad (1.1)$$

Cette formule s'exprime en temps discret par l'équation (1.2). Dans cette équation $x[n]$ et $g[n]$ représentent les discrétisations respectives des signaux $x(t)$ et $g(t)$. La fenêtre sur laquelle on regarde le signal a alors une longueur m .

$$D_m(f) = \sum_{n=-\infty}^{+\infty} x[n]g[n - m]e^{-j2\pi fn} \quad (1.2)$$

Le choix de la fenêtre g est très important. En effet, la longueur de cette fenêtre doit permettre de vérifier l'hypothèse de stationnarité du signal. De plus, cette longueur influe directement sur la résolution de la décomposition. Une fenêtre trop petite donnera une bonne résolution temporelle mais peu précise en fréquence. Une fenêtre trop grande aura l'effet inverse.

Nous appliquons la méthode du périodogramme de Welch ([Welch67]) pour calculer la transformée de Fourier sur la fenêtre. Le principe de la méthode de Welch est de calculer un périodogramme moyenné avec recouvrement, de manière à être plus robuste au bruit. Le signal est segmenté en plusieurs parties égales se chevauchant. Une fenêtre est alors appliquée sur chaque segment afin de calculer la transformée de Fourier (TF). Le résultat est alors la moyenne arithmétique des transformées des segments. Le principe de cette méthode, du signal temporel jusqu'à la représentation temps-fréquence, est illustré par le schéma de la figure 1.4.

Ici, la durée de la fenêtre utilisée pour la STFT est fixée à $2s$. Cette durée nous paraît adaptée à la bande étudiée : $[1-30]$ Hz tout en gardant une bonne résolution temporelle. On choisit de travailler uniquement sur la portion de bande $[1-30]$ Hz car l'activité EEG hors de cette bande est fortement sujette aux artéfacts dans des conditions normales d'enregistrement. Le spectre est calculé toutes les secondes avec un recouvrement de $1s$. La méthode de Welch est appliquée en séparant la portion de signal en onze segments de $1s$ avec 90% de recouvrement chacun. On choisit de travailler sur des segments de $1s$ afin d'observer les phénomènes fréquentiels dès 1Hz. Le choix du recouvrement est fait de manière arbitraire afin d'avoir suffisamment de mesures pour le périodogramme de Welch. La transformée de Fourier de chaque segment est calculée avec une fenêtre de Hamming.

Ainsi, la puissance relative de chaque bande est calculée comme le rapport entre la puissance dans la bande de fréquences en question et la puissance totale du spectre EEG sur la bande $[1-30]$ Hz. Par exemple, la puissance relative dans la bande α sera calculée de la manière suivante :

$$Puissance_relative(\alpha) = \frac{Puissance(\alpha)}{Puissance(EEG)} \quad (1.3)$$

Les puissances relatives des bandes α ($[8-12]$ Hz), θ ($[4-8]$ Hz) et β ($[16-26]$)Hz sont notées respectivement α_{rel} , θ_{rel} et β_{rel} . On choisit d'indexer le temps par l'indice i . Ainsi, $\alpha_{rel}(i)$ représente la puissance relative dans la bande α à l'instant i .

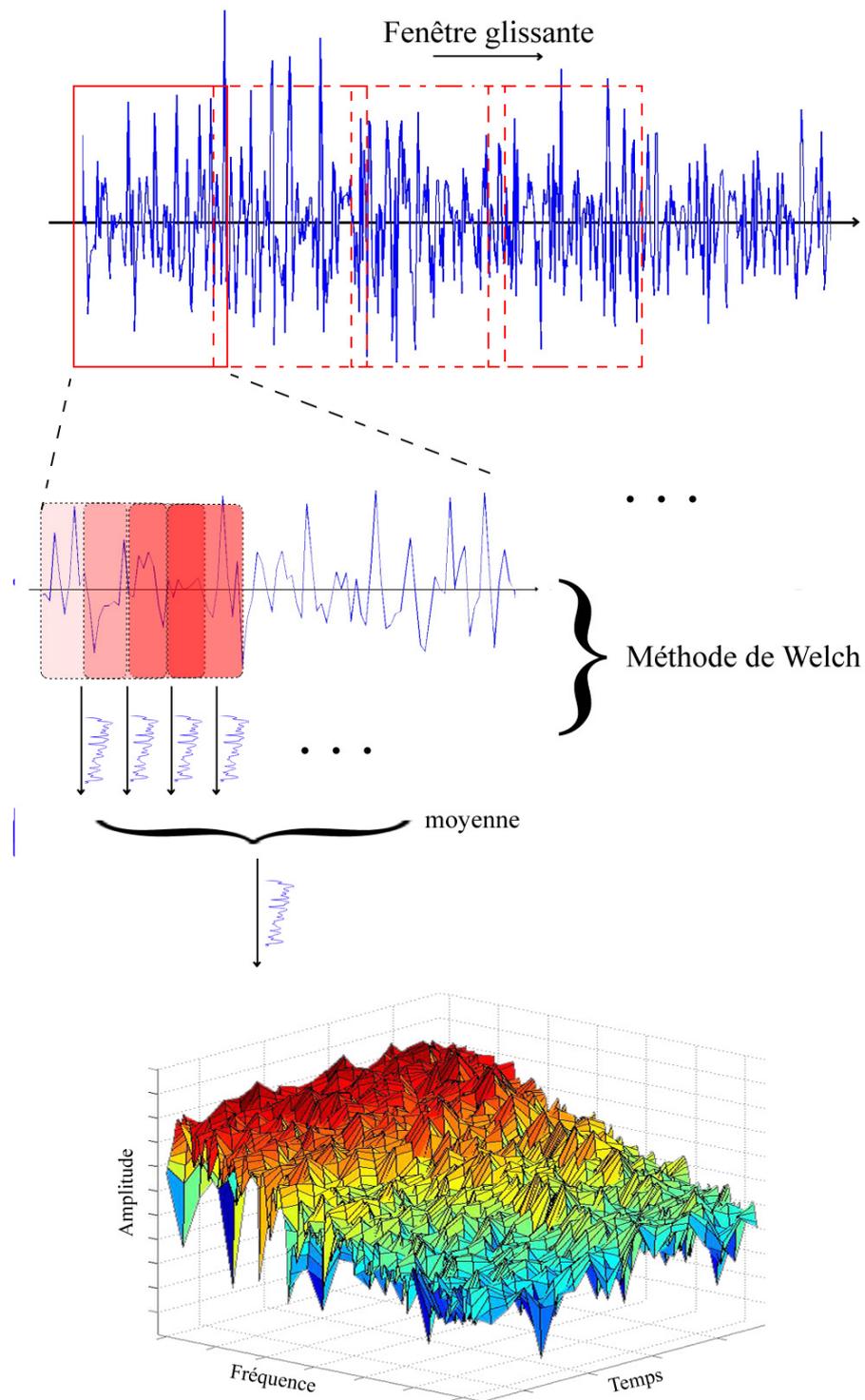


FIG. 1.4 – Transformée de Fourier avec fenêtre glissante et méthode de Welch

Filtrage médian

Un filtre médian est appliqué sur les signaux de puissance relative afin de lisser ces signaux et de rejeter les valeurs anormales, ponctuellement aberrantes. La médiane est la valeur qui divise en deux parties un échantillon telle que chaque partie contienne le même nombre de valeurs. Contrairement à la moyenne arithmétique, la valeur médiane permet d'atténuer l'influence perturbatrice des valeurs extrêmes enregistrées lors de circonstances exceptionnelles, du moment que le nombre de ces valeurs ne dépasse pas la moitié de la longueur de l'échantillon. On calcule ici la médiane des puissances relatives toutes les secondes en utilisant une fenêtre glissante de longueur $T = 10s$. Cette valeur est choisie de manière à lisser le signal sans pour autant altérer ses variations globales. Ce filtrage médian a lieu sur les puissances relatives avant d'appliquer la MCT. La valeur de la médiane de la puissance relative dans la bande α est noté m_{alpha} . Elle définie selon l'équation (1.4).

$$med_{\alpha}(i) = \text{médiane}(\alpha_{rel}(i - T + 1 : i)) \quad (1.4)$$

1.2.2 Comparaison de moyennes

La méthode de comparaison de moyennes (*MCT* pour *Means Comparison Test*) utilisée est inspirée de l'ouvrage de Ragot et al. [Ragot90] sur le diagnostic. La MCT est appliquée sur les signaux de puissance relative α_{rel} , θ_{rel} et β_{rel} après que ceux-ci aient été filtrés médians, notés respectivement m_{α} , m_{θ} et m_{β} . Le principe est de comparer une fenêtre glissante à une référence fixe comme présenté sur la figure 1.5.

La méthode de MCT classique possède des conditions très restrictives dues au fait que les variances théoriques sont inconnues. Soient deux populations indépendantes de longueur respective n_1 et n_2 , de moyenne \bar{x}_1 et \bar{x}_2 et de variance expérimentale s_1^2 et s_2^2 . Alors la variable :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (1.5)$$

suit une loi de Student à $n_1 + n_2 - 2$ degrés de liberté. L'hypothèse d'identité des espérances μ_1 et μ_2 est alors éprouvée par un test bilatéral au seuil de confiance λ :

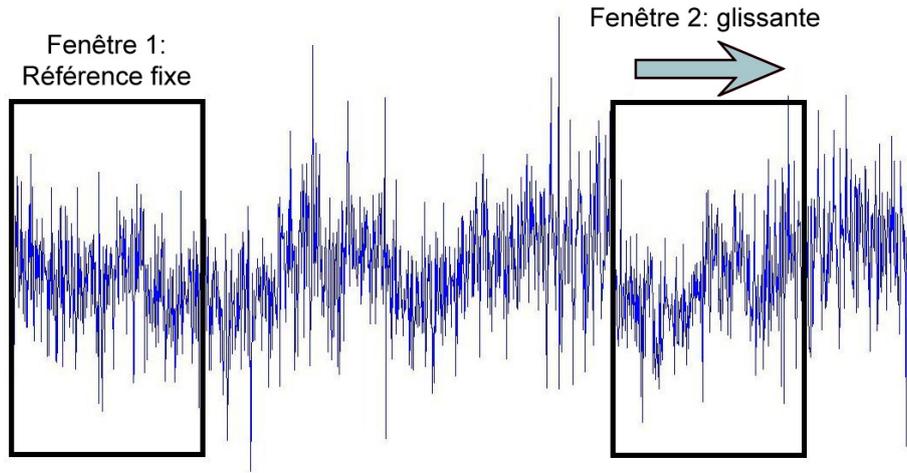


FIG. 1.5 – Illustration des fenêtres utilisées pour la MCT

$$-t_{1-\lambda/2} < t < t_{1-\lambda/2}.$$

Ce test est peu sensible à la non-normalité des distributions, du moment que celles-ci restent symétriques. Si les deux populations sont de même longueur n (i.e. $n_1 = n_2 = n$), le test devient alors également robuste à la non égalité des variances théoriques σ_1^2 et σ_2^2 . On peut alors exprimer l'équation 1.5 de la manière suivante :

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2 + s_2^2}{n-1}}} \quad (1.6)$$

Ainsi, la variable t suit une loi de Student à $n - 1$ degrés de liberté. De plus, si les populations n_1 et n_2 sont suffisamment nombreuses (au moins supérieures à 20 échantillons), on peut alors considérer que les variances expérimentales s_1^2 et s_2^2 estiment avec suffisamment de précision les variances théoriques σ_1^2 et σ_2^2 . Dans un pareil cas, le test peut alors être réalisé avec la variable :

$$u = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad (1.7)$$

qui suit une loi centrée réduite. La comparaison des moyennes est alors vérifiée par un test bilatéral de seuil λ : $-u_{1-\lambda/2} < u < u_{1-\lambda/2}$. L'égalité des variances théoriques n'est

alors plus nécessaire.

Ici, on réalise le test sur les signaux de puissance relative filtrés chaque seconde. La référence est calculée à partir d'une fenêtre fixe au début de l'enregistrement. On fait donc la supposition qu'avant de commencer à conduire, le conducteur n'est pas en état d'hypovigilance. La moyenne et la variance calculées sur cette période de temps fournissent alors une référence d'éveil et de vigilance du conducteur. Les valeurs de moyenne et de variance calculées par la suite en ligne sur la fenêtre glissante sont comparées à cette référence d'éveil afin d'évaluer le niveau de vigilance du conducteur.

La longueur utilisée pour la fenêtre de référence est $n_1 = 60s$ et celle utilisée pour la fenêtre glissante est $n_2 = 30s$. Ainsi, les populations n_1 et n_2 sont bien supérieures à 20 (car la puissance relative est calculée chaque seconde). On peut donc utiliser la formule (1.7). Le seuil λ fixe alors le pourcentage de fausses alarmes attendues. Plus le seuil λ est grand et plus bas sera le taux de fausses alarmes. En revanche, un seuil trop haut peut nous faire rater des détections. Il faudra donc trouver un compromis entre le taux de fausses alarmes toléré et le taux de bonnes détections voulu. Ici, le seuil λ sera choisi empiriquement et discuté dans la partie 1.3.3. On note ce seuil λ_{MCT} . Le résultat de la MCT dans la bande α noté MCT_α est défini selon l'équation (1.8).

$$MCT_\alpha(i) = \frac{M_\alpha(i) - M_{\acute{e}veil}}{\sqrt{\frac{\sigma_\alpha^2(i)}{l_{MCT}} - \frac{\sigma_{\acute{e}veil}^2}{L_{MCT}}}} \quad (1.8)$$

Dans cette équation $M_\alpha(i)$ et $\sigma_\alpha^2(i)$ représentent respectivement la moyenne et la variance de m_α à l'instant i calculées sur une fenêtre glissante de longueur $l_{MCT} = 30s$. Les variables $M_{\acute{e}veil}$ et $\sigma_{\acute{e}veil}^2$ représentent respectivement la moyenne et la variance de m_α calculées sur une fenêtre de référence de longueur $L_{MCT} = 60s$. Ces différents paramètres sont définis selon les équations (1.9), (1.10), (1.11) et (1.12).

$$M_\alpha(i) = \frac{1}{l_{MCT}} \sum_{j=i-l_{MCT}}^i med_\alpha(j) \quad (1.9)$$

$$\sigma_\alpha^2(i) = \frac{1}{l_{MCT} - 1} \sum_{j=i-l_{MCT}}^i (med_\alpha(j) - M_\alpha(i))^2 \quad (1.10)$$

$$M_{\acute{e}veil} = \frac{1}{L_{MCT}} \sum_{j=1}^{L_{MCT}} med_{\alpha}(j) \quad (1.11)$$

$$\sigma_{\acute{e}veil}^2(i) = \frac{1}{L_{MCT} - 1} \sum_{j=1}^{L_{MCT}} (med_{\alpha}(j) - M_{\acute{e}veil})^2 \quad (1.12)$$

1.2.3 La détection d'artéfacts

Un des problèmes majeurs de l'interprétation des signaux EEG est la présence de perturbations électriques dans le signal. Ces perturbations sont appelées *artéfacts*. Nous proposons ici une méthode pour détecter la présence d'artéfacts de haute amplitude dans le signal.

Définition

Dans le cadre de l'EEG, on appelle artéfact toute perturbation électrique du signal EEG d'origine non cérébrale. Le signal EEG est quasiment toujours perturbé par des artéfacts. L'amplitude de ces artéfacts peut être relativement élevée par rapport à l'amplitude des signaux corticaux étudiés. C'est pour cette raison que l'interprétation des signaux EEG est très délicate.

On distingue deux types d'artéfacts. Tout d'abord il y a les artéfacts biologiques. Ces artéfacts très communs sont des perturbations électriques générées par des mouvements du corps tels que les clignements d'yeux, les battements cardiaques ou les mouvements des muscles faciaux. Le second type d'artéfact est l'artéfact environnemental. Les artéfacts environnementaux sont souvent dus à des problèmes d'instrumentation tels que des faux contacts au niveau des électrodes, ou des problèmes de terre. Par exemple, il est très courant de retrouver des perturbations autour de 50Hz causées par les oscillations du courant électrique. C'est l'une des raisons, entre autres, qui fait que l'on étudie l'EEG seulement sur la bande [1-30]Hz.

Méthode de détection

La majorité des techniques de détection et de correction d'artéfacts récemment utilisées est basée sur des techniques de séparation de sources telles que l'analyse en composantes indépendantes (*ICA* pour *Independent Component Analysis*) ([Vigario97], [Joyce03]). Le principe de ces techniques est de séparer l'EEG en différentes composantes de manière à isoler les artéfacts et reconstruire le signal EEG sans ces perturbations. L'inconvénient de ce type de technique est qu'il nécessite de travailler sur un nombre important de canaux EEG pour pratiquer la séparation de sources.

Ici, on se propose de travailler uniquement sur le canal pariétal. La technique proposée pour détecter les artéfacts de haute amplitude fonctionne donc à partir d'un seul canal EEG. Les artéfacts de haute amplitude sont principalement dus à des problèmes d'électrodes. Ces artéfacts génèrent des valeurs de l'ordre du mV alors que le signal EEG est de l'ordre du μV . De plus, ils polluent toute la bande EEG comme le montre la zone rouge sur le spectrogramme de la figure 1.6. Il est alors quasiment impossible d'extraire une information fiable de l'EEG lorsqu'un artéfact se produit.

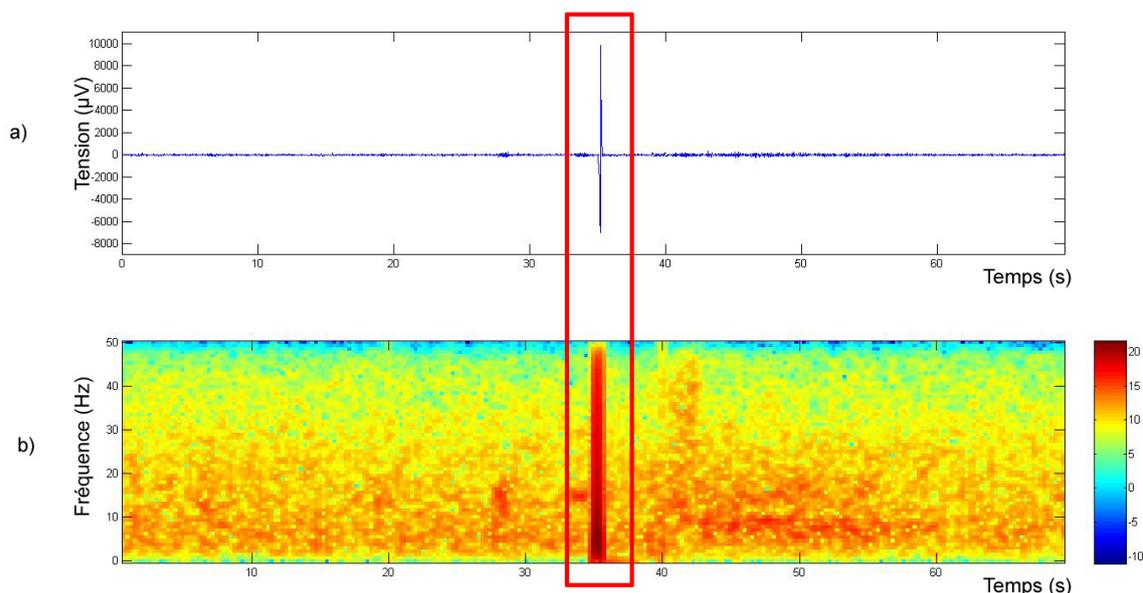


FIG. 1.6 – Exemple d'artéfact de haute amplitude sur l'EEG (a) et son spectrogramme (b)

La solution proposée est d'appliquer une comparaison de variances (*VCT* pour *Va-*

riances Comparison Test) sur le signal EEG non filtré. Là encore, cette méthode est inspirée de l'ouvrage de Ragot et al. ([Ragot90]) sur le diagnostic. Le principe est relativement similaire à celui de la MCT : la variance calculée sur une fenêtre glissante est comparée à une variance de référence calculée sur une fenêtre fixe considérée sans artéfact.

Considérons deux populations indépendantes et de distribution normale. Soient n_1 et n_2 leurs longueurs et s_1^2 et s_2^2 leurs variances expérimentales. Alors, la variable

$$F = \frac{s_1^2}{s_2^2} \quad (1.13)$$

suit une loi de Fisher à $k_1 = n_1 - 1$ et $k_2 = n_2 - 1$ degrés de liberté. L'égalité des variances peut alors être éprouvée par le test bilatéral de seuil de confiance λ_{art} : $F_{\lambda_{art}/2} < F < F_{1-\lambda_{art}/2}$.

Dans notre méthode, le test est réalisé toutes les 10s sur le signal EEG brut. La variance calculée sur une fenêtre glissante est comparée à celle calculée sur une fenêtre référence fixe, comme présenté sur la figure 1.5. La variance de référence est calculée durant la première minute de conduite, comme pour la MCT. Ensuite, la variance calculée sur une fenêtre glissante de 10s est comparée à cette référence toutes les 10s. Il n'y a donc pas de recouvrement. On a pu constater empiriquement que les artéfacts de haute amplitude sont des phénomènes ponctuels ne durant que quelques secondes. Nous avons donc choisi de travailler avec une fenêtre de 10s car cette durée nous a semblé bien adaptée pour détecter les artéfacts en question. Le seuil λ_{art} est choisi empiriquement et discuté dans la section 1.3.4. La sortie de cet algorithme est noté d_{art} . Elle est placée à "0" lorsqu'aucun artéfact n'a été détecté et à "1" lorsqu'un artéfact est détecté.

Le but de cet algorithme est de détecter la présence d'artéfacts, pas de les rejeter. En effet, nous souhaitons proposer un système qui fonctionne uniquement à partir du canal pariétal, il sera donc difficile de retrouver l'information EEG perdue au niveau de l'artéfact puisqu'il ne nous sera a priori pas possible de la récupérer ailleurs. Néanmoins, détecter la présence d'artéfacts nous apporte une information quant à la qualité du signal. Quand un artéfact est détecté, on ne prend pas en compte la puissance relative EEG sur le segment de 10s correspondant pour évaluer la vigilance du conducteur.

1.2.4 Intérêt de la méthode

L'algorithme de détection d'hypovigilance présenté peut fonctionner entièrement en temps réel. Cependant, une latence est induite par les différentes fenêtres de traitement. Le filtrage médian calculé sur 10s introduit un retard de 5s et la MCT calculée sur 30s un retard de 15s. Ainsi, la décision fournie par l'algorithme est retardée de 20s par rapport à l'acquisition des signaux. La détection d'artéfacts réalisée en parallèle avec une fenêtre glissante de 10s introduit un retard de 5s négligeable par rapport au retard de 20s engendré par la MCT. Cette latence devra être prise en compte au moment de la comparaison des résultats obtenus avec ceux de l'expert.

Le but général de l'algorithme proposé est la détection de l'hypovigilance chez le conducteur. La MCT détecte les bouffées d'ondes α chez le conducteur, indicateur d'hypovigilance comme indiqué dans la section 1.1.1. La référence est calculée sur une fenêtre fixe choisie au début du signal, en faisant l'hypothèse que le conducteur est complètement éveillé au moment de commencer à conduire. Ainsi, la moyenne calculée sur la fenêtre glissante est comparée à cette référence d'éveil. Si le test bilatéral est supérieur au seuil λ_{MCT} , le conducteur sera considéré comme "fatigué". Dans le cas inverse, il sera considéré comme "éveillé". La figure 1.7 présente le traitement réalisé sur le signal depuis l'EEG brut jusqu'à la prise de décision. Tout d'abord, la puissance relative dans la bande α (b) est calculée à partir du signal EEG (a). Le signal est ensuite lissé par un filtre médian (c). La MCT est appliquée (d) sur le signal filtré puis "seuillé" de manière à prendre la décision (e) quant au niveau d'éveil du conducteur ("0" représente un conducteur éveillé et "1" un conducteur en hypovigilance).

Des artéfacts de haute amplitude polluent le signal EEG et génèrent des valeurs aberrantes et isolées sur toute la bande EEG du spectre de puissance. En plus du filtre médian qui filtre le signal α_{rel} afin de rejeter les valeurs aberrantes pour éviter de fausses détections, une VCT est calculée sur l'EEG brut afin de détecter la présence d'artéfacts de haute amplitude. Cette détection d'artéfacts ne nous permet pas de récupérer l'information manquante aux endroits des artéfacts mais nous donne néanmoins une information quant à la qualité du signal autour de ce point. Si des artéfacts sont trouvés sur une portion du signal, les décisions quant à la vigilance du conducteur sur cette portion seront moins fiables que s'il n'y avait pas d'artéfacts.

Un des points essentiels concernant la détection de bouffées d'ondes α dans le signal EEG est la difficulté à définir des paramètres généraux, et notamment un seuil commun

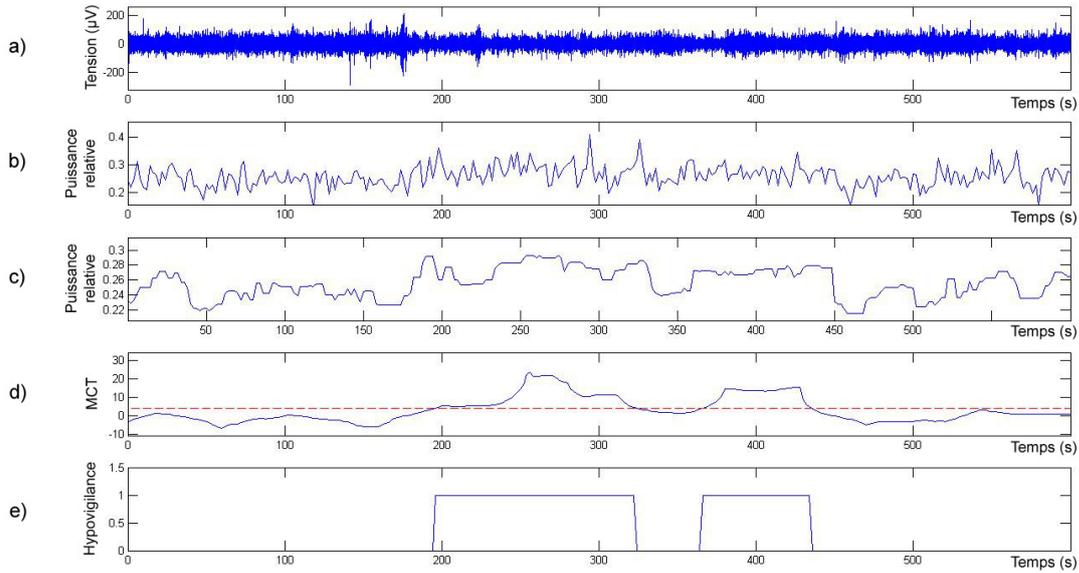


FIG. 1.7 – Traitement du signal depuis l’EEG brut jusqu’à la détection d’hypovigilance

pour tous les conducteurs à cause de fortes différences inter-conducteurs ([Karrer04]). Ici, les références d’éveil $M_{éveil}$ et $\sigma_{éveil}$ sont apprises pour chaque conducteur depuis une fenêtre de référence. De plus, comme la sortie de la MCT est une variable suivant une loi normale centrée réduite, le seuil de détection λ_{MCT} a une signification statistique et est le même pour chaque conducteur. De la même manière, la sortie de la VCT est une variable suivant une loi de Fisher. Le seuil λ_{art} utilisé pour la détection d’artéfacts de haute amplitude a donc également une signification statistique et est commun à tous les conducteurs.

1.3 Résultats et discussion

Cette partie présente les résultats obtenus à l’aide de notre méthode sur une base de données expertisée. Ces résultats seront comparés à ceux de la littérature et les différents choix de paramètres seront discutés.

1.3.1 Base de données

La base de données utilisée pour l'évaluation de notre méthode nous a été fournie par le Centre d'Études de Physiologie Appliquée (CEPA) de Strasbourg. Cette base de données a été réalisée sur le simulateur de conduite *PAVCAS* (Poste d'Analyse de la Vigilance Automobile Simulée). Ce simulateur dynamique se présente sous la forme d'une plateforme mobile à 4 degrés de liberté (mouvements longitudinal, vertical, roulis et tangage de l'habitacle), associée à une unité de visualisation interactive en temps réel. L'unité de visualisation reproduit dans de très bonnes conditions de vision un réseau autoroutier sous forme d'images de synthèse simulant les conditions diurnes, les conditions nocturnes ou les conditions de conduite dans le brouillard. Ces images sont projetées sur cinq écrans placés devant le véhicule et disposés sur un demi-cercle. Ce simulateur est présenté sur la figure 1.8.



FIG. 1.8 – Le simulateur dynamique PAVCAS

La base de données est composée de quarante enregistrements réalisés à partir de vingt sujets. Chaque sujet a été enregistré alors qu'il conduisait en conditions diurnes pendant 90 minutes, une première fois parfaitement reposé après une nuit normale de sommeil et une seconde fois en condition de privation de sommeil (le sujet ayant dormi quatre heures seulement). La base de données se compose donc de 60 heures de données de conduite. Chaque enregistrement inclut quatre canaux EEG (frontal (F3), central (C3), pariétal (P3) et occipital (O1) gauche, cf. figure 1.2), un EOG vertical ainsi que la vidéo

du conducteur depuis l'habitacle du véhicule (vidéo noir et blanc, échantillonnée à 25fps). Pour chaque session, la vigilance du conducteur a été évaluée par un médecin expert selon l'échelle d'évaluation OSS (cf. tableau 1.1 section 1.1.2). L'acquisition des signaux physiologiques (EEG et EOG) a été réalisée à 250Hz.

Ce genre de données est très difficile à obtenir. En effet, les campagnes d'acquisition sont très lourdes. L'acquisition EEG nécessite une grande précision dans le placement des électrodes. De plus, il est possible d'effectuer des expérimentations sur le sommeil humain uniquement dans des établissements agréés et spécialisés. Ces campagnes nécessitent parfois de garder les sujets une semaine complète de manière à éviter les perturbations extérieures et à assurer des conditions de sommeil équivalentes pour les différents sujets. Enfin, l'étiquetage des différents niveaux d'hypovigilance est réalisé manuellement par un spécialiste. Même si ce travail est réalisé relativement rapidement (20 minutes de travail pour l'expertise d'une heure d'enregistrement par un spécialiste entraîné), ce travail nécessite une grande expérience. La mise à disposition d'une base de données aussi complète fut une énorme opportunité pour nous. Nous remercions grandement le CEPA pour ces données sans laquelle il aurait été difficile de travailler.

1.3.2 Techniques de validation

La méthode que nous proposons fournit une décision binaire [éveillé ; fatigué] alors que la base de données a été manuellement expertisée sur cinq niveaux. De plus, l'expert classe par intervalle de 20s sans chevauchement alors que notre système prend des décisions toutes les secondes. Afin de pouvoir comparer nos résultats aux décisions expertes, la technique de validation suivante a été utilisée. Lorsqu'on analyse la première colonne de l'échelle OSS (cf. tableau 1.1) rappelée dans le tableau 1.2, on constate que les bouffées d'ondes α et θ apparaissent dès le niveau 1 d'hypovigilance.

Afin de convertir les cinq niveaux experts en une décision binaire, on considérera le conducteur comme étant dans l'état "fatigué" si la décision experte est supérieure ou égale à 1 comme présenté sur la figure 1.9. Ainsi, on garantit de détecter les états précoces d'hypovigilance et de pouvoir alerter le conducteur suffisamment tôt. La figure 1.9 présente pour un conducteur donné la décision experte à cinq niveaux (a) ainsi que sa binarisation (b). La détection d'hypovigilance obtenue par notre système est présentée en (c).

De plus, chaque portion de 20s classée par l'expert est directement comparée aux

Objective Sleepiness Score	Contenu EEG
0	Activité β continue en fond, pas d'activité α et θ
1	Apparition d'ondes α et/ou θ pendant au moins 5s cumulées
2	Apparition d'ondes α et/ou θ pendant au moins 5s cumulées ou Apparition d'ondes α et/ou θ pendant 5 à 10s cumulées
3	Apparition d'ondes α et/ou θ pendant 5 à 10s cumulées ou Apparition d'ondes α et/ou θ pendant plus de 10s cumulées
4	Apparition d'ondes α et/ou θ pendant plus de 10s cumulées

TAB. 1.2 – Critères OSS restreints au contenu EEG

décisions données par notre méthode : si durant l'intervalle de 20s, notre système classe au moins 2s dans l'état "fatigué", alors la décision sur tout l'intervalle est "fatigué". Dans le cas contraire, la décision sur l'intervalle est "éveillé". La latence engendrée par notre algorithme est évidemment prise en compte lors de la synchronisation avec les décisions expertes.

Les décisions sont ensuite comparées une à une et rangées selon la table de contingence du tableau 1.3.

		Décision experte	
		Éveillé	Fatigué
Décision automatique	Éveillé	Vrais Négatifs (TN)	Faux Négatifs (FN)
	Fatigué	Faux Positifs (FP)	Vrais Positifs (TP)

TAB. 1.3 – Table de contingence

Le taux de vrais positifs (TP_{rate} pour *True Positive Rate*) ou taux de bonnes détections est le rapport entre le nombre de bonnes détections de l'état "fatigué" et le nombre total de décisions "fatigué" prises par l'expert. Le taux de faux positifs (FP_{rate} pour *False Positive*

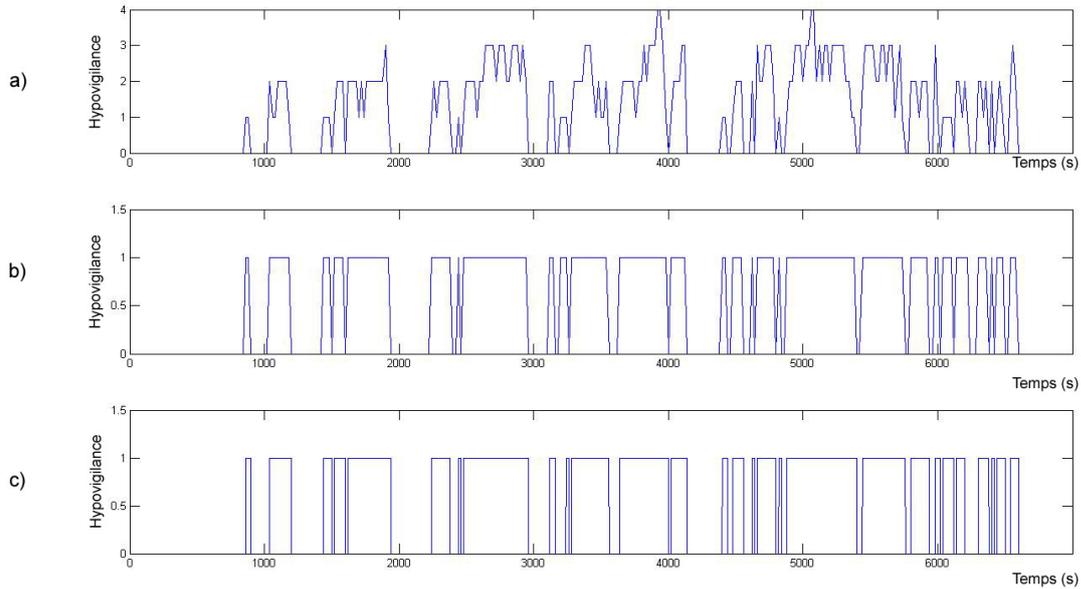


FIG. 1.9 – Comparaison entre la décision experte (a,b) et la décision de notre système (c)

Rate) ou fausses alarmes est le rapport entre le nombre de mauvaises détections “fatigué” et le nombre total de décisions “éveillé” prises par l’expert. Ces taux sont calculés selon les équations (1.14) and (1.15).

$$TP_{rate} = \frac{TP}{TP + FN} \quad (1.14)$$

$$FP_{rate} = \frac{FP}{FP + TN} \quad (1.15)$$

Les résultats sont présentés sous forme de *courbes ROC* (pour *Receiver Operating Characteristic*) ([Hanley82]). La *caractéristique de fonctionnement du récepteur* (traduction française de l’acronyme ROC) est une mesure de la performance d’un classifieur binaire, c’est-à-dire d’un système qui a pour objectif de catégoriser des entités en deux groupes distincts sur la base d’une ou plusieurs de leurs caractéristiques. Graphiquement, on représente souvent la mesure ROC sous la forme d’une courbe qui donne le taux de classifications correctes dans un groupe (dit taux de vrais positifs ou taux de détection, TP_{rate}) en fonction du nombre de classifications incorrectes (taux de faux positifs ou taux de fausses alarmes, FP_{rate}) pour ce même groupe. Le but est évidemment d’avoir le plus fort TP_{rate} et le plus faible FP_{rate} .

Nous utilisons également le test de McNemar ([McNemar47]) pour comparer les résultats des différents algorithmes proposés. Ce test permet la comparaison de deux algorithmes de classification sur un ensemble de validation S_V . Soit C_1 et C_2 deux classifieurs, on note alors :

- n_{00} le nombre d'échantillons mal classés par C_1 et C_2 ,
- n_{11} le nombre d'échantillons bien classés par C_1 et C_2 ,
- n_{01} le nombre d'échantillons mal classés par C_1 mais bien classés par C_2 ,
- n_{10} le nombre d'échantillons bien classés par C_1 mais mal classés par C_2 .

On peut représenter ces indices par la table de contingence 1.4.

		Classifieur C_2	
		Mal classés	Bien classés
Classifieur C_1	Mal classés	n_{00}	n_{01}
	Bien classés	n_{10}	n_{11}

TAB. 1.4 – Table de contingence pour le test de McNemar

Le chi-deux de McNemar $\chi_{McNemar}^2$ est alors calculé selon l'équation 1.16

$$\chi_{McNemar}^2 = \frac{(|n_{10} - n_{01}| - 1)^2}{n_{10} + n_{01}} \quad (1.16)$$

L'équivalence des performances des deux classifieurs est alors éprouvée par l'inégalité (1.17).

$$\chi_{McNemar}^2 < \chi^2(1) = 3,84 \quad (p = 0,05) \quad (1.17)$$

Si l'inégalité (1.17) est vérifiée, alors les deux classifieurs ont des performances équivalentes. Si ce n'est pas le cas, le classifieur le plus performant est celui qui classe bien le plus d'échantillons. On rappelle que la loi du chi-deux à un degré de liberté pour seuil de confiance $p = 0,05$ donne $\chi^2(1) = 3,84$.

1.3.3 Résultats à partir de la puissance relative alpha

L'algorithme de détection d'hypovigilance proposé a été appliqué sur toute la base de données, soient 60 heures de données de conduites de 20 conducteurs différents. Cela représente 5512 échantillons de 20s classés "éveillé" et 2045 classés "fatigués" classés par l'expert. On a testé notre algorithme sur chacun des quatre canaux EEG de la base de données (F3,C3,P3 et O1) en faisant varier le seuil de détection λ_{MCT} (défini dans la partie 1.2.2) de 1,5 à 5. Ces résultats sont présentés sur la figure 1.10. Ils ont été obtenus en appliquant la MCT sur la puissance relative dans la bande alpha (signal MCT_α) et sans considérer la détection d'artéfacts. Les marqueurs "étoiles" correspondent au canal P3, les "cercles" correspondent au canal F3, les "carrés" au canal C3 et les "triangles" au canal O1. La tête en bas à droite de la figure rappelle au lecteur la position de ces différents canaux. Pour chaque marqueur, les résultats qui sont les plus à droite correspondent au seuil de détection λ_{MCT} le plus bas et ceux représentés par les marqueurs les plus à gauche au λ_{MCT} le plus élevé. Ce résultat est cohérent : quand on augmente le seuil de détection, on diminue le nombre de fausses alarmes (et donc le FP_{rate}) tout en diminuant également le nombre de bonnes détections (et par conséquent le TP_{rate}).

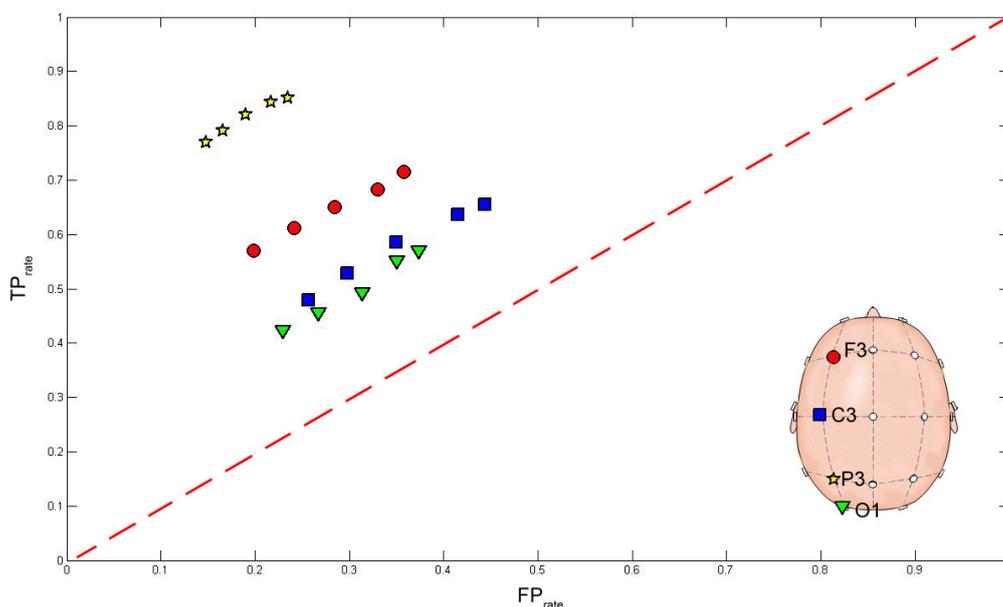


FIG. 1.10 – Résultats obtenus sur les différents canaux EEG

Il apparaît clairement sur la figure 1.10 que les résultats sont les meilleurs quand le canal pariétal P3 est utilisé. Cette constatation est en accord avec la littérature : l'hypovi-

gilance est caractérisée par une augmentation de l'activité dans la bande α principalement dans la région pariétale du cerveau. En effet, les résultats obtenus avec les autres canaux EEG enregistrés en C3, F3 et O1 sont à peine meilleurs que ceux qu'on aurait pu obtenir avec un classifieur aléatoire, représenté sur la figure par la ligne pointillée rouge. Dans la suite de cette partie, nous présenterons seulement les résultats obtenus au niveau du canal P3.

Le résultat optimal obtenu sur le canal P3 est $TP_{rate} = 82,1\%$ et $FP_{rate} = 19,2\%$ avec le seuil $\lambda_{MCT} = 3$. Cependant, en changeant λ_{MCT} on ne dégrade pas les performances ($TP_{rate} = 85,1\%$ et $FP_{rate} = 23,5\%$ pour $\lambda_{MCT} = 1,5$ et $TP_{rate} = 76,9\%$ et $FP_{rate} = 14,8\%$ pour $\lambda_{MCT} = 5$). Cela prouve que la méthode n'est pas sensible à la valeur du seuil de détection de la MCT.

1.3.4 La détection d'artéfacts

Un exemple de détection d'artéfacts est présenté sur la figure 1.11. Sur cette figure, le premier signal (a) est l'EEG brut. Le signal (b) est le résultat de la VCT. La ligne en pointillé représente le seuil $\lambda_{art} = 6$ de la détection d'artéfacts. Enfin, le signal (c) montre le résultat de la détection d'artéfacts (ligne en pointillé) : quand le signal est à "zéro", aucun artéfact de haute amplitude n'est détecté et quand il est à "un", un artéfact a été détecté.

On peut voir sur cet exemple que les artéfacts détectés correspondent bien à des perturbations électriques de haute amplitude sur le signal EEG. Comme ces artéfacts n'ont pas été recensés par un expert sur la base de données, il n'est pas possible de quantifier les performances de notre algorithme de détection d'artéfacts. Cependant, une analyse visuelle de tous les enregistrements montre que tous les artéfacts de haute amplitude apparents ont été détectés. La figure 1.12 nous présente le nombre d'artéfacts détectés dans la base de données (nombre total et pourcentage correspondant de la base de données) en fonction du seuil de la détection d'artéfacts λ_{art} (défini dans la section 1.2.3). La ligne pointillée rouge correspond au nombre effectif d'artéfacts de haute amplitude qui ont été recensés par une analyse visuelle. Cette ligne donne une indication du nombre d'artéfacts à détecter.

On peut voir sur cette figure que si le seuil de détection est trop faible ($\lambda_{art} \leq 5$), la détection d'artéfacts est trop sensible et que énormément de points sont rejetés. On détecte

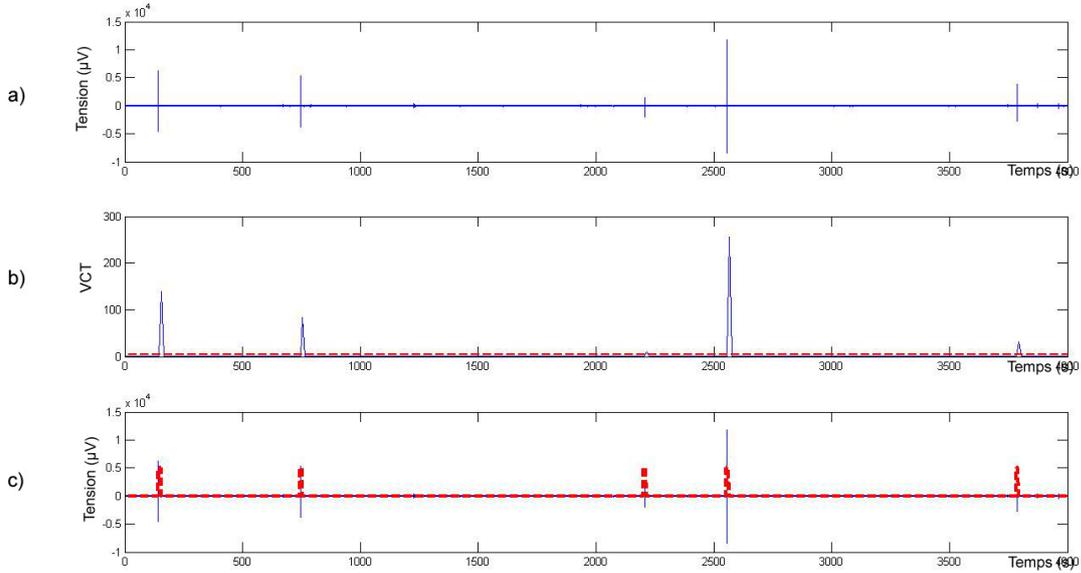


FIG. 1.11 – Exemple de détection d’artéfacts de haute amplitude

alors bien plus d’artéfacts qu’il n’y en a réellement. Visuellement, cela correspond à un nombre élevé de fausses alarmes. Quand on augmente le seuil λ_{art} , le nombre d’artéfacts détectés décroît rapidement jusqu’à $\lambda_{art} = 6$, qui correspond visuellement à un bon seuil de détection. En effet, pour cette valeur de seuil, tous les artéfacts de haute amplitude apparents sont détectés sans fausses alarmes. On vérifie visuellement que les artéfacts détectés par notre algorithme correspondent bien aux artéfacts de haute amplitude. Dans ce cas, on remarque que les artéfacts de haute amplitude représentent une faible part de la base de données : environ 2%.

Regardons à présent les résultats obtenus en prenant en compte cette détection d’artéfacts sur la figure 1.13. Ces résultats ont été obtenus en utilisant le signal MCT_α sur le canal P3. Sur cette figure, les “cercles” correspondent aux résultats obtenus en considérant uniquement les décisions prises sur des segments où aucun artéfact n’a été détecté. Le seuil de détection d’artéfacts utilisé est $\lambda_{art} = 6$. Les “carrés” représentent les résultats de la détection d’hypovigilance obtenus sans prendre en considération la détection d’artéfacts, tels qu’ils ont été présentés dans la section 1.3.3.

Il apparaît clairement sur la figure 1.13 que les résultats sont améliorés lorsque l’on prend la détection d’artéfacts en considération : le TP_{rate} augmente un peu alors que le FP_{rate} diminue légèrement. Si on utilise le seuil $\lambda_{MCT} = 3$, les résultats passent de

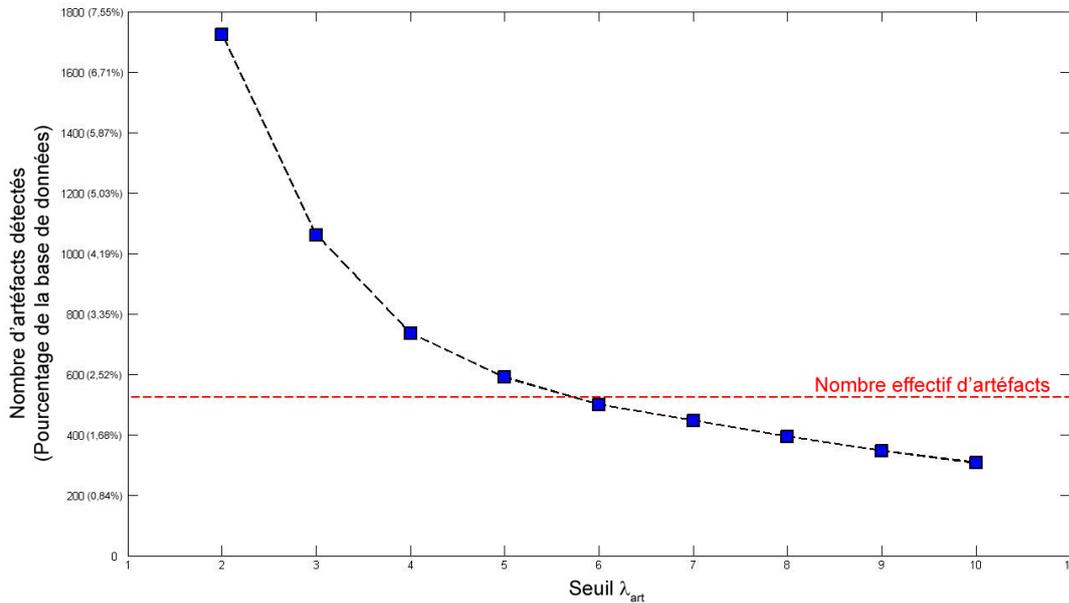


FIG. 1.12 – Nombre d’artefacts détectés en fonction du seuil λ_{art}

$TP_{rate} = 82,1\%$ et $FP_{rate} = 19,2\%$ à $TP_{rate} = 82,4\%$ et $FP_{rate} = 18,3\%$. Il est donc clair que la détection d’artefacts permet de diminuer le nombre de fausses alarmes. Le fait que les résultats sont seulement légèrement améliorés peut s’expliquer par le fait que les artefacts de haute amplitude représentent seulement 2% de la base de données.

La détection d’artefacts sera prise en compte dans toute la suite de cette partie.

1.3.5 Résultats à partir d’autres indices

Les résultats présentés dans la section 1.3.4, améliorés par la détection d’artefacts, vont maintenant être comparés aux résultats obtenus à partir d’autres indices, trouvés dans la littérature. Ces résultats sont présentés sur la figure 1.14. Sur cette figure, les résultats obtenus à partir de MCT_{α} sont représentés par les marqueurs “étoiles”. Il s’agit des résultats présentés dans les sections précédentes. Les “carrés” et “cercles” représentent les résultats obtenus avec MCT_{θ} et MCT_{β} . À noter que comme l’activité β est caractéristique de sollicitation cognitive et augmente avec cette dernière, l’hypovigilance devrait être marquée par une baisse de l’activité β . Ainsi, l’algorithme de détection appliqué sur MCT_{beta} considérera le conducteur comme “fatigué” quand la sortie de la MCT sera in-

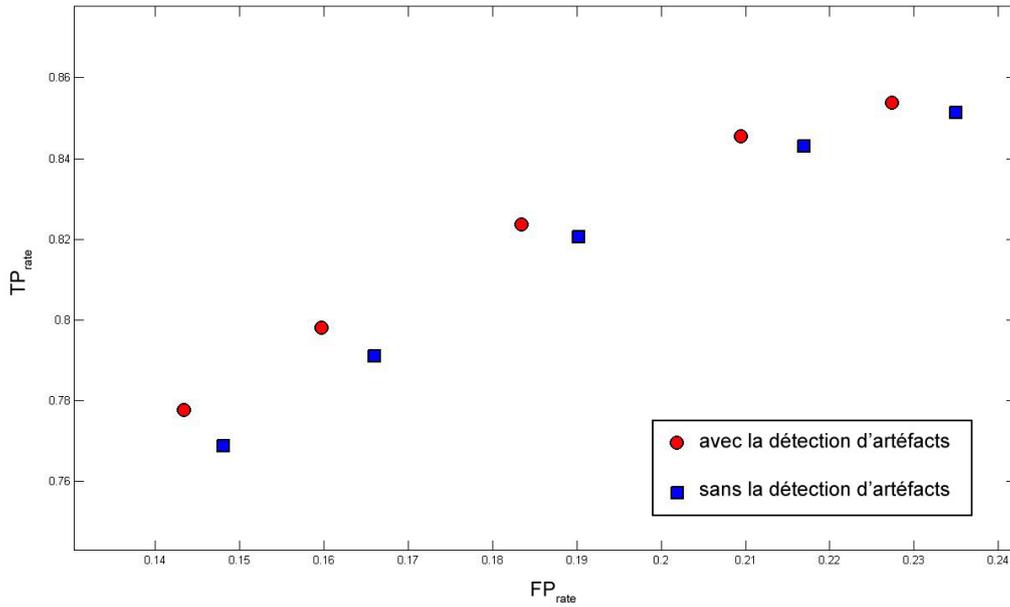


FIG. 1.13 – Courbes ROC de la détection d'hypovigilance lorsque la détection d'artefact est utilisée

férieure au seuil de détection $-\lambda_{MCT}$ variant de -5 à -1,5. Les “triangles” représentent les résultats obtenus par la combinaison des signaux $MCT_\alpha|MCT_\theta$. Dans ce cas les décisions sont prises indépendamment sur MCT_α et MCT_θ puis elles sont fusionnées à l'aide d'un “OU logique”. Le seuil optimum ($\lambda_{MCT} = 3$) trouvé précédemment a été utilisé pour la détection sur MCT_α . Les résultats présentés correspondent aux variations du seuil de détection utilisé sur le signal MCT_θ et variant de 1,5 à 5. L'idée d'utiliser à la fois les signaux extraits des bandes α et θ est inspirée du tableau 1.1. En effet, ce tableau relate que l'hypovigilance est caractérisée par une hausse de l'activité électrique dans chacune des deux bandes de fréquence α et θ . Les “hexagones” représentent les résultats obtenus en utilisant l'indice $\frac{\alpha+\theta}{\beta}$ en entrée de la MCT. Cet indice a été suggéré par De Waard & Brookhuis ([DeWaard91]). En effet, étant donnée que l'activité des bandes α et θ est censée augmenter avec l'hypovigilance alors que l'activité de la bande β décroît, ce paramètre devrait augmenter avec la baisse de vigilance. Dans ce cas, la MCT est réalisée directement sur le rapport entre la somme des signaux α_{rel} et θ_{rel} et le signal β_{rel} . Ainsi, un seul seuil est utilisé pour prendre les décisions.

On constate sur cette figure que les meilleurs résultats restent ceux obtenus en appliquant notre algorithme sur le signal MCT_α ($TP_{rate} = 82,4\%$ et $FP_{rate} = 18,3\%$ pour

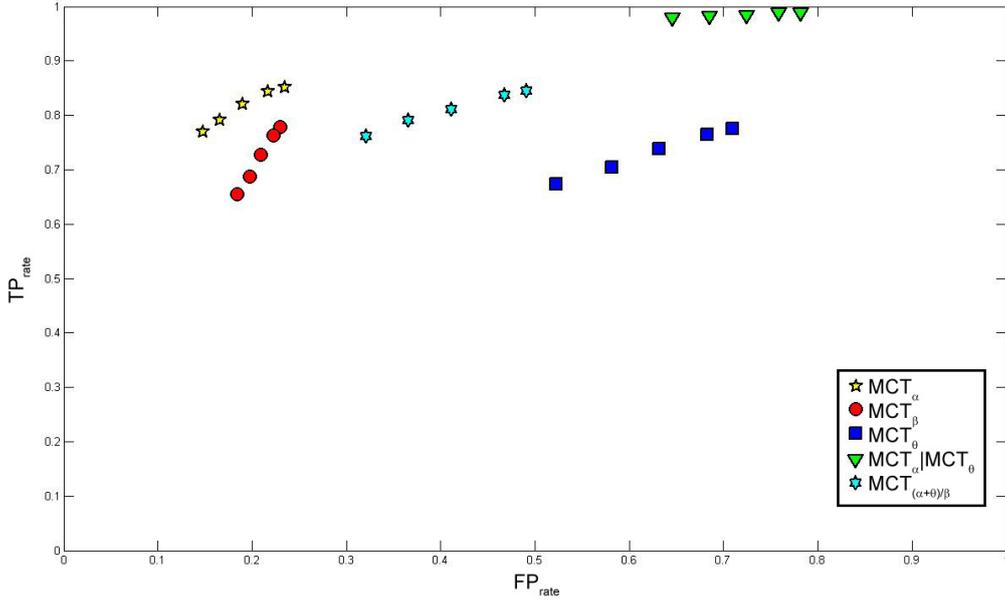


FIG. 1.14 – Courbes ROC obtenues en utilisant différents indices pour la détection d'hypovigilance

$\lambda_{MCT} = 3$). Comme notre algorithme a été testé avec le même seuil sur une base de données regroupant 20 sujets différents, cela montre que notre méthode peut être appliquée sur différentes personnes sans avoir à adapter ses paramètres.

Les résultats obtenus en utilisant l'indice $MCT_\alpha|MCT_\theta$ montrent que le signal MCT_θ n'apporte pas d'information pertinente quant à la détection d'hypovigilance. En effet, quand cette information est fusionnée avec MCT_α , le nombre de fausses alarmes augmente considérablement. Cela est confirmé par les résultats obtenus avec le signal MCT_θ uniquement. Les résultats obtenus avec le ratio $\frac{\alpha+\theta}{\beta}$ ainsi qu'avec le signal MCT_β seul sont corrects ($TP_{rate} = 76,2\%$ et $FP_{rate} = 32,1\%$ pour $MCT_{\frac{\alpha+\theta}{\beta}}$ et $TP_{rate} = 75,9\%$ et $FP_{rate} = 24,1\%$ pour MCT_β), mais ils restent cependant moins bons que ceux obtenus en se servant uniquement de l'information MCT_α .

1.3.6 Résultats de la fusion des informations α et β

On a pu voir dans la section précédente que les meilleurs résultats étaient ceux obtenus avec les signaux MCT_α et MCT_β . Naturellement, cela nous incite à fusionner ces

informations de manière à essayer d’augmenter encore la fiabilité de notre détecteur d’hypovigilance. Nous décidons donc d’utiliser la logique floue pour fusionner les informations MCT_α et MCT_β . Cette technique est basée sur la théorie des ensembles flous de Zadeh ([Zadeh65]) qui est développée plus en détails dans l’annexe B.

Soient $\mu_{Dr}(MCT_\alpha)$ et $\mu_{Dr}(MCT_\beta)$ deux fonctions d’appartenance représentant le degré d’appartenance des signaux MCT_α et MCT_β à la sortie de la comparaison de moyennes, considérés indépendants, à la classe “fatigué”. Le but est de prendre une décision $Dr(MCT_\alpha, MCT_\beta)$ en utilisant à la fois $\mu_{Dr}(MCT_\alpha)$ et $\mu_{Dr}(MCT_\beta)$. Le conducteur est considéré comme étant en état d’hypovigilance si les décisions faites sur MCT_α et MCT_β sont toutes deux “fatigué”. Cela peut être exprimé par le produit suivant :

$$Dr(MCT_\alpha, MCT_\beta) = \frac{\mu_{Dr}(MCT_\alpha)\mu_{Dr}(MCT_\beta)}{\mu_{Dr}(MCT_\alpha)\mu_{Dr}(MCT_\beta) + \mu_{Aw}(MCT_\alpha)\mu_{Aw}(MCT_\beta)} \quad (1.18)$$

Il faut noter que $\mu_{Aw}(\cdot)$ est la fonction d’appartenance à la classe “éveillé” et donc le complémentaire de $\mu_{Dr}(\cdot)$. Elle est définie selon l’équation (1.19). Le dénominateur est utilisé ici afin de normaliser $Dr(\alpha_{rel}, \beta_{rel})$ entre 0 et 1.

$$\mu_{Aw}(MCT_\alpha) = 1 - \mu_{Dr}(MCT_\alpha) \quad (1.19)$$

Il faut maintenant définir les fonctions d’appartenance $\mu_{Dr}(MCT_\alpha)$ et $\mu_{Dr}(MCT_\beta)$. Une étude de la probabilité d’être dans l’état “fatigué” en fonction du seuil λ_{MCT} sur MCT_α et MCT_β , $P(Dr|MCT_\alpha)$ et $P(Dr|MCT_\beta)$, est présentée sur la figure 1.15. La ligne pointillée avec les marqueurs “carrés” représente la probabilité expérimentale $P(Dr|MCT_\alpha)$ et celle avec les marqueurs “cercles” la probabilité expérimentale $P(Dr|MCT_\beta)$. Ces probabilités représentent le pourcentage de décisions “fatigué” correctes obtenu sur les parties du signal au dessus du seuil λ_{MCT} .

Les fonctions d’appartenance $\mu_{Dr}(MCT_\alpha)$ et $\mu_{Dr}(MCT_\beta)$ sont alors définies à partir des résultats de la figure 1.15. Comme $P(Dr|MCT_\alpha)$ et $P(Dr|MCT_\beta)$ ont des comportements relativement similaires, la fonction d’appartenance sera utilisée pour α_{rel} et $-\beta_{rel}$. Le degré d’appartenance est alors calculé sur MCT_α et MCT_β à l’aide de la fonction μ_{Dr} définie par l’équation (1.20) et représentée sur la figure 1.16.

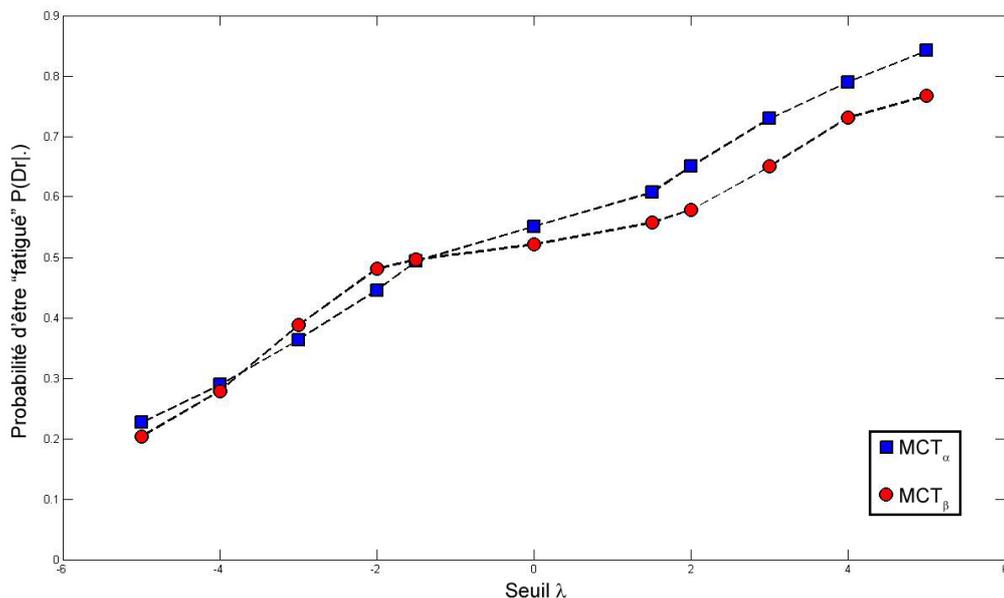


FIG. 1.15 – Probabilité expérimentale d'être "fatigué" en fonction du seuil λ

$$\mu_{Dr}(x) = \begin{cases} 0 & si & x < -5 \\ \frac{1}{2} \cdot \frac{x-a'}{b'-a'} & si & -5 \leq x < -1,5 \\ 0,5 & si & -1,5 \leq x \leq 1,5 \\ 0,5 + \frac{1}{2} \cdot \frac{x-a''}{b''-a''} & si & 1,5 < x \leq 5 \\ 1 & si & x > 5 \end{cases} \quad (1.20)$$

$$\begin{aligned} \text{avec} \quad a' &= -5 \\ b' &= -1,5 \\ a'' &= 1,5 \\ b'' &= 5 \end{aligned}$$

Le conducteur sera alors considéré comme "fatigué" si le résultat de l'équation (1.18) $Dr(MCT_\alpha, MCT_\beta)$ est supérieur à 0,5. Les résultats obtenus avec cette méthode sont présentés sur la figure 1.17 par le marqueur "cercle". Ils sont comparés aux résultats obtenus avec MCT_α seulement (marqueurs "carrés") ainsi qu'avec MCT_β (marqueurs "triangle"). Ces résultats sont calculés seconde par seconde et comparés par groupes de 20s aux décisions de l'expert comme présenté précédemment.

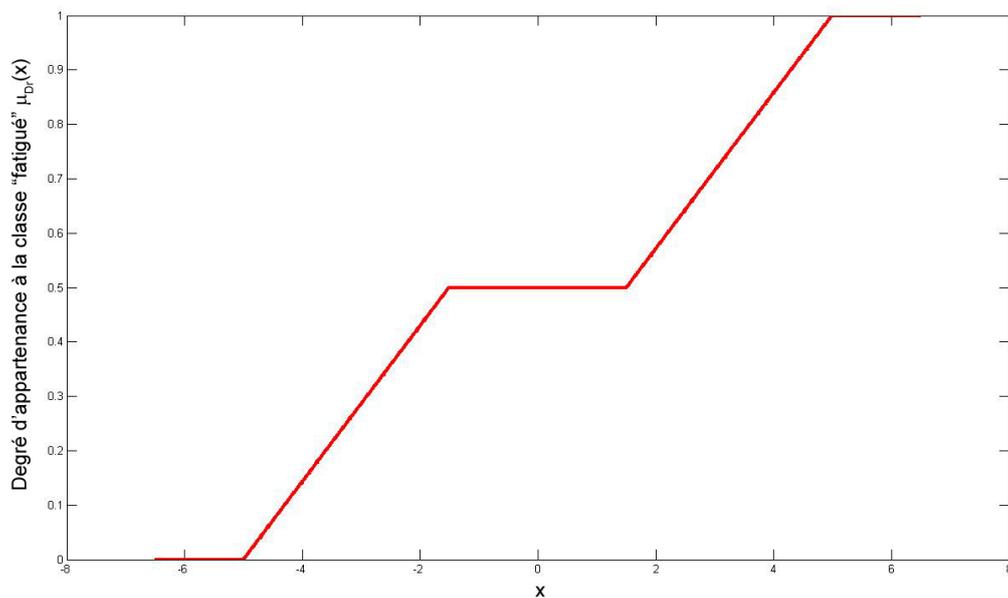
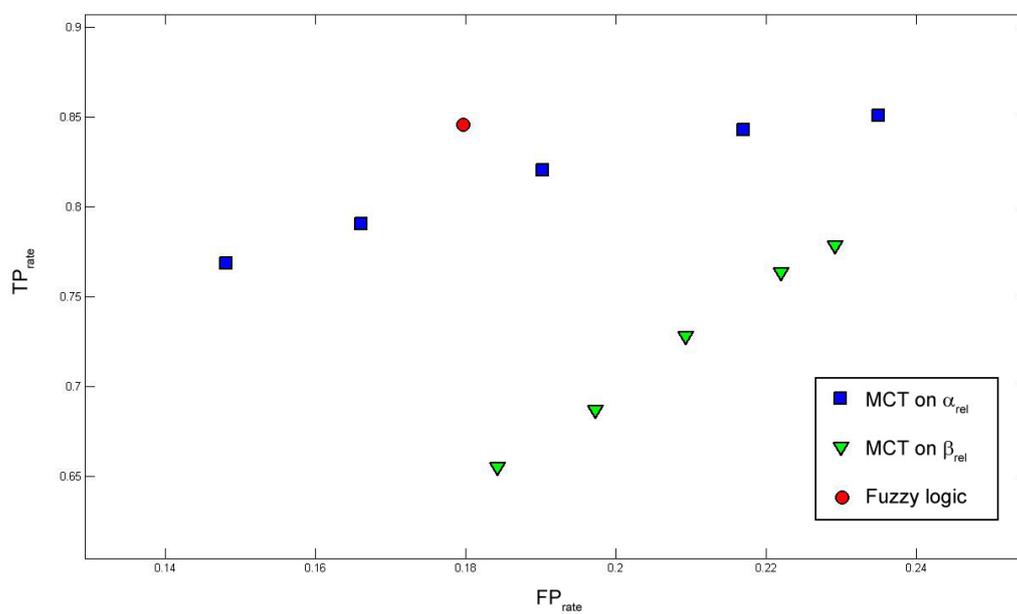


FIG. 1.16 – Fonction d'appartenance à la classe fatiguée

FIG. 1.17 – Performances obtenues en utilisant α_{rel} , β_{rel} et $(\alpha_{rel}, \beta_{rel})$ fusionnés par logique floue

Il apparaît sur cette figure que les résultats sont améliorés par la fusion des informations

α_{rel} et β_{rel} . En effet le TP_{rate} est augmenté à $TP_{rate} = 84,6\%$ alors que le FP_{rate} est diminué à $FP_{rate} = 17,9\%$ (rappelons que les résultats étaient $TP_{rate} = 82,4\%$ et $FP_{rate} = 18,3\%$ avec $\lambda_{MCT} = 3$ en utilisant seulement l'information MCT_{α}).

Afin de comparer les performances obtenues par fusion aux performances obtenues à partir de MCT_{alpha} uniquement, nous procédons à un test de McNemar. La table de contingence de la comparaison entre la détection par fusion et la détection utilisant seulement MCT_{alpha} pour $\lambda_{MCT} = 3$ est donnée dans le tableau 1.5.

		Détection à partir de MCT_{alpha} ($\lambda_{MCT} = 3$)	
		Mal classés	Bien classés
Détection par fusion	Mal classés	1201	0
	Bien classés	68	6188

TAB. 1.5 – Table de contingence pour la comparaison entre la détection par fusion avec celle utilisant MCT_{alpha} ($\lambda_{MCT} = 3$)

Le calcul du chi-deux de McNemar donne $\chi_{McNemar}^2 = 66,01$. On constate donc que $\chi_{McNemar}^2 \gg \chi^2(1)$. On en déduit que les performances des deux algorithmes ne sont pas équivalentes. En effet, on remarque dans le tableau 1.5 qu'il y a beaucoup plus d'échantillons bien classés par la détection par fusion et mal classés par la détection sur MCT_{α} ($n_{10} = 68$) que l'inverse ($n_{01} = 0$). On en conclut que la détection par fusion est effectivement plus performante.

Cela montre la pertinence de l'information MCT_{β} pour détecter l'hypovigilance, lorsqu'elle est associée à MCT_{α} . De plus, comparativement à la méthode présentée dans la section 1.3.3, il n'est plus nécessaire de choisir un seuil de détection approprié sur MCT_{α} et MCT_{β} . La fusion par logique floue a donc amélioré la fiabilité de notre méthode de détection d'hypovigilance.

Pour la suite, on note $d_{EEG}(i)$ la décision d'hypovigilance réalisée à partir de la fusion des informations MCT_{α} et MCT_{β} à l'instant i . Cette variable vaut "0" lorsque le conducteur est "éveillé" et "1" lorsqu'il est "fatigué", comme défini dans l'équation 1.21.

$$d_{EEG}(i) = \begin{cases} 0 & \text{si } Dr(MCT_{\alpha}(i), MCT_{\beta}(i)) \leq 0,5 \\ 1 & \text{si } Dr(MCT_{\alpha}(i), MCT_{\beta}(i)) > 0,5 \end{cases} \quad (1.21)$$

1.3.7 Comparaison à la littérature

Si on compare maintenant nos résultats à ceux trouvés dans la littérature, on constate que notre méthode est très performante. Il faut cependant garder une certaine réserve car les résultats présentés ici n'ont pas été obtenus sur les mêmes données. Lin et al. ([Lin05a]) ont proposé de surveiller les performances de conduite, c'est-à-dire la capacité du conducteur à maintenir son véhicule au milieu de la route, en utilisant un modèle de régression linéaire sur un EEG de deux canaux. Lorsque le modèle est entraîné et testé sur la même session, ils obtiennent une corrélation de $r = 0,88$ entre leur modèle et les performances de conduite. Cependant ce résultat retombe à $r = 0,7$ lorsque le modèle est entraîné et testé sur des sessions différentes. Ainsi, on constate que cette méthode nécessite d'être paramétrée pour chaque conducteur puisque le modèle estimé sur un conducteur ne fonctionne pas forcément très bien sur un autre. Lin et al. ont réussi à améliorer ces résultats en utilisant une ICA sur un EEG de 33 canaux ([Lin05b]) avant d'appliquer leur modèle de régression linéaire. Ils obtiennent alors une corrélation de $r = 0,88$ entre leur estimation et les performances réelles de conduite sur la session test. Néanmoins, deux réserves peuvent être émises à cette méthode. Elle nécessite un très grand nombre de données et de canaux pour être entraînée et elle n'a été testée que sur cinq sujets seulement. Ben Khalifa et al. ([BenKhalifa04]) obtiennent 92% de bonnes détections d'hypovigilance en entraînant un réseau de neurones sur le spectre de l'EEG enregistré sur le canal P4-O2. Cependant ce bon résultat est obtenu sur les données d'apprentissage du modèle et retombe à 76% de bonnes détections sur les données de validation. De plus, là encore la méthode n'a été testée que sur un nombre limité de conducteurs. Trejo et al. ([Trejo05]) obtiennent un impressionnant 97% de bonnes détections avec leur modèle de régression par moindres carrés partiels. Leur technique nécessite cependant 30 canaux EEG. Enfin, Rosipal et al. ([Rosipal07]) arrivent à 77% de bonnes détections des états d'hypovigilance en utilisant un hGMM sur le contenu spectral d'un EEG converti sous la forme de coefficients d'un modèle autorégressif. Cette étude a été menée sur un nombre conséquent de conducteurs mais nécessite elle aussi une longue période d'entraînement. Concernant les méthodes sans entraînement et indépendantes du conducteur, Pal et al. ([Pal08]) obtiennent une corrélation de $r = 0,78$ entre leurs indices obtenus par la distance de Mahalanobis et les performances de conduite du conducteur. Une synthèse de ces

résultats est présentée dans le tableau 1.6.

Auteurs	Résultats	Technique	Nombre de canaux	Nombre de sujets
Picot et al., 2009	82,5%	MCT et fusion floue sur α et β	1	20
Ben Khalifa et al., 2004	76%	Réseaux de neurones	1	4
Lin et al., 2005	$r=0,88$	Régression linéaire	33	10
Rosipal et al., 2007	76%	Gaussian Mixture Model	3	32
Pal et al., 2008	$r=0,88$	Distance de Mahalanobis sur α et θ	1	13

TAB. 1.6 – Synthèse de la comparaison de nos résultats à ceux de la littérature concernant la détection de l'hypovigilance à partir des signes cérébraux

L'avantage de notre méthode, en plus d'un bon taux de détection des états d'hypovigilance, est qu'elle n'a pas besoin d'être adaptée ou entraînée. Le seuil utilisé pour la détection reste le même pour chaque conducteur. Notre méthode a le grand avantage de fonctionner à partir d'un seul canal EEG. Le détecteur d'artéfacts permet de déterminer la qualité du canal afin d'éviter les fausses détections dues à des perturbations. De plus les résultats sont fiables car la méthode a été testée sur un nombre conséquent de conducteurs.

1.4 Conclusion du chapitre

On a pu constater dans ce chapitre que le problème de l'hypovigilance n'est pas un problème simple. En effet, du fait que l'hypovigilance est une transition entre les deux états que sont l'éveil et le sommeil, il est délicat de définir des caractéristiques précises. Malgré cela, des caractéristiques physiologiques de l'hypovigilance ont été dégagées, donnant lieu à des règles de notation de différents niveaux de vigilance, bien que ces règles ne soient pas standardisées et restent relativement subjectives.

Un algorithme de détection en ligne d'hypovigilance chez le conducteur à partir d'un seul canal EEG a été présenté dans ce chapitre. Cet algorithme est basé sur un test

de comparaison de moyennes appliqué aux puissances relatives des bandes α et β de l'EEG. Les résultats obtenus sur ces deux canaux sont ensuite fusionnés en utilisant la logique floue. Cet algorithme, appliqué en ligne, génère une latence de 20s. Il ne nécessite aucun réglage. Les performances obtenues sur une base de données réalisée à partir de 20 conducteurs différents sont **84,6% de bonnes détections et 17,9% de fausses alarmes en utilisant uniquement le canal pariétal P3 de l'EEG**. Un système de détection des artéfacts de haute amplitude est également présenté. Il est combiné au détecteur d'hypovigilance de manière à détecter en temps réel les périodes de temps où le signal EEG n'est pas fiable (car pollué par des artéfacts). Le détecteur d'artéfacts utilise un seuil unique qui est indépendant du conducteur.

On a pu voir dans la section 1.1.1 que l'hypovigilance pouvait être caractérisée par l'EEG afin de détecter des signes de défaillance de la capacité à traiter l'information et par l'étude des clignements afin de détecter les signes de baisse des capacités perceptives. Il semble maintenant judicieux d'étudier comment l'hypovigilance peut être caractérisée au niveau des clignements d'yeux. La fusion des deux approches devrait permettre d'obtenir un système de détection d'hypovigilance complet.

Chapitre 2

Caractérisation vidéo des clignements : comparaison avec l'EOG

D'après l'échelle OSS de caractérisation de l'hypovigilance, il apparaît que les clignements sont des indices visuels très pertinents. À l'heure actuelle, les informations relatives aux clignements (tels que fréquence, durée, etc) sont extraites à partir de l'analyse de signaux EOG dans la plupart des systèmes existants. Le but de ce chapitre est de voir dans quelle mesure une caractérisation des clignements est possible à partir de l'analyse d'un signal vidéo en lieu et place des signaux EOG. En effet, l'acquisition d'une vidéo du visage est moins gênante pour le conducteur que les acquisitions EOG. Nous commençons par une présentation des indices extraits sur les signaux EOG relatifs aux clignements dans le but de caractériser l'hypovigilance. Ensuite, nous examinons en quelle mesure et avec quelle précision ces mêmes indices peuvent être obtenus à partir d'un signal vidéo. En particulier, nous verrons qu'il est nécessaire d'envisager l'utilisation d'une caméra rapide (d'une fréquence d'acquisition supérieure à 25Hz). Nous terminons par la comparaison entre la caractérisation des clignements par EOG et la caractérisation des clignements par analyse vidéo. Une méthode originale de détection et de caractérisation des clignements à partir d'un signal vidéo est proposée ici.

2.1 État de l'art

Dans cette partie, nous définissons ce qu'est l'EOG et son utilité quant à la caractérisation de la baisse de vigilance. Après une revue des différentes techniques pour caractériser la baisse de vigilance à travers les indices visuels, nous verrons la nécessité de développer une approche par caméra rapide.

2.1.1 Hypovigilance et EOG

Même si l'EOG ne semble pas la méthode la plus intuitive pour analyser les signes visuels (on pourrait d'abord penser à utiliser la vidéo), cette méthode est très utilisée par les médecins spécialistes de l'hypovigilance car elle permet une très grande précision quant à la détection et la caractérisation des clignements.

Definition de l'EOG

L'*électrooculographie* (EOG) est la mesure de l'activité électrique de l'œil. Ainsi, l'EOG peut servir à relever les mouvements de l'œil et les clignements. L'œil peut être considéré comme un dipôle électrique, le pôle positif étant la cornée et le pôle négatif la rétine. L'EOG correspond en fait à la mesure de la différence de potentiel électrique entre ces deux pôles, qui peut varier entre 0,4 et 1mV. Lorsque les yeux sont au repos, l'EOG mesure une tension constante, appelée ligne de base. Lorsque les yeux bougent, les pôles s'éloignent ou se rapprochent des électrodes, ce qui provoque un changement du potentiel mesuré comme le montre la figure 2.1. Le signe du changement dépend de la direction du mouvement. De plus, l'amplitude de la mesure est proportionnelle à l'amplitude du mouvement de l'œil ([Andreassi00]).

On mesure l'EOG en plaçant des électrodes autour des yeux. Il est nécessaire de pouvoir différencier les mouvements horizontaux et verticaux de l'œil car les clignements sont des mouvements uniquement verticaux des paupières. On place donc les électrodes verticalement d'un côté et de l'autre de l'œil pour l'EOG vertical (EOG-V). Pour l'enregistrement de l'EOG horizontal (EOG-H), on place les électrodes sur les cotés extérieurs des yeux. On place également une électrode au milieu du front qui servira de référence ([Marmor93]). Le schéma du placement des électrodes est présenté sur la figure 2.2. Par

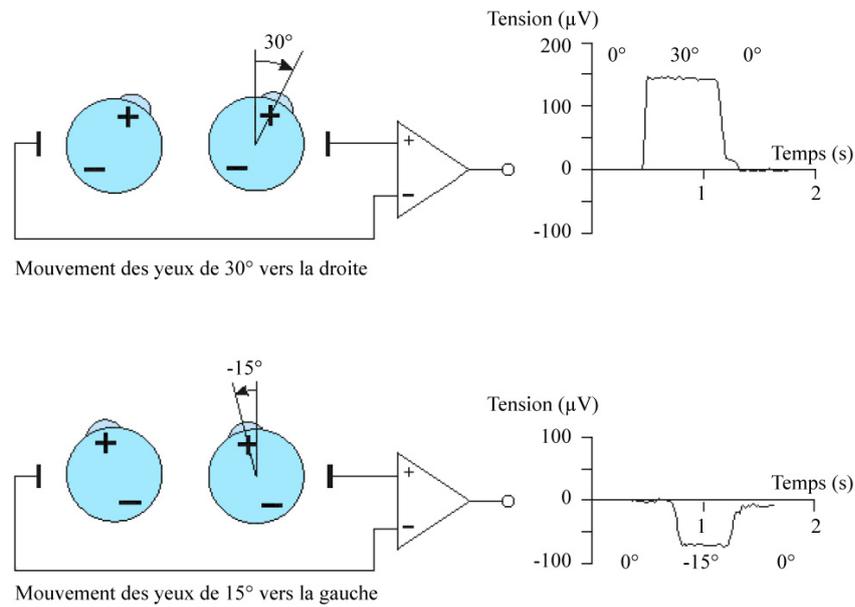


FIG. 2.1 – Changements du potentiel EOG liés aux mouvements des yeux

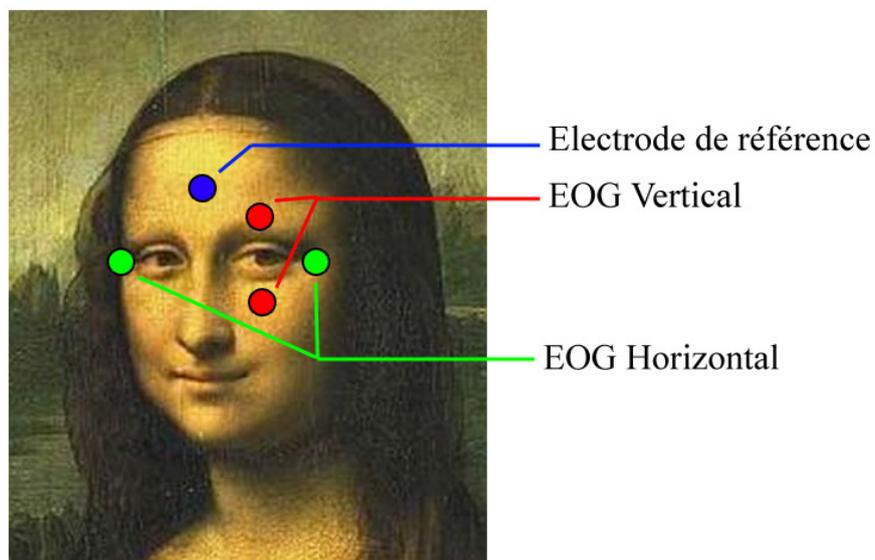


FIG. 2.2 – Placement des électrodes pour l'enregistrement EOG

la suite nous ne nous occuperons que de l'EOG vertical.

Enfin, afin d'avoir la meilleure résolution possible, les enregistrements sont réalisés avec une fréquence d'échantillonnage élevée (dans notre cas $f_e = 250Hz$). L'enregistrement de

la composante continue (DC) de l'EOG permet la prise en compte de très longs clignements. Cependant, il s'accompagne d'un problème de dérive lente de la ligne de base, ce qui peut provoquer la saturation de l'amplificateur ([Marmor93]). L'enregistrement en mode courant alternatif (AC) corrige ce problème de dérive mais filtre la composante continue, ce qui rend la détection de longs clignements délicate. L'utilisation de ces modes d'enregistrement dépend de l'application.

Les paramètres de l'EOG

D'après plusieurs études ([Santamaria87, Renner97, Galley04]), l'analyse des clignements est un bon indicateur pour caractériser l'hypovigilance. Selon Renner & Merhing ([Renner97]), la dynamique oculaire (clignements et fermeture des yeux) est révélatrice de la dégradation de nos capacités perceptives. On définit le clignement comme la fermeture d'un œil durant moins de 0,5s. La durée minimale d'un clignement est de l'ordre de 150ms. Les clignements peuvent être caractérisés par les différents paramètres présentés figure 2.3 :

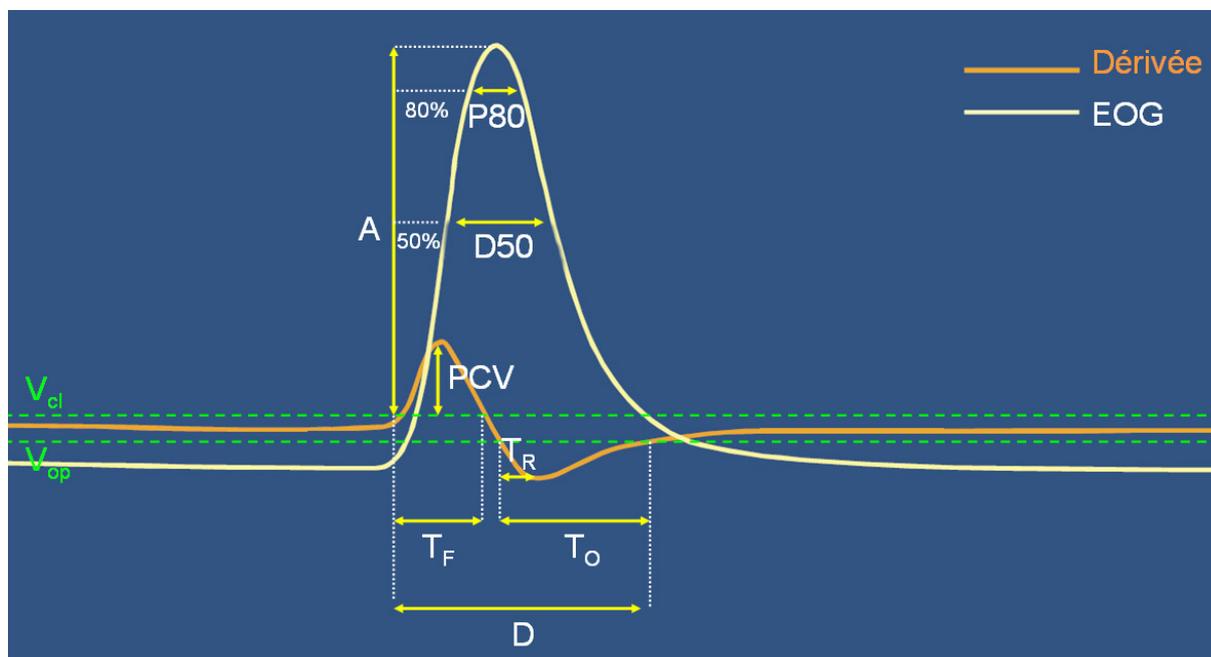


FIG. 2.3 – Exemple EOG d'un clignement, sa dérivée et les paramètres considérés

Durée du clignement (D). Elle est calculée depuis le début du clignement *debh* jusqu'à sa fin *fnb*. En pratique, cette variable est très sensible aux seuils de détection V_{Cl} et

V_{Op} . Ces seuils sont appliqués sur la dérivée du signal EOG (cf. figure 2.3) afin de déterminer les périodes de fermeture et d'ouverture de l'oeil et donc le début et la fin du clignement. Une légère variation de ces seuils peut introduire de fortes variations de la durée du clignement. C'est pourquoi on préfère souvent lui substituer la mesure suivante : la durée à 50%.

Durée à 50% (D_{50}). Elle est calculée comme la durée qui s'écoule entre la moitié de l'amplitude de fermeture et la moitié de l'amplitude d'ouverture comme présenté sur la figure 2.4. Elle permet ainsi d'éviter le problème délicat de la détection précise du début et de la fin du clignement.

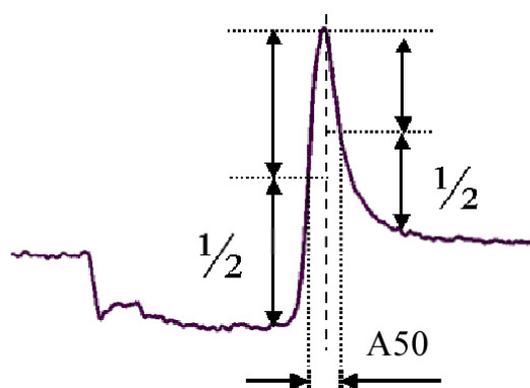


FIG. 2.4 – Définition de la durée à 50%

Fréquence (F). Elle correspond au nombre de clignements sur une fenêtre de temps donnée.

Amplitude du clignement (A). Elle est calculée depuis le début du clignement jusqu'au maximum du clignement.

Durées de fermeture (T_F) et d'ouverture (T_O). Elles correspondent respectivement aux temps mis par la paupière pour se fermer et pour se réouvrir entièrement.

Vitesses de fermeture (V_F) et d'ouverture (V_O). Elles correspondent respectivement aux paramètres dynamiques de fermeture et d'ouverture. Elles sont calculées comme la valeur moyenne de la dérivée sur les durées de fermeture et d'ouverture de l'œil.

Délai de réouverture (T_R). C'est le temps mis par la paupière pour atteindre sa vitesse maximale au moment de l'ouverture. Il est calculé entre le début de la réouverture de l'œil et le moment où la vitesse d'ouverture (calculée à partir de la dérivée du signal) est maximum.

Vitesse maximale de fermeture (PCV pour *Peak Closing Velocity*). Il s'agit de la vitesse maximale atteinte par la paupière au moment de la fermeture de l'œil.

Pourcentage de fermeture à 80%. Ce paramètre est également appelé *PERCLOS80* pour *PERcentage of eye CLOSure*. Il s'agit de la durée pendant laquelle l'œil est fermé à au moins 80%. Il est donc calculé à 80% de l'amplitude maximale. De plus il est calculé sur une fenêtre de temps donnée. Ce paramètre a été proposé par Wierwille ([Wierwille94]).

Ces différentes grandeurs sont calculées à l'aide du signal EOG du clignement ainsi que de sa dérivée, comme présenté sur la figure 2.3. Le modèle présenté ici décrit le clignement à partir de sa dérivée. En effet, le début et la fin du clignement sont définis à partir de la dérivée et des seuils V_{cl} et V_{op} .

Parmi ces différentes grandeurs, certaines sont révélatrices de la baisse de vigilance. Celle-ci est principalement associée à l'augmentation de la durée des clignements ([Morris96, Stern00, Caffier03]) ainsi qu'à l'augmentation de leur fréquence ([Hoffman46, Stern94]). Cependant, les études de Caffier et al. ([Caffier03]) nous mettent en garde sur l'utilisation de la fréquence des clignements pour détecter l'hypovigilance. En effet, l'augmentation du nombre de clignements peut être due à des facteurs environnementaux (air sec, poussière, etc). Concernant la durée des clignements, elle semble être un indicateur fiable. Les travaux de Damousis et al. ([Damousis07]) ont montré que la détection des très longs clignements était un excellent indicateur. Morris & Miller ont suggéré que la diminution de l'amplitude des clignements pouvait être associée à la baisse de vigilance. Cependant Johns ([Johns03]) a proposé d'utiliser plutôt le rapport amplitude-vélocité A/PCV comme mesure universelle de la baisse de vigilance. La pertinence de cet indice a été confirmée par Damousis et al. ([Damousis07]). Le pourcentage de fermeture de l'œil à 80% semble également être un indice très pertinent ([Knipling98]). En effet, Dinges et al. ([Dinges98]) ont montré que cet indice est fortement corrélé avec la baisse de vigilance. Les paramètres de fermeture et d'ouverture de l'œil semblent également être de bons paramètres ([Caffier05]), notamment les paramètres dynamiques que sont les vitesses de fermeture et d'ouverture ([Galley04, Damousis07]) ou le temps de réouverture de la paupière ([Galley04, Schleicher08]). La pertinence de ces paramètres a été évaluée sur des bases de données conséquentes en comparant leur évolution à l'évolution de l'hypovigilance.

Les clignements des yeux subissent donc des modifications dues à la baisse de vigilance. La figure 2.5 montre la dégradation des clignements récupérés par l'EOG de l'état éveillé vers un état d'hypovigilance avancée. Cette figure présente cinq clignements (numérotés de 1 à 5) dont l'évolution correspond à une baisse de vigilance. On constate sur cette

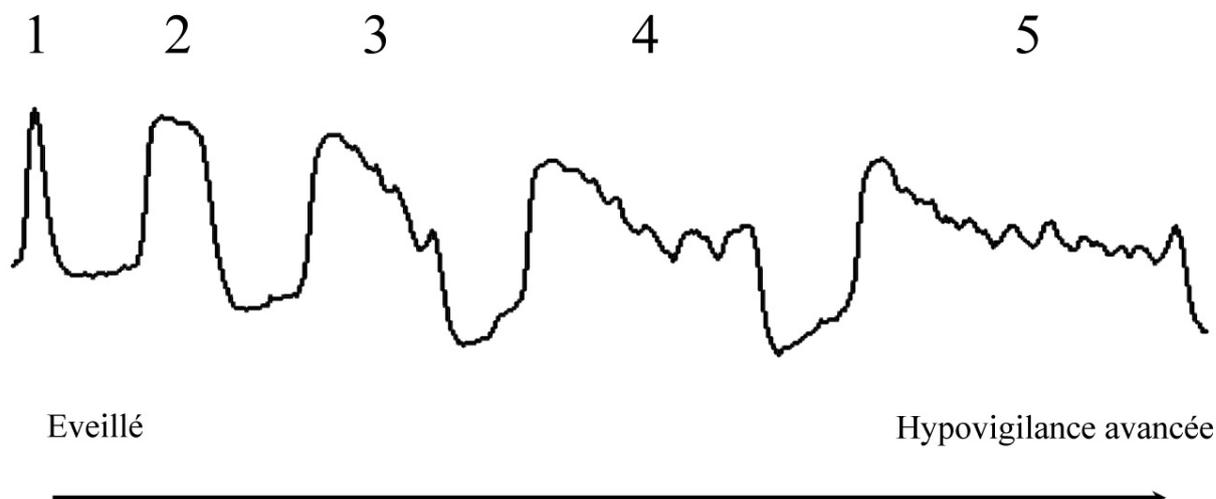


FIG. 2.5 – Enregistrement EOG de la baisse de vigilance ([Svensson04])

figure que la durée des clignements augmente avec l'hypovigilance. On remarque aussi une dégradation de la forme du clignement avec apparition d'une zone de fluctuation qui correspond à l'œil complètement fermé. Comme nous avons pu le voir, il existe une multitude de paramètres permettant de caractériser la dégradation des clignements. Les études de Caffier et al. ([Caffier03]) ont montré que seule la fréquence est sujette au problème de différences interindividuelles.

2.1.2 Les systèmes de détection d'hypovigilance par analyse des clignements

Au regard des différentes études présentées dans la section 2.1.1, il apparaît clairement que les informations oculaires récupérées par l'EOG sont de bons indicateurs de l'hypovigilance. En effet, on peut rappeler que l'EOG sert aux médecins experts à évaluer les signes visuels de l'hypovigilance. Il a été présenté dans la partie 1.1.2 que les indices visuels servaient à l'évaluation de l'hypovigilance au même titre que l'EEG à travers les échelles KDS ([Gillberg96]) mais également OSS ([Muzet03]).

L'EOG a donc naturellement fait l'objet de systèmes de détection automatique d'hypovigilance. Svensson ([Svensson04]) a proposé un système de détection du niveau de vigilance fonctionnant à partir des paramètres de durée et de fréquence de clignement

ainsi que du rapport amplitude-vélocité. Le niveau moyen de ces variables est appris au début de l'enregistrement, le système détecte les clignements qui s'éloignent trop de ces niveaux. À noter que les clignements inférieurs à 0,5s ne sont pas pris en compte. Des techniques avec apprentissage ont également été utilisées pour classer les différents niveaux de vigilance à partir de l'EOG telles que l'analyse par partitionnement de données (ou *clustering* en anglais) dont l'utilisation a été suggérée par Ohsuga et al. ([Ohsuga07]) ou les machines à vecteur support (*Support Vector Machine* en anglais) proposées par Hu & Zheng ([Hu09]). D'autres approches plus originales ont été proposées. Virkkala et al. ([Virkkala07]) ont suggéré d'étudier le contenu spectral de deux canaux EOG. Cette technique permet de détecter des états d'hypovigilance qui correspondent plutôt au stade 1 de sommeil (cf. échelle de Rechtschaffen & Kales, [Rechtschaffen68]). Hussain et al. ([Hussain08]) ont proposé d'utiliser un modèle de mélanges gaussiens sur des courbes de Lissajous obtenues à partir des canaux horizontaux et verticaux de l'EOG ainsi que des mouvements de la tête.

Enregistrer un EOG sur un conducteur n'est pas forcément évident, notamment car les électrodes doivent être placées très proches des yeux ce qui peut gêner le conducteur. Des approches de détection de clignements utilisant la vidéo ont donc été développées pour détecter les mouvements oculaires. L'utilisation de la vidéo introduit cependant des problèmes de traitement d'images comme la localisation du visage ou des yeux. Pour palier ce problème, certaines études ont eu recours à des éclairages infra-rouge (IR) qui produisent un reflet très facilement détectable au niveau des yeux. Il existe d'ailleurs des techniques de détection d'hypovigilance basées uniquement sur l'oculographie IR ([Johns07]). Nous ne détaillerons pas ici les différentes techniques de détection de visages ou des yeux mais des revues de littérature existent à ce sujet ([Hjelmas01, Yang02, Ji05]).

Des techniques vidéo relativement simples ont d'abord été proposées ([Singh99, Parmar02, Brandt04]) pour détecter l'hypovigilance. Ces techniques se contentent de détecter si l'œil est ouvert ou fermé. Une alerte est alors envoyée au conducteur s'il garde ses yeux fermés trop longtemps. Plus récemment, des techniques prenant en compte les multiples paramètres des clignements ont été proposées. C'est le cas de la méthode de Noguchi et al. ([Noguchi07]) qui ont suggéré l'application de modèles de Markov cachés (HMM) sur les paramètres de durée, d'amplitude et de vitesse des clignements afin d'obtenir un classificateur à neuf niveaux de vigilance. Il s'agit également du cas de Omi et al. ([Omi08]) qui ont proposé une fonction d'estimation du niveau de vigilance résultant de l'analyse par régression multiple sur différents paramètres tels que le PERCLOS80, la fréquence des clignements et la durée de fermeture de l'œil. Des techniques d'estimation de la di-

rection du regard du conducteur ont également été développées comme celle de Smith et al. ([Smith03]). Leur méthode estime la direction du regard du conducteur à l'aide d'une seule caméra à partir de la position des yeux et des coins de la bouche. Le conducteur est averti lorsque son regard est détaché de la route pendant trop de temps. Il en est de même pour la méthode de Ji & Yang ([Ji02]) qui ont utilisé deux caméras afin d'estimer la direction du regard et la position du visage par stéréoscopie. Cette approche est couplée à l'analyse du PERCLOS80 afin d'estimer le niveau de vigilance du conducteur. Cudalbu et al. ([Cudalbu05]) ont proposé une approche relativement semblable mais à l'aide d'une caméra haute vitesse, c'est-à-dire une caméra utilisant un échantillonnage supérieur aux 30fps (30 images par seconde) classiques, afin d'avoir une plus grande précision sur la mesure du PERCLOS80. Grace et al. ([Grace98]) avaient déjà eu l'idée d'augmenter l'échantillonnage de la caméra afin d'obtenir une meilleure résolution sur les paramètres relevés. Pour ce faire, ils ont utilisé deux caméras à 30fps synchronisées entre elles de manière à doubler la cadence d'acquisition et obtenir du 60fps . Il est également possible que la caméra puisse servir à récupérer les expressions du visage ([Vural07]). Ces derniers ont proposé d'analyser les mouvements de la tête ainsi que les expressions du conducteur à l'aide du système de codage des actions faciales (*FACS* pour *Facial Action Coding System*) afin de classer le niveau de vigilance du conducteur.

Enfin, différentes études ([Popieul00, Simon03, Schleicher08]) ont proposé d'analyser les saccades oculaires (i.e., mouvements rapides de l'œil permettant de changer la direction du regard vers un nouveau point de fixation) ainsi que les durées de fixation. Une augmentation du nombre de saccades accompagne la baisse de vigilance mais la durée des fixations est indépendante de la baisse de vigilance. Ces études ont été réalisées à l'aide d'un oculomètre, système permettant de récupérer avec précision la direction du regard.

2.1.3 Problématique

Il apparaît donc que les informations oculaires de clignements sont des informations essentielles de l'évaluation du niveau de vigilance. Cependant, si l'EOG permet d'étudier les différents paramètres oculaires avec une grande précision, il semble peu ergonomique voire dans une certaine mesure dangereux s'il est appliqué au conducteur.

L'approche vidéo semble d'une part plus intuitive et d'autre part plus adaptée à la détection de l'hypovigilance chez le conducteur car moins contraignante. En effet, l'approche vidéo permet de retrouver les caractéristiques oculaires de manière non-intrusive.

De plus elle permet, en plus des paramètres classiques de clignements d'yeux, d'analyser d'autres paramètres comme les expressions ou la direction du regard. En revanche, le faible échantillonnage des systèmes vidéo ne permet pas de récupérer les clignements avec une grande précision. Si l'on considère que la durée minimale d'un clignement est de l'ordre de $150ms$, cela représente 5 images pour une acquisition classique à $30fps$ ce qui est insuffisant pour estimer avec précision les paramètres de durée et de vitesse. De plus, cette fréquence d'acquisition de $30fps$ est vraiment faible par rapport aux $250Hz$ de l'enregistrement EOG. Damousis et al. ([Damousis07]) ont suggéré dans leur étude sur les différentes caractéristiques de la baisse de vigilance chez le conducteur que les paramètres oculaires pourraient être obtenus à l'aide d'une caméra haute fréquence. Ce type de caméra peut en effet réaliser des acquisitions jusqu'à $200fps$, ce qui est beaucoup plus proche des $250Hz$ utilisés pour l'EOG. Il existe quelques études réalisées à l'aide de caméras rapides ([Cudalbu05, Ohzeki06]) mais aucune d'entre elles ne proposent des résultats pertinents quant à la détection de la baisse de vigilance. De plus, aucune étude ne s'est intéressée au problème de la caractérisation de clignements à partir du signal vidéo à la place du signal EOG.

Nous proposons ici d'utiliser une caméra haute fréquence pour la caractérisation des clignements. Une étude comparative d'une part à partir de l'EOG et d'autre part à partir de vidéos rapides est réalisée afin de voir dans quelle mesure une caméra rapide peut effectivement remplacer l'EOG en terme de précision et de caractérisation des clignements. Le but de cette étude est la conception d'un système vidéo de caractérisation des clignements en cas d'hypovigilance à partir de la caméra rapide.

2.2 Caractérisation des clignements par analyse vidéo

Dans cette partie nous présentons un algorithme de détection des clignements oculaires puis une caractérisation quantitative des clignements détectés à partir de signaux d'énergie calculés à partir des données vidéo.

Le fonctionnement global de l'algorithme de détection et de caractérisation des clignements est présenté figure 2.6. La vidéo est d'abord traitée à l'aide de filtres rétinien bio-inspirés afin de réhausser les contours statiques et dynamiques. Puis le visage est détecté et les yeux sont localisés à l'aide d'une méthode originale basée des gradients et des projections au niveau du visage. Enfin, des signaux d'énergie sont calculés sur la zone des

yeux. Ce sont ces signaux énergétiques qui nous permettent de détecter et de caractériser les clignements.

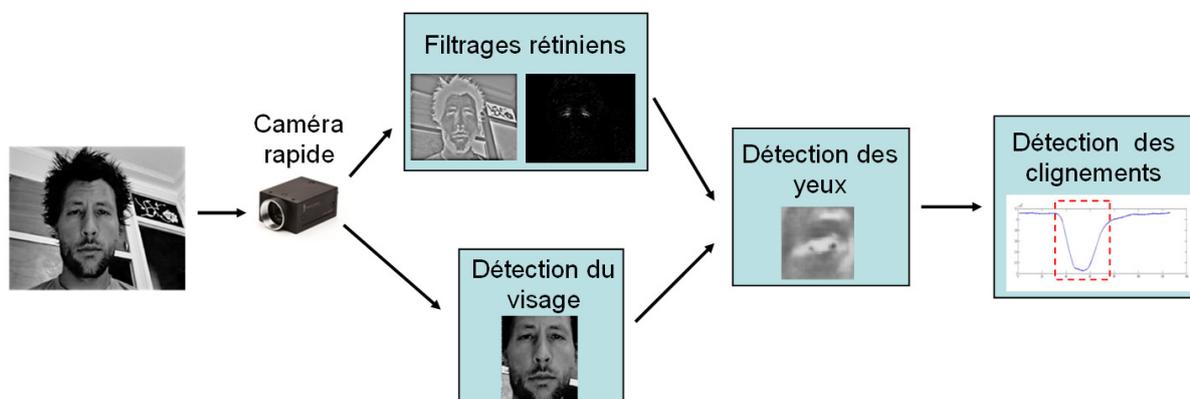


FIG. 2.6 – Méthode de détection des clignements

Ce travail s’inscrit dans la continuité de celui effectué par Benoit et al. ([Benoit07a]) sur la détection de vigilance à partir d’indices visuels. Notons que l’apport de notre travail sur cette méthode se situe au niveau de la détection des yeux et de la caractérisation des clignements. Nous avons développé une méthode originale permettant la détection des yeux peu importe leur état ouvert ou fermé. De plus, une technique originale de caractérisation des clignements à partir de l’énergie des contours statiques et dynamiques des yeux a également été développée. Les différentes étapes de l’algorithme permettant l’extraction des signaux énergétiques sont d’abord présentées puis nous exposons comment ces signaux nous permettent de caractériser les clignements.

2.2.1 Filtrages rétiniens

La première étape de notre méthode consiste en un pré-traitement du flux vidéo à l’aide des filtres rétiniens bio-inspirés développés par Benoit et al. ([Benoit07b]). Ces filtres ont pour but de réduire le bruit, réhausser les contours et renforcer les contours mobiles.

Filtre OPL

Le filtre OPL est un filtre bio-inspiré qui a pour but de modéliser la couche plexiforme externe (*OPL* pour *Outer Plexiform Layer*) de la rétine. Il s’apparente à un filtre spatio-

temporel non séparable dont la fonction de transfert est représentée figure 2.7. Ce filtre modélise l'action du réseau synaptique des photorécepteurs et des cellules horizontales de la rétine.

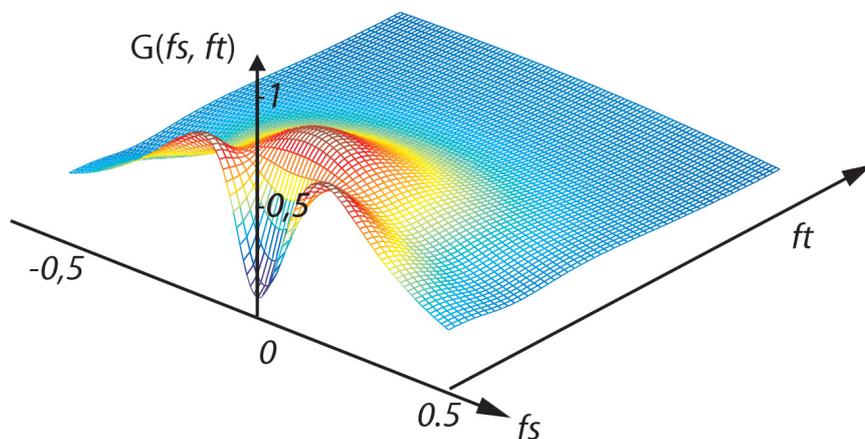


FIG. 2.7 – Fonction de transfert du filtre OPL

Ce filtre a une tendance passe-bas temporelle avec les hautes fréquences spatiales ainsi qu'un gain élevé pour une bande de fréquence définie autour de la fréquence temporelle nulle. Son effet passe-bande spatial dans les basses fréquences temporelles permet un rehaussement des contours de l'image. La tendance passe-bas au niveau des hautes fréquences spatiales et temporelles minimise le bruit spatio-temporel. Enfin, l'effet passe-bande large temporel au niveau des basses fréquences spatiales, accompagné d'une compression logarithmique en entrée de filtre réalise une adaptation locale de luminance. Cette adaptation est très appréciable car elle permet d'éviter les problèmes dus aux variations d'éclairage. La figure 2.8 montre un exemple du filtrage OPL sur une image. On constate sur cet exemple que les différents contours de l'image ont bien été rehaussés et ce indépendamment des variations d'éclairage. En effet, on peut observer des variations d'éclairage au niveau de l'arrière plan sur la gauche de l'image non filtrée ou au niveau du visage de la personne qui n'apparaissent plus sur l'image filtrée.

Filtre IPL

Le filtre IPL est un filtre bio-inspiré qui a pour but de modéliser la couche plexiforme interne (*IPL* ou *Inner Plexiform Layer*) de la rétine, dédiée à la détection de stimuli mobiles. Ce filtre agit comme un opérateur de dérivation temporelle. Il met en évidence les contours mobiles tout en éliminant les contours statiques. La figure 2.9 montre un exemple

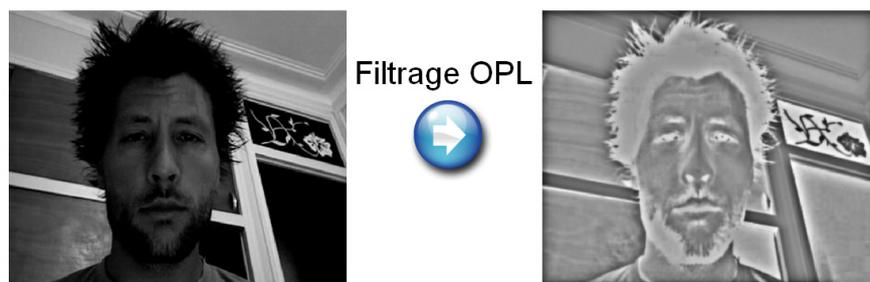


FIG. 2.8 – Exemple de filtrage OPL

de filtrage IPL sur une image extraite d'une séquence durant laquelle une personne ferme les yeux. La sortie IPL met bien en évidence les contours mobiles (ici les paupières) et l'amplitude de la sortie dépend de l'amplitude du mouvement.



FIG. 2.9 – Exemple de filtrage IPL lors d'une fermeture des yeux

2.2.2 Extraction de zones d'intérêt

Détection du visage

La localisation de visage est réalisée à l'aide de la Toolbox MPT ([MPTSite]) basée sur les travaux de Viola & Jones ([Viola02]). Ces détecteurs utilisent une cascade de filtres chargés d'extraire des indices caractéristiques du visage. Cette cascade de filtres est optimisée par une phase d'apprentissage par méthode dite "boostée" ([Bartlett03]). Notons que d'autres algorithmes présentant des performances très intéressantes peuvent également être utilisés pour la détection de visage. Duffner & Garcia ([Duffner05]) ont proposé une méthode hiérarchique comprenant trois étapes : une détection rapide du visage, une seconde étape de détection des caractéristiques plus fines du visage (les yeux,

la bouche, le nez) et enfin, une analyse fine de ces traits (contours des yeux, de la bouche, ...). Cette méthode est très précise, mais est plus coûteuse en temps de calcul que celle énoncée précédemment. Le choix du détecteur de visage s'est orienté vers un algorithme de type Viola & Jones pour sa rapidité d'exécution. Cette méthode est efficace lorsque le visage fait face à la caméra, les performances baissent de façon significative lorsque le visage s'incline de plus de 20° dans le plan de la caméra ou effectue une rotation 3D de plus de 25° . De plus, la localisation du visage reste grossière puisque d'une image à la suivante, la boîte englobante donnée par le détecteur peut subir un changement de taille de l'ordre de 20% et un décalage spatial de l'ordre de 10% de la taille du visage. Cela est peu gênant dans le cadre de notre travail car nous travaillons sur des conducteurs filmés de face. De plus, les bonnes performances globales ainsi que la rapidité d'exécution sont des avantages dans le but de réaliser un algorithme fonctionnant en temps réel. La figure 2.10 présente des exemples de détection de visage réalisés à l'aide de la Toolbox MPT.



FIG. 2.10 – Exemples de détection de visages

Détection des yeux

Une fois le visage détecté, il faut maintenant détecter automatiquement la boîte englobante de chaque œil. C'est sur ces boîtes que l'analyse automatique de clignement est menée.

On suppose que chaque œil se trouve dans le quart supérieur du visage (droit ou gauche). La conséquence immédiate de cette hypothèse est que le visage doit être vertical dans le plan de la caméra. Ainsi le système peut tolérer des rotations jusqu'à 45° dans le plan de la caméra. Chacun des yeux est recherché de manière indépendante. On cherche donc à définir une boîte englobante autour de l'œil à partir de la zone de recherche prédéfinie. On choisit de prendre une boîte englobante carrée représentant $1/16^{\text{ème}}$ de la taille de la boîte englobante visage comme présenté sur la figure 2.11. Ce choix est

conforme aux travaux de Hammal ([Hammal06]). Sur la figure 2.11, la boîte englobante du visage est représentée en pointillés vert, la zone de recherche de l'œil est en rouge et la boîte englobante de l'œil est en jaune.

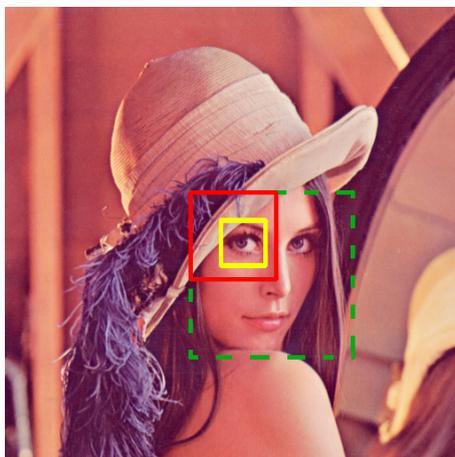


FIG. 2.11 – Illustration des différentes boîtes englobantes pour la détection des yeux

Dans le visage, l'œil ouvert peut se caractériser par la richesse de ses contours. En effet, des contours horizontaux sont présents entre l'œil et la paupière ainsi que des contours verticaux au niveau de la frontière entre l'iris et la cornée (blanc de l'œil). Des méthodes de gradient peuvent ainsi permettre de détecter l'emplacement de l'œil ([Kothari96]). Dans notre cas cependant, l'œil doit être détecté qu'il soit ouvert ou fermé. Lorsque l'œil est fermé, seuls les contours horizontaux persistent. La paupière présente alors les mêmes caractéristiques que le sourcil qui présente essentiellement des contours horizontaux.

La méthode que nous proposons pour la détection des yeux est présentée figure 2.12. Un rehaussement des contours est réalisé sur la zone de recherche de l'œil à l'aide du filtre OPL. Le gradient horizontal de cette image est alors calculé puis fusionné à cette image par multiplication de manière à mettre en avant les jonctions horizontales. On réalise alors les projections horizontale et verticale de l'image de manière à détecter la position de l'œil ([Zhou04, Peng05]). L'abscisse de l'œil correspond au maximum global de la projection horizontale. Pour l'ordonnée de l'œil, on s'intéresse aux maxima locaux de la projection verticale. On choisira alors le maximum correspondant à la partie inférieure de la zone de recherche. En utilisant cette méthode, nous faisons en fait l'hypothèse que s'il existe deux maxima locaux, celui de la partie supérieure de l'image correspond au sourcil et celui de la partie inférieure à l'iris. Une fois le centre de l'œil trouvé, une boîte englobante peut alors être définie autour de l'œil. L'originalité de cette méthode est qu'on se sert des paupières pour situer les yeux. Il est ainsi possible de détecter les yeux même s'ils sont fermés, ce

qui n'est pas le cas des méthodes qui se basent uniquement sur les caractéristiques de l'iris pour détecter les yeux.

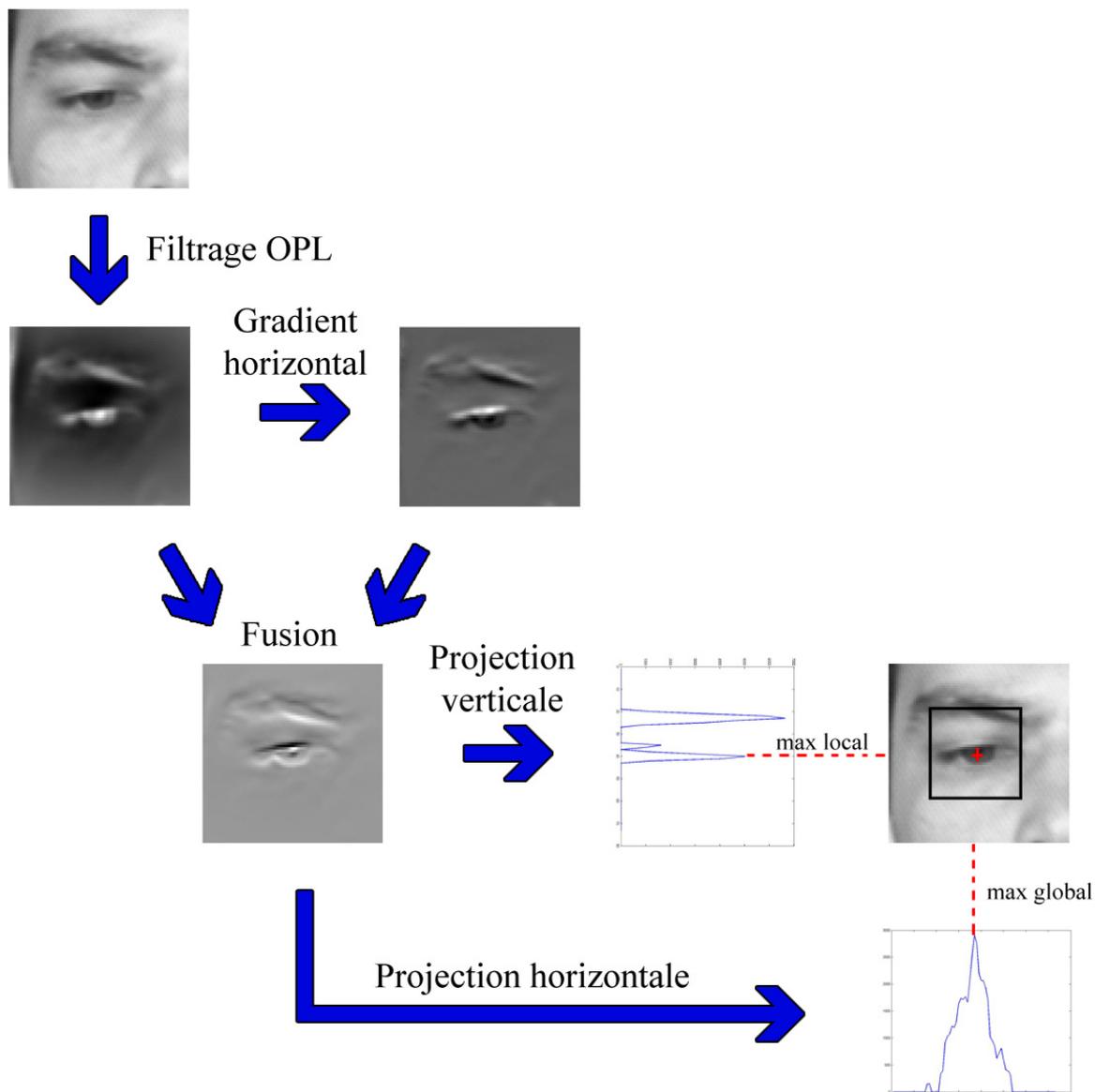


FIG. 2.12 – Méthode de détection de l'œil

Cet algorithme a été testé sur la base de visages BioID ([BioIDSite]). Cette base comprend 1521 visages vus de face, avec et sans lunettes, pris sous diverses conditions d'éclairage. De plus la position de chaque œil a été étiquetée manuellement. Cette base propose des images de résolution moyenne (286×384 pixels). Nous avons également testé notre algorithme sur cette même base de visages mais en réduisant de moitié la taille des images (soit 143×192 pixels) de manière à observer les performances de notre algorithme

sur des images de plus basse résolution. Des exemples des différents résultats obtenus sont présentés figure 2.13.

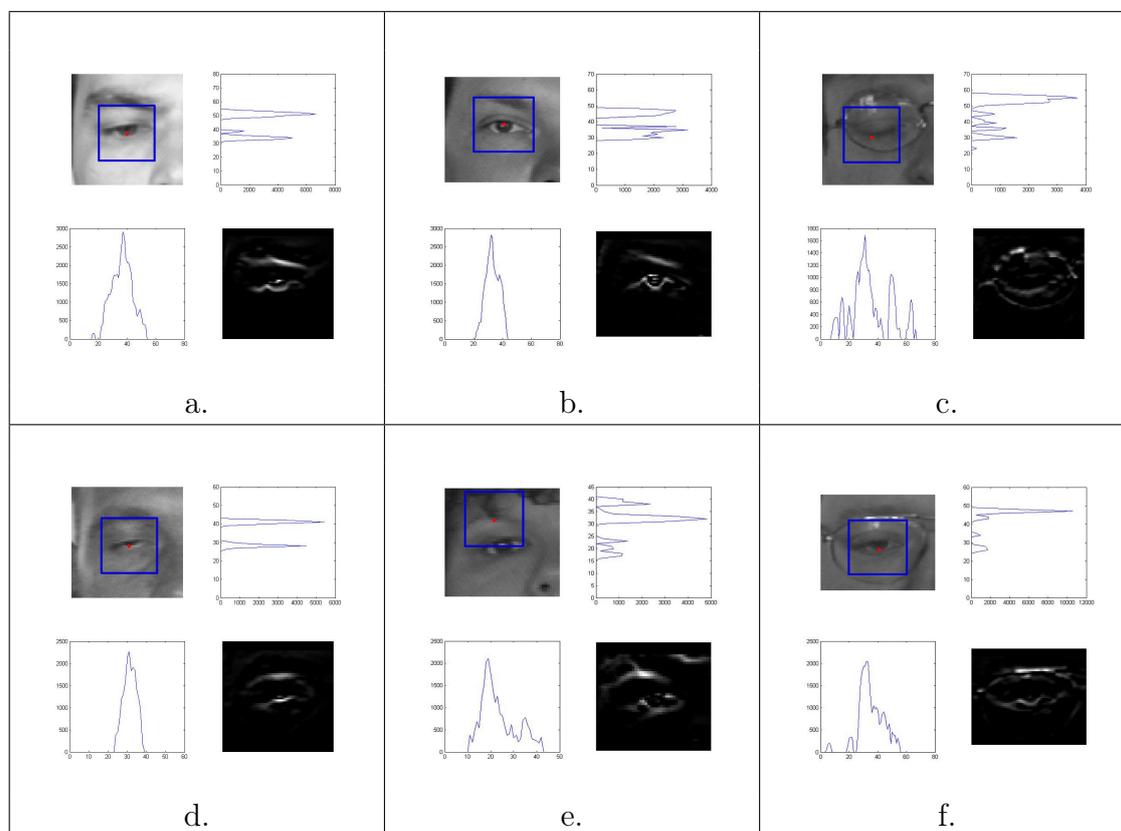


FIG. 2.13 – Exemples de détections de la position de l'œil

Sur cette figure, dans chaque groupement de 4 images, on a représenté l'image originale avec la boîte englobante trouvée par notre algorithme en haut à gauche, les images en haut à droite et en bas à gauche représentent respectivement les projections verticales et horizontales de l'image résultant de la fusion entre le filtrage OPL et le gradient horizontal, représenté en bas à droite. On remarque que l'œil est bien détecté (images a, b, c, d, et f) même si la personne porte des lunettes (images c et f) ou si la personne a les yeux clos (images c et d). La détection peut cependant être faussée dans le cas où des contours horizontaux (générés par des mèches de cheveux par exemple) sont détectés au dessus du sourcil comme dans le cas de l'image e.

Le tableau 2.1 présente une synthèse quantitative des résultats de la détection des yeux obtenue sur la base de visages BioID et sur la base BioID en réduisant la taille des images. On définit le taux de succès comme le pourcentage d'images pour lesquelles la position trouvée est comprise dans la zone réelle de l'œil (i.e., dans l'iris ou le blanc de

Taille des images	Taille moyenne de l'iris	Taux de succès	Écart moyen (pixels)	Écart type (pixels)
286 × 384 pixels	8 pixels	97%	3,74	2,74
143 × 192 pixels	4 pixels	93%	2,91	1,49

TAB. 2.1 – Performances de l'algorithme de détection des yeux sur la base BioID (1521 images)

l'œil). On donne également la moyenne et l'écart type de la distance absolue en pixels qui sépare la position de l'œil trouvée par notre algorithme et celle étiquetée manuellement.

On constate sur le tableau 2.1 que la détection des yeux fonctionne très bien. Le fait d'utiliser le filtrage OPL sur l'image pour la détection de contours nous affranchit des problèmes de variations d'éclairage. La détection de l'œil peut donc être réalisée sur toute la base de données sans avoir à adapter les réglages de notre algorithme. On remarque également que la présence de personnes portant des lunettes ne dégrade pas les performances de notre algorithme.

En comparant ces résultats à ceux obtenus par la méthode de Hammal ([Hammal06]), on note que notre algorithme a une précision légèrement inférieure quant à la localisation de l'iris (précision de l'ordre de 4 pixels versus 2 pixels pour la méthode de Hammal). Cela nous importe peu car nous cherchons seulement à placer une boîte englobante autour de l'œil afin de réaliser l'analyse des clignements. Un des avantages de notre méthode est qu'elle est plus tolérante vis-à-vis de la qualité des images traitées. En effet, elle fonctionne avec une résolution plus basse et peut détecter de manière correcte des diamètres d'iris de l'ordre de 4 pixels contrairement au minimum de 7 pixels annoncé pour la méthode de Hammal.

L'autre avantage de cette méthode est sa capacité à détecter l'œil même s'il est fermé, ce qui n'est pas le cas avec les méthodes classiques type gradient qui ont besoin de l'iris pour fonctionner correctement. Ce fait est illustré par la figure 2.14. Sur cette figure on peut constater que la détection "classique" par gradients horizontaux et verticaux ne permet pas de détecter l'œil fermé car la paupière n'est pas assez riche en contours. En effet, cette méthode est basée sur la recherche du maximum de contours horizontaux et

verticaux, elle détecte dans cet exemple la narine en bas à droite, plus riche en contours que l'œil fermé. En revanche, la méthode que nous proposons met clairement en évidence la paupière fermée et la détection de l'œil est correcte. Ceci est un avantage non négligeable pour la suite de nos travaux, notamment pour suivre et analyser l'œil lors des clignements.

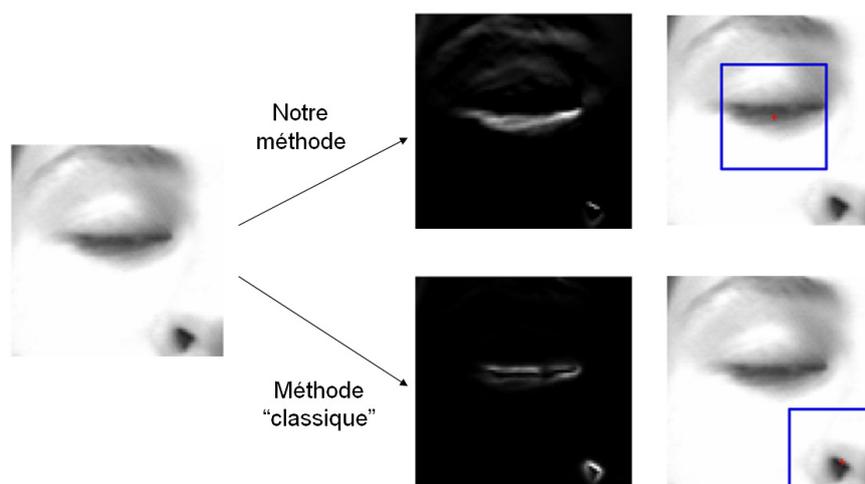


FIG. 2.14 – Comparaison de détection d'un œil fermé

2.2.3 Caractérisation des clignements

Les énergies des images filtrées OPL et IPL sont calculées sur les boîtes englobantes des yeux pour chaque image. Ces énergies sont obtenues en calculant le spectre de puissance de la zone de l'image considérée par FFT. Il en résulte deux signaux d'énergie. Le signal d'énergie lié à l'image filtrée OPL, noté E_{OPL} , correspond à la quantité de contours au niveau des yeux. Le signal d'énergie lié à l'image filtrée IPL, noté E_{IPL} , correspond à la quantité de contours en mouvement au niveau des yeux. Ces deux signaux sont filtrés à l'aide d'un filtre moyenneur avant leur analyse afin de les lisser. La taille de la fenêtre moyennuse est de 0,05s. Cette durée a été choisie afin de lisser le signal sans faire disparaître les clignements. Dans la suite, on suppose que les mouvements des deux yeux sont similaires. On utilisera donc les signaux issus d'un seul des deux yeux.

Détection des clignements

Nous sommes partis du principe que la quantité de contours contenus dans la boîte englobante de l'œil est moins importante lorsque l'œil est fermé que lorsque l'œil est ouvert. En effet, l'œil fermé possède uniquement des contours horizontaux liés à la paupière alors que l'œil ouvert possède des contours horizontaux (jonction iris-paupière) et verticaux (jonction iris-blanc de l'œil). On devrait donc observer une baisse de l'énergie E_{OPL} lors de la fermeture de l'œil. De plus, un clignement est délimité par une fermeture et une ouverture plus ou moins espacées de la paupière, c'est-à-dire deux mouvements de la paupière. On devrait donc observer deux pics d'énergies correspondant à ces mouvements sur le signal E_{IPL} . Cela est bien confirmé sur la figure 2.15 qui présente l'évolution des signaux E_{OPL} (a) et E_{IPL} (b) lors d'un clignement. Les vignettes en haut de la figure correspondent à l'évolution du clignement.

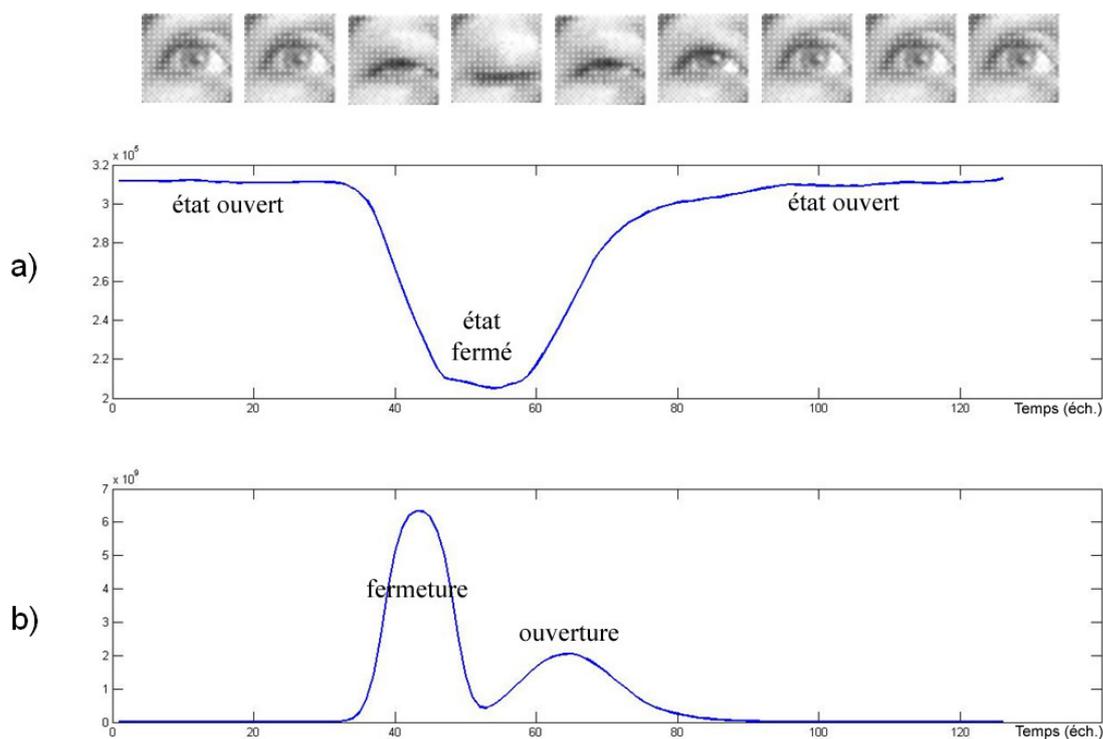


FIG. 2.15 – Exemple des signaux E_{OPL} (a) et E_{IPL} (b) lors d'un clignement

On constate effectivement sur cette figure que l'énergie E_{OPL} diminue avec la fermeture de l'œil, le niveau minimum étant atteint lorsque l'œil est complètement fermé. De plus, la fermeture et l'ouverture de la paupière font effectivement apparaître des pics d'énergie

sur le signal E_{IPL} . Nous nous basons donc sur ces changements caractéristiques pour détecter les clignements. De plus, même si les énergies sont très différentes entre l'état "œil fermé" et l'état "œil ouvert", leur valeur absolue dépend des individus. Il sera donc nécessaire d'adapter le seuil de détection à chaque conducteur par une phase d'apprentissage automatique.

Le principe de notre méthode est de détecter la présence d'éventuels clignements sur E_{IPL} par la détection des pics d'énergie correspondant à la fermeture et à l'ouverture de la paupière puis de valider et affiner la détection de ces clignements à l'aide du signal E_{OPL} . Pour cela, nous avons commencé par seuiller le signal E_{IPL} pour détecter les pics d'énergie. La valeur du seuil de détection est calculée à partir de la valeur moyenne de E_{IPL} et de son écart type calculé sur une période d'apprentissage où aucun clignement ne se produit. Expérimentalement, il s'avère que quelques secondes ($2s$ à $200fps$) d'apprentissage suffisent largement. Soient m la valeur moyenne et σ l'écart type, on définit expérimentalement le seuil de détection S_{IPL} selon l'équation (2.1).

$$S_{IPL} = m(E_{IPL}) + 5 \cdot \sigma(E_{IPL}) \quad (2.1)$$

Une fois les pics détectés, on vérifie sur le signal E_{OPL} qu'ils correspondent bien à une chute ou une augmentation relativement forte du signal. Ainsi, on peut avoir confirmation que les pics correspondent bien à des périodes d'ouverture ou de fermeture de la paupière. Cette vérification se fait en comparant la différence entre les valeurs maximale et minimale de E_{OPL} sur une fenêtre correspondant aux pics détectés sur E_{IPL} à l'écart type du signal lorsqu'aucun clignement n'a lieu. Le clignement est validé si la différence entre les valeurs maximale et minimale est supérieure à cinq fois l'écart type du signal sans clignement comme l'exprime la condition (2.2). Ce critère a été choisi empiriquement. Le fort écart qui existe cependant entre les valeurs "œil fermé" et "œil ouvert" implique que les modifications de ce critère n'ont que peu d'influence.

$$\max(E_{OPL}) - \min(E_{OPL}) > 5 \cdot \sigma(E_{OPL}) \quad (2.2)$$

Cette méthode de détection est illustrée sur la figure 2.16. On détecte d'abord les pics d'énergie par seuillage sur le signal E_{IPL} . Les minima (m_1 et m_2) et maxima (M_1 et M_2) locaux sont calculés sur les zones correspondant aux pics. Le clignement est alors accepté

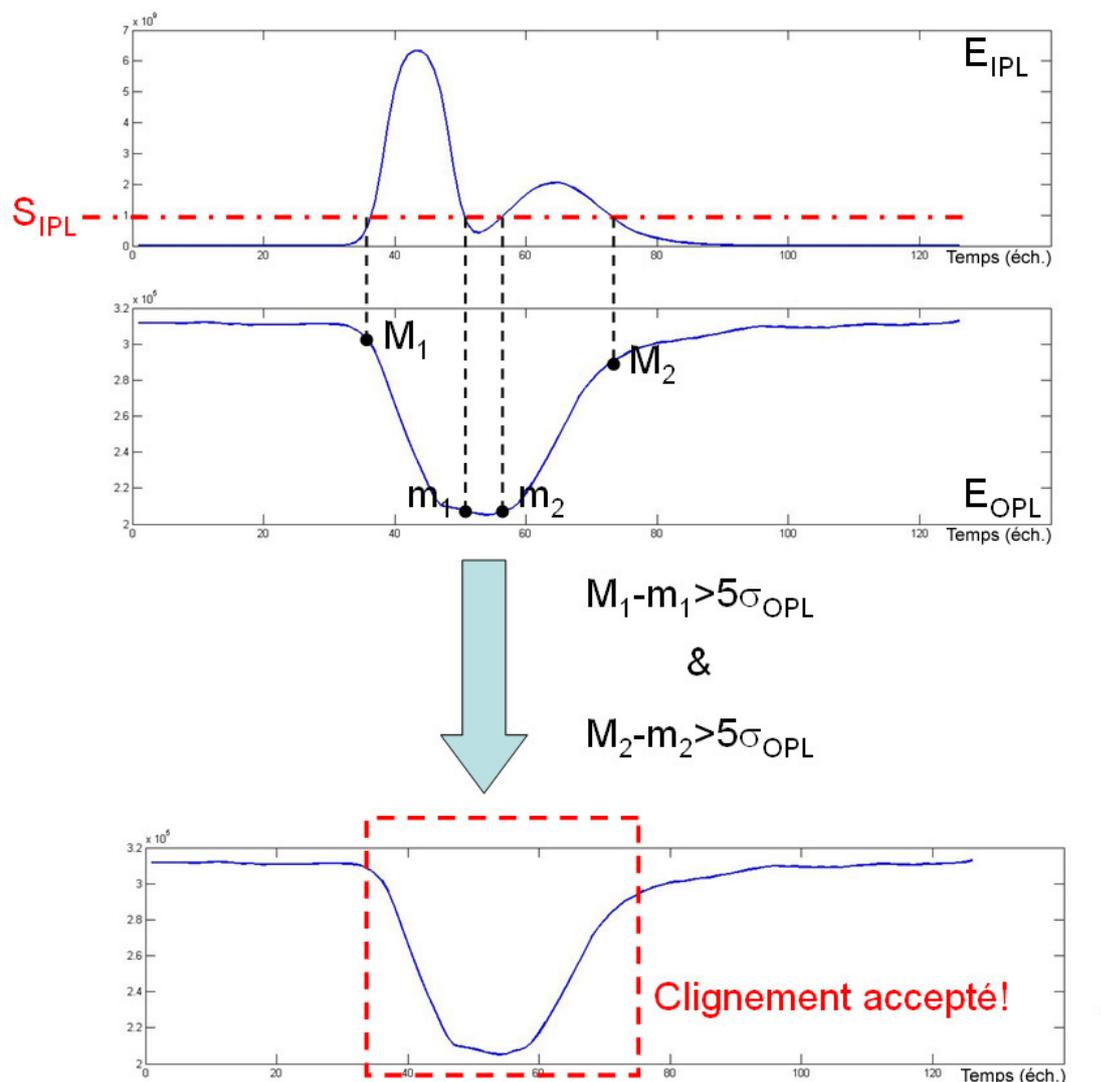


FIG. 2.16 – Méthode de détection des clignements

car les minima et maxima locaux vérifient la condition (2.2).

L'avantage de cette méthode par rapport à une méthode qui se servirait seulement du signal E_{OPL} en le seuillant réside dans son indépendance vis-à-vis des baisses d'énergie dues à du bruit de mesure ou à des artefacts tels que la perte de l'œil au niveau de l'algorithme de détection des yeux. On évite ainsi d'effectuer de fausses détections. La figure 2.17 illustre l'effet d'une mauvaise détection de l'œil sur les signaux E_{OPL} et E_{IPL} . Le cadre vert correspond à la mauvaise détection de l'œil. On constate que sa forme peut ressembler à un clignement sur le signal E_{OPL} . On constate cependant qu'il n'y a aucune

hausse d'énergie sur le signal E_{IPL} , contrairement à un vrai clignement (cadre en pointillés rouge). Cet évènement ne sera donc pas détecté par notre méthode alors qu'il l'aurait été si on avait effectué la détection uniquement à partir des variations de E_{OPL} .

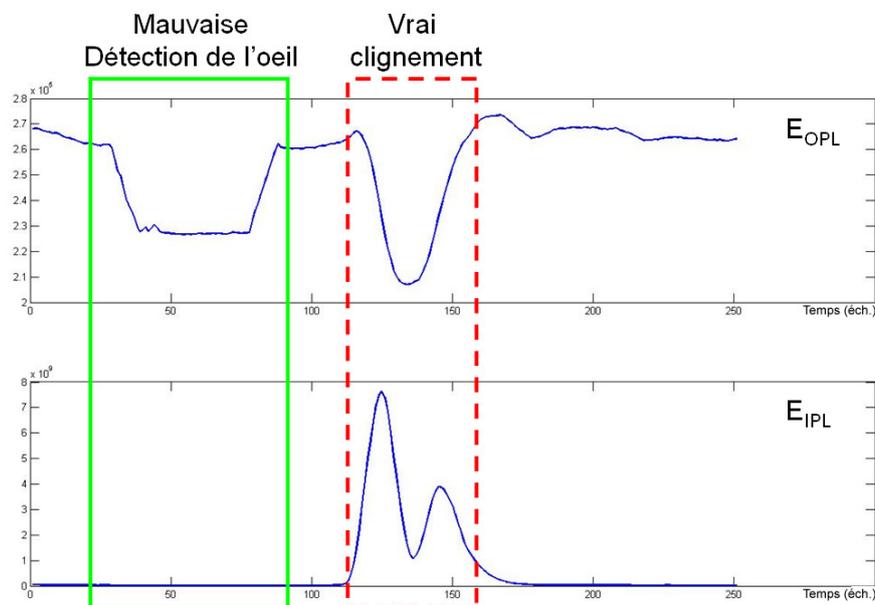


FIG. 2.17 – Exemple de mauvaise détection de l'œil sur les signaux E_{OPL} et E_{IPL}

Extraction de caractéristiques

Sur chaque clignement détecté on peut alors extraire différents paramètres des signaux E_{OPL} et E_{IPL} . Le signal E_{OPL} nous donne des informations statiques sur le clignement. Ce signal ressemble au signal EOG (au signe près). Les mêmes caractéristiques que celles extraites de l'EOG (cf. partie 2.1.1) sont extraites de ce signal :

- Durée du clignement (D), calculée depuis le début du clignement jusqu'à sa fin. Le début et la fin du clignement sont déterminées sur le signal E_{IPL} à l'aide du seuil S_{IPL} (en rouge sur la figure 2.18)
- Durée à 50% (D_{50}), calculée entre la moitié de l'amplitude de fermeture et la moitié de l'amplitude d'ouverture ;
- Fréquence (F) des clignements ;
- Amplitude du clignement (A) ;
- Durée de fermeture (T_F) et d'ouverture (T_O) ;

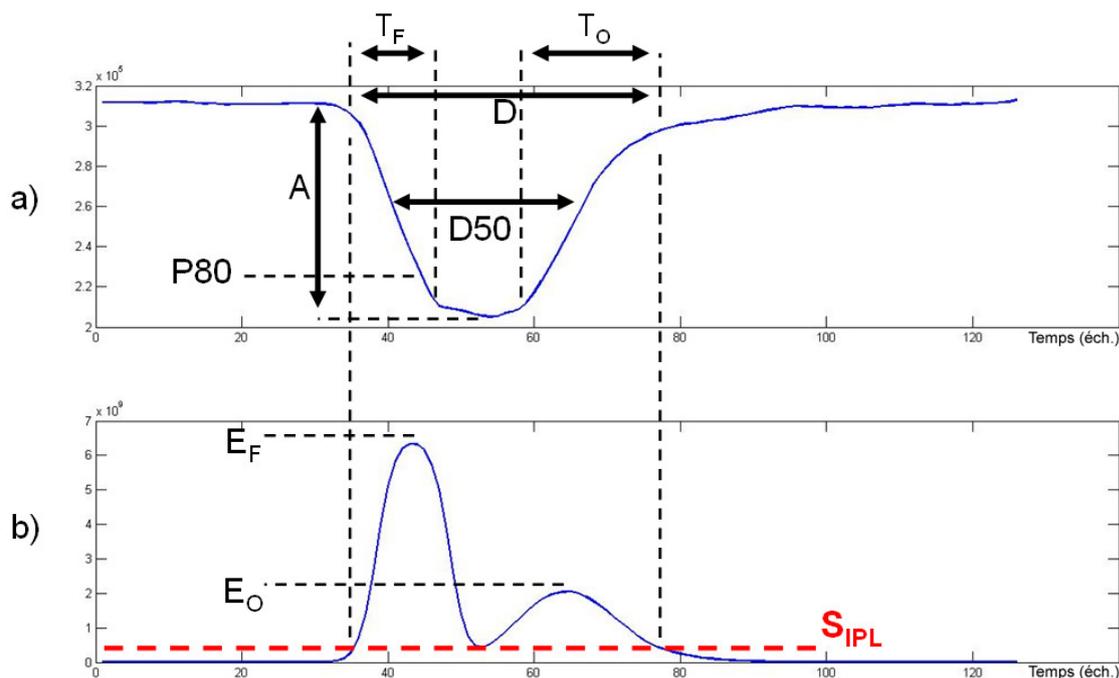


FIG. 2.18 – Paramètres extraits des signaux E_{OPL} (a) et E_{IPL} (b) lors d'un clignement

- Vitesse de fermeture (V_F) et d'ouverture (V_O), calculées à partir de la dérivée du signal E_{OPL} ;
- Délai de réouverture (T_R) ;
- Vitesse maximale de fermeture (PCV), calculée à partir de la dérivée du signal E_{OPL} ;
- Pourcentage de fermeture à 80% ($PERCLOS80$) ;
- Maxima d'énergie lors de la fermeture (E_F) et d'ouverture (E_O), calculés à partir du signal E_{IPL} et correspondant aux paramètres dynamiques de fermeture et d'ouverture de l'œil.

Les différents paramètres extraits des signaux E_{OPL} (a) et E_{IPL} (b) lorsqu'un clignement a lieu sont présentés sur la figure 2.18.

2.3 Comparaison des indices extraits par EOG avec les indices extraits par la vidéo

Cette partie présente une comparaison entre les caractérisations de clignements calculées à l'aide de l'EOG et celles calculées par analyse vidéo. Cette étude a été réalisée à partir d'une base de données comprenant des enregistrements vidéo et EOG synchronisés. Le but est d'étudier à quelles conditions on peut obtenir la même précision par analyse vidéo que par analyse EOG quant à la caractérisation des clignements afin de détecter la baisse de vigilance. Ce travail a été présenté dans [Picot09b].

2.3.1 Base de données

Pour pouvoir comparer l'analyse de clignements par analyse EOG et vidéo, nous avons réalisé une campagne d'acquisition de données contenant des enregistrements EOG et vidéo synchronisés. Cette base de données contient cinquante-six enregistrements provenant de quatorze sujets différents (onze hommes et trois femmes). Trois des sujets portent des lunettes. Chaque enregistrement contient un EOG et une séquence vidéo haute fréquence, tous deux acquis de manière synchronisée. Chaque sujet a été enregistré quatre fois à quatre échantillonnages différents : $30fps$, $100fps$, $150fps$ et $200fps$. L'enregistrement EOG est un enregistrement horizontal réalisé autour d'un seul œil (le gauche) comme présenté sur la figure 2.19 et échantillonné à $250Hz$. L'enregistrement vidéo a été réalisé à l'aide d'une caméra Point Grey "Dragonfly Express". La résolution vidéo utilisée est de 320×240 pixels. L'EOG a été enregistré à l'aide d'une station d'acquisition Biopac "MP150".



FIG. 2.19 – Extraits de la base de données

Chaque enregistrement dure environ $1min$ durant laquelle on a demandé au sujet de cligner des yeux de manière complètement aléatoire. L'EOG a été enregistré sur un seul

œil pour deux raisons. La première est que l'on suppose que les clignements d'yeux sont parfaitement symétriques. La seconde raison est que l'on a besoin de garder un œil sans électrodes autour afin que notre algorithme de détection d'œil et de caractérisation de clignements ne soit pas perturbé. On demande au sujet de réaliser un long clignement au début de chaque enregistrement afin de pouvoir réaliser la synchronisation entre EOG et vidéo au moment du traitement. Les clignements du signal EOG ont été analysés à l'aide d'un algorithme de détection semi-automatique inspiré des travaux de Sharabaty et al. ([Sharabaty08b]) et décrit dans l'annexe A. Les clignements ont été analysés dans les séquences vidéo à l'aide de notre algorithme. La base de données regroupe environ 2800 clignements répartis par échantillonnages selon le tableau 2.2.

Taux d'échantillonnage	<i>30fps</i>	<i>100fps</i>	<i>150fps</i>	<i>200fps</i>
Nombre de clignements	634	767	634	773

TAB. 2.2 – Nombre de clignements

2.3.2 Résultats et discussion

La comparaison entre EOG et vidéo a été réalisée sur toute la base de données pour les quatre fréquence d'échantillonnage vidéo.

Détection des clignements

Les performances de notre algorithme quant à sa capacité à détecter les clignements ont été évaluées. Les résultats sont présentés dans le tableau 2.3. Une détection est considérée comme correcte quand le clignement a été détecté à la fois sur la vidéo et sur l'EOG. La détection est considérée comme une non-détection lorsque le clignement est détecté sur l'EOG mais qu'il n'est pas détecté sur la vidéo. Enfin, une détection est considérée comme une fausse alarme lorsqu'un clignement a été détecté sur la vidéo mais pas sur l'EOG. Les résultats présentent le nombre de clignements dans chaque classe ainsi que le pourcentage de la base d'étude que cela représente (entre parenthèses). L'EOG est donc considéré comme la référence. Le pourcentage est calculé par rapport au nombre de clignements effectivement présents dans la base.

Fréquence d'échantillonnage	Détections correctes	Non-détection	Fausse alarmes
30fps	611 (96,4%)	23 (3,6%)	21
100fps	738 (96,2%)	29 (3,8%)	13
150fps	612 (96,5%)	22 (3,5%)	6
200fps	758 (98,1%)	15 (1,9%)	3

TAB. 2.3 – Résultats de la détection vidéo de clignements

On constate sur ce tableau que les clignements sont correctement détectés quelle que soit la fréquence d'échantillonnage vidéo utilisée. On remarque cependant que le nombre d'erreurs (non-détections et fausses détections) décroît lorsque la fréquence d'échantillonnage augmente. Les non-détections sont dues principalement à des clignements successifs très rapides détectés comme un seul clignement par vidéo. Un exemple de deux clignements rapides successifs détecté par EOG (a) mais pas par la vidéo (b) est présenté sur la figure 2.20.

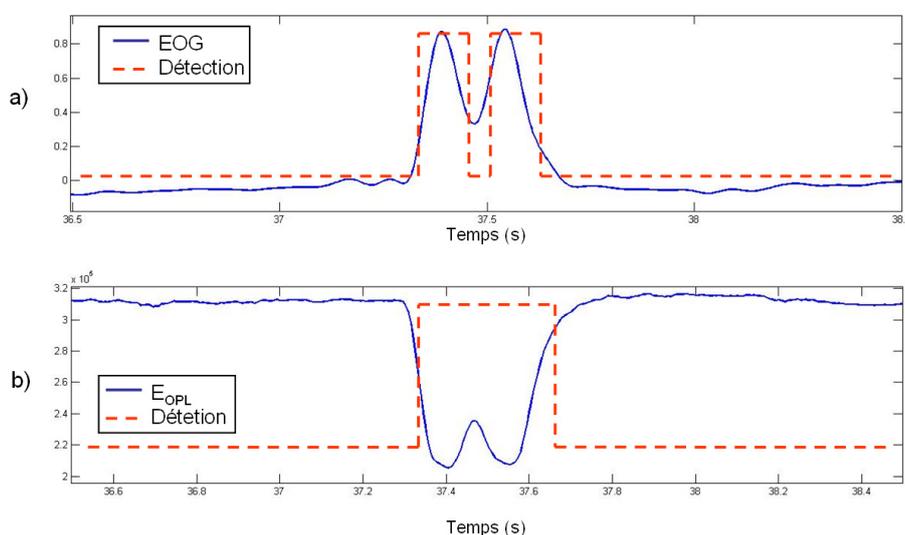


FIG. 2.20 – Exemple de détection EOG (a) et vidéo (b) de deux clignements rapides successifs

Sur cette figure, on constate que les deux clignements sont trop proches l'un de l'autre pour que le signal vidéo ait le temps de remonter à son niveau de référence "ouvert". Lorsque l'on regarde la durée moyenne des clignements manqués présentée dans le tableau

2.4, il s'avère effectivement que les clignements manqués sont des clignements très rapides.

Taux d'échantillonnage	<i>30fps</i>	<i>100fps</i>	<i>150fps</i>	<i>200fps</i>
Durée moyenne (ms)	183,72	149,53	142,29	125,60

TAB. 2.4 – Durée moyenne des clignements manqués

Il semble donc logique que ce type d'erreur décroisse avec l'augmentation du taux d'échantillonnage vidéo. En effet, plus l'échantillonnage vidéo est élevé, plus les clignements sont décrits par un grand nombre d'images et plus il est possible de distinguer des clignements rapides très proches. Ce type d'erreur peut être très ennuyeux car les longs clignements sont souvent associés à la baisse de vigilance ([Damousis07]). Si un long clignement est détecté à la place de deux clignements successifs, il peut en résulter une fausse alarme lors de la détection d'hypovigilance.

Pour comprendre l'origine des fausses alarmes, nous avons appliqué notre algorithme de détection sur la base de données en réalisant la détection de l'œil manuellement afin d'enlever toutes les erreurs qui pourraient être dues à une mauvaise localisation de l'œil par l'algorithme. Les résultats sont donnés dans le tableau 2.5.

Fréquence d'échantillonnage	Détections correctes	Non détections	Fausse alarmes
<i>30fps</i>	616	18	3
<i>100fps</i>	752	22	1
<i>150fps</i>	628	6	3
<i>200fps</i>	759	14	0

TAB. 2.5 – Résultats de la détection vidéo de clignements avec localisation manuelle de l'œil

Celui-ci confirme que les fausses détections sont dues à de mauvaises détections de la zone de l'œil par l'algorithme puisque le taux de fausses alarmes diminue fortement lorsque la zone de l'œil est détectée manuellement. Ces mauvaises détections induisent des baisses d'énergie E_{OPL} et sont considérées par l'algorithme comme des clignements. Ce type d'erreur décroît lorsque l'échantillonnage vidéo augmente. Cela semble logique

car ces erreurs sont ponctuelles et étant donné que le nombre d'échantillons utilisés par le filtre moyenné augmente avec l'échantillonnage, ces erreurs sont plus facilement lissées pour un échantillonnage élevé que pour un échantillonnage faible. On voit également sur ce tableau que le nombre de détections manquées ne varie quasiment pas par rapport au nombre trouvé avec la localisation automatique de l'œil. Cela confirme donc que ces détections manquées sont bien dues à la rapidité des clignements et qu'il ne s'agit pas d'un problème de localisation de l'œil. Enfin, ce tableau représente en quelques sortes les résultats optimaux de détection pouvant être obtenus avec notre algorithme si la détection automatique de l'œil était parfaite. On constate donc que pour un échantillonnage vidéo de $200fps$, les performances sont extrêmement proches des performances optimales.

Comparaison des paramètres des clignements

Nous allons maintenant comparer un à un les différents paramètres extraits de la vidéo à ceux extraits de l'EOG pour les clignements détectés. On analyse la corrélation entre les paramètres vidéo et EOG plutôt que la différence de moyennes. En effet, les paramètres étant extraits de signaux de nature différente, il semble utopique de vouloir retrouver les mêmes valeurs. De plus, l'hypovigilance est caractérisée par l'évolution dynamique des différents paramètres et non leurs valeurs absolues. On cherche donc à déterminer si les paramètres vidéo et EOG évoluent de la même manière. Le coefficient de corrélation donne une indication sur l'erreur commise lorsqu'on représente la liaison entre deux variables par une droite. Le tableau 2.6 présente les coefficients de corrélation obtenus entre les paramètres extraits du signal EOG et ceux estimés à partir des signaux d'énergie extraits de la vidéo. Plus le coefficient de corrélation est proche de 1 (ou de -1), plus la représentation de la liaison par une droite est juste. Un fort coefficient de corrélation montre donc que les deux paramètres sont fortement liés et évoluent linéairement.

On constate sur ce tableau que les paramètres de durée (D , $D50$ et $P80$) sont fortement corrélés quelle que soit la fréquence d'échantillonnage vidéo. Ces corrélations augmentent avec la fréquence d'échantillonnage car on augmente la résolution temporelle. La fréquence (F) est elle aussi parfaitement corrélée. Cela s'explique par les très bonnes performances obtenues pour la détection vidéo des clignements présentée dans le tableau 2.3. En revanche, les temps de fermeture et d'ouverture (T_F et T_O) ont des taux de corrélation vraiment faibles. De plus, les paramètres dynamiques concernant la vitesse des clignements (V_F , V_O , PCV et T_R) ainsi que l'amplitude (A) des clignements ont des taux de corrélation très moyens.

Paramètre	30fps	100fps	150fps	200fps
Durée (D)	0,84	0,91	0,91	0,96
Durée à 50% (D_{50})	0,72	0,96	0,93	0,97
Temps de fermeture (T_F)	0,26	0,36	0,31	0,40
Temps d'ouverture (T_O)	0,32	0,36	0,36	0,33
Amplitude (A)	0,56	0,45	0,38	0,45
Fréquence (F)	0,97	0,99	0,99	0,99
PERCLOS80 (P_{80})	0,78	0,90	0,91	0,93
Vitesse de fermeture (V_F)	0,34	0,46	0,48	0,61
Vitesse d'ouverture (V_O)	0,27	0,47	0,49	0,60
Vitesse max de fermeture (PCV)	0,34	0,43	0,49	0,60
Temps de réouverture (T_R)	0,39	0,41	0,38	0,42

TAB. 2.6 – Coefficients de corrélation pour les différents paramètres

Concernant les temps d'ouverture et de fermeture, on peut trouver l'explication sur la figure 2.21. Cette figure représente l'EOG, le signal E_{OPL} issu de l'analyse vidéo et leurs dérivées respectives pour le même clignement. L'EOG et sa dérivée sont représentés par le trait plein bleu. Le signal E_{OPL} et sa dérivée sont représentés en pointillés rouges. On constate sur cette figure que même si le début et la fin du clignement correspondent bien pour les deux signaux comme le montrent leurs dérivées, la période de fermeture de l'œil calculée à partir du signal E_{OPL} est plus courte que celle calculée sur l'EOG. De même, la période de réouverture commence plus tard sur le signal issu de la vidéo. Le signal E_{OPL} est également plus bruité que le signal EOG lorsque l'œil est fermé. En fait, le signal vidéo semble saturer lorsque l'œil est fermé, ce qui n'est pas le cas du signal EOG qui continue à suivre une légère courbure. Cela explique pourquoi les temps de fermeture et d'ouverture calculés à partir du signal E_{OPL} ne sont que très faiblement corrélés à ceux calculés à partir de l'EOG.

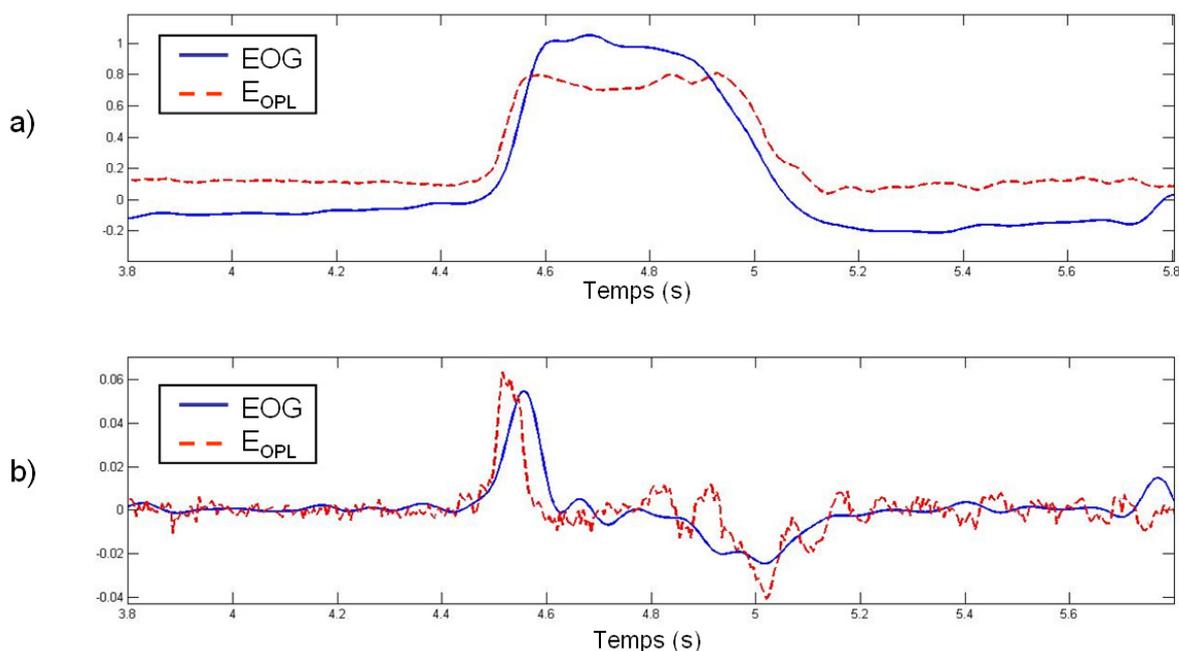


FIG. 2.21 – Exemple des signaux EOG et E_{OPL} (a) et leur dérivées (b) lors d'un clignement

Concernant les corrélations moyennes pour l'amplitude et les paramètres dynamiques (V_F , V_O et PCV), nous avons regardé la corrélation qui pouvait exister pour le rapport amplitude-vitesse (A/PCV) utilisé pour la détection de l'hypovigilance ([Johns03]). Nous avons également regardé la corrélation entre les paramètres issus du signal E_{IPL} qui représente la dynamique du clignement d'un point de vue vidéo et les paramètres de vitesse de fermeture et d'ouverture (V_F et V_O) issus de l'EOG. Les paramètres issus de

Paramètre	30fps	100fps	150fps	200fps
A/PCV	0,67	0,76	0,82	0,78
E_F et V_F	0,46	0,68	0,68	0,74
E_O et V_O	0,35	0,61	0,58	0,44

TAB. 2.7 – Coefficients de corrélation entre l’EOG et la vidéo

E_{IPL} sont les maxima d’énergie sur les périodes de fermeture (E_F) et d’ouverture (E_O). Ils sont comparés respectivement à V_F et V_O , les vitesses de fermeture et d’ouverture de l’œil calculées à partir de l’EOG. Les résultats de la corrélation de ces différents paramètres sont présentés dans le tableau 2.7.

On constate tout d’abord que le taux de corrélation pour le ratio A/PCV est bien meilleur que pour l’amplitude ou la vitesse maximale de fermeture seules. De plus, la corrélation entre la vitesse de fermeture de la paupière (V_F) calculée à partir de l’EOG et le maximum d’énergie E_{IPL} durant la période de fermeture (E_F) est également élevée pour un échantillonnage rapide ($\geq 100fps$). En effet, le signal E_{IPL} est proportionnel à l’amplitude des mouvements détectés sur la vidéo. Ce résultat semble donc logique. En revanche, le taux de corrélation entre la vitesse d’ouverture (V_O) calculée à partir de l’EOG et le maximum d’énergie E_O calculé à partir de la vidéo reste faible. Cela peut éventuellement s’expliquer par les différences qui existent durant l’ouverture de l’œil entre le signal EOG et les signaux d’énergie issus de la vidéo observés sur la figure 2.21.

En ce qui concerne le taux d’échantillonnage vidéo, les résultats s’avèrent significativement meilleurs pour un échantillonnage rapide ($\geq 100fps$) que pour un échantillonnage classique à $30fps$. Les résultats sont souvent meilleurs avec un taux d’échantillonnage à $200fps$, mais pas de manière significative comparé aux autres taux d’échantillonnage rapide. Cependant la détection des clignements s’est avérée bien plus efficace à cette échantillonnage avec notamment un nombre de fausses détections très faible. Ces fausses détections pouvant être très gênantes quant à la détection d’hypovigilance comme nous l’avons précisé plus haut, **un échantillonnage vidéo à $200fps$ semble donc la solution** la mieux adaptée dans notre cas.

2.4 Conclusion du chapitre

Une méthode originale de détection et de caractérisation des clignements oculaires par analyse vidéo a été présentée dans ce chapitre. La détection des yeux est réalisée par une méthode utilisant à la fois gradient et projections. Cette méthode permet la détection des yeux quel que soit leur état. La détection des clignements est alors réalisée à partir de signaux d'énergie statique et dynamique calculés sur la zone de l'oeil.

La comparaison entre la caractérisation des clignements par EOG et par analyse vidéo montre qu'une approche par caméra rapide peut effectivement remplacer l'analyse traditionnelle par EOG. L'approche par caméra rapide ($\geq 100fps$) est également justifiée par l'amélioration significative des résultats obtenus lors de son utilisation comparative-ment à ceux obtenus à l'aide d'une caméra à échantillonnage standard ($30fps$). Il a été montré que plusieurs paramètres liés à la détection de l'hypovigilance tels que **la durée, la durée à 50%, la fréquence, le PERCLOS80 et la vitesse de fermeture de l'oeil peuvent être extraits de l'analyse vidéo avec la même précision que sur l'EOG**. En revanche, certains paramètres extraits de la vidéo ne semblent pas corrélés avec leurs homologues extraits de l'EOG. Cependant, en fonction de l'application qu'on en fait, ils pourraient tout de même servir à détecter la baisse de vigilance. C'est le cas du rapport amplitude-vitesse (A/PCV) qui est fortement corrélé alors que ni l'amplitude, ni la vitesse ne présentent de corrélation significative. De la même manière, même si les performances sont légèrement améliorées lorsqu'on augmente l'échantillonnage vidéo, l'amélioration obtenue n'est pas suffisamment significative pour déterminer le choix de l'échantillonnage. Une fois de plus, c'est l'application qui déterminera le choix de ce paramètre. Il a cependant été montré dans cette étude que **l'utilisation d'une caméra rapide présente un net avantage par rapport à l'utilisation d'un échantillonnage classique à $30fps$. Dans notre cas, l'utilisation d'un taux d'échantillonnage de $200fps$ semble le meilleur choix** pour réaliser un système de détection d'hypovigilance.

Ainsi, nous avons montré que le signal EOG pouvait être remplacé par un signal vidéo haute fréquence pour la détection et la caractérisation quantitative des clignements. Nous allons maintenant étudier comment utiliser ces différents indices vidéo pour détecter la baisse de vigilance. Nous nous intéresserons à savoir quels sont les paramètres les plus significatifs parmi ceux extraits avec forte précision. Le but est de proposer un algorithme pour détecter la baisse de vigilance à partir de différents indices issus d'une analyse vidéo.

Chapitre 3

Détection d'hypovigilance par analyse quantitative de vidéo des clignements

Un algorithme de détection d'hypovigilance à partir d'une analyse vidéo est présenté dans ce chapitre. La comparaison entre analyse vidéo et analyse EOG réalisée dans le chapitre précédent a permis de mettre en évidence plusieurs indices visuels pouvant être extraits sur la vidéo avec la même précision que sur les EOG. Nous cherchons donc ici à concevoir un système de détection d'hypovigilance basé sur ces indices. Pour cela, la pertinence de ces différents indices dans le but d'évaluer le niveau de vigilance est étudiée afin de choisir les indices les plus adaptés. Les différents indices sélectionnés sont alors fusionnés par logique floue afin d'obtenir un algorithme de détection d'hypovigilance ne nécessitant aucun apprentissage et fonctionnant quel que soit le conducteur. Cette méthode est ensuite validée sur une base de données expertisée puis comparée à la littérature.

3.1 Revue des techniques existantes

Comme nous l'avons vu précédemment, il existe plusieurs techniques permettant d'extraire des informations à partir des clignements en utilisant l'EOG ou la vidéo. Celles-ci ont été présentées dans la section 2.1.2. Nous nous intéressons ici au lien qui existe entre les différentes grandeurs estimées et les signes d'hypovigilance, c'est-à-dire comment utiliser ces grandeurs dans le cadre de la détection d'hypovigilance.

Un des signes visuels de l'hypovigilance relativement simple à utiliser est la détection des temps trop longs de fermeture des yeux. Il s'agit d'alerter le conducteur lorsque celui-ci ferme les yeux trop longtemps. Plusieurs études ont proposé de détecter les fermetures en segmentant l'iris de l'oeil sur la vidéo ([Ueno94, Rimini01, Horng04, Dong05]). En effet, la taille de l'iris diminue jusqu'à disparaître lorsque l'oeil est fermé. D'autres travaux ont proposé de détecter les clignements en suivant l'évolution de la paupière haute et de la paupière basse, les clignements correspondant à la diminution de l'écart entre les deux paupières ([Ito02]). Enfin, des méthodes ont suggéré de détecter les clignements par corrélation entre la boîte englobante de l'oeil et un modèle d'oeil fermé ([Singh99]). Ainsi la corrélation est maximale quand l'oeil est complètement fermé. Ces différentes techniques ont toutes en commun d'alerter le conducteur uniquement en cas de fermeture prolongée des yeux, c'est-à-dire lorsque l'oeil est considéré fermé pendant plus de cinq images d'affilée. Cela représente environ $0,2s$ dans le cadre d'une fréquence d'échantillonnage standard ($30fps$). Certains obtiennent de bons taux de détection (corrélation moyenne de $r = 0,77$ entre la détection vidéo et l'évaluation de l'hypovigilance par données physiologiques pour Ueno et al. [Ueno94], 88,9% de bonnes détections pour Horng et al. [Horng04] et jusque 100% de bonnes détections pour Dong & Wu [Dong05]). Cependant nous n'avons pas d'explication concernant l'expertise "référence" de l'hypovigilance et certains de ces résultats ont été obtenus sur des données peu nombreuses (quatre sujets pour Horng et al. et seulement deux pour Dong & Wu). De plus, ces systèmes alertent le conducteur lorsque ce dernier garde les yeux fermés trop longtemps, c'est-à-dire lorsque le conducteur est déjà dans un état d'hypovigilance avancé (voire endormi au volant), ce qui semble un peu tard. En effet, un système de détection d'hypovigilance est plus efficace s'il avertit le conducteur avant que celui-ci ne s'endorme au volant, même s'il est important de le réveiller si cela arrive.

Le second type de technique consiste en des approches multivariées plus complètes. Une étude de Johns ([Johns03]) a proposé d'utiliser le rapport amplitude-vélocité (A/PCV) pour détecter les moments de défaillance du conducteur jusqu'à 60s à l'avance. Ce rapport est calculé à partir d'une détection par caméra IR. Il a été proposé afin de normaliser la vélocité (PCV). En effet, la vélocité diminue avec la baisse de vigilance alors que l'amplitude change très peu. L'avantage de l'utilisation de ce rapport est qu'il n'a pas besoin d'être calibré. Cette étude a mené à l'élaboration d'une échelle de mesure de l'hypovigilance (*Johns Drowsiness Scale*, [Johns07]) utilisant le rapport précédent ainsi que la durée des clignements pour effectuer une analyse par régression multiple. Cette échelle propose dix niveaux de vigilance mais prend une décision binaire ("éveillé" ou "fatigué")

dans la pratique. Cette méthode obtient de bons résultats dans la prédiction des écarts de conduite (i.e. au moins deux roues et la moitié de la voiture hors de la voie) avec un taux de détection de 75%. En revanche, elle génère 30% de fausses alarmes. Les travaux de Omi et al. ([Omi08]) ont également suggéré d'utiliser un modèle de régression multiple. Un large nombre de paramètres est utilisé dans le modèle : l'aire sous le niveau d'ouverture de l'œil à 100%, le nombre de clignements par minute, la durée des clignements, l'intervalle entre les clignements, le pourcentage d'ouverture de l'œil, l'écart-type des durées des clignements... Ces paramètres sont extraits à l'aide d'une caméra IR. Le modèle de régression multiple évalue le niveau de vigilance du conducteur avec succès. L'étude menée sur dix cas présente un taux de bonnes détections de 84% et seulement 9% de fausses alarmes. Le point faible de cet algorithme reste cependant qu'il nécessite un apprentissage sur un nombre important d'exemples.

Une analyse fonctionnelle du problème de la caractérisation de l'hypovigilance à partir des clignements (détectés sur l'EOG) a été réalisé par Svensson ([Svensson04]) débouchant sur un algorithme décisionnel. Cet algorithme effectue une estimation de la valeur moyenne m et de l'écart-type σ du rapport amplitude-vélocité, de la durée entre les clignements ainsi que leur fréquence. Un seuil est ensuite défini à partir de ces données sous la forme :

$$seuil = m + k \cdot \sigma$$

où k est un coefficient multiplicatif. Les clignements sont alors analysés par groupement de dix. L'algorithme effectue une classification à quatre niveaux de vigilance. Le niveau d'hypovigilance dépend du nombre de clignements dont les caractéristiques sont au-dessus du seuil pour chaque indice. Une correspondance de 57% est obtenue entre le score de l'algorithme et l'expertise KSS réalisée sur une base de données de 20 conducteurs. Ce résultat augmente, cependant, jusqu'à 70% de correspondance lorsque les seuils sont réajustés pour certains conducteurs, en ajustant la valeur de référence de l'état "éveillé" et en retirant les clignements de courte durée ($< 300ms$) principalement générés par des mouvements des yeux.

Damousis et al. ([Damousis09]) ont proposé l'utilisation d'un système expert flou (*FES* pour *Fuzzy Expert System*). Ce système effectue une prédiction d'hypovigilance par fusion de différents indices extraits de l'analyse des clignements à partir d'une caméra à 60Hz. Les paramètres utilisés sont la durée moyenne des longs clignements ($\geq 300ms$), la fréquence des clignements ainsi que l'intervalle maximum entre deux clignements successifs. Tous ces indices sont calculés sur une fenêtre de 20s. Chaque variable est séparée en trois

classes floues, donnant lieu à la génération de 27 règles logiques décrivant le système de détection. Une représentation schématique en trois-dimensions est présentée sur la figure 3.1. Le système prend une décision binaire entre les états “alert” (ou “éveillé”) et “warning” (ou “fatigué”). Ce système a été entraîné puis testé sur les données de 30 conducteurs avec un taux de bonnes détections de 90% entre la sortie de l’algorithme et l’expertise des données. Cependant, le système génère 30% de fausses alarmes.

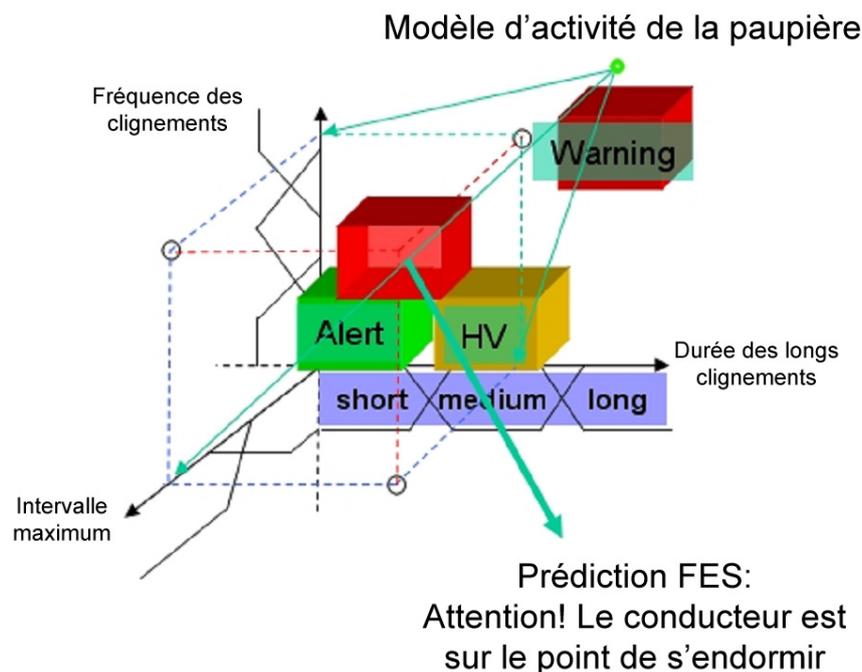


FIG. 3.1 – Représentation schématique du système expert flou proposé par Damousis et al. ([Damousis09])

D’autres techniques de classification non supervisée ont été utilisées pour détecter l’hypovigilance. Ohsuga et al. ([Ohsuga07]) ont proposé d’utiliser une technique de clustering k -means sur l’amplitude et les temps de fermeture et d’ouverture des clignements. Ces données sont normalisées pour éviter les différences interindividuelles. La méthode des k -means a pour but de partager un ensemble de points en k classes dont la variance intra-classe est minimale. Ohsuga et al. isolent ainsi cinq classes de clignements dont la fréquence d’apparition semble liée à l’apparition de l’hypovigilance. Les performances de cet algorithme n’ont néanmoins pas été quantifiées. Récemment, Hu & Zheng ([Hu09]) ont suggéré l’utilisation d’une machine à vecteur de support (*SVM* pour *Support Vector Machine*) sur différents indices extraits des clignements sur l’EOG afin d’obtenir une prédiction de l’hypovigilance à trois niveaux. Les indices utilisés sont la durée, l’amplitude, les vitesses moyennes et maximales de fermeture et d’ouverture, le PERCLOS80 et les

temps de fermeture et d'ouverture du clignement. Ces paramètres sont moyennés puis analysés sur une fenêtre de 20s. Tous les indices sont préalablement normalisés avant le traitement afin d'éviter les différences interindividuelles. L'algorithme a été testé par validation croisée sur les données de cinq sujets. Les performances obtenues sont 80,7% de détections correctes et 16,7% de fausses alarmes. Cette technique obtient de bons résultats mais utilise l'EOG et non la vidéo.

Dans ce contexte, l'approche multivariable semble donc la plus adaptée au problème de la détection de l'hypovigilance. De plus, il apparaît nécessaire dans l'intention d'obtenir un système fonctionnant en ligne de proposer un algorithme ne nécessitant que peu ou pas d'apprentissage. C'est à partir de ces remarques que nous élaborons notre système de détection vidéo d'hypovigilance.

3.2 Analyse mono-variable

Nous commençons notre étude par analyse mono-variable du problème afin de trouver quels sont les indices quantitatifs les plus adaptés à la détection d'hypovigilance. Nous nous intéressons également à la question du choix de la taille de la fenêtre d'analyse.

3.2.1 Préambule

Nous effectuons notre analyse à partir des données EOG de la base de données présentée dans la section 1.3.1 du premier chapitre. En effet, nous ne possédons pas de données vidéo expertisées. L'expertise de vigilance réalisée sur la base de données EOG (représentant 60 heures de données de conduite à partir de vingt conducteurs différents) nous permet de quantifier les performances des différents algorithmes proposés. Cependant, par souci de cohérence, nous restreignons notre étude aux indices dont la valeur obtenue par acquisition par caméra rapide présente une forte corrélation avec leurs homologues acquises par EOG. Au regard de l'étude comparative entre EOG et vidéo présentée dans le chapitre 2, nous travaillons donc à partir des paramètres suivants :

- la durée à 50% ($D50$),
 - le pourcentage de fermeture de l'œil à 80% ($PERCLOS80$),
 - la fréquence des clignements (F),
-

- le rapport amplitude-vélocité (A/PCV),
- le maximum d'énergie lors de la fermeture (E_{cl}).

Ainsi, même si nous travaillons à partir des grandeurs mesurées sur les données EOG, il a été montré dans le chapitre précédent que ces mêmes grandeurs peuvent être obtenues à partir des données vidéo avec la même précision. Le système de détection développé devrait donc fonctionner aussi bien avec des données provenant d'une analyse vidéo à $200Hz$.

Enfin, on rappelle que les données ont été expertisées selon l'échelle OSS proposée par Muzet et al. (tableau 1.1) comprenant 5 niveaux de vigilance. Les algorithmes présentés ici sont des algorithmes donnant une information binaire sur l'état du conducteur : "éveillé" ou "fatigué". Il est donc nécessaire de binariser les décisions expertes. Nous proposons et évaluons ici deux binarisations possibles des décisions expertes. La première consiste à considérer comme "fatigué" toute décision de l'expert supérieure ou égale à 1 comme nous l'avons fait pour valider notre algorithme de détection EEG dans la section 1.3.2. Nous noterons cette binarisation $Dr1$. L'autre binarisation possible consiste à considérer comme "fatigué" toute décision de l'expert supérieure ou égale à 2. En effet, si l'on regarde la colonne concernant les indices visuels de l'hypovigilance de l'échelle OSS et présentée dans le tableau 3.1, on remarque que les signes visuels apparaissent à partir du niveau 2 d'hypovigilance. Nous noterons cette binarisation $Dr2$. Par la suite, nous comparons nos résultats à ces deux types de binarisation afin d'apprécier laquelle est la plus en adéquation avec notre système de détection vidéo.

Ainsi, dans le cas de la binarisation $Dr1$ on dispose de 5512 échantillons "éveillé" et 2045 échantillons "fatigué" et dans le cas de la binarisation $Dr2$ on dispose de 6461 échantillons "éveillé" et 1096 échantillons "fatigué".

3.2.2 Techniques décisionnelles

La technique de décision retenue pour étudier la pertinence de chaque indice consiste à placer un seuil sur chaque indice afin d'effectuer une détection binaire. Pour cela, on calcule les indices moyens sur une fenêtre glissante de longueur l de l'ordre de quelques dizaines de seconde (cette longueur sera discutée section 3.2.4). On extrait en fait les indices de chaque clignement présent dans la fenêtre d'étude puis on effectue la moyenne indice par indice. Le déplacement de la fenêtre temporelle s'effectue avec un recouvrement de $l - 1$

Objective Sleepiness Score	Clignements et mouvements d'yeux
0	Clignements et mouvements d'yeux normaux
1	Clignements et mouvements d'yeux normaux
2	Clignements et mouvements d'yeux lents ou Clignements et mouvements d'yeux normaux
3	Clignements et mouvements d'yeux lents ou et clignements et mouvements d'yeux normaux
4	Clignements et mouvements d'yeux lents

TAB. 3.1 – Critères OSS restreints aux indices visuels

secondes. On obtient donc un relevé seconde par seconde pour chaque indice. On prend alors une décision sur ce signal représentant l'évolution de la valeur moyenne de chaque indice par seuillage. La détection d'hypovigilance est donc réalisée toutes les secondes. Le schéma de la figure 3.2 illustre ce fonctionnement.

Nous proposons ici deux manières de fixer un seuil pour effectuer la détection d'hypovigilance sur chacune des variables étudiées. La première consiste à utiliser un seuil fixe s_F commun pour tous les conducteurs. Le choix de la valeur de ce seuil sera discuté et décidé expérimentalement dans la section 3.2.3. On vérifiera a posteriori que les valeurs de seuil trouvées pour chacune des variables ne sont pas aberrantes. La seconde manière consiste en l'apprentissage du seuil au début de l'enregistrement, en supposant que le conducteur est complètement éveillé lorsqu'il commence à conduire comme nous l'avons fait pour la détection EEG (cf. section 1.2.4). On propose de calculer le seuil selon l'équation (3.1) comme suggéré dans [Svensson04].

$$Seuil = m + k \cdot \sigma \quad (3.1)$$

Dans cette équation, m et σ représentent respectivement la moyenne et l'écart-type

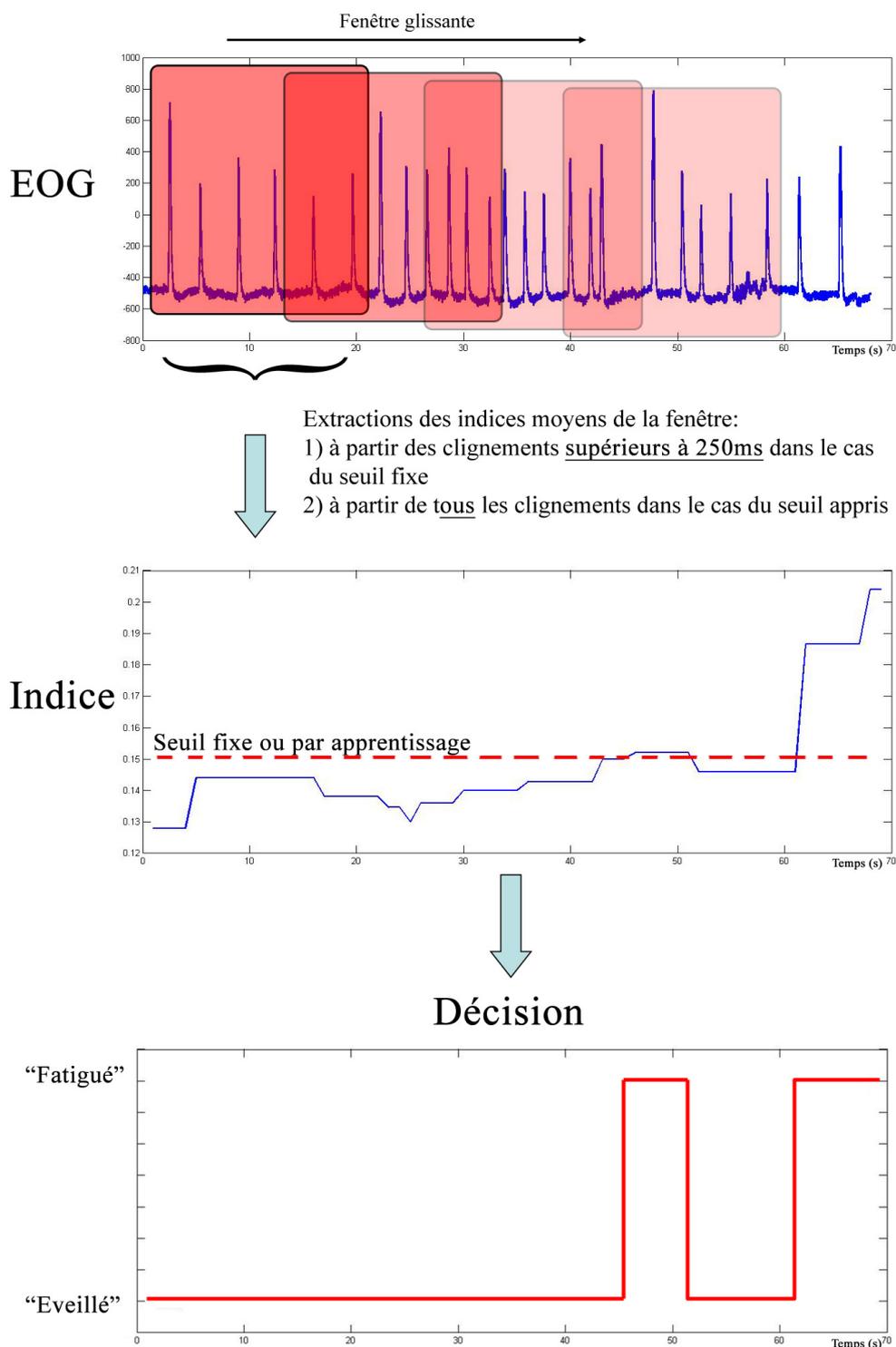


FIG. 3.2 – Technique décisionnelle mono-variable

calculés sur le début de l’enregistrement. Le paramètre k est un coefficient multiplicatif dont la valeur sera discutée et décidée expérimentalement dans la section 3.2.3. Pour apprendre les valeurs de m et σ , on essaie deux durées d’apprentissage différentes : 1min et 5min. On note respectivement s_{1min} et k_{1min} le seuil et le coefficient multiplicatif correspondant à une période d’apprentissage de 1min. Le seuil et le coefficient multiplicatif correspondant à une période d’apprentissage de 5min sont notés s_{5min} et k_{5min} . Le choix de la durée d’apprentissage est discuté dans la section 3.2.3.

Il a été suggéré dans la littérature ([Svensson04, Damousis09]) de se restreindre à l’étude des clignements longs. Nous définissons la limite des clignements longs comme ayant une durée supérieure ou égale à $250ms$. Cela signifie que nous calculons les indices moyens sur la fenêtre glissante uniquement à partir des clignements dont la durée est supérieure ou égale à $250ms$. Cependant, cela pose un problème dans le cadre d’un apprentissage sur le début de l’enregistrement. En effet, le retrait des clignements courts ($< 250ms$) implique qu’il ne reste parfois plus aucun clignement sur la période d’apprentissage. Il est alors impossible de calculer la moyenne et l’écart-type qui doivent nous permettre de définir le seuil. Nous nous restreignons donc à l’étude des longs clignements uniquement dans le cadre de la détection par seuil fixe. En revanche, pour la détection à partir d’un seuil appris, nous considérons tous les clignements.

3.2.3 Sélection des variables les plus pertinentes

Nous comparons ici les différents types de seuils sur la totalité de la base de données expertisée mise à disposition par le CEPA et présentée section 1.3.1. La longueur de la fenêtre d’étude utilisée est de $20s$ avec un recouvrement de $19s$. Nous discuterons ce choix dans la section 3.2.4. L’utilisation d’une fenêtre d’étude induit un retard de la moitié de la durée de la fenêtre (i.e. $10s$ dans ce cas). Ce retard est évidemment pris en compte lors de la comparaison de nos résultats à ceux de l’expert. Nous avons vu dans la section précédente que la détection d’hypovigilance est réalisée toutes les secondes. En revanche, les décisions de l’expert sont prises toutes les $20s$ sans recouvrement. Il est donc nécessaire de comparer les deux classifications $20s$ par $20s$. Nous considérons que le conducteur est dans l’état “fatigué” sur un intervalle de $20s$ si notre système a classé au moins $10s$ de cet intervalle comme “fatigué”. Sinon, le conducteur est considéré comme “éveillé”.

Les plages de variation du seuil fixe s_F ainsi que des coefficients multiplicatifs k_{1min} et k_{5min} dépendent de la variable. Les résultats sont représentés sous forme de courbes

ROC (cf. section 1.3.2). Pour chaque marqueur, les résultats qui sont les plus à droite correspondent au seuil de détection le plus bas et ceux représentés par les marqueurs les plus à gauche au seuil le plus élevé. Ce résultat est cohérent : quand on augmente le seuil de détection, on diminue le nombre de fausses alarmes (et donc le FP_{rate}) tout en diminuant également le nombre de bonnes détections (et par conséquent le TP_{rate}). La comparaison à la binarisation $Dr1$ est présentée à gauche et la comparaison à la binarisation $Dr2$ est à droite. Les résultats obtenus avec un seuil fixe sont représentés par les carrés bleus, ceux obtenus avec un seuil appris sur la première minute par les cercles rouges et ceux obtenus avec seuil appris sur les cinq premières minutes par les triangles. Nous commençons par présenter les différents résultats pour chaque variable, puis nous les discutons de manière synthétique par la suite.

Durée à 50% ($D50$)

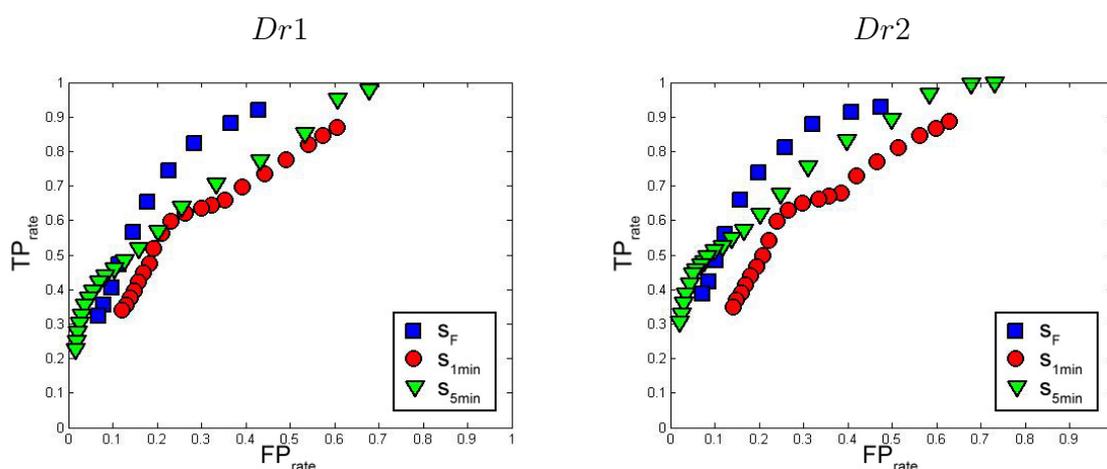


FIG. 3.3 – Résultats obtenus pour la durée à 50%

Les résultats obtenus pour la durée à 50% sont présentés sur la figure 3.3. On fait varier le seuil fixe s_F de 0,1s à 0,2s par pas de 0,01s. Les coefficients multiplicatifs k_{1min} et k_{5min} varient de 0,2 à 4 par pas de 0,2. Les résultats et seuils “optimaux” (i.e. dont le rapport entre bonnes détections et fausses alarmes a été jugé le plus acceptable dans le cadre de notre application) sont présentés dans le tableau 3.2.

On constate sur la figure 3.3 que l’on obtient de meilleurs résultats en utilisant un seuil fixe plutôt qu’un seuil appris. De plus, dans le cas d’un seuil par apprentissage les résultats sont meilleurs en effectuant l’apprentissage sur les cinq premières minutes. Enfin,

	Dr1	Dr2
Seuil fixe	$TP_{rate} = 74,3\%$ $FP_{rate} = 22,7\%$ $s_F = 130ms$	$TP_{rate} = 81,3\%$ $FP_{rate} = 25,9\%$ $s_F = 130ms$
Seuil appris sur 1min	$TP_{rate} = 62,2\%$ $FP_{rate} = 26,3\%$ $k_{1min} = 2$	$TP_{rate} = 62,9\%$ $FP_{rate} = 26,5\%$ $k_{1min} = 2,2$
Seuil appris sur 5min	$TP_{rate} = 63,8\%$ $FP_{rate} = 25,6\%$ $k_{5min} = 1,2$	$TP_{rate} = 67,8\%$ $FP_{rate} = 24,9\%$ $k_{5min} = 1,4$

TAB. 3.2 – Résultats optimaux obtenus pour la durée à 50%

on remarque que les résultats sont meilleurs pour $Dr2$ que pour $Dr1$ sauf dans le cas du seuil appris sur la première minute où les résultats restent inchangés. **Le résultat optimal obtenu est $TP_{rate} = 81,3\%$ et $FP_{rate} = 25,9\%$ avec un seuil fixe $s_F = 130ms$ pour $Dr2$.**

Pourcentage de fermeture de l'œil à 80%

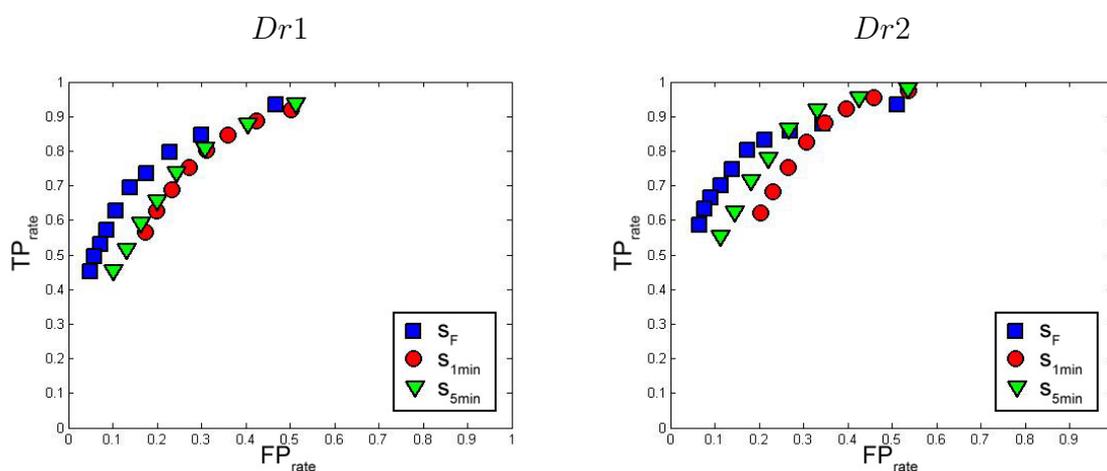


FIG. 3.4 – Résultats obtenus pour le PERCLOS80

Les résultats obtenus pour le PERCLOS80 sont présentés sur la figure 3.4. On fait varier le seuil fixe s_F de 0,25% à 2,5% par pas de 0,25%. Les coefficients multiplicatifs k_{1min} et k_{5min} varient de 0,5 à 4 par pas de 0,5. Les résultats et seuils “optimaux” sont présentés dans le tableau 3.3.

	Dr1	Dr2
Seuil fixe	$TP_{rate} = 79,8\%$ $FP_{rate} = 22,8\%$ $s_F = 0,75\%$	$TP_{rate} = 83,2\%$ $FP_{rate} = 21,3\%$ $s_F = 0,75\%$
Seuil appris sur 1min	$TP_{rate} = 74,6\%$ $FP_{rate} = 27,2\%$ $k_{1min} = 2,5$	$TP_{rate} = 74,5\%$ $FP_{rate} = 26,7\%$ $k_{1min} = 3$
Seuil appris sur 5min	$TP_{rate} = 72,9\%$ $FP_{rate} = 24,5\%$ $k_{5min} = 2$	$TP_{rate} = 77,4\%$ $FP_{rate} = 22,1\%$ $k_{5min} = 2,5$

TAB. 3.3 – Résultats optimaux obtenus pour le PERCLOS80

On remarque sur la figure 3.4 que les résultats sont meilleurs en utilisant un seuil fixe. Les résultats sont équivalents entre l'apprentissage sur une et cinq minutes pour *Dr1*. En revanche pour *Dr2*, l'apprentissage sur cinq minutes donne de meilleurs résultats que l'apprentissage sur une minute. Concernant le seuil fixe et l'apprentissage de cinq minutes, les résultats sont meilleurs pour *Dr2* que pour *Dr1*. **Le résultat optimal obtenu est $TP_{rate} = 83,2\%$ et $FP_{rate} = 21,3\%$ avec un seuil fixe $s_F = 0,75$ pour *Dr2*.**

Fréquence (F)

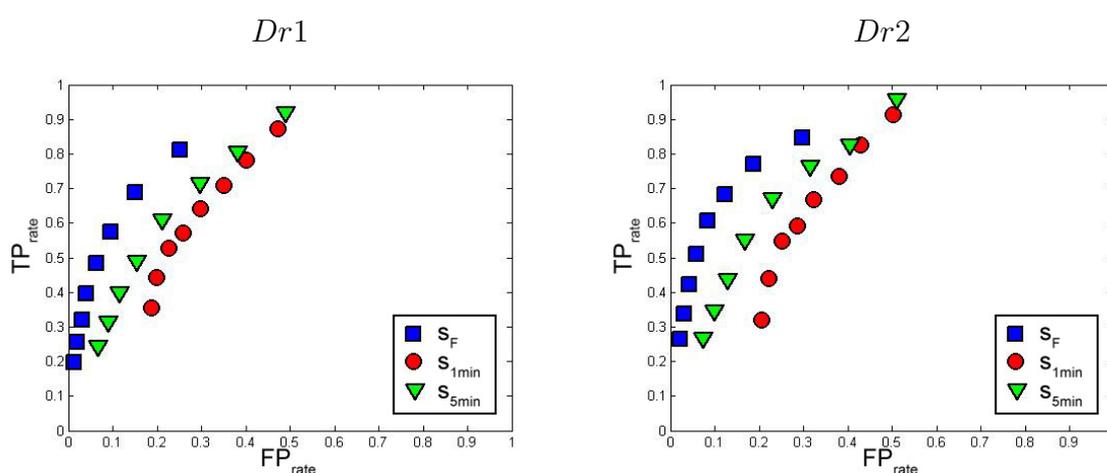


FIG. 3.5 – Résultats obtenus pour la fréquence

Les résultats obtenus pour la fréquence sont présentés sur la figure 3.5. On fait varier le seuil fixe s_F de 0,05 à 0,4 par pas de 0,05. Les coefficients multiplicatifs k_{1min} et k_{5min} varient de 0,5 à 4 par pas de 0,5. Les résultats et seuils “optimaux” sont présentés dans le

tableau 3.4.

	Dr1	Dr2
Seuil fixe	$TP_{rate} = 69,0\%$ $FP_{rate} = 15,0\%$ $s_F = 0,1$	$TP_{rate} = 76,9\%$ $FP_{rate} = 18,7\%$ $s_F = 0,1$
Seuil appris sur 1min	$TP_{rate} = 63,5\%$ $FP_{rate} = 29,7\%$ $k_{1min} = 2$	$TP_{rate} = 65,7\%$ $FP_{rate} = 32,4\%$ $k_{1min} = 2$
Seuil appris sur 5min	$TP_{rate} = 70,6\%$ $FP_{rate} = 29,7\%$ $k_{5min} = 1,5$	$TP_{rate} = 75,4\%$ $FP_{rate} = 31,6\%$ $k_{5min} = 1,5$

TAB. 3.4 – Résultats optimaux obtenus pour la fréquence

La figure 3.5 présente des résultats meilleurs dans le cadre du seuil fixe. L'apprentissage sur cinq minutes donne ici encore de meilleurs résultats que celui sur la première minute. Enfin, on constate une amélioration pour $Dr2$ par rapport à $Dr1$, sauf pour l'apprentissage du seuil sur la première minute. **Le résultat optimal obtenu est $TP_{rate} = 76,9\%$ et $FP_{rate} = 18,7\%$ avec un seuil fixe $s_F = 0,1$ pour $Dr2$.**

Rapport amplitude-vélocité (A/PCV)

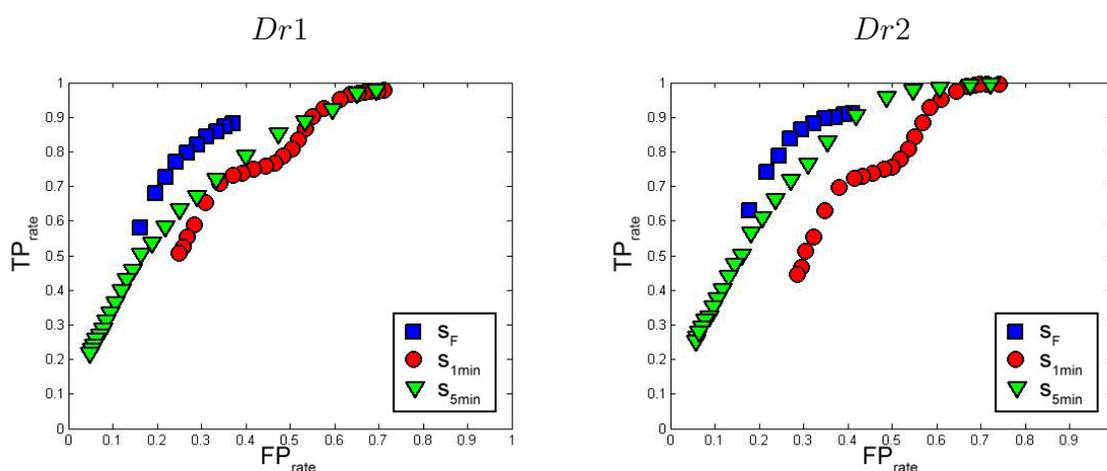


FIG. 3.6 – Résultats obtenus pour le rapport amplitude-vélocité

Les résultats obtenus pour le rapport amplitude-vélocité sont présentés sur la figure 3.6. On fait varier le seuil fixe s_F de 15,5 à 18 par pas de 0,25. Les coefficients multiplicatifs

k_{1min} et k_{5min} varient de 0,2 à 2,5 par pas de 0,1. Les résultats et seuils “optimaux” sont présentés dans le tableau 3.5.

	Dr1	Dr2
Seuil fixe	$TP_{rate} = 77,1\%$ $FP_{rate} = 24,2\%$ $s_F = 17$	$TP_{rate} = 78,7\%$ $FP_{rate} = 24,4\%$ $s_F = 17,25$
Seuil appris sur 1min	$TP_{rate} = 71,0\%$ $FP_{rate} = 34,0\%$ $k_{1min} = 2$	$TP_{rate} = 69,8\%$ $FP_{rate} = 38,1\%$ $k_{1min} = 2$
Seuil appris sur 5min	$TP_{rate} = 72,1\%$ $FP_{rate} = 33,5\%$ $k_{5min} = 0,8$	$TP_{rate} = 76,5\%$ $FP_{rate} = 31,1\%$ $k_{5min} = 0,9$

TAB. 3.5 – Résultats optimaux obtenus pour le rapport amplitude-vélocité

Dans le cas du rapport amplitude-vélocité, les résultats sont meilleurs pour l'utilisation d'un seuil fixe comme le montre la figure 3.6. Là encore, l'apprentissage sur les cinq premières minutes est plus efficace que sur la première minute. En revanche, à l'exception des résultats obtenus avec un apprentissage sur cinq minutes, on ne note pas d'amélioration entre $Dr1$ et $Dr2$. **Le résultat optimal obtenu est $TP_{rate} = 78,7\%$ et $FP_{rate} = 24,4\%$ avec un seuil fixe $s_F = 17,25$ pour $Dr2$.**

Énergie de fermeture de l'œil (E_{cl})

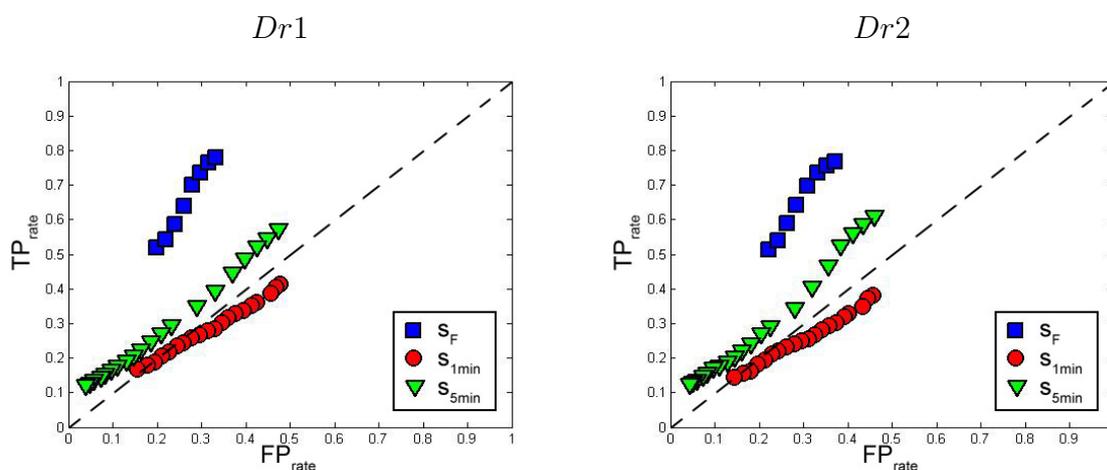


FIG. 3.7 – Résultats obtenus pour l'énergie de fermeture de l'œil

Les résultats obtenus pour l'énergie de fermeture de l'œil sont présentés sur la figure 3.7. On fait varier le seuil fixe s_F de 160 à 230 par pas de 10. Les coefficients multiplicatifs k_{1min} et k_{5min} varient de -0,1 à -2 par pas de 0,1. Les résultats et seuils "optimaux" sont présentés dans le tableau 3.6. Contrairement aux autres variables, on cherche ici à détecter une baisse de l'activité E_{cl} .

	Dr1	Dr2
Seuil fixe	$TP_{rate} = 63,9\%$ $FP_{rate} = 26,0\%$ $s_F = 190$	$TP_{rate} = 69,7\%$ $FP_{rate} = 31,0\%$ $s_F = 190$
Seuil appris sur 1min	$TP_{rate} = 16,4\%$ $FP_{rate} = 15,6\%$ $k_{1min} = -2$	$TP_{rate} = 14,5\%$ $FP_{rate} = 14,4\%$ $k_{1min} = -2$
Seuil appris sur 5min	$TP_{rate} = 12,2\%$ $FP_{rate} = 4,0\%$ $k_{5min} = -2$	$TP_{rate} = 12,3\%$ $FP_{rate} = 4,4\%$ $k_{5min} = -2$

TAB. 3.6 – Résultats optimaux obtenus pour l'énergie de fermeture de l'œil

La figure 3.7 montre une fois encore que le seuil fixe est plus adapté. En revanche on constate que les résultats sont assez moyens, notamment pour l'apprentissage sur la première minutes dont les résultats sont légèrement moins bons que ceux d'une classification aléatoire, représenté par le trait en pointillé. Aucune amélioration n'a lieu entre *Dr1* et *Dr2*. **Le résultat optimal obtenu est $TP_{rate} = 63,9\%$ et $FP_{rate} = 26,0\%$ avec un seuil fixe $s_F = 190$ pour *Dr1*.**

Discussion

Tout d'abord, l'analyse des différentes figures (fig. 3.3, 3.4, 3.5, 3.6, 3.7) fait ressortir que toutes les variables à l'exception de l'énergie de fermeture de l'œil (E_{cl}) sont des paramètres pertinents pour la détection de l'hypovigilance. Cela se constate très clairement sur la figure 3.8 qui présente les résultats optimaux pour chaque variable en fonction de la binarisation utilisée. Sur cette figure, l'indice $D50$ est représenté par les carrés bleus, l'indice $PERCLOS80$ par les ronds rouges, l'indice F par les triangles verts, l'indice A/PCV par les losanges jaunes et l'indice E_{cl} par les étoiles bleues claires.

Les résultats obtenus pour la variable E_{cl} sont trop moyens (meilleur résultat $TP_{rate} = 63,9\%$ et $FP_{rate} = 26,0\%$ avec un seuil fixe) pour utiliser la variable dans la suite de

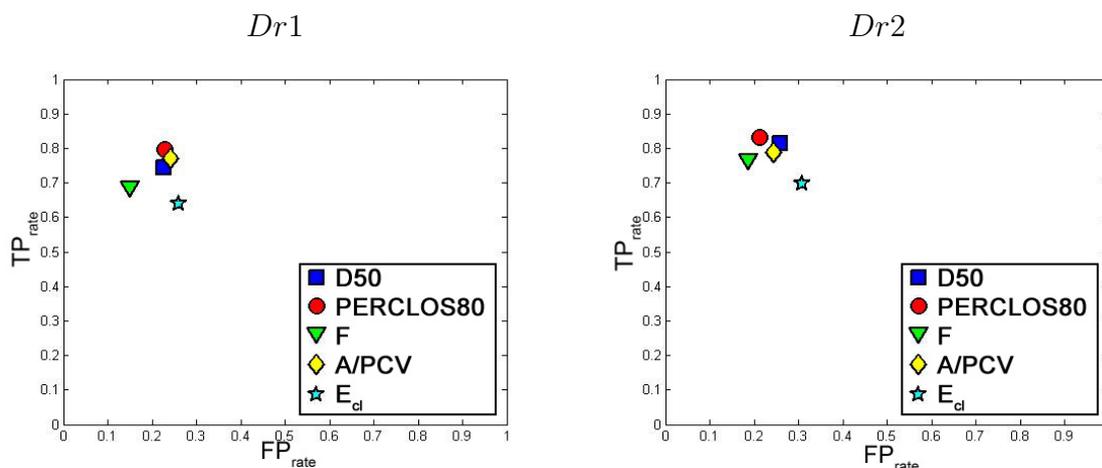


FIG. 3.8 – Résultats optimaux obtenus pour chaque indice

l'étude. Les meilleurs résultats sont obtenus en réalisant la détection à l'aide du PERCLOS80 ($TP_{rate} = 83,2\%$ et $FP_{rate} = 21,3\%$ avec un seuil fixe). Cela est cohérent avec la littérature qui présente ce paramètre comme particulièrement pertinent pour détecter l'hypovigilance ([Knipling98, Dinges98, Galley04, Cudalbu05]). **On utilisera donc les paramètres suivants dans la suite de notre étude :**

- la durée à 50% ($D50$),
- le PERCLOS80 ($P80$),
- la fréquence des clignements (F),
- le rapport amplitude-vélocité (A/PCV).

La comparaison entre les résultats obtenus pour les binarisations $Dr1$ et $Dr2$ montre que l'on obtient de meilleurs résultats ($D50$, $PERCLOS80$, F) avec $Dr2$. Cette binarisation revient à considérer le conducteur comme "fatigué" si son expertise d'hypovigilance est supérieure ou égale à deux (≥ 2) selon le tableau 1.1 proposé par Muzet et al. ([Muzet03]). Cela est cohérent avec la structure du tableau qui définit le stade 2 d'hypovigilance par l'apparition de clignements longs. Par la suite, nous considérerons uniquement cette binarisation.

Enfin, il apparaît également lors de l'analyse des différentes figures, que **l'apprentissage du seuil sur les cinq premières minutes donne de meilleurs résultats que l'apprentissage sur la première minute uniquement**. Cela est le cas pour tous les paramètres. On en déduit donc que la première minute n'est pas suffisamment représentative du comportement oculaire du conducteur. De plus, il ressort de cette analyse que

les résultats obtenus avec l'utilisation d'un seuil fixe sont meilleurs que ceux obtenus avec un seuil défini par apprentissage sur le début de l'enregistrement. Outre la définition du seuil, les deux techniques diffèrent quant à la nature des données sur lesquelles elles sont appliquées. En effet, le seuil fixe est appliqué sur des données oculaires où l'on a retiré les clignements courts ($< 250ms$) comme le suggère la littérature ([Svensson04, Damousis09]). En revanche, le seuil par apprentissage est appliqué sur les données oculaires complètes. Ce choix a été fait afin de garantir la présence de clignements à traiter dans la fenêtre d'apprentissage et de pouvoir ainsi définir le seuil. La figure 3.9 présente les résultats obtenus sur les données oculaires complètes avec un seuil fixe. Ces résultats sont représentés par les ronds rouges. Ces résultats sont comparés à ceux obtenus sur les données oculaires où l'on a retiré les clignements courts, représentés par les carrés bleus, et aux résultats obtenus avec un apprentissage du seuil sur les cinq premières minutes, représentés par les triangles verts.

On constate sur la figure 3.9 que la prise en compte des clignements courts ($< 250ms$) ne change ni les résultats obtenus pour la durée à 50% ni ceux obtenus pour le rapport amplitude-vélocité. Ceux-ci restent meilleurs que ceux obtenus avec le seuil appris sur les cinq premières minutes. En revanche, la prise en compte des clignements courts dégrade les résultats obtenus pour la fréquence et le PERCLOS80. Ces derniers deviennent alors moins bons que ceux obtenus avec l'apprentissage du seuil. Cela peut s'expliquer par l'étude de Caffier et al. ([Caffier03]) sur les paramètres oculaires de l'hypovigilance. Il est expliqué dans cette étude que la fréquence des clignements est le principal paramètre touché par les différences interindividuelles. Cela explique que, lorsqu'elle est appliquée sur les mêmes données, la détection utilisant un seuil appris sur le début de l'enregistrement est plus efficace que celle utilisant un seuil fixe. L'apprentissage prend alors en compte les différences interindividuelles. Il en est de même pour le PERCLOS80 dont la valeur dépend du nombre de clignements sur la fenêtre d'étude (donc de la fréquence). Cela n'est pas le cas pour les paramètres $D50$ et A/PCV qui sont des valeurs moyennes calculées sur la fenêtre d'étude. Le rapport amplitude-vélocité est d'ailleurs présenté dans [Johns03] comme un paramètre universel. Il semble donc que c'est l'étude des clignements longs qui est la plus pertinente pour la détection de l'hypovigilance, comme suggéré dans [Damousis09]. De plus, la restriction à ces clignements longs permet de s'affranchir des différences individuelles concernant la fréquence des clignements. **Par la suite, on utilisera des seuils fixes pour la détection en se restreignant à l'étude des clignements longs ($\geq 250ms$).**

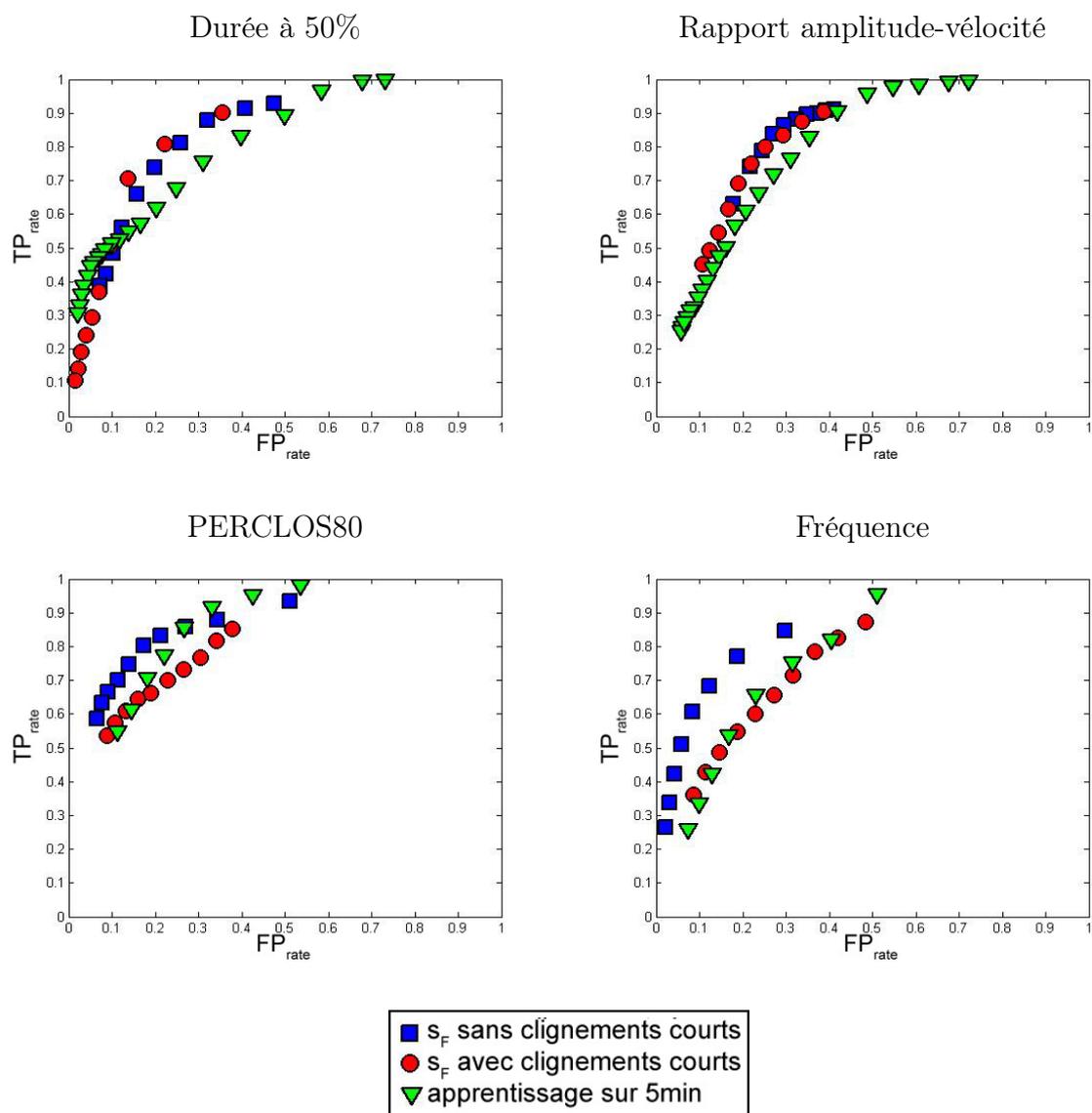


FIG. 3.9 – Résultats obtenus avec et sans considérer les courts clignements pour les différentes variables

3.2.4 Choix de la fenêtre de traitement

Nous comparons ici différentes longueurs pour la fenêtre d'étude. Les différents paramètres sont calculés sur les clignements présents dans la fenêtre d'étude. Il s'agit d'une fenêtre glissante se déplaçant seconde par seconde. Ainsi, si la fenêtre a une longueur de l secondes, on aura un recouvrement de $l - 1$ secondes. De plus, l'utilisation de cette fenêtre d'étude induit un retard de la moitié de sa longueur $\frac{l}{2}$ secondes. Les trois longueurs de fenêtre comparées sont 20s, 40s et 60s. Ces fenêtres induisent des retards respectifs de

10s, 20s et 30s. Ces retards sont évidemment pris en compte lors de la comparaison de nos résultats avec ceux de l'expert. La comparaison est présentée dans la figure 3.10 pour les quatre variables étudiées. La fenêtre de 20s est représentée par les carrés bleus, celle de 40s par les ronds rouges et celle de 60s par les triangles verts.

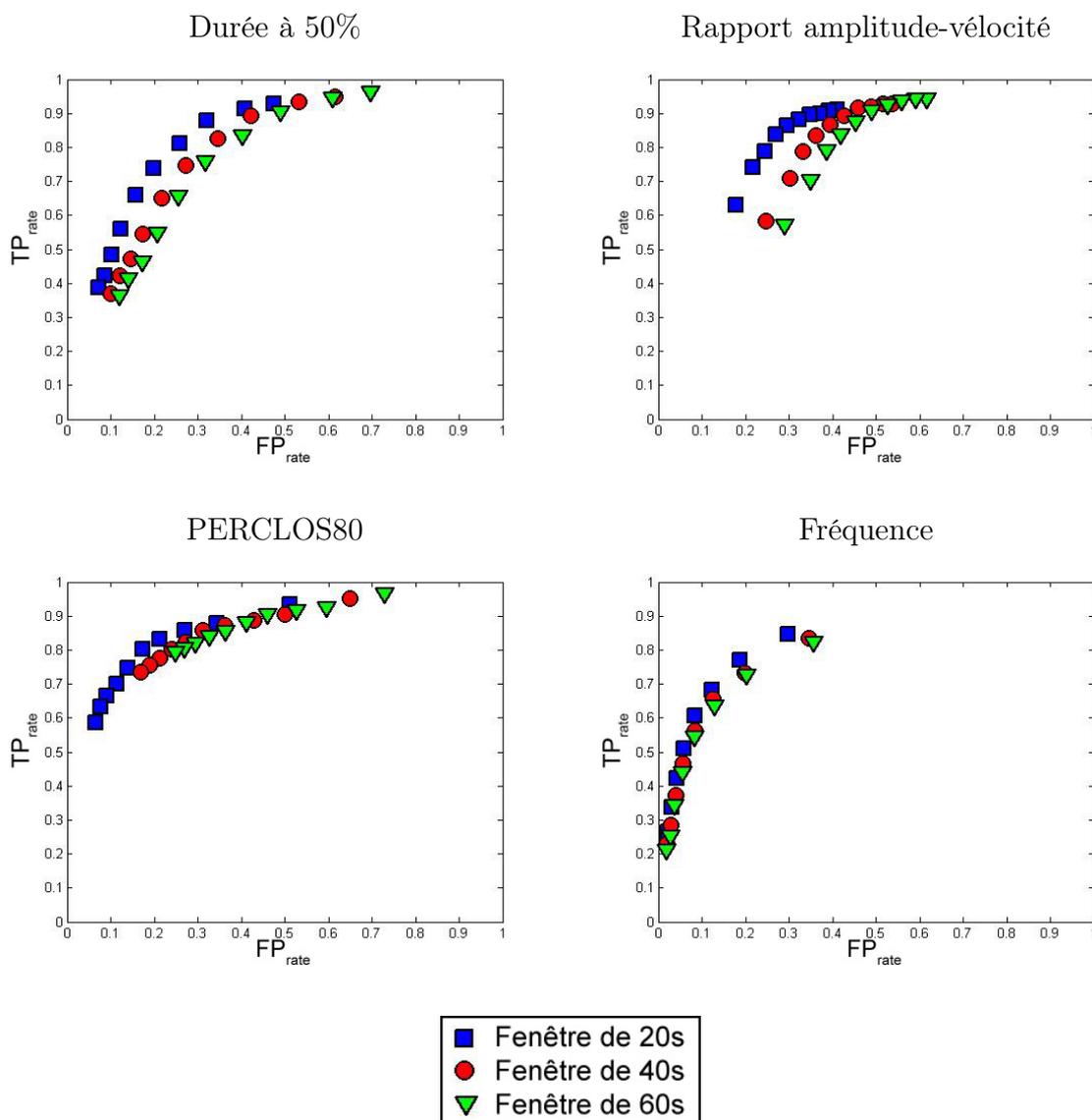


FIG. 3.10 – Résultats obtenus pour les différentes longueurs de la fenêtre d'étude

On constate sur la figure 3.10 que les meilleurs résultats sont obtenus avec une fenêtre d'étude de longueur 20s, quelle que soit la variable utilisée. De plus, les résultats se dégradent au fur et à mesure qu'on augmente la longueur de la fenêtre d'étude. **On choisit donc d'utiliser une fenêtre de 20s pour la suite de notre travail.** Cette fenêtre introduit une latence de 10s.

3.2.5 Synthèse

Nous présentons ici une synthèse de l'étude réalisée sur les différentes variables. Concernant les variables, nous utiliserons quatre variables sur les cinq étudiées. **Ces variables sont :**

- **la durée à 50% (D_{50})**,
- **le PERCLOS80% (P_{80})**,
- **la fréquence des clignements (F)**,
- **le rapport amplitude-vélocité (A/PCV)**.

En effet, ces variables ont montré une plus grande pertinence pour la détection d'hypovigilance, contrairement à l'énergie de fermeture de l'oeil (E_{cl}) qui a donné des résultats très moyens. Le paramètre le plus pertinent qui est ressorti de cette étude est le PERCLOS80.

Ces variables sont calculées chaque seconde sur **une fenêtre glissante de longueur 20s** avec un recouvrement de 19s. C'est en effet pour cette longueur que nous avons obtenu les meilleurs résultats, toutes variables confondues. De plus, les variables sont calculées sur les données oculaires **restreintes aux longs clignements ($\geq 250ms$)** ce qui permet de s'affranchir des problèmes de différences interindividuelles.

La détection d'hypovigilance est faite à l'aide d'un seuil fixe. Cette solution a donné de meilleurs résultats que l'apprentissage du seuil sur les premières minutes de l'enregistrement. Il s'agit d'une détection binaire ["éveillé" ; "fatigué"]. **L'état "fatigué" correspond à un score 2 d'hypovigilance** (cf. tableau 1.1). Cela est cohérent car ce niveau est défini par l'apparition de clignements longs. Le tableau 3.7 présente une synthèse des meilleurs résultats obtenus variable par variable ainsi que le seuil fixe s_F utilisé pour la détection. Ces résultats sont représentés sur la figure 3.11. L'indice D_{50} est représenté par le carré bleu, le PERCLOS80 par le rond rouge, la fréquence par le triangle vert et le rapport A/PCV par le losange jaune.

On remarque sur le tableau 3.7 que les résultats sont bons avec **en moyenne 80% de bonnes détections et 22% de fausses alarmes**. De plus les seuils sont cohérents par rapport aux plages de valeurs de chacune des variables.

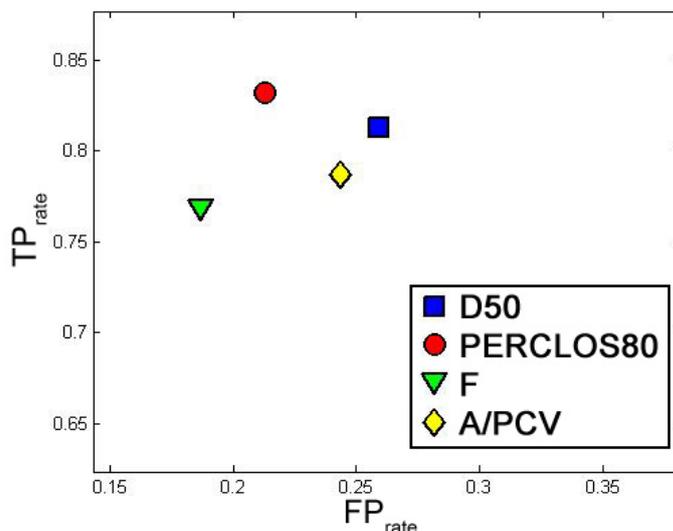


FIG. 3.11 – Représentation des résultats optimaux obtenus avec une détection par seuil fixe

	Résultats	Seuil s_F
D50	$TP_{rate} = 81,3\%$ $FP_{rate} = 25,9\%$	130ms
PERCLOS80	$TP_{rate} = 83,2\%$ $FP_{rate} = 21,3\%$	0,75%
F	$TP_{rate} = 76,9\%$ $FP_{rate} = 18,7\%$	0,1
A/PCV	$TP_{rate} = 78,7\%$ $FP_{rate} = 24,4\%$	17,25

TAB. 3.7 – Synthèse des résultats obtenus avec une détection par seuil fixe

3.3 Méthode proposée de détection d'hypovigilance

Dans cette partie, nous présentons l'algorithme que nous avons développé afin de réaliser la détection d'hypovigilance sur la base des indices vidéo. Cet algorithme utilise les quatre variables pertinentes présentées dans la section précédente pour effectuer la détection. On note respectivement $D50_m$, $P80_m$, F_m et A/PCV_m les valeurs moyennes sur une fenêtre de 20s des indices D50, PERCLOS80, F et A/PCV calculés sur le signal vidéo. L'équation (3.3) présente comment est calculé l'indice $D50_m$. Les autres indices sont calculés de la même manière. Dans cette équation k représente le nombre de clignements présents dans la fenêtre d'étude de longueur D à l'instant i et $D50_j$ représente la durée à

50% du j^{e} clignement.

$$k = \text{nombre de clignements présents sur la fenêtre } [i - D; i] \quad (3.2)$$

$$D50_m(i) = \frac{1}{k} \sum_{j=1}^k D50_j \quad (3.3)$$

Une vue d'ensemble de notre méthode est présentée sur la figure 3.12. Le principe de cette méthode est de fusionner les différents paramètres par logique floue afin d'améliorer la fiabilité de la décision. Pour cela, on procède tout d'abord à une transformation des variables réelles en variables floues : c'est la *fuzzification*. Les variables floues sont alors fusionnées par une somme pondérée. On effectue alors la détection d'hypovigilance sur le résultat de la fusion. Cette technique est basée sur la théorie des ensembles flous de Zadeh ([Zadeh65]) qui est développée plus en détails dans l'annexe B.

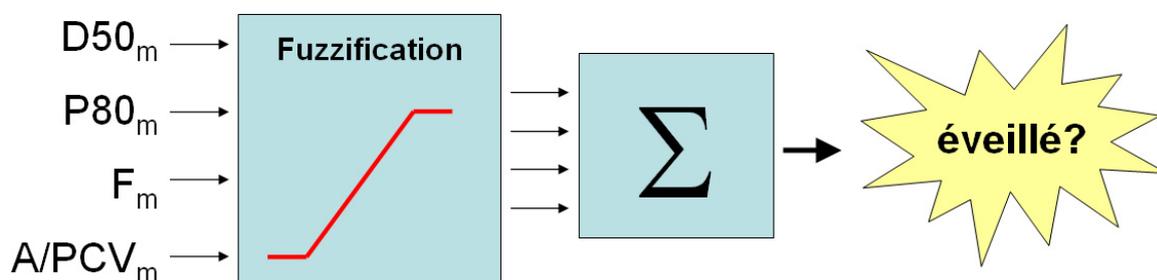


FIG. 3.12 – Méthode de détection d'hypovigilance à partir des données oculaires

3.3.1 Description de la méthode

La fuzzification

La fuzzification (de l'anglais *fuzzy* qui signifie *flou*) consiste en la transformation des variables réelles en variables floues dont la valeur varie entre 0 et 1. Il s'agit alors de définir les fonctions d'appartenance de chaque variable, c'est-à-dire attribuer à chaque variable un degré d'appartenance aux différents états. La fonction de fuzzification utilisée est une

fonction linéaire bornée définie selon l'équation (3.4).

$$f_{a,b}(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases} \quad (3.4)$$

De manière à standardiser cette fonction, les seuils a et b sont définis en fonction du seuil s_F (défini dans le paragraphe 3.2.5) selon les équations (3.5) et (3.6).

$$a = s_F - 0,25 \cdot s_F \quad (3.5)$$

$$b = s_F + 0,25 \cdot s_F \quad (3.6)$$

Les valeurs $-0,25$ et $+0,25$ ont été choisies de manière à centrer la fonction d'appartenance sur le seuil s_F considéré. Cette fonction est représentée par la figure 3.13. La fonction finale ainsi définie est notée $f_{vidéo}$.

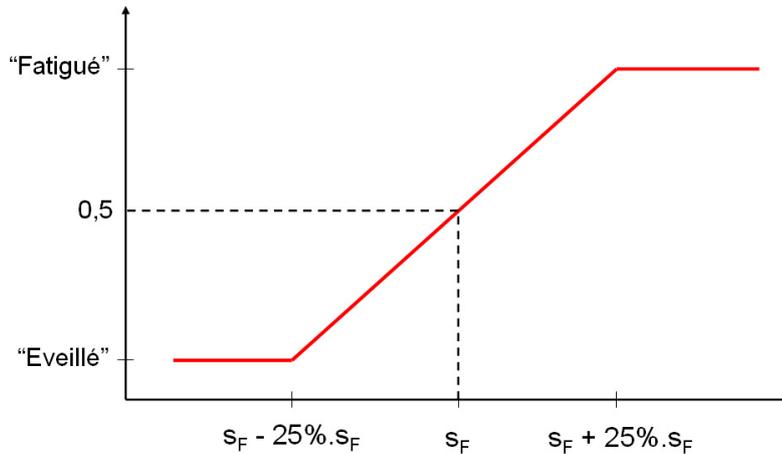


FIG. 3.13 – Fonction d'appartenance en fonction du seuil s_F de la variable considérée

On peut comprendre cette fonction de la manière suivante : plus la variable est supérieure au seuil s_F , plus son degré d'appartenance à la classe "fatigué" est fort, et inversement. Le degré d'appartenance varie alors entre 0 et 1, la valeur 0 correspondant à la classe "éveillé" et la valeur 1 à la classe "fatigué". Le degré d'appartenance à la classe "fatigué" des variables $D50_m$, $P80_m$, F_m et A/PCV_m est noté respectivement μ_{D50} , μ_{P80} , μ_F et $\mu_{A/PCV}$. Il est calculé selon l'équation (3.7).

$$\mu_{D50}(i) = f_{EOG}(D50_m(i)) \quad (3.7)$$

Fusion par logique floue

Une fois les indices transformés par fuzzification, leurs valeurs sont dans l'intervalle $[0;1]$. Il est alors possible de les fusionner par logique floue. La fusion est réalisée en effectuant la somme pondérée des différentes variables. Soit x_1, \dots, x_n n indices de degré d'appartenance respectif μ_1, \dots, μ_n . La fusion de ces n variables est calculée selon l'équation 3.8.

$$\mu_{fusion}(i) = \frac{1}{n} \sum_j \mu_j(i), \quad j \in \{D50, P80, F, A/PCV\} \quad (3.8)$$

Le résultat de la fusion est noté μ_{fusion} et sa valeur se situe évidemment dans l'intervalle $[0;1]$. La pertinence de la détection est alors décidée à partir de la valeur de μ_{fusion} : plus sa valeur est proche de 1, plus la détection est pertinente. La valeur 1 correspond à la classe "fatigué" et 0 à la classe "éveillé". On considère alors que le conducteur est dans l'état "fatigué" si cette variable est supérieure à 0,5. Sinon, le conducteur est considéré comme "éveillé". On note $d_{vidéo}(i)$ la décision d'hypovigilance réalisée à partir de l'analyse vidéo à l'instant i . Cette variable vaut "0" quand le conducteur est "éveillé" et "1" quand il est "fatigué". Cela est résumé par l'équation (3.9).

$$d_{vidéo}(i) = \begin{cases} 0 & si \quad \mu_{fusion}(i) \leq 0,5 \\ 1 & si \quad \mu_{fusion}(i) > 0,5 \end{cases} \quad (3.9)$$

Description algorithmique

Algorithme 1 Détection d'hypovigilance par fusion floue

ENTRÉES: $D50_m, P80_m, F_m$ et A/PCV_m

SORTIES: $d_{vidéo}$ dont la valeur est 0 si le conducteur est “éveillé” et 1 s’il est “fatigué”

À chaque instant i
pour $j \in \{D50, P80, F, A/PCV\}$ **faire**
 Fuzzification : $\mu_j(i) = f_{vidéo}(j_m(i))$
fin pour

Fusion : $\mu_{fusion}(i) = \frac{1}{n} \sum_j \mu_j(i), j \in \{D50, P80, F, A/PCV\}$

Décision : $d_{vidéo}(i) = \begin{cases} 0 & \text{si } \mu_{fusion}(i) \leq 0,5 \\ 1 & \text{si } \mu_{fusion}(i) > 0,5 \end{cases}$

3.3.2 Intérêt de la méthode

Le principal problème lié à l'utilisation de seuils de détection fixes est leur impossibilité à s'adapter au conducteur. Même si les différences inter-individuelles semblent négligeables comme nous l'avons vu dans la section 3.2, il se peut que la valeur moyenne d'une variable varie légèrement d'un conducteur à l'autre. La fuzzification permet de remédier à ce problème en définissant une plage de transition entre un état et un autre plutôt que de réaliser une décision binaire. En centrant cette plage de transition sur les seuils optimaux que nous avons trouvés (cf. 3.2.5), on s'affranchit alors des petites variations que pourrait subir la valeur moyenne des variables d'un conducteur à l'autre.

3.4 Résultats et discussion

Cette partie présente les résultats obtenus par notre méthode sur une base de données expertisée. Ces résultats sont comparés à ceux de la littérature. Enfin, nous vérifions le choix des variables réalisé en le comparant à celui réalisé par l'induction d'un arbre de décision. Le calcul de $d_{vidéo}$ a lieu chaque seconde. Cependant, les décisions de l'expert ont lieu toutes les 20s. On compare donc nos décisions aux décisions de l'expert par bloc de 20s. Le conducteur est alors considéré comme “fatigué” si durant ces 20s, au moins 10s ont été classées comme “fatigué” par notre algorithme.

3.4.1 Résultats obtenus par notre méthode

L'algorithme de détection a été appliqué sur l'ensemble de la base de données fournie par le CEPA et présentée dans la section 1.3.1. On rappelle que ces données représentent 60 heures de données de conduite provenant de vingt conducteurs différents. De plus, les données ont toutes été expertisées par un médecin spécialisé.

Les différents résultats obtenus sont présentés sur la figure 3.14. Le carré bleu correspond au résultat obtenu par notre méthode utilisant la logique floue. À titre indicatif, le résultat obtenu par un compromis en logique booléenne est représenté par le rond rouge. On effectue la détection à partir de chaque variable et on considère le conducteur comme "fatigué" si au moins trois des quatre sorties sont dans l'état "fatigué". Les autres marqueurs représentent les résultats obtenus par détection sur les variables traitées de manière indépendante comme présenté dans la section 3.2. La durée à 50% ($D50$) est représentée par le triangle vert, le PERCLOS80 par l'étoile jaune, la fréquence (F) par l'hexagramme cyan et le rapport amplitude-vitesse (A/PCV) par le losange violet.

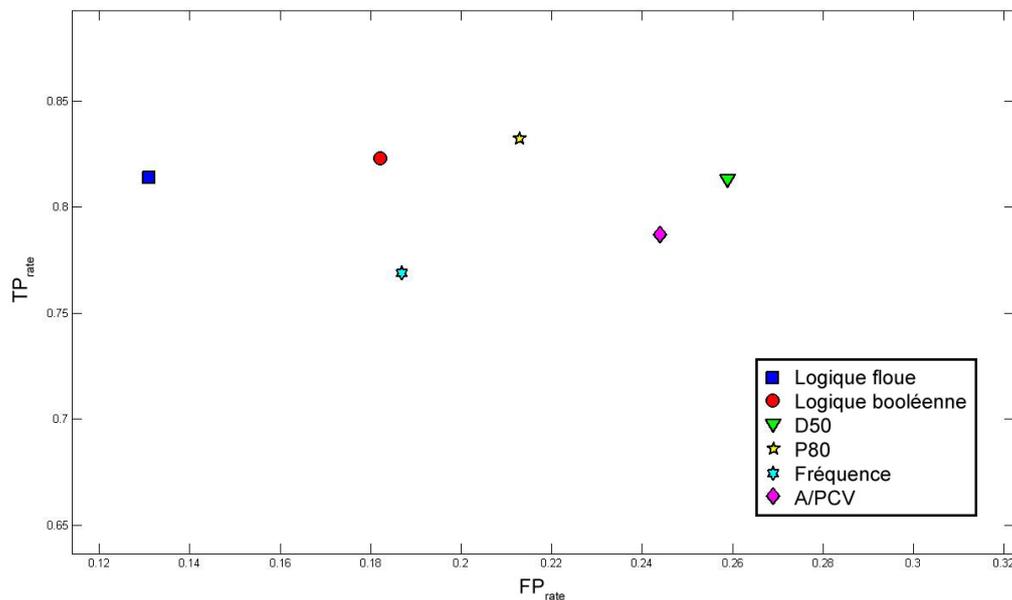


FIG. 3.14 – Résultats par fusion des différentes variables comparés à ceux obtenus variable par variable

On remarque tout d'abord sur la figure 3.14 que les résultats sont meilleurs lorsqu'on fusionne les variables que lorsqu'elles sont traitées indépendamment. En effet on constate

que le taux de bonnes détections augmente et le nombre de fausses alarmes diminue lorsqu'on fusionne les différents indices. On passe de $TP_{rate} = 80\%$ en moyenne pour les indices traités indépendamment à $TP_{rate} = 82,3\%$ en booléen et $TP_{rate} = 81,4\%$ en flou, ce qui est très bon. De plus, la fusion permet de décroître le taux de fausses alarmes. Le taux moyen de fausses alarmes obtenu pour les indices traités indépendamment est $FP_{rate} = 22,6\%$. Dans le cadre de la fusion, on obtient $FP_{rate} = 18,2\%$ en booléen et $FP_{rate} = 13,1\%$ en flou. On en déduit que l'utilisation de plusieurs variables permet de valider les détections d'hypovigilance et d'améliorer la fiabilité de l'algorithme de détection.

Ce résultat est confirmé lorsqu'on effectue un test de McNemar (cf. section 1.3.2) afin de comparer les performances obtenues par fusion floue et les performances obtenues en utilisant le PERCLOS80 (meilleures performances obtenues à l'aide d'une variable unique). La table de contingence de cette comparaison est donnée dans le tableau 3.8.

		Détection à partir du PERCLOS80	
		Mal classés	Bien classés
Fusion floue	Mal classés	712	86
	Bien classés	540	6219

TAB. 3.8 – Table de contingence pour la comparaison entre la détection par fusion floue avec celle utilisant uniquement le PERCLOS80

Le calcul du chi-deux de McNemar donne $\chi_{McNemar}^2 = 327,81$. On constate donc que $\chi_{McNemar}^2 \gg \chi^2(1)$. On en déduit que les performances des deux algorithmes ne sont pas équivalentes. En effet, on remarque dans le tableau 3.8 qu'il y a beaucoup plus d'échantillons bien classés par la détection par fusion floue et mal classés par la détection à partir du PERCLOS80 seul ($n_{10} = 540$) que l'inverse ($n_{01} = 86$). On en conclut que la détection par fusion floue est effectivement plus performante.

De plus, il apparaît sur cette figure que l'approche par logique floue est plus performante que par logique booléenne. En effet, le taux de bonnes détections est équivalent dans les deux cas ($TP_{rate} = 82,3\%$ en booléen et $TP_{rate} = 81,4\%$ en flou). En revanche, l'approche floue permet de diminuer fortement le taux de fausses alarmes qui passe à $FP_{rate} = 13,1\%$ contre $FP_{rate} = 18,2\%$ en booléen. On en conclut que **les meilleurs**

résultats sont obtenus en utilisant la logique floue pour fusionner les différentes variables. Les résultats sont alors 81,4% de bonnes détections et 13,1% de fausses alarmes.

Ce résultat se vérifie par un test de McNemar comparant les performances obtenues par fusion floue et par fusion booléenne. La table de contingence de cette comparaison est donnée dans le tableau 3.9.

		Fusion booléenne	
		Mal classés	Bien classés
Fusion floue	Mal classés	764	34
	Bien classés	180	6579

TAB. 3.9 – Table de contingence pour la comparaison entre la détection par fusion floue avec celle par fusion booléenne

Le calcul du chi-deux de McNemar donne $\chi_{McNemar}^2 = 97,79$. On constate donc que $\chi_{McNemar}^2 \gg \chi^2(1)$. On en déduit que les performances des deux algorithmes ne sont pas équivalentes. En effet, on remarque dans le tableau 3.9 qu'il y a beaucoup plus d'échantillons bien classés par la détection par fusion floue et mal classés par la détection à partir du PERCLOS80 seul ($n_{10} = 180$) que l'inverse ($n_{01} = 34$). On retrouve bien que la fusion par logique floue donne de meilleurs résultats qu'en utilisant un "OU logique".

Afin d'évaluer la sensibilité de notre méthode de détection aux seuils de réglage s_F , on réalise un tirage aléatoire des seuils s_F de chaque variable uniformément de plus ou moins 40% réparti autour des valeurs de seuils optimales utilisées. Les fonctions d'appartenance sont alors centrées sur ces nouveaux seuils. On réalise une vingtaine de tirages aléatoires qui sont tous représentés par des carrés bleus. La détection optimale précédemment trouvée est représentée par le rond rouge. Les résultats sont présentés dans la figure 3.15. À titre indicatif, les résultats monovariables ont également été placés sur la figure.

On constate sur la figure 3.15 que les résultats subissent peu de variations lorsqu'ils sont déterminés par tirage aléatoire des seuils (centré autour des valeurs précédemment trouvées). En effet, le TP_{rate} varie de 76,9% à 83,4% et le FP_{rate} de 11,4% à 14,3%. On constate que, dans tous les cas, les performances restent très correctes. De plus, cela montre que les seuils peuvent être définis *a priori* sans apprentissage (en restant cohérent avec

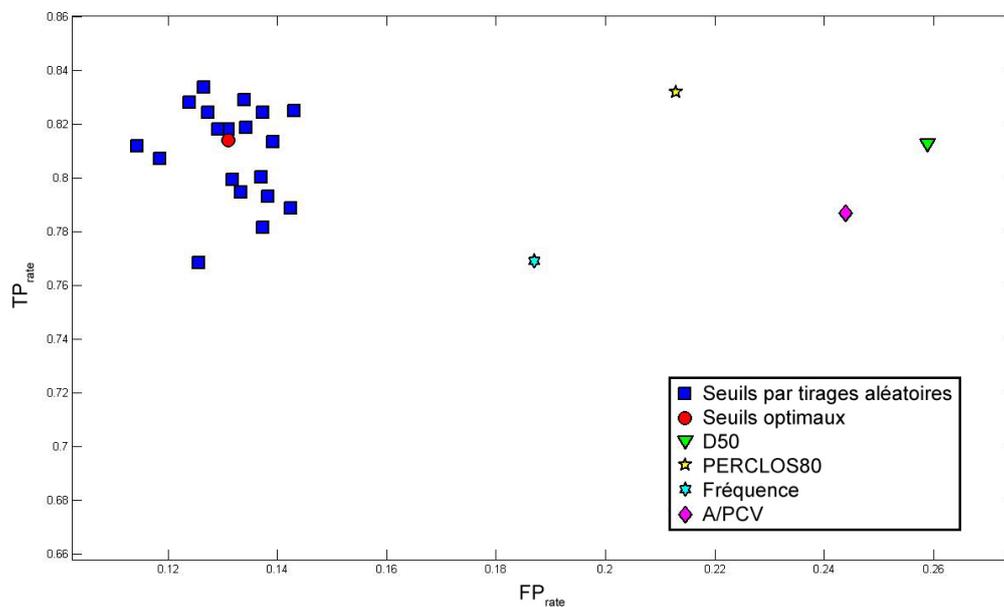


FIG. 3.15 – Résultats obtenus par tirage aléatoire des seuils de détections

les valeurs attendues pour chaque variable). En effet, l'utilisation de plages de transition floues au lieu de seuils binaires apporte une certaine tolérance quant au centrage de ces plages : les résultats restent bons même si les valeurs des seuils qui nous servent à définir les fonctions d'appartenance ne sont pas optimales.

3.4.2 Comparaison à la littérature

Lorsque nous comparons nos résultats à ceux de la littérature, on constate que notre algorithme possède de très bonnes performances, avec notamment un taux de fausses alarmes relativement faible. Il faut cependant garder certaines réserves car les résultats présentés ici n'ont pas été obtenus sur les mêmes données.

Johns ([Johns07]) avait obtenu 75% de bonnes détections et 30% de fausses alarmes en utilisant uniquement le rapport amplitude-vélocité. Ces résultats sont cohérents avec ceux que nous avons trouvés en utilisant ce paramètre. En effet, nous trouvons 78,7% de bonnes détections pour 24,4% de fausses alarmes, ce qui est relativement semblable. De plus, Johns a présenté le rapport amplitude-vélocité comme un paramètre universel indépendant des différences inter-individuelles, ce qui semble être confirmé par le fait que l'on n'obtienne

pas de meilleurs résultats en réalisant un apprentissage du seuil de détection sur les premières minutes de conduite. Svensson ([Svensson04]), qui a proposé un système de détection combinant différents indices dont le seuil est calculé par apprentissage sur le début de la conduite, n'obtient d'ailleurs que 70% de correspondance entre son système et les relevés experts.

Si l'on compare maintenant aux résultats obtenus par des méthodes d'apprentissage supervisé, on constate que notre algorithme reste très performant. Hu & Zheng ([Hu09]) ont obtenu 80,7% de bonnes détections et 16,7% de fausses alarmes avec leur méthode utilisant une SVM sur une dizaine de paramètres extraits de l'EOG, ce qui est légèrement moins bon que ce que nous obtenons. Ces résultats ont été obtenus sur une base de données réalisée à partir de 37 sujets. Il faut noter que leur méthode réalise une détection à trois niveaux de vigilance alors que notre système prend une décision binaire. Cependant, leur méthode nécessite un apprentissage sur un grand nombre de données. Les données complètes de conduite de cinq conducteurs différents sont utilisées. Omi et al. ([Omi08]) obtiennent de bonnes performances ($TP_{rate} = 84\%$ et $FP_{rate} = 9\%$) en utilisant un modèle à régression multiple sur une dizaine de paramètres. Leur technique a été validée sur deux groupes de cinq conducteurs, l'un des deux groupes servant à l'entraînement de la méthode. De plus, les auteurs restent assez évasifs sur les détails de leur méthode.

Enfin, Damousis et al. ([Damousis09]) ont proposé un système basé également sur de la logique floue pour réaliser la détection de l'hypovigilance. Leur technique utilise trois paramètres et obtient 90% de bonnes détections. Cependant, le taux de fausses alarmes est de 30% ce qui est très élevé.

Une synthèse de ces résultats est présentée dans le tableau 3.10

Au regard de ces différents résultats, on constate que notre méthode obtient effectivement de très bonnes performances. Le taux de fausses alarmes est assez faible ce qui est une bonne chose. De plus, l'avantage de cette méthode est que les seuils peuvent être définis a priori et donc ne nécessitent pas d'apprentissage. Enfin, les résultats obtenus sont fiables car l'algorithme a été testé sur un grand nombre de conducteurs.

Auteurs	TP_{rate}	FP_{rate}	Technique	Signal
Picot et al., 2009	81,4%	13,1%	Fusion floue des paramètres D50, P80, F et P80	EOG/ Vidéo
Johns, 2007	75%	30%	A/PCV uniquement	Vidéo
Omi et al., 2008	84%	9%	Régression multiple sur une dizaine de paramètres	EOG
Hu & Zheng, 2008	81%	17%	SVM sur une dizaine de paramètres	EOG
Damousis et al., 2009	90%	30%	Fusion floue des durées, fréquences et intervalles des clignements	Vidéo

TAB. 3.10 – Synthèse de la comparaison de nos résultats à ceux de la littérature concernant la détection d’hypovigilance à partir d’indices visuels

3.4.3 Approche par les arbres de décision

Les techniques d’induction d’arbres de décision sont des techniques d’extraction de connaissances à partir de données qui permettent de générer automatiquement un ensemble de règles permettant une décision. Ces techniques permettent d’acquérir une connaissance explicite à partir d’un ensemble de données et d’extraire les variables pertinentes pour la décision. Ces techniques sont expliquées plus en détail dans l’annexe C. L’objectif est ici de déterminer, à partir des données, quelles sont les variables les plus pertinentes pour détecter l’hypovigilance.

La construction de l’arbre de décision est réalisée en testant toutes les variables potentielles pour la discrimination des sous groupes afin de choisir celle qui maximise un critère donné. Dans notre cas, chaque discrimination est binaire et le critère utilisé est la minimisation du coût de mauvais classement. Il s’agit de l’algorithme *Cost Sensitive C4.5* proposé par Chauchat et al. ([Chauchat01]). L’arbre de décision est réalisé à l’aide du logiciel *Sipina* ([Sipina]).

Les variables utilisées sont celles utilisées pour l’analyse mono-variable (cf. section 3.2). Il s’agit de la durée à 50% ($D50$), du pourcentage de fermeture de l’œil à 80% ($PERCLOS80$), de la fréquence (F), du rapport amplitude-vitesse (A/PCV) et de l’énergie de fermeture de l’œil (E_{Cl}). Les moyennes et les maxima de ces indices sont calculés toutes les 20s sans recouvrement. Les moyennes sont notées *mean* et les maxima *max*. Le tableau 3.11 présente les différentes données d’entrée utilisées pour la construction

de l'arbre de décision. La figure 3.16 représente les résultats de la classification sur un sous-échantillon de la base de données prêtée par le CEPA comprenant 1907 cas "éveillé" et 1289 cas "fatigué".

Indices	<ul style="list-style-type: none"> - $D50$ moyen ($meanA50$) et maximal ($maxA50$) - $PERCLOS80$ moyen ($meanP80$) et maximal ($maxP80$) - F moyen ($meanF$) et maximal ($maxF$) - A/PCV moyen ($meanAPCV$) et maximal ($maxAPCV$) - E_{Cl} moyen ($meanEcl$) et maximal ($maxEcl$)
Fenêtre d'étude	20s sans recouvrement
Données	1907 cas "éveillé" 1289 cas "fatigué"

TAB. 3.11 – Données d'entrée utilisées pour la construction de l'arbre de décision

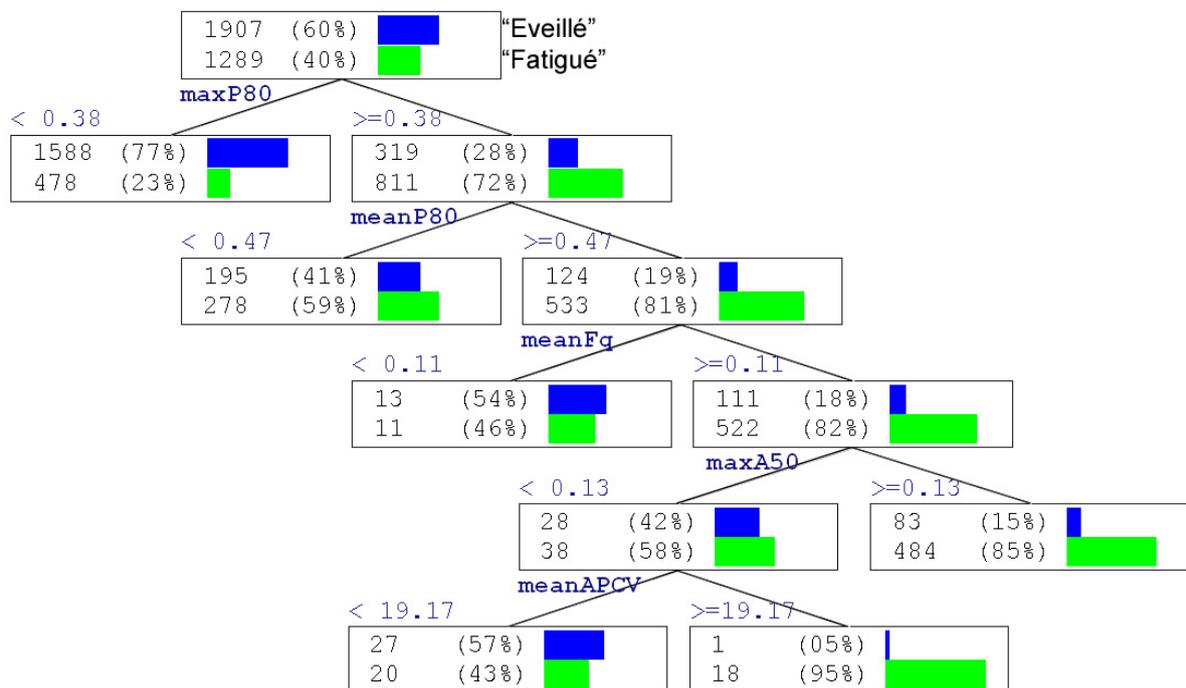


FIG. 3.16 – Arbre de décision d'hypovigilance

On constate sur la figure 3.16 que l'algorithme se sert principalement du PERCLOS80. Cela confirme ce que nous avons vu dans la section 3.2 : le PERCLOS80 est le paramètre le plus pertinent pour la détection de l'hypovigilance. La valeur des seuils utilisée pour le

PERCLOS80 est assez haute mais elle peut s'expliquer par le fait que l'on travaille sur le maximum du PERCLOS80. Concernant la valeur des seuils de D50 et de la fréquence, on constate que la valeur des seuils utilisée est relativement similaire. En revanche, le seuil utilisé pour le rapport amplitude-vélocité est un peu plus élevé mais il intervient alors que la classification est déjà bien avancée. Enfin, on constate que l'indice E_{CI} n'a pas été utilisé. Cela confirme ce que nous avons trouvé : cet indice n'est pas pertinent pour détecter l'hypovigilance.

Cette modélisation par arbre de décision nous permet de vérifier par une autre méthode le choix des indices précédemment établi : la durée D50, le PERCLOS80, la fréquence et le rapport A/PCV sont les plus pertinents pour détecter la baisse de vigilance. L'indice E_{CI} n'apporte pas d'information. De plus, on retrouve bien que le PERCLOS80 est l'indice le plus pertinent.

3.5 Conclusion du chapitre

Ce chapitre présente un système de détection d'hypovigilance par analyse vidéo des clignements. Il ressort que les paramètres les plus pertinents sont la durée des clignements à 50%, le pourcentage de fermeture de l'oeil à 80%, la fréquence et le rapport amplitude-vélocité. Ces paramètres sont étudiés uniquement sur les longs clignements ($\geq 250ms$) ce qui permet de s'affranchir des différences inter-individuelles. Des seuils fixes indépendants des conducteurs ont ainsi pu être mis en avant.

Un système de détection fusionnant les quatre paramètres par logique floue a été élaboré. Ce système a été testé sur une base de données réalisée à partir de vingt conducteurs différents et donne des performances très satisfaisantes : **81,4% de bonnes détections et 13,1% de fausses alarmes**. La fuzzification des variables est réalisée par des fonctions d'appartenance centrées sur les valeurs optimales de seuil de détection des différentes variables. Cependant, il a été montré que le système reste performant même si ces valeurs de seuil sont légèrement différentes tout en restant cohérentes avec les valeurs des variables étudiées. On peut donc ainsi définir les fonctions d'appartenance *a priori* afin d'obtenir un système performant qui ne nécessite pas d'apprentissage sur le conducteur.

Nous disposons désormais d'un système de détection d'hypovigilance par analyse de l'activité cérébrale et d'un système de détection d'hypovigilance par analyse vidéo des

clignements. Chacun de ces algorithmes s'est avéré performant sur une base de données conséquente. Nous allons maintenant tenter de fusionner ces deux approches afin d'améliorer les performances et la fiabilité de notre système de détection.

Chapitre 4

Vers une fusion des approches physiologiques et vidéo

Le but de ce chapitre est de poser les bases d'un système de fusion des données physiologiques et vidéo pour la détection d'hypovigilance. Ce système se base sur les deux algorithmes de détection développés dans les chapitres précédents. Nous cherchons ici à comprendre comment faire interagir ces deux algorithmes de manière à augmenter la fiabilité de la détection d'hypovigilance réalisée. Pour cela nous utilisons des techniques de fusion de données. Ce chapitre commence par un exposé de différentes techniques de fusion de données et un tour d'horizon des systèmes de fusion existant pour la détection de la baisse de vigilance. Un système expert permettant la fusion des approches EEG et vidéo est ensuite présenté. Enfin, les performances de ce système sont évaluées sur une base de données réalisée à partir d'une vingtaine de conducteurs différents et comparées aux performances des systèmes de la littérature.

4.1 État de l'art

Le but de cette partie est de définir ce qu'est la fusion de données. Nous exposons les avantages et les problèmes qu'elle pose. Nous présentons également les grandes lignes des différentes méthodes de fusion de données. Enfin, une revue des différents systèmes de détection d'hypovigilance qui utilisent des techniques de fusion de données est réalisée.

4.1.1 La fusion de données

Principes

La fusion de données est une technique permettant de traiter simultanément des informations provenant de différentes sources. Ceci permet d'améliorer la pertinence d'une décision par rapport à celle obtenue à partir de chaque donnée traitée de manière indépendante ([Desodt96]). Cette technique est apparue dans les années 70-80 avec la multiplication des capteurs et des sources d'information ainsi qu'avec les progrès informatiques et l'augmentation des capacités à traiter l'information.

Dans le cadre de la prise de décision, le but de la fusion de données est d'intégrer soit des informations provenant de sources multiples à un instant donné, soit d'une source unique sur une fenêtre de temps afin que l'information finale soit plus riche que la somme des informations provenant des différentes sources. Pour cela, on cherche à utiliser chaque source de manière à exploiter au mieux ses possibilités tout en palliant ses inconvénients. Ainsi, on pourra améliorer les capacités de décision d'un système ainsi que la pertinence de la décision prise.

Cependant, la diversité des informations à traiter engendre plusieurs problèmes dont il faut tenir compte.

Incertitude. Le problème d'incertitude est lié à l'acquisition des différentes informations. En effet, il arrive au cours du temps que l'acquisition subisse des perturbations. L'information acquise est alors déformée, voire incomplète. L'intégration de cette information, lorsqu'elle reste possible, est susceptible de générer des erreurs dans le système. C'est le cas par exemple lors de l'occurrence d'artéfacts sur le signal EEG. De plus, plus on a de sources d'informations, plus on a de probabilités qu'une de ces sources soit perturbée.

Perception. Le problème de perception concerne le réglage des différentes sources d'acquisition. En effet, l'information provenant de la fusion ne sera pas forcément la plus fiable si on utilise les réglages correspondant à la meilleure fiabilité des capteurs isolés. Par exemple, il est parfois nécessaire d'avoir un taux de fausses alarmes plus important sur un capteur de manière à assurer un taux de non-détections final plus faible qu'en détection isolée.

Diversité. Le problème de la diversité des types d'informations est lié à la multitude des capteurs. Les données à traiter peuvent alors être de natures très différentes : signaux physiologiques, vidéos, etc. De plus, la résolution et l'interprétation de ces signaux peuvent également être fortement hétérogènes. Il est alors parfois nécessaire d'effectuer un prétraitement local sur chaque source de manière à diminuer le nombre de données et faciliter l'intégration de celles-ci. Les problèmes de fiabilité et de dépendance des sources doivent également être gérés.

Temporalité. Le problème de temporalité vient du fait que toutes les informations ne sont pas acquises à la même vitesse. De plus, les prétraitements appliqués (filtrage, estimation, etc) génèrent un retard par rapport à l'acquisition des données brutes. Un recalage temporel est alors nécessaire pour permettre à l'algorithme de fusion de travailler conjointement sur les différentes informations.

Les différentes méthodes

De nombreuses méthodes sont utilisées pour fusionner des données. Le type de méthode utilisé est indépendant de l'application. Il est également possible de croiser différentes méthodes permettant d'utiliser de manière simultanée les avantages de chacune. Nous proposons ici un aperçu des différents types de méthodes existants.

Le raisonnement bayésien. Le raisonnement de type bayésien est une technique basée sur la théorie des probabilités et la théorie des graphes ([Jaynes03]). Elle tient son nom du type d'inférence utilisé, basé sur le théorème de Bayes. Le but de cette méthode est de mettre à jour la vraisemblance de différentes hypothèses à partir des observations de manière à trouver l'hypothèse qui semble la plus juste. Cette méthode nécessite une connaissance des probabilités *a priori* qui n'est pas toujours facile à obtenir. De plus, elle ne modélise pas l'ignorance et considère les différentes hypothèses comme mutuellement exclusives, ce qui n'est pas toujours réaliste.

Les estimateurs. L'utilisation d'estimateurs consiste à estimer un état à partir de plusieurs observations passées. Il s'agit d'une méthode d'inférence statistique. On cherche à minimiser l'erreur entre la prédiction de l'état et son observation *a posteriori*. L'estimateur

le plus connu est sans doute le filtre de Kalman ([Kalman60]) qui estime les états d'un système dynamique à partir d'une prédiction et d'une série de mesures incomplètes ou bruitées. Cette technique ne permet pas de modéliser l'incertitude. Son exactitude dépend fortement du modèle employé. Cependant il est possible d'utiliser plusieurs estimateurs en cascade afin d'améliorer les prédictions.

La théorie de l'évidence. Cette méthode est également appelée *théorie des fonctions de croyance* ou *théorie de Dempster Shafer* ([Shafer76]). Cette méthode consiste en une généralisation de l'inférence bayésienne au traitement de l'incertain. Elle permet de manipuler des évènements non nécessairement exclusifs. Il est ainsi possible de représenter explicitement l'incertitude d'un évènement, ce qui est parfois nécessaire afin de déterminer la vraisemblance d'une hypothèse basée sur des évidences multiples.

Les systèmes experts. Le but d'un système expert est de reproduire les mécanismes cognitifs d'un expert par un ensemble de règles ([Jackson98]). Il s'agit d'une méthode classique d'intelligence artificielle qui peut également servir d'outil d'aide à la décision. Un système expert est composé d'une *base de données* qui enregistre la dynamique du système, d'une *base de connaissance* qui contient les règles de décision et d'un *moteur d'inférence* qui permet de tirer des conclusions à partir des faits et des règles. La plupart des systèmes experts reposent sur des mécanismes de logique formelle et utilisent le raisonnement déductif. Le système expert peut associer différentes techniques de représentation de l'incertitude : probabilité, fonction de croyance, logique floue, etc.

Les méthodes d'optimisation. Le problème de la fusion de données peut être vu comme un problème d'optimisation combinatoire. C'est le cas des *réseaux de neurones* qui permettent la fusion d'informations de nature différente ([Lippmann87]). Cette méthode nécessite un apprentissage. Son fonctionnement est inspiré de celui des neurones biologiques. Dans sa phase d'association, des techniques de traitements parallèles sont mises en œuvre ce qui en fait une méthode très rapide pouvant être appliquée en temps réel. C'est également le cas des *arbres de décision* qui permettent de générer automatiquement un ensemble de règles permettant une décision. On peut ainsi acquérir une connaissance explicite à partir d'un ensemble de données et extraire les variables pertinentes pour la décision. Ces méthodes sont très utilisées pour résoudre des problèmes de classification.

4.1.2 Revue des systèmes existants pour la détection d'hypovigilance

Nous nous consacrons ici à la présentation des différents systèmes de la littérature qui utilisent des techniques de fusion de données pour détecter l'hypovigilance. Nous nous restreignons aux systèmes fonctionnant sur des données hétérogènes. Ainsi les systèmes utilisant la fusion de données sur plusieurs indices issus du même signal (EEG, EOG ou video) ne sont pas présentés ici, puisqu'ils l'ont déjà été dans les chapitres relatifs à ces signaux (cf. paragraphes 1.1.3, 2.1.2 et 3.1).

Des réseaux bayésiens ont été utilisés par Ji et al. pour surveiller l'hypovigilance ([Ji06]). Ils ont proposé un système utilisant à la fois des réseaux bayésiens statiques et dynamiques. Le réseau statique sert à estimer la prédisposition du conducteur à être en état d'hypovigilance en fonction de plusieurs facteurs externes tels que les conditions atmosphériques, le moment de la journée, la chaleur, les conditions de travail, la qualité du sommeil, etc. En parallèle, un réseau dynamique estime l'hypovigilance du conducteur à partir de l'évolution d'indices physiologiques tels que le mouvement des yeux, de la tête, les baillements, de la direction du regard, etc. L'estimation du niveau de vigilance du conducteur est réalisée à partir de ces deux réseaux statique et dynamique. Toutes les probabilités sont définies *a priori*, ce qui nécessite une bonne connaissance du conducteur afin d'évaluer ses conditions de sommeil ou de travail par exemple. Une période de test est cependant utilisée pour affiner les différents paramètres en fonction du conducteur. Ce système a été testé sur huit sujets. Le niveau de vigilance a été testé par un test psychomoteur *TOVA* (*Test Of Variables of Attention*, [Tova]) qui consiste à mesurer le temps de réponse d'un sujet à des stimuli visuels ou auditifs. Une corrélation $r = 0,95$ a été trouvée entre les performances moyennes au test TOVA et l'estimation moyenne du niveau de vigilance pour les huit sujets. Cette technique semble très efficace mais nécessite cependant une connaissance complète de l'environnement du conducteur ce qui n'est pas toujours possible.

Hussain et al. ([Hussain08]) ont proposé d'utiliser un modèle de mélanges gaussiens (*GMM* pour *Gaussian Mixture Model*) sur des signaux EOG et les mouvements de la tête. Un GMM est une technique de classification qui considère qu'un échantillon de données suit une loi qui est une combinaison linéaire de plusieurs fonctions de densité. Un apprentissage statistique est donc réalisé sur les différentes entrées afin de définir ces fonctions. Une fois l'apprentissage réalisé, la prise de décision va s'effectuer de manière

similaire au réseau de Bayes : on cherche à vérifier la vraisemblance d'une hypothèse de manière statistique. Hussain et al. utilise en entrée de leur modèle des courbes de Lissajous provenant de la combinaison des canaux horizontaux et verticaux de l'EOG ainsi que du signal fourni par un accéléromètre fixé au niveau de la tête du conducteur. Le GMM classe ensuite l'état du conducteur selon deux classes : "éveillé" ou "endormi". Le système a été testé sur les données de dix sujets sur des tâches simples telles que lire un livre. Un quart des données a servi à l'apprentissage du modèle. Le système a été validé sur le reste des données avec un taux de succès de 89,5% entre les états détectés par le GMM et une expertise réalisée à partir de l'enregistrement vidéo du comportement des sujets. On peut supposer que ce système détecte un stade déjà avancé d'hypovigilance car les signes d'hypovigilance au niveau comportemental interviennent relativement tard. De plus, les deux états du systèmes sont "éveillé" et "endormi", ce qui représente deux états fort distincts. Le système ne semble donc pas détecter la transition qui nous intéresse entre ces deux états.

Des systèmes de détection utilisant la théorie de l'évidence ont également été proposés. Gruyer et al. ([Gruyer05]) ont montré la faisabilité et l'efficacité d'une telle approche pour fusionner des données orientées sur le conducteur (détection par EEG, EOG, vidéo, ...), orientées sur la dynamique du véhicule (mouvements du volant, temps de croisement de ligne, position latéral du véhicule, ...) et orientées sur l'environnement du conducteur (température, heure de la journée, luminosité, etc...). Yang et al. ont également proposé un système basé sur la théorie de l'évidence ([Yang05]). Leur système utilise la vidéo pour détecter les mouvements de la tête et des yeux et les expressions faciales, des indices physiologiques comme le rythme cardiaque ainsi que des paramètres contextuels. Les trois états de sortie du système sont "éveillé", "fatigué" et "inconnu". Effectivement, l'intérêt de la théorie de l'évidence est la possibilité de considérer les hypothèses de manière non mutuellement exclusives. Ainsi on admet qu'on peut être dans un état que le système ne connaît pas et ne peut pas classer. Cependant, les deux systèmes présentés n'ont pas été validés et n'ont été testés que conceptuellement. Il est donc difficile d'évaluer leur efficacité. Plus récemment, Su & Zheng ([Su07]) ont proposé un système de détection d'accidents liés à l'hypovigilance basé sur la théorie de l'évidence. Cette méthode suggère d'utiliser les changements de direction du regard, la durée des clignements ainsi que la position latérale du véhicule pour détecter l'hypovigilance. Elle a été testée sur les données de conduite de onze conducteurs sur un simulateur de conduite. Ce système peut prédire l'occurrence d'accidents dans les 20s suivantes avec un taux de succès de 71%. Cependant, il génère 42% de fausses alarmes. De plus, la durée de 20s entre la prédiction et l'occurrence

de l'accident est trop courte. En effet, il semble plus efficace d'avertir le conducteur avant que l'hypovigilance ne soit trop avancée et que l'accident soit inévitable.

Un système expert pour l'évaluation du niveau de vigilance a été proposé par Boverie et al. ([Boverie08]). Ce système propose de fusionner des informations physiologiques du conducteur, des informations concernant le véhicule ainsi que des informations contextuelles par logique floue afin de donner une évaluation à quatre niveaux de la vigilance du conducteur. Les indices utilisés sont les clignements des yeux, la position latérale du véhicule, les mouvements du volant, le moment de la journée ainsi que la durée de conduite. Les clignements des yeux et la position du véhicule sont détectés par caméra. Le système a été testé sur les données de conduite de douze conducteurs. Bien qu'il semble efficace sur les exemples présentés, les auteurs ne donnent pas de données précises quant à la validation de leur système.

Enfin, Khardi & Hernandez-Gress ont proposé un système hybride de diagnostic de l'état de vigilance du conducteur très complet ([Khardi00]). Ce système fonctionne à partir des paramètres mécaniques, physiologiques et comportementaux. Ainsi les mouvements des yeux et de la tête, la pression sur les freins, la position de l'accélérateur, la position des lignes blanches de la route, le régime moteur, la force exercée sur le volant, la vitesse du véhicule, l'angle du volant ainsi que la présence d'obstacles sur la route sont enregistrés en continu et fusionnés par réseau de neurones de manière à évaluer le niveau de vigilance du conducteur. Une analyse en composantes indépendantes (ICA) est réalisée à partir des différents indices de manière à obtenir des variables synthétiques à partir des variables d'entrée. Un apprentissage est alors réalisé par réseaux de neurones gaussiens généralisés. Aucun détail n'est donné sur la période d'apprentissage. Cette méthode a été testée sur les données de conduites de dix conducteurs et comparée à l'expertise des signaux physiologiques EEG et EOG. Cet algorithme obtient un taux de bonnes détections de 93,5% et seulement 0,5% de fausses alarmes. On peut cependant s'interroger sur le temps d'apprentissage ainsi que sur la robustesse de cette méthode face aux différences inter-individuelles.

L'utilisation de techniques de fusion de données pour détecter l'hypovigilance chez le conducteur à partir de données issues de différentes approches (physiologiques, comportementales, vidéo, ...) est relativement récente. Il existe encore peu de systèmes essayant d'associer différentes approches pour améliorer les résultats. De plus, parmi les systèmes existants, des techniques de fusion très différentes sont utilisées. Nous proposons de développer un système de fusion des approches EEG et vidéos présentées dans les chapitres 1

et 3. Ces deux approches utilisant la logique floue, on s'oriente vers un système de fusion du type "système expert" pour réaliser la détection d'hypovigilance.

4.2 Notre méthode

Nous présentons dans cette partie le système que nous avons développé afin de fusionner les approches de détection d'hypovigilance par EEG et par vidéo. Selon le tableau 1.1 du chapitre 1 qui a servi à l'expertise des données, l'approche par EEG effectue une détection du niveau 1 d'hypovigilance caractérisé par des signes de baisse de vigilance au niveau de l'activité cérébrale mais pas de l'activité oculaire. Par contre, l'approche par vidéo effectue une détection du niveau 2 d'hypovigilance caractérisé par des signes de baisse de vigilance à la fois dans l'activité cérébrale et dans l'activité oculaire du conducteur. La détection par caméra rapide effectue donc une détection plus tardive de l'hypovigilance. Cela est rappelé dans le tableau 4.1 qui présente l'échelle OSS restreinte aux niveaux 1 et 2 d'hypovigilance.

Objective Sleepiness Score	Contenu EEG	Clignements et mouvements d'yeux
0	Activité β continue en fond, pas d'activité α et θ	et clignements et mouvements d'yeux normaux
1	Apparition d'ondes α et/ou θ pendant au moins 5s cumulées	et clignements et mouvements d'yeux normaux
2	Apparition d'ondes α et/ou θ pendant au moins 5s cumulées	et clignements et mouvements d'yeux lents

TAB. 4.1 – Critères OSS restreints aux niveaux 1 et 2 d'hypovigilance

On décide d'effectuer d'abord la détection EEG puis d'effectuer la détection vidéo si des signes d'hypovigilance ont été détectés à partir de l'EEG. Ainsi la détection prise à partir de la vidéo dépendra de celle prise à partir de l'EEG. En effet, la décision prise par la vidéo est censée rendre compte d'un niveau d'hypovigilance où à la fois les activités cérébrales et oculaires présentent des signes de l'hypovigilance. Ainsi, l'architecture générale de notre système de fusion pour la détection de l'hypovigilance correspond au schéma de la figure 4.1.

Sur cette figure, les valeurs 0 et 1 en sortie des modules de détection EEG et vidéo correspondent au fait d'avoir détecté des signes d'hypovigilance sur 20s de signal considéré.

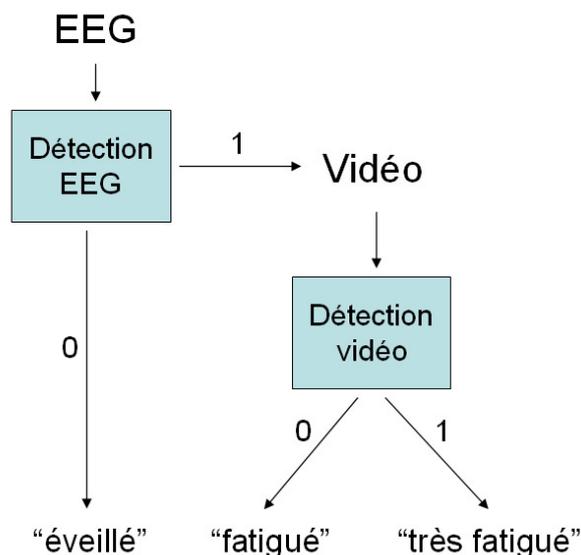


FIG. 4.1 – Architecture générale du système de fusion

Afin de comparer nos résultats aux résultats de l'expert, nous choisissons d'effectuer la fusion des signaux EEG et vidéo par blocs de 20s. Ainsi, si aucun signe d'hypovigilance n'est détecté sur les 20s considérées, la sortie est à 0 et si des signes d'hypovigilance sont détectés, elle est à 1. La méthode de fusion peut alors se comprendre de la manière suivante. On effectue d'abord la détection sur le signal EEG. Si aucun signe de baisse de vigilance n'est détecté sur l'EEG, alors le conducteur est considéré comme "éveillé". Si des signes d'hypovigilance sont détectés, on regarde alors la détection vidéo. Si aucun signe n'est détecté sur la vidéo, alors on est dans un niveau 1 d'hypovigilance car on observe des signes au niveau de l'EEG mais pas au niveau de la vidéo. Le conducteur est alors considéré comme "fatigué". Sinon, on est dans un niveau 2 d'hypovigilance car on observe des signes à la fois sur l'EEG et la vidéo. Le conducteur est alors considéré comme "très fatigué". L'ensemble des mécanismes de décision est détaillé dans les paragraphes suivants.

4.2.1 Les détections EEG et video

Le but de cette section est de rappeler brièvement les techniques de détection que nous avons développées pour l'EEG et sur la vidéo. Ces techniques ont été présentées dans les chapitres précédents.

Détection EEG

Le système de détection EEG a été présenté dans la section 1.3.6. Ce système fonctionne à partir des signaux α_{rel} et β_{rel} calculés à partir de l'EEG. Ces signaux sont filtrés médian (med_α et med_β) puis on effectue un test de comparaison de moyenne (MCT) sur chacun d'eux (MCT_α et MCT_β). On calcule le degré d'appartenance à la classe "fatigué" des résultats de la MCT pour chacun des deux signaux ($\mu_{Dr}(MCT_\alpha)$ et $\mu_{Dr}(MCT_\beta)$). On effectue alors la fusion de ces signaux par logique floue comme présenté dans le paragraphe 1.3.6. Le résultat de cette fusion d_{EEG} permet de détecter l'hypovigilance de niveau 1 chez le conducteur. On considère alors la détection par blocs de 20s. La décision sur 20s est notée D_{EEG} . Elle est définie par l'équation (4.1), ce qui correspond à la technique de validation utilisée dans le chapitre 1.

$$D_{EEG}(i) = \begin{cases} 1 & \text{si } Card_{j \in \{i-20:i\}} \{d_{EEG}(j) = 1\} \geq 2 \\ 0 & \text{sinon} \end{cases} \quad (4.1)$$

Un retard de 20s est introduit sur la décision par les différentes fenêtres de traitement. De plus, ce système est couplé à un détecteur d'artéfacts qui permet d'évaluer la qualité du signal EEG. La sortie du détecteur d'artéfacts est la variable d_{art} qui vaut "1" lorsqu'un artéfact est détecté et "0" dans le cas contraire. Cette variable est calculée toutes les 10s sans recouvrement. On définit D_{art} la détection d'artéfact sur 20s. Elle est calculée selon l'équation (4.2). Si, sur les 20s considérées, au moins un artéfact a été détecté, alors D_{art} vaut 1.

$$D_{art}(i) = \begin{cases} 1 & \text{si } d_{art}(i-20:i-10) = 1 \mid d_{art}(i-10:i) = 1 \\ 0 & \text{sinon} \end{cases} \quad (4.2)$$

Détection vidéo

Le système de détection vidéo a été présenté dans la partie 3.3. Ce système utilise les indices suivant : la durée à 50% des clignements, le pourcentage de fermeture de l'œil à 80%, la fréquence des clignements et le rapport amplitude-vélocité. Ces indices sont calculés sur une fenêtre de 20s et notés respectivement $D50_m$, $P80_m$, F_m et A/PCV_m . On

calcule alors le degré d'appartenance des différents indices (μ_{D50} , μ_{P80} , μ_F et $\mu_{A/PCV}$) à la classe "fatigué". Les indices sont fusionnés par logique floue à l'aide d'une somme pondérée. Le résultat de cette fonction $d_{vidéo}$ permet de détecter l'hypovigilance de niveau 2 chez le conducteur. Un retard de 10s est introduit sur la décision par la chaîne de traitement. La détection est également ici considérée par bloc de 20s. Elle est notée $D_{vidéo}$ et est définie par l'équation (4.3) en accord avec la technique de validation utilisée dans le chapitre 3.

$$D_{vidéo}(i) = \begin{cases} 1 & \text{si } Card_{j \in \{i-20:i\}} \{d_{vidéo}(j) = 1\} \geq 10 \\ 0 & \text{sinon} \end{cases} \quad (4.3)$$

4.2.2 Le système de fusion

Nous présentons ici l'algorithme complet de détection d'hypovigilance par fusion des signaux EEG et vidéo. Le schéma de principe est présenté sur la figure 4.2.

La première étape de notre algorithme consiste à effectuer la détection d'artéfacts sur le signal EEG ainsi que les détections d'hypovigilance sur les signaux EEG et vidéo à partir des différents indices présentés précédemment. Les différentes fenêtres de calcul utilisées pour ces différentes détections induisent des latences respectives de 20s pour la détection EEG et de 10s pour la détection vidéo. Il est donc nécessaire d'effectuer une synchronisation temporelle des différentes détections afin que, lorsqu'on effectue la fusion des décisions d_{art} , d_{EEG} et $d_{vidéo}$ à l'instant i , elles fassent toutes références à la même portion de signal. Cette synchronisation est expliquée dans le paragraphe 4.2.3. Une fois ces indices synchronisés, on effectue la fusion des détections par blocs de 20s de la manière suivante. On commence par regarder le résultat de la détection d'artéfacts D_{art} pour évaluer la qualité du signal EEG. Si celui-ci ne présente pas d'artéfact, on regarde le résultat de la détection EEG. Dans le cas contraire, on ne prend pas en compte la détection EEG (car elle risque d'être erronée à cause de la présence d'artéfacts) et on passe directement à la détection vidéo. Dans le cas où l'on considère la détection EEG, on regarde si des signes d'hypovigilance ont été trouvés dans le signal EEG. Si aucun signe n'est trouvé, on considère le conducteur comme "éveillé" ce qui correspond au niveau 0 de l'échelle OSS. Dans ce cas, on ne considère pas la détection vidéo puisque selon cette échelle, les signes oculaires apparaissent en présence de signes EEG. Si des signes d'hypovigilance sont détectés sur l'EEG, on considère alors la détection vidéo. Si aucun signe visuel d'hypovigilance n'est détecté, on est dans le cas du niveau 1 de

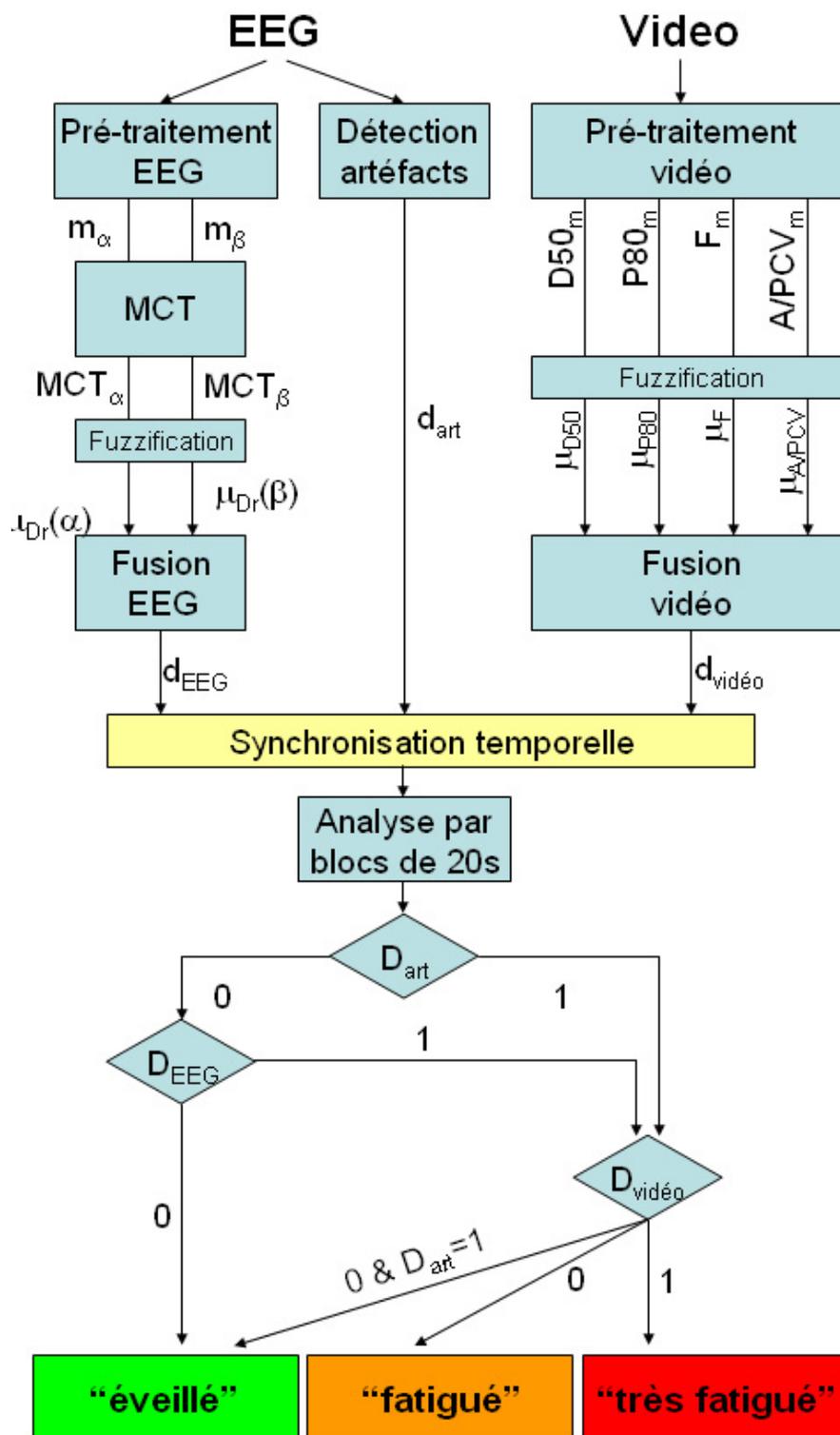


FIG. 4.2 – Architecture complète du système de fusion

l'échelle OSS. On considère alors le conducteur comme "fatigué". Dans le cas où aucun signe d'hypovigilance n'est détecté par l'analyse vidéo et qu'un artefact a été détecté (on ne prend donc pas en compte la détection EEG), on considère que le conducteur est "éveillé" (niveau 0 de l'échelle OSS). Enfin, dans le cas où l'on détecte des signes d'hypovigilance sur le signal vidéo et sur le signal EEG, on est alors dans le cas du niveau 2 de l'échelle OSS (présence de signes cérébraux et visuels d'hypovigilance) et on considère le conducteur comme "très fatigué". L'algorithme 2 décrit le processus de décision.

4.2.3 Synchronisation temporelle

Cette section présente la synchronisation temporelle qui est réalisée sur les différentes détections afin de réaliser leur fusion. Nous commençons par présenter la synchronisation des détections EEG et vidéo puis celle avec la détection d'artefacts.

On rappelle que la détection EEG est réalisée toutes les secondes. Les fenêtres utilisées pour cette détection introduisent une latence de 20s (cf. section 1.2.4). La détection vidéo est également calculée toutes les secondes. La fenêtre de calcul utilisée introduit une latence de 10s (cf. section 3.2.4). On souhaite lors de la fusion des détections d_{EEG} et $d_{vidéo}$, qu'elles fassent toutes deux référence à un même instant t . L'instant i correspond à l'instant où est calculé d_{EEG} . Ainsi, la détection $d_{EEG}(i)$ fait référence à l'instant $t = i - 20s$ du signal EEG. On souhaite avoir la détection vidéo $d_{vidéo}$ correspondant au même instant t . On souhaite donc trouver l'instant j tel que la détection $d_{vidéo}(j)$ corresponde à l'instant t du signal EOG. On sait que $t = j - 10s$ puisque la latence est de 10s pour la détection vidéo. On peut ainsi exprimer j en fonction de i et on trouve $j = i - 10s$. Ainsi pour synchroniser les détections EEG et vidéo, il faudra fusionner $d_{EEG}(i)$ avec $d_{vidéo}(i - 10)$.

Les détections EEG et vidéo sont ensuite analysées au moment de la fusion par blocs de 20s sans recouvrement comme le fait l'expert. On travaille donc sur les détections $D_{EEG} = d_{EEG}(i - 20 + 1 : i)$ et $D_{vidéo}(i - 10) = d_{vidéo}(i - 30 + 1 : i - 10)$. La détection d'artefacts a lieu toutes les 10s sans recouvrement. Afin de synchroniser la détection d'artefacts D_{art} avec les détections EEG et vidéo sur les 20s étudiées, on considère les deux détections d'artefacts correspondant à ce segment de 20s. Il s'agit donc des détections $D_{art}(i - 30) = d_{art}(i - 40 + 1 : i - 30)$ et $D_{art}(i - 20) = d_{art}(i - 30 + 1 : i - 20)$ puisque la détection d'artefacts a lieu directement sur le signal EEG.

En réalisant ainsi la synchronisation temporelle, on est sûr que les informations fu-

Algorithme 2 Détection d'hypovigilance par fusion EEG et vidéo

ENTRÉES: EEG et vidéo

SORTIES: $D \in \{\text{“éveillé”}; \text{“fatigué”}; \text{“très fatigué”}\}$

$d_{art} \leftarrow \text{detection_artifact}(EEG)$

Prétraitement EEG : calcul de m_α et m_β

$MCT_\alpha \leftarrow MCT(m_\alpha)$

$MCT_\beta \leftarrow MCT(m_\beta)$

Fuzzification : $\mu_{Dr}(\alpha)$ et $\mu_{Dr}(\beta)$

$d_{EEG} \leftarrow \text{fusion}(\mu_{Dr}(\alpha), \mu_{Dr}(\beta))$

Prétraitement vidéo : calcul de $D50_m$, $P80_m$, F_m et A/PCV_m

Fuzzification : μ_{D50} , μ_{P80} , μ_F , et $\mu_{A/PCV}$

$d_{vidéo} \leftarrow \text{fusion}(\mu_{D50}, \mu_{P80}, \mu_F, \mu_{A/PCV})$

Synchronisation temporelle (cf. paragraphe 4.2.3)

Calcul de D_{EEG} , $D_{vidéo}$ et D_{art}

si $D_{art} = 0$ **alors**

si $D_{EEG} = 0$ **alors**

$D \leftarrow \text{“éveillé”}$

sinon

si $D_{vidéo} = 0$ **alors**

$D \leftarrow \text{“fatigué”}$

sinon

$D \leftarrow \text{“très fatigué”}$

fin si

fin si

sinon

si $D_{vidéo} = 0$ **alors**

$D \leftarrow \text{“éveillé”}$

sinon

$D \leftarrow \text{“très fatigué”}$

fin si

fin si

Retourner D

sionnées font référence à la même période des signaux EEG et vidéo.

4.2.4 Intérêt de la méthode

Le principal avantage de notre système de fusion est que la décision prise sur la vidéo affine la décision prise à partir du signal EEG, permettant ainsi d’obtenir un système de détection à trois niveaux [“éveillé” ; “fatigué” ; “très fatigué”]. De plus, comme la vidéo permet la détection d’un niveau d’hypovigilance où des signes cérébraux et visuels de l’hypovigilance sont présents, l’apport de la détection EEG devrait permettre de confirmer la détection vidéo et de diminuer le nombre de fausses alarmes générées par cette dernière.

En effet, la détection vidéo n’est prise en compte que lorsque des signes d’hypovigilance sont détectés dans l’EEG, conformément à l’échelle d’évaluation OSS. Ainsi, la détection vidéo est conditionnée par la détection EEG. En revanche, nous ne sommes pas en mesure de valider les décisions EEG à partir d’autres mesures. Les performances de notre système sont donc limitées par celles du système de détection EEG, c’est-à-dire que tout ce qui aura été mal classé par la détection EEG restera mal classé.

Le détecteur d’artéfacts permet d’évaluer la qualité du signal EEG. Quand un artéfact est détecté sur l’EEG, cela signifie que la détection EEG n’est pas fiable. La décision d’hypovigilance est alors réalisée uniquement à partir du signal vidéo. Cela permet de diminuer les fausses classifications dues à la dégradation du signal EEG par des artéfacts, comme nous l’avons montré dans le paragraphe 1.3.4 du chapitre.

La synchronisation temporelle permet de prendre en compte le décalage temporel qui existe entre les différents indices. Des retards de durées différentes sont induits sur les indices et il est absolument nécessaire de tout synchroniser de manière à ce que les décisions prises sur chacun des signaux EEG et vidéo correspondent aux mêmes instants.

De plus, notre système bénéficie des avantages des algorithmes de détection EEG et vidéo. Ces avantages sont la robustesse aux différences inter-individuelles qui est caractéristique des deux systèmes de détection ainsi que le faible temps d’apprentissage. En effet, seule une minute d’apprentissage est nécessaire pour la détection EEG.

4.3 Résultats et discussions

Nous présentons dans cette partie les résultats obtenus par notre système sur une large base de données expertisée. Ces résultats sont commentés et comparés à ceux de la littérature. Enfin, la méthode utilisée est comparée à la modélisation obtenue à l'aide d'un arbre de décision.

4.3.1 La base de données

La méthode a été testée sur la base de donnée fournie par le CEPA et présentée dans la section 1.3.1. Comme pour le chapitre précédent, nous ne disposons pas de données vidéo d'hypovigilance expertisées et synchronisées avec des signaux EEG. Nous avons donc décidé de tester notre système de fusion en remplaçant la vidéo haute fréquence par le signal EOG. En effet, il a été montré dans le chapitre 3 que comme notre système de détection vidéo utilise des indices qui peuvent être récupérés avec la même précision par EOG et par vidéo. Ainsi, nous pouvons travailler sur la base de données du CEPA qui représente 60 heures de conduites enregistrées à partir de vingt conducteurs différents. L'avantage, en plus du nombre considérable de données, est que toutes ces données ont été expertisées par un spécialiste.

4.3.2 Résultats obtenus

Notre système est testé sur toute la base de données du CEPA soit 7557 échantillons expertisés. Les résultats sont comparés à ceux de l'expert. On considère comme "fatigué" les décisions expertes égales à 1 et comme "très fatigué" celles supérieures ou égales à 2. Les résultats sont présentés dans le tableau 4.2. Les colonnes correspondent au classement de l'expert et les lignes au classement fait par notre système. Les pourcentages ont été calculés sur chaque colonne.

La première chose que l'on constate sur ce tableau est que les résultats sont bons avec environ 80% de classifications correctes. Comme prévu, les performances sont bornées par la détection EEG. Le taux de bonnes classifications de l'état "éveillé" est de 82,1%, ce qui correspond à un taux de fausses alarmes de 17,9% (échantillons classés "fatigué" ou "très fatigué" au lieu de "éveillé"). On retrouve bien le taux de fausses alarmes optimal obtenu

		Expert		
		“éveillé”	“fatigué”	“très fatigué”
Notre système	“éveillé”	82,1%	14,0%	3,7%
	“fatigué”	14,5%	72,6%	16,7%
	“très fatigué”	3,4%	13,4%	79,6%
Nombre d'échantillons		5512	949	1096

TAB. 4.2 – Résultats obtenus par notre système de détection

pour la détection EEG (cf. chapitre 1). Le taux de bonnes classifications de l'état “fatigué” est de 72,6%. C'est un peu faible mais c'était prévisible. En effet, le taux de bonnes détections EEG est de 84,6% et le taux de fausses alarmes de la détection vidéo est de 13,1% (cf. chapitre 3). Cela signifie que parmi les 84,6% de bonnes détections, seuls 86,9% ($= 100\% - 13,1\%$) seront bien détectés dans l'état “fatigué”. Cela correspond à 73,5% ($= 84,6\% \times 86,9\%$) des états classés “fatigué” par l'expert, ce qui n'est pas très éloigné des 72,6% obtenus. Ainsi on pourrait s'attendre à observer le même phénomène pour la détection des états “très fatigué”. On trouve cependant un taux de bonnes détections de 79,6% pour l'état “très fatigué”, ce qui est proche des 81,4% obtenus pour la détection vidéo. On en déduit que la validation par l'EEG est efficace. Ainsi, les non-détections d'état d'hypovigilance semblent être principalement des états “fatigué” plutôt que des “très fatigué”.

Le tableau 4.3 présente les résultats obtenus en effectuant une fusion par “OU logique” entre les détections EEG et vidéo. On peut qualifier cette fusion d' “aveugle” car elle est réalisée sans prendre en compte les liens qui peuvent exister entre les différentes approches. Cela signifie que l'on considère les détections EEG et vidéo indépendamment l'une de l'autre.

On constate dans ce tableau que l'on retrouve le taux optimal de 81,4% de détection vidéo pour les états “très fatigué”. Cela est cohérent avec la fusion aveugle puisque dès que des signes d'hypovigilance sont détectés par vidéo, le conducteur est considéré comme “très fatigué” quelque soit le résultat de la détection EEG. Ce taux de bonnes dé-

		Expert		
		“éveillé”	“fatigué”	“très fatigué”
Notre système	“éveillé”	75,2%	10,9%	1,8%
	“fatigué”	14,9%	72,6%	16,8%
	“très fatigué”	9,9%	16,4%	81,4%

TAB. 4.3 – Résultats obtenus par fusion “aveugle”

tections est légèrement meilleur que celui obtenu avec notre système de détection (79,6%). En revanche, la fusion aveugle augmente fortement le nombre de fausses alarmes car on constate que le taux de bonnes détections de l'état “éveillé” est fortement diminué. Il passe de 82,1% pour notre système de fusion à 75,2% pour le système de fusion aveugle. Le taux de bonnes détections de l'état “fatigué” reste inchangé entre les deux systèmes de fusion. Cela semble logique puisque dans les deux systèmes ces taux correspondent aux cas où des signes d'hypovigilance ont été détectés sur l'EEG mais pas sur la vidéo. À la vue de ces différentes remarques, il apparaît que notre système est plus performant qu'un système de fusion aveugle.

Nous procédons à un test de McNemar (cf. section 1.3.2) pour comparer les performances de notre algorithme de fusion à celles de l'algorithme de fusion “aveugle”. La table de contingence de cette comparaison est donnée dans le tableau 4.4.

		Algorithme de fusion “aveugle”	
		Mal classés	Bien classés
Notre algorithme	Mal classés	1397	27
	Bien classés	378	5755

TAB. 4.4 – Table de contingence pour la comparaison entre notre algorithme avec un algorithme de fusion “aveugle”

Le calcul du chi-deux de McNemar donne $\chi_{McNemar}^2 = 302,47$. On constate donc que $\chi_{McNemar}^2 \gg \chi^2(1)$. On en déduit que les performances des deux algorithmes ne

sont pas équivalentes. En effet, on remarque dans le tableau 4.4 qu’il y a beaucoup plus d’échantillons bien classés par notre algorithme et mal classés par l’algorithme de fusion “aveugle” ($n_{10} = 378$) que l’inverse ($n_{01} = 27$). On en conclut que notre algorithme est plus performant qu’un algorithme de fusion “aveugle”.

Pour vérifier l’apport de l’information EEG pour la détection vidéo, nous calculons les coefficients ROC de la détection du niveau “très fatigué” (par rapport aux états “éveillé” et “fatigué” confondus) par fusion vidéo et EEG. Cette détection peut ainsi être comparée à la détection vidéo présentée dans le chapitre 3. On trouve $TP_{rate} = 79,6\%$ et $FP_{rate} = 5,2\%$. Ce résultat est représenté par le rond rouge sur la figure 4.3. Il est comparé au résultat optimal obtenu pour la détection vidéo (cf. section 3.4.1) et représenté par le carré bleu.

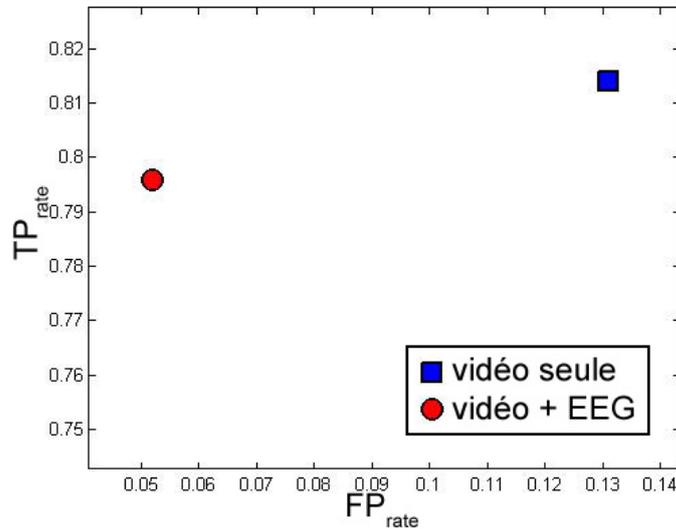


FIG. 4.3 – Comparaisons des résultats obtenus pour la détection vidéo validée par la détection EEG et la détection vidéo seule

On remarque sur cette figure que le taux de bonnes détections est légèrement diminué lorsqu’on valide les résultats par la détection EEG : on passe de $TP_{rate} = 81,4\%$ à $TP_{rate} = 79,6\%$. En revanche, la validation par l’EEG permet de diminuer fortement les fausses alarmes. On passe de $FP_{rate} = 13,1\%$ lorsqu’on effectue la détection à partir de la vidéo seule à $FP_{rate} = 5,2\%$ quand on valide cette dernière par la détection EEG. On peut donc conclure que l’EEG permet de diminuer considérablement le nombre de fausses alarmes de la détection vidéo. Cela permet également d’expliquer les meilleures performances observées pour notre système par rapport à un système de fusion aveugle. On en déduit que les signes visuels d’hypovigilance apparaissent effectivement après les signes cérébraux détectés par l’EEG, comme présenté dans le tableau 1.1 du chapitre 1.

De la même manière que précédemment, on peut procéder à un test de McNemar de manière à comparer les performances obtenues en utilisant la vidéo seule à celles obtenues en validant la détection vidéo par la détection EEG. La table de contingence de cette comparaison est donnée dans le tableau 4.5.

		Détection vidéo seule	
		Mal classés	Bien classés
Notre algorithme	Mal classés	528	21
	Bien classés	490	6583

TAB. 4.5 – Table de contingence pour la comparaison de la fusion des détection vidéo et EEG à la détection vidéo seule

Le calcul du chi-deux de McNemar donne $\chi_{McNemar}^2 = 428,62$. On constate donc que $\chi_{McNemar}^2 \gg \chi^2(1)$. On en déduit que les performances des deux algorithmes ne sont pas équivalentes. En effet, on remarque dans le tableau 4.5 qu'il y a beaucoup plus d'échantillons bien classés lorsqu'on fusionne vidéo et EEG et mal classés par la détection vidéo seule ($n_{10} = 490$) que l'inverse ($n_{01} = 20$). On en conclut que la détection est plus performante lorsqu'on fusionne EEG et vidéo que lorsqu'on utilise la vidéo seule.

Afin d'illustrer les différentes remarques, les figures 4.4 et 4.5 montrent des exemples de détection d'hypovigilance réalisées respectivement sur les sujets n° 12 et 8. Sur ces figures, les deux premières courbes représentent les détections EEG et vidéo. Le troisième graphique présente une comparaison entre le relevé réalisé par l'expert (en rouge) et le résultat de notre système de détection (en pointillés verts). On rappelle que le niveau 2 de notre détection correspond aux niveaux 2, 3 et 4 de l'expert.

On remarque sur la figure 4.4 que la détection d'hypovigilance de notre système correspond avec le relevé effectué par le médecin au niveau de la forme générale. On remarque cependant au niveau du cadre en pointillés rouge que nous sommes moins précis que l'expert pour détecter les nuances comme en atteste la non-détection de plusieurs oscillations de 20s entre l'état 1 et l'état 2 d'hypovigilance. Cela est dû aux fenêtres d'étude utilisées qui lissent les signaux EEG et vidéo. Les variations ponctuelles sont alors estompées et seules les fortes variations sont détectées.

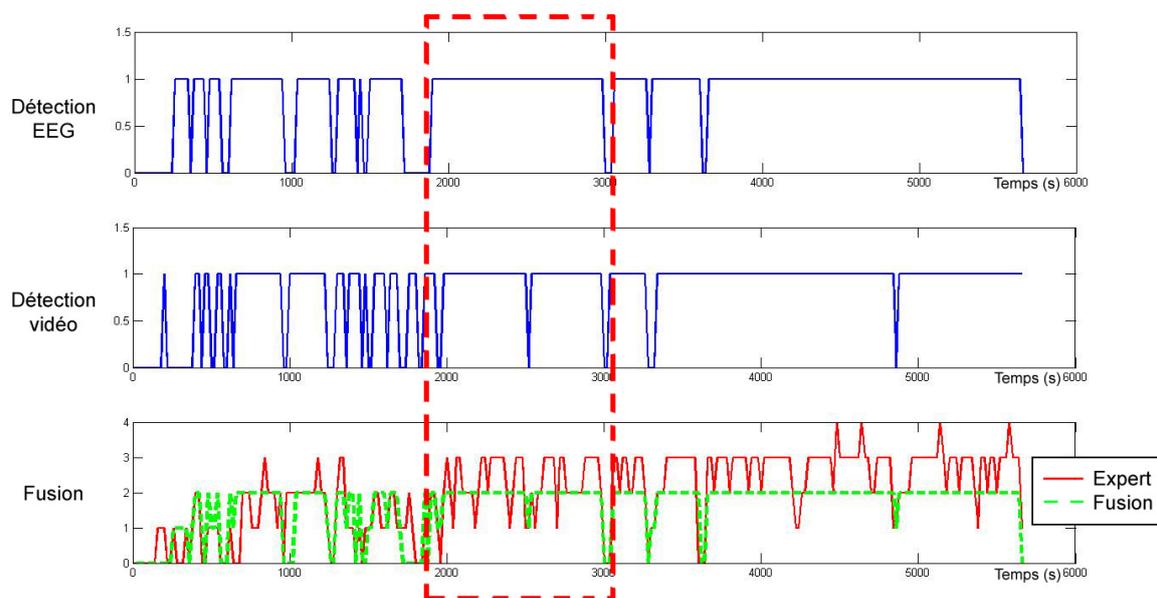


FIG. 4.4 – Detections d'hypovigilance pour le sujet n° 12

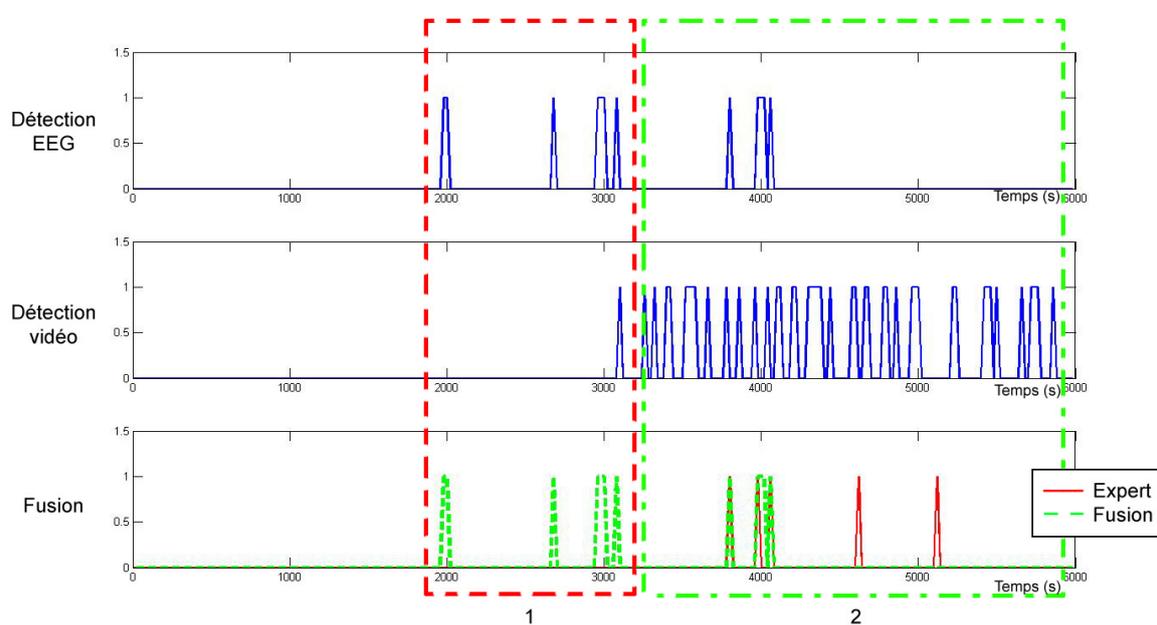


FIG. 4.5 – Detections d'hypovigilance pour le sujet n° 8

On peut voir sur la figure 4.5 que notre système produit des fausses alarmes par rapport au relevé de l'expert. Cela est visible dans le cadre 1 en pointillés réguliers rouges. On constate que comme on ne dispose pas d'informations pour valider la détection EEG, si

cette dernière est fausse, cela produira des fausses alarmes dans notre système final. En revanche, la validation de la détection vidéo par la détection EEG permet d'éviter des fausses alarmes qui auraient pu être générées par des détections vidéo erronées, comme on peut le constater dans le cadre 2 en pointillés irréguliers verts. On constate, dans ce cadre, que de nombreuses fausses alarmes sont produites par la détection vidéo (deuxième relevé) mais elles n'interviennent pas dans la détection finale grâce à la validation par la détection EEG.

4.3.3 Comparaison avec la littérature

La revue de la littérature sur les systèmes de détection d'hypovigilance par fusion de données hétérogènes présentée dans la partie 4.1.2 a montré qu'il existait assez peu de systèmes de ce type. De plus, seule la moitié de ces systèmes a été testée sur des bases de données d'hypovigilance afin d'attester de leurs performances. Enfin, il est nécessaire de garder une certaine réserve car les résultats présentés ici n'ont pas été obtenus sur les mêmes données que nous. Le fait qu'il n'existe pas de base de données de référence rend cette comparaison subjective.

Ji et al. ([Ji06]) obtiennent une corrélation $r = 0,95$ entre l'estimation moyenne d'hypovigilance de leur système et les performances moyennes de huit sujets à un test psychomoteur d'hypovigilance. On rappelle que leur système effectue une estimation du niveau de vigilance à partir de réseaux de Bayes. Ces réseaux utilisent en entrée des indices dynamiques liés au sujet comme la direction du regard, les baillements ou les mouvements des yeux ainsi que des indices statiques très complets de l'environnement du sujet (heure du jour, conditions de sommeil, luminosité, etc). Ce système semble plus performant que le nôtre mais il nécessite une connaissance très complète de l'environnement du conducteur, ce qui n'est pas toujours réalisable.

L'utilisation de modèles de mélanges gaussiens sur des données EOG et les mouvements de la tête a permis à Hussain et al. ([Hussain08]) d'obtenir un taux de bonnes détections de 89,5% sur une dizaine de sujets. Cependant, leur système détecte un état d'endormissement des sujets avec notamment la détection d'éventuels "décrochages" de la tête au moment de l'endormissement. Cet état correspond à un état d'hypovigilance très avancé. Il semble indispensable d'être capable de prévenir le conducteur suffisamment longtemps avant que celui ne s'endorme. Ce système semble donc effectuer une détection trop tardive. Notre système a pour objectif la détection des signes précoces d'hypovigilance. Il

semble donc délicat de comparer les deux systèmes.

Khaldi & Hernandez-Gress ([Khaldi00]) ont proposé un système utilisant des réseaux de neurones sur un ensemble très complet d'indices mécaniques (pression sur le frein et sur l'accélérateur), physiologiques (mouvements des yeux et de la tête) et comportementaux (position de la voiture par rapport à la route, présence d'obstacle). Ce système obtient 93,5% de bonnes détections d'états d'hypovigilance et seulement 0,5% de fausses alarmes. Ce système est très performant, cependant peu d'informations sont données par rapport à la phase d'apprentissage. En effet, un tel système peut nécessiter un apprentissage relativement long. De plus, rien n'est dit sur sa robustesse aux différences inter-individuelles. Notre système, bien que moins performant, ne nécessite pas de long apprentissage (seul un apprentissage en ligne d'une minute est effectué sur le conducteur pour la détection EEG). De plus, il est robuste aux différences inter-individuelles puisque les résultats ont été obtenus sur une vingtaine de conducteurs différents sans avoir à changer les paramètres d'un conducteur à l'autre.

Enfin, la théorie de l'évidence a été utilisée par Su & Zheng ([Su07]). Ce système effectue une détection d'hypovigilance à partir de la direction du regard, des clignements des yeux et de la position latérale du véhicule. Ce système a été testé sur une dizaine de conducteurs mais n'obtient que 71% de bonnes détections des états d'hypovigilance. De plus il génère un taux de 42% de fausses alarmes, ce qui est élevé. Les performances moyennes de ce système sont sans doute dues au nombre très limité d'indices utilisés.

Une synthèse de ces résultats est présentée dans le tableau 4.6

Auteurs	Résultats	Données	Technique	Niveaux d'hypovigilance
Picot et al., 2009	80,6%	EEG et vidéo	Fusion floue et combinaison de décisions	3
Khaldi & H.-Gress, 2000	93%	Mécaniques et vidéo	Réseaux de neurones	2
Ji et al., 2006	$r=0,95$	Vidéo et environnement	Réseaux bayésiens	%
Su & Zheng 2007	$TP_{rate} = 71\%$ $FP_{rate} = 42\%$	Vidéo et mécaniques	Théorie de l'évidence	2

TAB. 4.6 – Synthèse de la comparaison de nos résultats à ceux de la littérature concernant la détection de l'hypovigilance par fusion d'informations

Malgré ses bonnes performances, notre système de détection semble un peu moins performant que certains systèmes de la littérature ([Khardi00, Ji06]). Ces systèmes nécessitent cependant un nombre très important d'indices, ce qui peut rendre délicat leur mise en place. L'avantage de notre système est sa simplicité et sa robustesse aux différences inter-individuelles. En effet, il a pu être appliqué sur vingt conducteurs différents sans avoir à changer les paramètres de l'algorithme. De plus, ses performances restent très correctes avec 80% de bonnes classifications. Enfin, la pertinence de cette comparaison reste relative car toutes ces techniques ont été testées sur des données différentes.

4.3.4 Approche par les arbres de décisions

De la même manière que dans la section 3.4.3, on utilise l'induction d'arbres de décision afin de vérifier par une autre méthode, à partir des données, quelle est l'utilisation la plus pertinente des systèmes de détection EEG et vidéo afin de détecter l'hypovigilance.

Pour construire l'arbre, le critère utilisé est la minimisation du coût de mauvais classement. Il s'agit de l'algorithme *Cost Sensitive C4.5* proposé par Chauchat et al. ([Chauchat01]). L'arbre de décision est réalisé à l'aide du logiciel *Sipina* ([Sipina]). Les variables utilisées ici sont les détections D_{EEG} et $D_{vidéo}$ sur 20s réalisées à partir de nos systèmes de détection EEG et vidéo. On souhaite observer comment organiser ces variables pour effectuer la détection d'hypovigilance. Le tableau 4.7 présente les différentes données d'entrées utilisées pour la construction de l'arbre de décision. La figure 4.6 représente les résultats de la classification sur un sous-échantillon de la base de données prêtée par le CEPA comprenant 1907 cas "éveillé" et 949 cas "fatigué" et 1096 cas "très fatigué".

Indices	<ul style="list-style-type: none"> – Détection EEG D_{EEG} – Détection vidéo $D_{vidéo}$
Fenêtre de calcul	Décision prise sur 20s
Données	1907 cas "éveillé" 949 cas "fatigué" 1096 cas "très fatigué"

TAB. 4.7 – Données d'entrée utilisées pour la construction de l'arbre de décision

On constate sur cette figure que l'algorithme se sert d'abord de la détection EEG pour

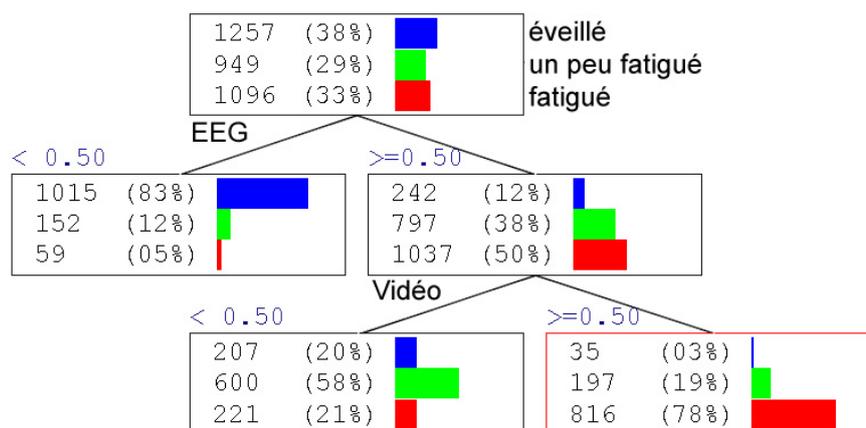


FIG. 4.6 – Arbre de décision d’hypovigilance

différencier les états “éveillé” et “fatigué”. Il effectue ensuite une deuxième classification à partir de la détection vidéo afin de différencier les états “fatigué” et “très fatigué”. De manière automatique, le système a généré les mêmes règles que nous avons utilisées pour notre système. Cela est cohérent avec les règles utilisées par l’expert pour analyser les données et présentées dans le tableau 1.1 du chapitre 1. Cette modélisation nous permet de valider les traitements bas niveau de nos détection EEG et vidéo puisqu’on retrouve bien que la détection EEG correspond au niveau 1 d’hypovigilance de l’échelle OSS et la détection vidéo au niveau 2. De plus, cette modélisation valide également notre système de fusion puisque les mêmes règles de décision que nous avons utilisées sont générées ici de manière automatique.

4.3.5 Vers une validation des résultats

Les résultats présentés dans la partie 4.3.2 ont été obtenus à partir des indices optimisés des autres chapitres. Ils ne constituent donc pas une validation de notre système. Une vraie validation nécessiterait de nouvelles données. Une analyse en partitionnant les données disponibles peut également être envisagée.

Nous avons tenté au cours de la thèse d’effectuer une campagne d’acquisition de données EEG et vidéo d’hypovigilance afin de valider nos algorithmes. L’EEG était enregistré à l’aide de la station d’acquisition Biopac “MP150”. La vidéo était enregistré à une cadence de 200fps par une caméra rapide Point Grey “Dragonfly Express”. L’hypovigilance

était évaluée à l'aide d'une tâche psychomotrice de vigilance (*PVT* pour *Psychomotor Vigilance Task*) de 10 minutes. Cette tâche mesure le temps de réaction d'un individu à un stimulus visuel spécifique ([Dinges85]).

La campagne fut réalisée sur six sujets (quatre hommes et deux femmes). Chaque sujet était enregistré deux fois : une première fois en fin d'après-midi (vers 18h) et une seconde fois aux alentours de 1h du matin. Le premier enregistrement devait servir de référence pour l'éveil. Les sujets étaient priés de rester éveillés jusqu'à 1h du matin afin d'observer des états d'hypovigilance. La figure 4.7 présente l'exemple d'un sujet durant un de ces enregistrements.

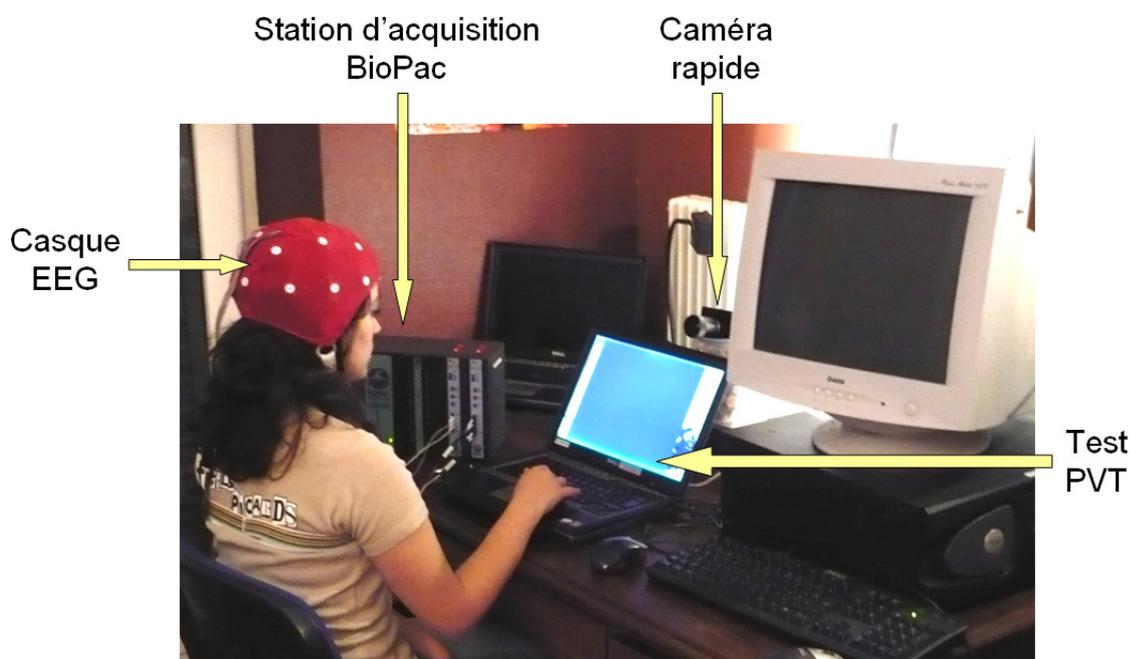


FIG. 4.7 – Exemple de sujet durant un enregistrement

Malheureusement, cette campagne d'acquisition fut avant tout l'occasion de mettre en évidence de nombreux points délicats d'un tel travail. Tout d'abord, il est difficile de trouver des sujets qui acceptent de passer cette expérience. En effet, le fait de devoir rester éveillé jusque tard dans la nuit n'est pas agréable et les sujets n'étaient pas rémunérés. La totalité des sujets ayant accepté de participer étaient des amis qui connaissaient les paradigmes de l'expérience, ce qui peut constituer un biais. Ces expériences nous ont également permis de mettre en avant la difficulté de réaliser des enregistrements EEG fiables. Le placement des électrodes est très délicat et nous ne disposons d'aucune expertise en la matière. Des problèmes d'éclairage au niveau de la caméra ont également été mis

en avant. En effet, les enregistrements étaient réalisés de nuit. Nous ne disposons pas d'éclairage infrarouge et nous avons essayé de nous contenter de l'éclairage de la pièce dans laquelle était réalisée l'expérience. La mauvaise qualité lumineuse a engendré des vidéos très bruitées. De plus, la cadence d'acquisition de la vidéo (200fps) étant bien supérieure à la fréquence du courant électrique (50Hz), on observe des variations de lumière correspondant à l'oscillation du courant électrique. Enfin, nous nous sommes également interrogés sur la validité et la longueur du test PVT car l'analyse des résultats n'a mis en avant que peu de périodes d'hypovigilance.

Ainsi, les données acquises sont quasiment inexploitable mais ouvrent la perspective d'une nouvelle campagne d'acquisition de meilleure qualité. On peut envisager d'effectuer la prochaine campagne plus tôt dans la journée. Cela permettrait d'avoir plus de sujets. On pourrait se servir du cycle circadien qui diminue la plupart de nos capacités en début d'après-midi (après le repas). Ainsi, les enregistrements pourraient être réalisés de jour, ce qui permettrait de supprimer le problème d'éclairage. On envisage également d'effectuer un test PVT plus long afin de garantir la présence de signes d'hypovigilance lors du test. Enfin, concernant l'acquisition EEG, il faut que nous continuions à en faire afin d'acquérir une certaine expérience en la matière.

4.4 Conclusion

Ce chapitre est une première étape vers un système de détection d'hypovigilance par fusion de données EEG et vidéo. Un système de fusion des deux algorithmes de détection d'hypovigilance, présentés dans les chapitres précédents, utilisant l'EEG et la caméra rapide est proposé. Il se présente sous la forme d'un système expert qui se sert de la détection EEG pour valider la détection vidéo.

Notre système obtient de bonnes performances avec environ 80% de bonnes classifications. L'avantage principal de la validation de la détection vidéo par la détection EEG est la forte diminution du taux de fausses alarmes engendrées par la détection vidéo qui passe de 13,1% à 5,2%. De plus, ce système conserve tous les avantages des algorithmes de détection EEG et vidéo présentés dans les chapitres précédents. Ces avantages sont la robustesse aux différences inter-individuelles, la faible période d'apprentissage ainsi que l'évaluation de la qualité du signal EEG par le détecteur d'artéfacts. Les résultats ont d'ailleurs été obtenus sur une base de données de vingt conducteurs différents sans avoir à

changer les paramètres de détection. Il a été montré que le choix d'utiliser d'abord l'EEG puis la vidéo est cohérent avec les règles utilisées par l'expert pour analyser les données d'hypovigilance.

Cependant, notre système est limité par le fait que nous ne disposons pas d'informations pour valider la détection EEG. Il serait donc nécessaire d'avoir un système de détection EEG plus fin de manière à améliorer les performances. De plus, il semble nécessaire d'effectuer l'acquisition d'une base de données de conduite contenant des signaux EEG synchronisés avec une vidéo haute fréquence de manière à valider notre système. En effet, les résultats ont été obtenus en utilisant l'EOG à la place de la caméra rapide comme cela avait été le cas dans le chapitre précédent. L'échec de notre tentative de réaliser une telle campagne d'acquisition nous a montré la difficulté d'un tel travail. Cependant, l'étude comparative entre EOG et caméra rapide réalisée (chapitre 2) et les précautions prises lors de l'élaboration de notre système de détection vidéo devraient nous permettre d'obtenir des performances similaires en appliquant notre système de fusion sur des données de conduite EEG et vidéo synchronisées.

“- *J’ai peur de m’endormir au volant.*
- *Ah... Tu sais ce que tu fais ?*
T’attaches tes cheveux au plafond.
- *... ?*
- *T’attaches tes cheveux au plafond.*
Comme ça, si tu baisses la tête,
ça va te tirer les cheveux
et ça va te réveiller.”

Extrait du film Eldorado de Bouli Lanners (2008)

Conclusions et perspectives

Conclusions

Ce travail de thèse a permis suite à une réflexion sur la définition et la caractérisation de l'état d'hypovigilance de proposer un système permettant sa détection par fusion d'informations EEG et vidéo. Ce système effectue une détection automatique de la baisse de vigilance chez le conducteur à l'aide d'un système expert fusionnant les approches physiologique et vidéo développées.

Ce système s'avère performant avec environ 80% de bonnes détections des états d'hypovigilance. De plus, ce travail a permis de mettre en avant l'intérêt d'utiliser l'information EEG en complément de la vidéo puisque cette dernière permet de faire chuter le taux de fausses alarmes autour de 5% (contre 13% en détection vidéo seule). Outre ses bonnes performances, un grand avantage de ce système est sa robustesse aux différences inter-individuelles. En effet, ce système ne nécessite pas de changer les réglages en fonction du conducteur et seul un apprentissage d'une durée d'une minute est réalisé pour effectuer la détection EEG. Les performances présentées ont d'ailleurs été obtenues sur une base de données d'hypovigilance réalisée à partir de vingt conducteurs différents sans avoir eu à changer les paramètres.

Notre système de détection s'appuie sur une détection d'hypovigilance à l'aide d'un unique canal EEG. La détection EEG est réalisée par comparaison de moyennes sur les ondes α et β . Ces deux informations sont alors fusionnées par logique floue afin d'augmenter la fiabilité du détecteur. Un détecteur d'artéfacts a également été développé de manière à évaluer la qualité du signal EEG. Ainsi, l'information EEG est utilisée par notre système uniquement lorsqu'elle est de bonne qualité afin de réduire le nombre de fausses alarmes. En plus de fonctionner à partir d'un seul canal EEG, cette méthode est robuste aux différences entre conducteurs et ne nécessite pas de réglage.

Notre système est également basé sur une détection d'hypovigilance par caméra rapide. Une méthode originale de caractérisation des clignements a été proposée. Cette caractérisation utilise les signaux d'énergie statique et dynamique générés au niveau de l'oeil. Cette approche a été comparée à l'approche traditionnelle par EOG. Cela a permis d'une part de mettre en évidence l'avantage de l'utilisation d'une caméra rapide par rapport à une caméra classique. D'autre part, nous avons pu analyser dans quelle mesure l'EOG pouvait être remplacé par une caméra pour caractériser les clignements des yeux, signes visuels de l'hypovigilance. Ainsi, un système de détection d'hypovigilance a pu être élaboré en fusionnant les indices visuels les plus pertinents.

Tous ces travaux ont pu être évalués sur une base de données de conduite en état d'hypovigilance réalisée à partir de vingt conducteurs différents. La richesse de cette base de données a permis d'évaluer la robustesse et les performances des algorithmes développés. Elle a également permis de faire apparaître certaines limites dévoilant de nouveaux problèmes et donc de nouveaux axes de recherche.

Enfin, le problème de l'hypovigilance est un problème très sérieux. Même s'il est possible de détecter cet état accidentogène à l'aide de systèmes de détection automatique, les plus performants soient-ils, le meilleur indicateur d'hypovigilance reste encore le conducteur qui doit apprendre à connaître ses limites. C'est pourquoi ces systèmes de détection ne sont utiles que s'ils s'accompagnent d'une sensibilisation et d'une responsabilisation du conducteur.

Perspectives

La perspective principale à l'heure actuelle est la réalisation d'une campagne d'acquisition de données d'hypovigilance EEG et vidéo haute-fréquence synchronisées. Une telle base de données nous permettrait de valider notre algorithme de détection vidéo d'hypovigilance ainsi que notre système final. Ces deux systèmes de détection n'ont été testés pour l'instant que sur l'EOG. Cependant la comparaison entre EOG et caméra rapide a montré que les résultats devraient être similaires en utilisant une caméra rapide. De plus, la première tentative que nous avons effectuée afin de réaliser l'acquisition d'une telle base de données nous a permis de mettre en évidence les problèmes engendrés. Ainsi, avec le recul vis-à-vis de cette expérience, nous devrions être capable de réaliser cette nouvelle campagne d'acquisition dans de meilleures conditions. Parmi les différents points à améliorer, l'utilisation d'un test PVT suffisamment long devrait permettre d'assurer

l'apparition d'hypovigilance chez les sujets. Une amélioration des conditions de luminosité semble également nécessaire afin qu'elles ne constituent pas une limite pour notre système. Il reste toutefois possible de dégrader artificiellement la luminosité des enregistrements par la suite afin d'évaluer son impact. Enfin, une plus grande expérience dans l'acquisition de signaux EEG est nécessaire afin que les signaux relevés soient exploitables.

Au niveau de la détection EEG, nous avons pu voir que c'est elle pour l'instant qui limite les performances de notre système final. On peut donc envisager à plus long terme de travailler à l'amélioration de ce système afin d'améliorer les performances globales du système de détection final. L'utilisation de la transformée d'Hilbert Huang pour la détection d'ondes α dans l'EEG ([Sharabaty08]) semble, à cet égard, une option intéressante.

Nous pouvons également envisager d'extraire d'autres indices vidéo. En effet, il semble que les informations telles que les bâillements ([Mohanty09]) ou la direction du regard ([Smith03]) du conducteur puissent nous aider à évaluer son niveau de vigilance. De même, on peut envisager l'utilisation de la vidéo pour détecter différentes émotions du conducteur telles que stress ou anxiété ([Healey99, Bartlett03]) afin de les prendre compte pour évaluer sa capacité à conduire.

Enfin, les informations extérieures sur l'environnement du conducteur semblent également donner des indications quant à l'hypovigilance probable du conducteur ([Ji06]). Une prise en compte de cet environnement pourrait être ajoutée à notre système de détection pour améliorer sa fiabilité.

Ainsi, ce système de détection d'hypovigilance par fusion d'informations physiologiques et vidéo n'est qu'un premier pas vers un système plus complet. L'hypovigilance est un état délicat à définir et détecter. La fusion d'informations de nature différente a permis d'améliorer sa détection. Ce n'est qu'en multipliant les sources d'informations, nous semble-t-il, que nous améliorerons ce type de système.

Annexe A

Algorithme de détection des clignements à partir de l'EOG

Cette annexe présente l'algorithme de détection des clignements à partir de l'EOG qui a été utilisé. Cet algorithme est issu du travail de thèse de Sharabaty ([Sharabaty08c]). Il s'agit d'un algorithme semi-automatique de détection de clignements car les seuils de détection sont choisis par l'utilisateur. Cet algorithme a été développé dans le cadre du projet SENSATION ([Sensation]), un projet européen dont le but était le développement de capteurs permettant la détection et la prédiction des états d'hypovigilance et de stress.

A.1 Le modèle de clignement utilisé

L'algorithme de détection utilisé propose de détecter les clignements à partir de leur dérivée. La figure A.1 présente le comportement de manière idéale du signal EOG en blanc et de sa dérivée en orange lors d'un clignement.

Afin d'effectuer la détection des clignements, on définit un seuil de vitesse de fermeture V_F et un seuil de vitesse d'ouverture V_O . Le début du clignement *debh* correspond alors à l'instant où la dérivée du signal EOG passe au dessus du seuil V_F . La dérivée repasse sous ce seuil à l'instant *finh* qui marque la fin de la phase de fermeture. La phase d'ouverture commence à l'instant *debb* lorsque la dérivée du signal EOG passe en dessous du seuil V_O . La fin du clignement est marquée par l'instant *finb* où la dérivée repasse au dessus de ce

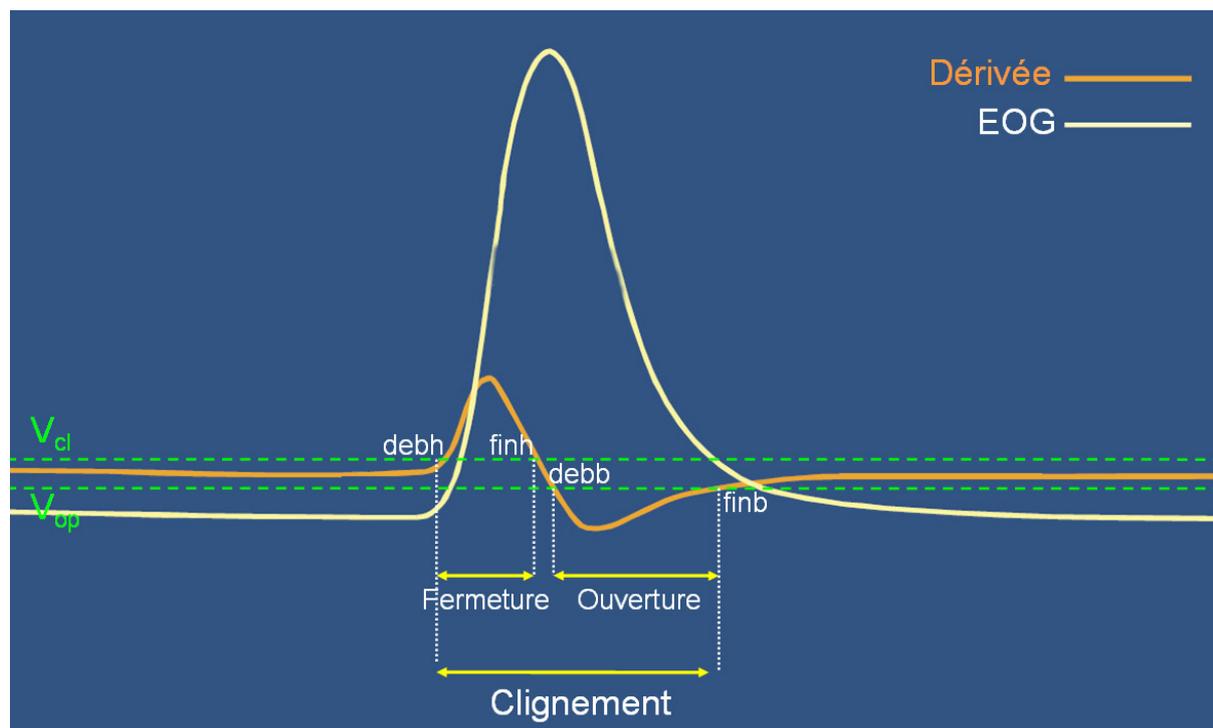


FIG. A.1 – Modèle de clignement et sa dérivée

seuil.

L'avantage d'effectuer la détection des clignements à partir de la dérivée du signal EOG, et non directement à partir de l'EOG, est d'éviter la détection de mouvements lents des yeux. Les seuils V_F et V_O sont définis de manière empirique en fonction du conducteur, c'est pourquoi on parle d'algorithme de détection semi-automatique.

A.2 L'algorithme de détection

L'algorithme de détection a été développé par Sharabaty à partir du modèle de clignement présenté dans la figure A.1. L'organigramme de cet algorithme est présenté sur la figure A.2. L'algorithme commence par localiser les mouvements des paupières à partir de la dérivée du signal EOG comme présenté dans le paragraphe A.1. Une validation des clignements a alors lieu après vérification de leurs propriétés.

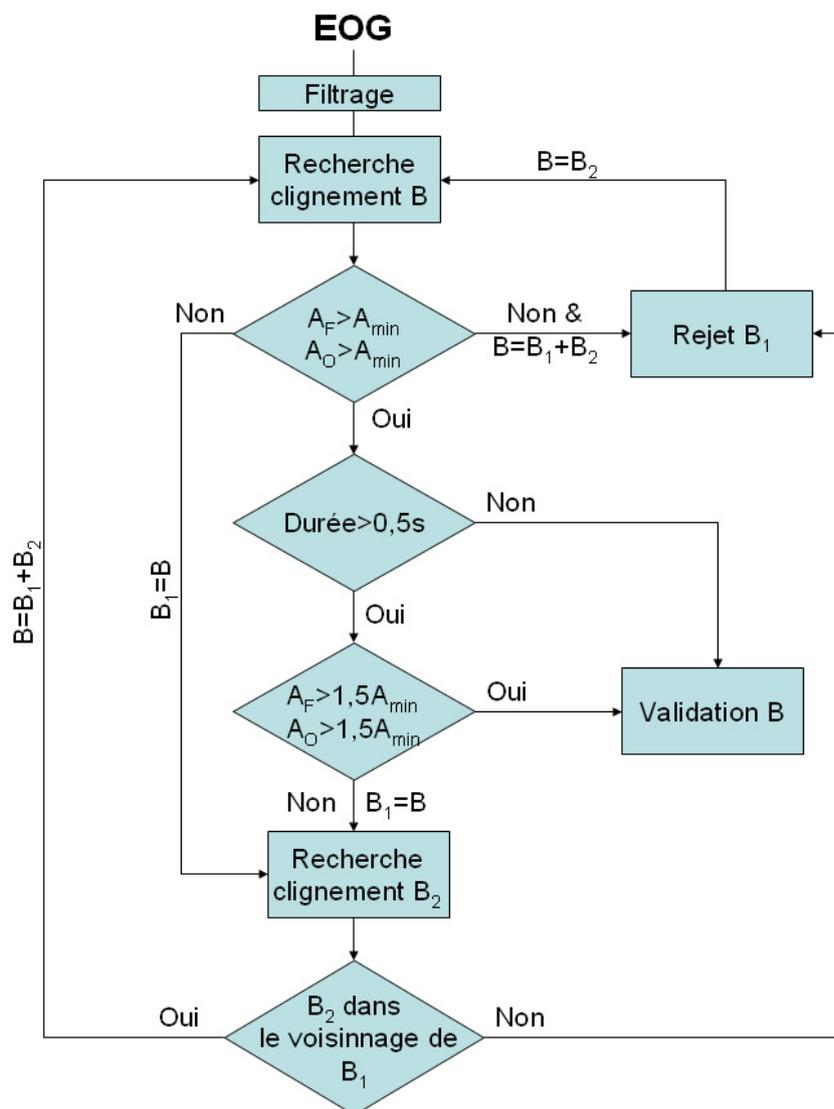


FIG. A.2 – Algorithme de détection des clignements

Pré-traitements. On effectue d'abord un filtrage des hautes fréquences ($> 10Hz$) du signal EOG vertical. La dérivée de l'EOG filtré est ensuite calculée. On sélectionne alors les périodes de l'EOG pour lesquelles la dérivée passe d'abord au dessus du seuil V_F puis en dessous du seuil V_O .

Validation des amplitudes. Les amplitudes de fermeture A_F et d'ouverture A_O sont calculées sur le clignement potentiel B . Cet algorithme traite de manière différente les clignements longs et courts. Dans le cas d'un clignement court, c'est-à-dire de durée inférieure à $0,5s$ (durée maximale d'un clignement proposée par les experts en accord avec la

définition de la durée d'un clignement), l'amplitude de celui-ci devra être supérieure au seuil A_{min} . Ce paramètre est propre à chaque individu et est défini de manière manuelle de la même manière que V_F et V_O . Dans le cas d'un clignement long ($\geq 0,5s$), son amplitude doit être supérieure à $1,5\dot{A}_{min}$ pour être validé comme fermeture prolongée.

Possibilité de regrouper des clignements contigus. Certains clignements longs présentent une amorce de réouverture avant l'ouverture définitive. L'algorithme offre donc la possibilité de regrouper deux clignements contigus dans le cas où ceux-ci ne passent pas la validation des amplitudes. Avant de rejeter ces mouvements, il est vérifié que les contraintes d'amplitude restent non vérifiées après regroupement.

A.3 Calcul des caractéristiques des clignements

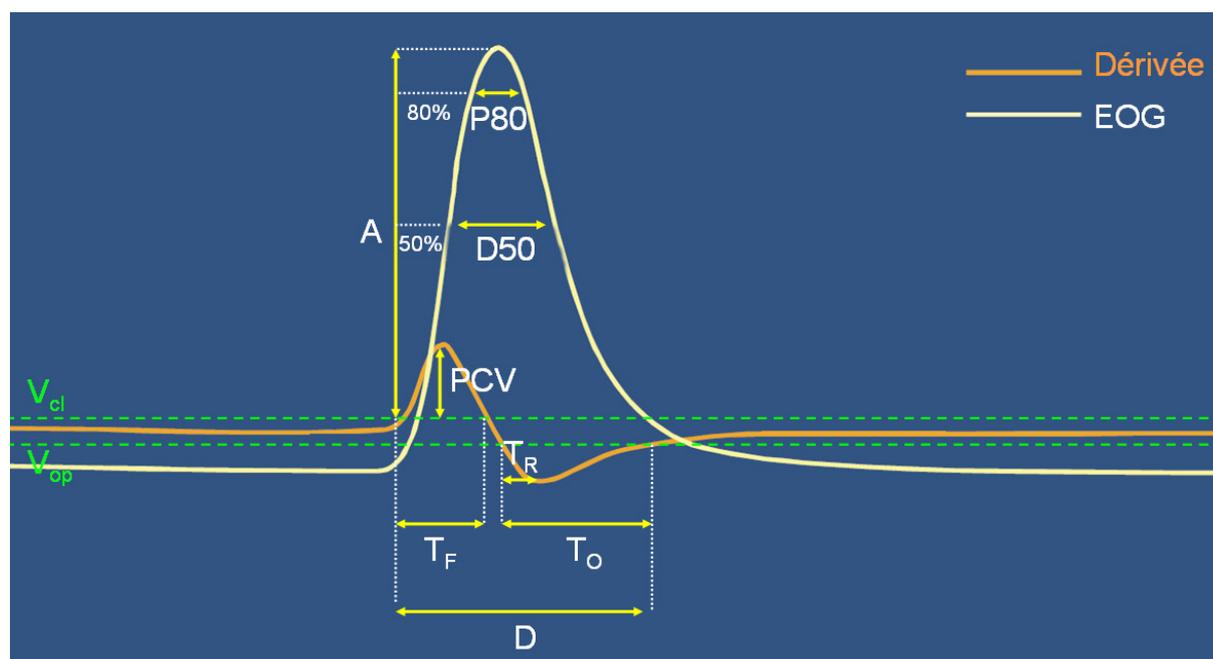


FIG. A.3 – Définition des caractéristiques d'un clignement

Une fois les clignements localisés, l'algorithme calcule les paramètres sélectionnés (cf. paragraphe 2.1.1). Tous ces paramètres sont représentés sur la figure A.3. Cette figure présente le signal EOG en blanc et sa dérivée en orange lors de l'occurrence d'un clignement. La liste des paramètres calculés est la suivante :

- Durée du clignement (D),
 - Durée à 50% (D_{50}),
 - Fréquence (F) des clignements,
 - Amplitude du clignement (A),
 - Durée de fermeture (T_F) et d'ouverture (T_O),
 - Vitesse de fermeture (V_F) et d'ouverture (V_O),
 - Délai de réouverture (T_R),
 - Vitesse maximale de fermeture (PCV),
 - Pourcentage de fermeture à 80% ($PERCLOS_{80}$).
-

Annexe B

Introduction à la logique floue

Cette annexe présente une introduction à la logique floue. Le but de cette partie est de définir ce qu'on appelle la logique floue et de donner les bases mathématiques permettant son utilisation. Des informations beaucoup plus complètes peuvent être trouvées dans l'ouvrage de Bouchon-Meunier ([Bouchon95]).

B.1 Les sous-ensembles flous

La logique floue est une technique qui a été développée par Zadeh et sa théorie sur les sous-ensembles flous ([Zadeh65]) dans les années soixante à l'Université de Californie. Cette théorie généralise la théorie des ensembles classiques. Le principe de cette théorie est qu'un élément peut appartenir plus ou moins à un ensemble. On peut ainsi modéliser les imprécisions dues à une connaissance globale d'un système très complexe ce qui permet une certaine flexibilité du raisonnement que la logique classique ne permet pas.

Un sous-ensemble flou A d'un ensemble X est défini par une fonction d'appartenance μ_A sur X à valeur dans l'intervalle $[0;1]$. Cette fonction est définie selon l'équation (B.1).

$$\begin{aligned} \mu_A &: X \rightarrow [0;1] \\ x &\mapsto \mu_A(x) \end{aligned} \tag{B.1}$$

Cette fonction d'appartenance peut être définie arbitrairement. Dans le cas d'appli-

cations pratiques, on fait généralement appel à des données statistiques ou à l'avis d'un expert. Le fait de projeter une variable dans l'intervalle $[0;1]$ à l'aide de la fonction μ_A est appelé *fuzzification* (de l'anglais *fuzzy* qui signifie *flou*). La valeur prise par la fonction d'appartenance en un point donné est appelé *degré d'appartenance*.

Pour illustrer ce concept, utilisons l'exemple de la caractérisation de la température de la mer dans une station balnéaire de la côte d'Opale. La température est caractérisée par les termes "froid" et "chaud" dont on place la limite 15°C . La figure B.1 présente une représentation graphique des fonctions d'appartenance à ces deux classes dans le cadre de la logique classique (a) et de la logique floue (b).

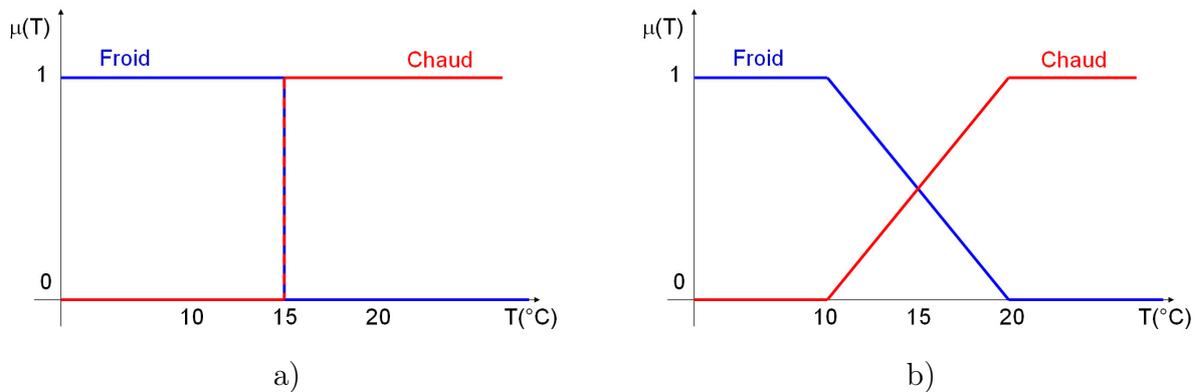


FIG. B.1 – Caractérisation de la température de l'eau en logique binaire (a) et en logique floue (b)

Ainsi, dans le cas de la logique classique, une température $T=17,5^\circ\text{C}$ est nécessairement dans le sous-ensemble "chaud". Pour la logique floue, cette température appartient à la fois aux sous-ensembles "chaud" et "froid" avec des degrés respectifs de 0,75 et 0,25 comme montré sur la figure B.2.

On constate donc que la logique floue nous permet de superposer sur des plages de température données les qualificatifs "chaud" et "froid" afin de produire des nuances au niveau de la caractérisation. On se rapproche ainsi de la perception humaine.

B.2 Opérations sur les sous-ensembles flous

De nombreuses opérations sur les sous-ensembles flous et les fonctions d'appartenances ont été définies par la théorie mathématique. On présente ici les opérations de base de

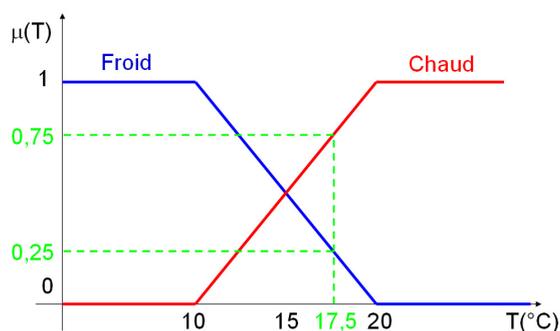


FIG. B.2 – Caractérisation d'une température $T=17,5$ ° C en logique floue

la théorie des ensembles flous. Soient A et B deux sous-ensembles flous de fonctions d'appartenance respectives μ_A et μ_B . On définit alors les fonctions suivantes.

B.2.1 Opérateur NON

Cet opérateur est également appelé *complément*. Le complémentaire d'un sous-ensemble A est l'ensemble des éléments qui n'appartiennent pas à ce sous-ensemble. Il est noté \bar{A} . La définition mathématique est donnée par l'équation (B.2). La figure B.3 donne une représentation graphique de cette notion.

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad (\text{B.2})$$

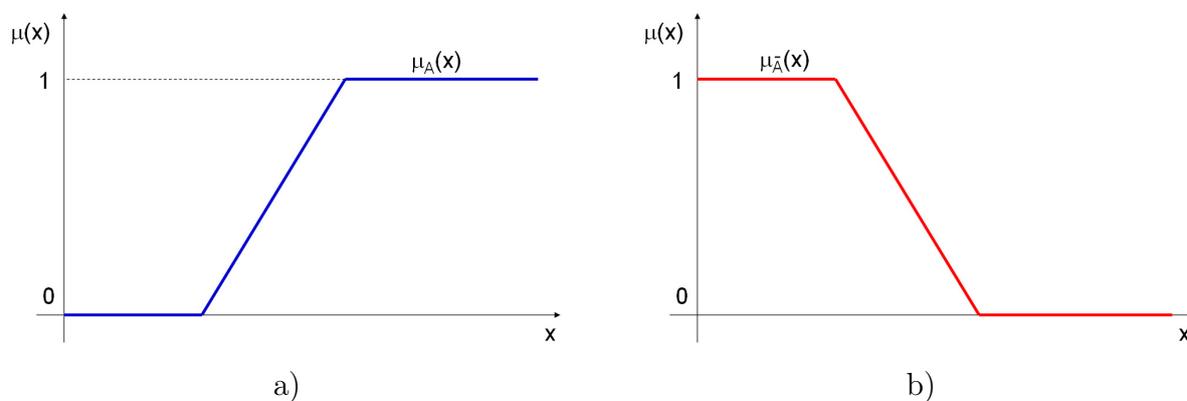


FIG. B.3 – Fonction d'appartenance au sous-ensemble A (a) et son complémentaire \bar{A} (b)

B.2.2 Opérateur ET

Cet opérateur est également appelé *intersection*. L'intersection de deux sous-ensembles A et B est l'ensemble des éléments appartenant à la fois à A et à B . Elle est noté $A \cap B$ et est définie par l'équation (B.3). La figure B.4 donne une représentation graphique de cette notion.

$$\begin{aligned} \mu_{A \cap B}(x) &= \mu_A(x) \cap \mu_B(x) \\ &= \min(\mu_A(x), \mu_B(x)) \end{aligned} \quad (\text{B.3})$$

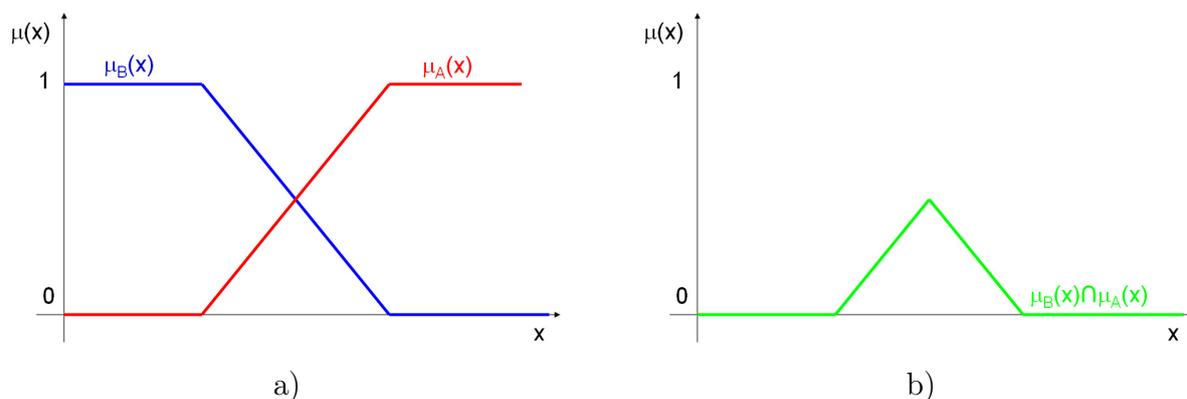
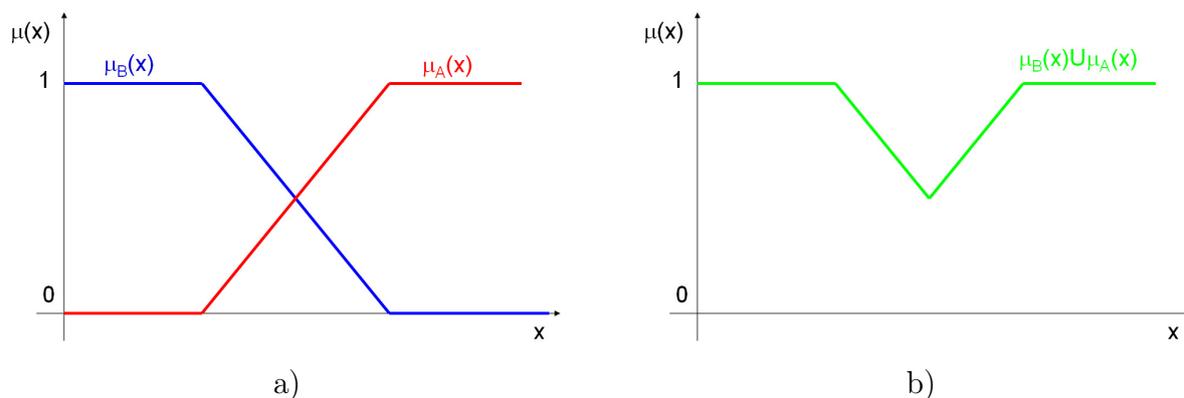


FIG. B.4 – Fonctions d'appartenance aux sous-ensembles A et B (a) et leur intersection (b)

B.2.3 Opérateur OU

Cet opérateur est également appelé *union*. L'union de deux sous-ensembles A et B est l'ensemble des éléments appartiennent soit à A , soit à B . Il est noté $A \cup B$ et est défini par l'équation (B.4). La figure B.5 donne une représentation graphique de cette notion.

$$\begin{aligned} \mu_{A \cup B}(x) &= \mu_A(x) \cup \mu_B(x) \\ &= \max(\mu_A(x), \mu_B(x)) \end{aligned} \quad (\text{B.4})$$

FIG. B.5 – Fonctions d'appartenance aux sous-ensembles A et B (a) et leur union (b)

B.2.4 Compromis logiques

D'autres fonctions peuvent également être définies pour manipuler les fonctions d'appartenance. On parle de compromis logique lorsque le résultat de l'association de plusieurs fonctions d'appartenance est borné entre le minimum et le maximum de ces fonctions d'appartenance. Ainsi, la fonction C définie par l'équation (B.5) est un compromis si elle respecte le critère (B.6).

$$\begin{aligned}
 C & : [0; 1] \times [0; 1] \rightarrow [0; 1] \\
 \mu_A, \mu_B & \mapsto C(\mu_A, \mu_B)
 \end{aligned}
 \tag{B.5}$$

$$\forall(\mu_A, \mu_B), \min(\mu_A, \mu_B) \leq C(\mu_A, \mu_B) \leq \max(\mu_A, \mu_B)
 \tag{B.6}$$

Annexe C

Introduction à l'induction d'arbres de décision

Cette annexe propose une introduction à la classification par induction d'arbres de décision. Elle commence par une présentation générale sur les principes de l'apprentissage par arbre de décision. Nous présentons ensuite l'algorithme C4.5 dont nous nous sommes servis au cours de la thèse. Il s'agit d'un résumé du cours de F. Denis et R. Gilleron ([CoursDG]) sur les méthodes d'apprentissage.

C.1 Présentation

Les techniques d'induction d'arbres de décision sont des techniques d'extraction de connaissances à partir de données qui permettent de générer automatiquement un ensemble de règles permettant une décision. Ces techniques permettent d'acquérir une connaissance explicite à partir d'un ensemble de données et d'extraire les variables pertinentes pour la décision.

Un arbre de décision est une représentation graphique d'une procédure de classification. Il est composé de *feuilles*, de *branches* et de *nœuds*. Chaque feuille de l'arbre correspond à une classe possible de l'ensemble de données à classer. À chaque nœud de l'arbre est associé un test. Les branches qui en découlent correspondent aux valeurs possibles à ce test. Le résultat de l'arbre de décision est un ensemble de règles sous la forme “*si ... alors*”

permettant la classification. Les règles de décision sont mutuellement exclusives. Ainsi, à une description complète est associée une seule feuille de l'arbre de décision.

C.2 L'apprentissage par arbre de décision

L'idée générale de l'apprentissage par arbre de décision est de diviser récursivement et le plus efficacement possible les exemples de l'ensemble d'apprentissage par des tests définis à l'aide des attributs de ces exemples. Le but est d'obtenir des sous-ensembles ne contenant (presque) plus que des exemples appartenant à la même classe.

Trois opérateurs sont nécessaires pour cela :

1. Décider si un nœud est terminal : le nœud est alors étiqueté comme une feuille.
2. Sélectionner un test à associer à un nœud.
3. Affecter une classe à une feuille.

Le schéma général de la construction d'un arbre de décision est présenté dans l'algorithme 3.

Algorithme 3 Apprentissage par arbre de décision

ENTRÉES: Échantillon S et sa description

SORTIES: Arbre de décision

Initialiser arbre à vide

tantque il existe un nœud non exploré **faire**

 Décider si le nœud courant est terminal

si le nœud est terminal **alors**

 Affecter une classe

sinon

 Sélectionner un test et créer le sous-arbre

fin si

 Passer au nœud suivant non exploré s'il existe

fin tantque

L'objectif est de construire l'arbre d'erreur de classification la plus faible possible. Cela n'est pas un problème facile car même si les feuilles sont étiquetées de manière à ce que l'erreur apparente soit minimale, ce n'est pas forcément le cas de l'erreur réelle. De plus, il faut proposer un critère qui permette d'arrêter la croissance de l'arbre au bon moment. Comme ce critère d'arrêt est difficile à définir, les algorithmes d'apprentissage par arbre de décision procèdent souvent en deux phases. La première est appelée *phase d'expansion*.

Elle correspond au mécanisme de construction de l'arbre présenté dans l'algorithme 3. La seconde phase correspond à la *phase d'élagage*. Elle consiste à supprimer certains sous-arbres de manière à faire diminuer l'erreur réelle.

Il existe de nombreux algorithmes pour construire des arbres de décision. Ces différents algorithmes se distinguent par les techniques utilisées pour l'expansion et l'élagage de l'arbre. Parmi eux, on peut citer :

- l'algorithme *CART* ([Breiman84]),
- l'algorithme *ID3* ([Quinlan86]),
- l'algorithme *CHAID* ([Kass80]),
- l'algorithme *QUEST* ([Breiman84]),
- l'algorithme *C4.5* ([Loh97]).

Nous présentons dans la suite l'algorithme C4.5 car c'est celui qui nous a servi au cours de la thèse. Cet algorithme a été choisi pour sa facilité d'application.

C.3 L'algorithme C4.5

On suppose que l'on travaille à partir d'un échantillon S . Le langage de représentation de cet échantillon est composé d'un certain nombre d'attributs. Soient (c_1, \dots, c_n) les n classes de S . On note t l'arbre de décision et p la position dans l'arbre (ou nœud). À chaque position p correspond un sous-ensemble. On note $N(p)$ le cardinal du sous-ensemble de la position p et $N(k, p)$ le cardinal des éléments de ce sous-ensemble ayant la classe k . La proportion d'éléments de la classe k à la position p est alors définie par l'équation (C.1).

$$P(k, p) = \frac{N(k, p)}{N(p)} \quad (\text{C.1})$$

Afin de définir l'algorithme C4.5, nous définissons d'abord les trois opérateurs utilisés pour construire l'arbre de décision lors de la phase d'expansion, puis nous verrons la phase d'élagage. L'ensemble A constitue l'ensemble d'apprentissage.

C.3.1 Algorithme de base

Afin de comparer les différents choix possibles, nous utilisons la fonction d'entropie définie par l'équation (C.2).

$$Entr(p) = - \sum_{k=1}^n P(k, p) \cdot \log(P(k, p)) \quad (C.2)$$

Décision du caractère terminal d'un nœud. On considère qu'un nœud p est terminal si tous les éléments du sous-ensemble associé appartiennent à la même classe ou si aucun test n'a pu être sélectionné.

Sélection du test à associer à un nœud. Parmi tous les tests envisageables, ne sont considérés que ceux pour lesquels il existe au moins deux branches ayant au moins deux éléments. Si aucun test ne satisfait cette condition alors le nœud est terminal. Dans le cas contraire, on choisit le test t qui maximise le gain défini par l'équation (C.3). Dans cette équation P_k représente la proportion d'éléments de A à la position p qui iront en position p_j , c'est-à-dire qui satisferont la $j^{\text{ème}}$ branche du test t .

$$Gain(p, t) = Entr(p) - \sum_{k=1}^n P_k \cdot Entr(p_j) \quad (C.3)$$

Le gain ainsi défini utilise l'entropie afin de mesurer le degré de mélange. Cependant, il privilégie les attributs ayant un grand nombre de valeurs. On le pondère donc par une fonction *SplitInfo* qui pénalise les tests qui répartissent les éléments en un trop grand nombre de sous-classes. Cette mesure est définie par l'équation (C.4). Dans cette équation, $P'(j, p)$ représente le nombre d'éléments présents à la position p prenant la $j^{\text{ème}}$ valeur du test t .

$$SplitInfo(p, t) = - \sum_{k=1}^n P'(j, p) \cdot \log(P'(j, p)) \quad (C.4)$$

La nouvelle fonction de gain est appelée ratio de gain et notée *GainRatio*. Elle est

définie par l'équation (C.5).

$$GainRatio(p, t) = \frac{Gain(p, t)}{SplitInfo(p, t)} \quad (C.5)$$

Le test t choisi pour la position p est celui qui maximise le ratio de gain.

Affectation d'une classe à une feuille. On attribue la classe majoritaire.

C.3.2 Élagage

L'algorithme C4.5 utilise l'ensemble d'apprentissage pour élaguer l'arbre obtenu. Le critère d'élagage est basé sur une heuristique permettant d'estimer l'erreur réelle sur un sous-arbre donné. Bien qu'il semble peu pertinent d'estimer l'erreur réelle sur l'ensemble d'apprentissage, il semble que la méthode donne des résultats corrects.

C.3.3 Cost Sensitive C4.5

Cette méthode est une variante de l'algorithme C4.5 intégrant explicitement la matrice de coût de mauvais classement lors de l'exportation de l'espace de solutions. Elle a été développée par Chauchat et al. ([Chauchat01]).

Par rapport à l'algorithme de base C4.5, les coûts de mauvais classements sont pris en compte de deux manières. Premièrement durant la phase d'expansion de l'arbre, ils sont pris en compte lors du calcul du critère de segmentation. Ils sont également pris en compte durant la phase d'élagage. L'erreur pessimiste n'est plus calculée pour décider de la suppression des feuilles, mais plutôt un coût pessimiste, qui tient compte du poids des sommets.

Bibliographie

- [Andreassi00] J. L. Andreassi. *Psychophysiology : Human Behavior & Physiological Response*. Lawrence Erlbaum Associates, 2000.
- [Bartlett03] M. Bartlett, G. Littlewort, I. Fasel, and J. Movellan. Real time face detection and facial expression recognition : Development and applications to human computer interaction. In *Proc. of the Computer Vision and Pattern Recognition Workshop*, Madison, USA, 2003.
- [BenKhalifa04] K. Ben Khalifa, M.H. Bédoui, M. Dogui, and F. Alexandre. Alertness states classification by SOM and LVQ neural networks. *International Journal of Information Technology*, 4 :228–231, 2004.
- [Benoit07a] A. Benoit, L. Bonnaud, A. Caplier, Y. Damousis, F. Jourde, L. Lawson, L. Nigay, M. Serrano, and D. Tzovaras. Multimodal signal processing and interaction for a driving simulator : Component-based architecture. *Journal of Multimodal User Interface*, 1 :49–58, 2007.
- [Benoit07b] A. Benoit. *Le système visuel humain au secours de la vision par ordinateur*. PhD thesis, Institut Polytechnique de Grenoble, 2007.
- [BioIDSite] <http://www.bioid.com/>.
- [Bittner01] R. Bittner, P. Smrcka, M. Pavelka, P. Vysoky, and L. Pousek. Fatigue indicators of drowsy drivers based on analysis of physiological signals. In *Proc. of the international symposium on medical data analysis ISMDA*, volume 2199, pages 62 – 68, Madrid, Spain, 2001.
- [Blinowska06] K. Blinowska and P. Durka. *Wiley Encyclopedia of Biomedical Engineering*, chapter Electroencephalography (EEG). Wiley, 2006.
- [Bouchon95] B. Bouchon-Meunier. *La logique floue et ses applications*. Adison Wesley France, 1995.
- [Boverie08] S. Boverie, A. Giralt, and J.-M. Le Quellec. Diagnostic fusion for in vehicle driver vigilance assessment. In *Proc. of the 17th IFAC World Congress*, Seoul, South Korea, 2008.
- [Brandt04] T. Brandt, R. Stemmer, and A. Rakotonirainy. Affordable visual driver monitoring system for fatigue and monotony. In *Proc. of the IEEE conference on Systems, Man and Cybernetics*, The Hague, Netherlands, 2004.
- [Breiman84] L. Breiman, J. H. Friedman, R. Olshen, and C. Stone. *CART : Classification and Regression Trees*. Wadsworth International Group, 1984.
-

-
- [Caffier03] P. Caffier, U. Erdmann, and P. Ullsperger. Experimental evaluation of eye-blink parameters as a drowsiness measure. *European Journal of Applied Physiology*, 89 :319–325, 2003.
- [Caffier05] P. Caffier, U. Erdmann, and P. Ullsperger. The spontaneous eye-blink as sleepiness indicator in patients with obstructive sleep apnea syndrome - a pilot study. *Sleep medicine*, 6 :155–162, 2005.
- [Cajochen00] C. Cajochen, D. P. Brunner, K. Kräuchi, P. Graw, and A. Wirtz-Justice. EEG and subjective sleepiness during extended wakefulness in seasonal affective disorder : circadian and homeostatic influences. *Biological Psychiatry*, 47(7) :610–617, 2000.
- [Chauchat01] J. H. Chauchat, R. Rakotomalala, M. Carloz, and C. Pelletier. Targeting customer groups using gain and cost matrix : a marketing application. In *Proc. of Data Mining for Marketing Applications Workshop, PKDD'2001*, pages 1–13, Freiburg, Germany, 2001.
- [CoursDG] F. Denis and R. Gilleron. Apprentissage à partir d'exemples - Notes de cours.
- [Cudalbu05] C. Cudalbu, B. Anastasiu, R. Radu, R. Cruceanu, E. Schmidt, and E. Barth. Driver monitoring with a single high-speed camera and IR illumination. In *International Symposium on Signals, Circuits and Systems*, volume 1, pages 219–222, IASI, Romania, 2005.
- [Damousis07] I. Damousis, I. Cester, S. Nikolaou, and D. Tzovaras. Psychological indicators based sleep onset prediction for the avoidance of driving accidents. In *Proc. of the 29th IEEE EMBS conference*, Lyon, France, 2007.
- [Damousis09] I. Damousis, D. Tzovaras, and M. Strintzis. A fuzzy expert system for the early warning of accidents due to driver hypo-vigilance. *Personal and Ubiquitous Computing*, 13(1) :43–49, 2009.
- [Desodt96] A.-M. Desodt-Lebrun. Fusion de données. *Techniques de l'ingénieur*, S1 :R7226.1–R7226.9, 1996.
- [DeWaard91] D. De Waard and K. A. Brookhuis. Assessing driver status : a demonstration experiment on the road. *Accident Analysis and Prevention*, 23(4) :297–307, 1991.
- [Dinges85] D. Dinges and J. Powell. Microcomputer analysis of performance on a portable, simple visual RT task sustained operations. *Behavioral Research Methods, Instrumentation, and Computers*, 17 :652–655, 1985.
- [Dinges95] D. Dinges. An overview of sleepiness and accidents. *Journal of sleep research*, 4(2) :4–14, 1995.
- [Dinges98] D. Dinges, M. MAllis, G. Maislin, and J. W. Powell. Final report : Evaluation of techniques for ocular measurement as an index of fatigue and as the basis for alertness management. Technical report, National Highway Traffic Safety Administration, 1998.
- [Dong05] W.H. Dong and X. Wu. Driver fatigue detection based on the distance of eyelid. In *Proc. IEEE VLSI Design and Video Technology*, Suzhou, China, 2005.
-

-
- [Duffner05] S. Duffner and C. Garcia. A hierarchical approach for precise facial feature detection. In *Coresa'05*, Rennes, France, 2005.
- [Galley04] N. Galley, R. Schleicher, and L. Galley. Blink parameter as indicators of driver's sleepiness - possibilities and limitations. *Vision in Vehicles*, 10 :189–196, 2004.
- [Gillberg96] M. Gillberg, G. Kecklund, and T. Akerstedt. Sleepiness and performance of professional drivers in a truck simulator - comparisons between day and night driving. *Journal of Sleep Research*, 5 :12–15, 1996.
- [Grace98] R. Grace, V. E. Byrne, S. M. Bierman, J.-M. Legrand, D. Gricourt, R. K. Davis, J. J. Staszewski, and B. Carnahan. A drowsy driver detection system for heavy vehicles. In *Proc. of the 17th IEEE Digital Avionics Systems Conference*, Seattle, USA, 1998.
- [Gruyer05] D. Gruyer, A. Rakotonirainy, and J. Vrignon. The use of belief theory to assess driver's vigilance. In *Australasian Road Safety Research, Policing and Education Conference*, Wellington, New Zealand, 2005.
- [Hammal06] Z. Hammal. *Segmentation des traits du visages, analyse et reconnaissance d'expressions faciales par le modèle de croyance transferable*. PhD thesis, Université J. Fourier, Grenoble, 2006.
- [Hanley82] J.A. Hanley and B.J. McNeil. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143(1) :29–36, 1982.
- [Healey99] J. Healey, J. Seger, and R. Picard. Quantifying driver stress : developping a system for collecting and processing bio-metric signals in natural situations. In *Proc. of the Rocky Mountain Bio-Engineering Symposium*, Copper Mountain, USA, 1999.
- [Hjelmas01] H. Hjelmäs and B. Low. Face detection : a survey. *Computer Vision and Image Understanding*, 83 :236–274, 2001.
- [Hoffman46] A. C. Hoffman. Eye movements during prolonged reading. *Journal of experimental psychology*, 36 :95–118, 1946.
- [Horng04] W.-B. Horng, C.-Y. Chen, Y. Chang, and C.-H. Fan. Driver fatigue detection based on eye tracking and dynamk, template matching. In *Proc. of the IEEE Int. Conf. on Networking, Sensing and Control*, Taipei, Taiwan, 2004.
- [Hu09] S. Hu and G. Zheng. Driver drowsiness detection with eyelid related parameters by support vector machine. *Expert Systems with Applications*, 36(4) :7651–7658, 2009.
- [Huang09] R.-S. Huang, T.-P. Jung T-P, and S. Makeig. Tonic changes in EEG power spectra during simulated driving. In *Proc. of the 13th int. Human-Computer Interface conference*, San Diego, USA, 2009.
- [Huang96] R.S. Huang, C.J. Kuo, and L.L. Tsai. EEG pattern recognition-arousal states detection and classification. In *Proc. IEEE International Conference on Neural Network (ICNN)*, Washington, USA, 1996.
-

-
- [Hussain08] A. Hussain, B. Bais, and S.A. Samad and S. Farshad Hendi. Novel data fusion approach for drowsiness detection. *Information Technology Journal*, 7(1) :48–55, 2008.
- [Ito02] T. Ito, S. Mita, K. Kozuka, T. Nakano, and S. Yamamoto. Driver blink measurement by the motion picture processing and its application to drowsiness detection. In *Proc of the 5th IEEE Int. Conf. on Intelligent Transportation Systems*, Singapore, 2002.
- [Jackson98] P. Jackson. *Introduction to expert systems*. Addison-Wesley Longman Publishing Co, 1998.
- [Jaynes03] Edwin T. Jaynes. *Probability theory : The logic of science*. Cambridge University Press, 2003.
- [Ji02] Q. Ji and X. Yang. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, 8 :357–377, 2002.
- [Ji05] Q. Ji, H. Wechsler, A. Duchowski, and M. Flickner. Special issue : eye detection and tracking. *Computer Vision and Image Understanding*, 98 :1–3, 2005.
- [Ji06] Q. Ji, P. Lan, and C. Looney. A probabilistic frame work for modeling and real-time monitoring human fatigue. *IEEE Transactions on systems, man, and cybernetics – Part A : systems and humans*, 6 :862–875, 2006.
- [Johns03] M.W. Johns. The amplitude-velocity ratio of blinks : a new method for monitoring drowsiness. *Sleep*, 26 :A51–A56, 2003.
- [Johns07] M. W. Johns, A. Tucker, R. Chapman, K. Crowley, and N. Michael. Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers. *Somnologie*, 11 :234–242, 2007.
- [Joyce03] C. A. Joyce, I. F. Gorodnitsky, and M. Kutas. Automatic removal of eye movement and blink artifacts from EEG data using blind component separation. *Psychophysiology*, 41(2) :313–325, 2003.
- [Jung00] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. J. McKeown, V. Iragui, and T. J. Sejnowski. Removing electroencephalographic artifacts from by blind source separation. *Psychophysiology*, 37 :163–178, 2000.
- [Jung97] T.-P. Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski. Estimating alertness from the EEG power spectrum. *IEEE Transactions on Biomedical Engineering*, 44(1) :60–69, 1997.
- [Kaida07] K. Kaida, T. Akerstedt, G. Kecklund, J. P. Nilsson, and J. Axelsson. Use of subjective and physiological indicators of sleepiness to predict performance during vigilance task. *Industrial health*, 45 :520–526, 2007.
- [Kalman60] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1) :35–45, 1960.
- [Karrer04] K. Karrer, T. Vöhringer-Kuhnt, T. Baumgarten, and S. Briest. The role of individual differences in driver fatigue prediction. In *3rd International Conference on Traffic & Transport Psychology*, pages 5–9, Nottingham , UK, 2004.
-

-
- [Kass80] G. V. Kass. An exploratory technique for investigating large quantities of categorical data. *Journal of Applied Statistics*, 29(2) :119–127, 1980.
- [Khardi00] S. Khardi and N. Hernandez-Gress. Validation d'un nouveau système hybride de diagnostic de l'état de vigilance du conducteur. *Recherche - Transport - Sécurité*, 66 :101–112, 2000.
- [Kiyimik05] M. K. Kiyimik, I. Güler, A. Dizibüyük, and M. Akin. Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application. *Computers in Biology and Medicine*, 35(7) :603–616, 2005.
- [Knipling98] R. Knipling. PERCLOS : a valid psychophysiological measure of alertness as assessed by psychomotor vigilance. Technical report, Federal Highway Administration (FHWA), 1998.
- [Kothari96] R. Kothari and J. L. Mitchell. Detection of eye locations in unconstrained visual images. In *Proc of the IEEE ICIP conference*, Lausanne, Switzerland, 1996.
- [Lin05a] C-T. Lin, R-C. Wu, T-P. Jung, S-F. Liang, and T-Y. Huang. Estimating driving performance based on EEG spectrum analysis. *EURASIP Journal on Applied Signal Processing*, 19 :3165–3174, 2005.
- [Lin05b] C. Lin, R. Wu, S. Liang, W. Chao, Y. Chen, and T. Jung. EEG -based drowsiness estimation for safety driving using independent component analysis. *IEEE Transactions on Circuits and Systems*, 52(12) :2726–2738, 2005.
- [Lippmann87] R. Lippmann. An introduction to computing with neural nets. *IEEE ASSP Magazine*, 4(2) :4–22, 1987.
- [Loh97] W.-Y. Loh and Y.-S. Shih. Split selection methods for classification trees. *Statistica Sinica*, 7 :815–840, 1997.
- [Makeig00] S. Makeig, T.-P. Jung, and T. J. Sejnowski. Awareness during drowsiness : Dynamics and electrophysiological correlates. *Canadian Journal of Experimental Psychology*, 54(4) :266–273, 2000.
- [Mallat89] S. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transaction on pattern analysis and machine intelligence*, 11(7) :674–693, 1989.
- [Malmivuo95] J. Malmivuo and R. Plonsey. *Bioelectromagnetism - Principles and Applications of Bioelectric and Biomagnetic Fields*. Oxford University Press, 1995.
- [Marmor93] M. F. Marmor and E. Zrenner. Standard for clinical electro-oculography. *Documenta Ophthalmologica*, 85 :115–124, 1993.
- [Matousek83] M. Matousek and I. Petersén. A method for assessing alertness fluctuations from EEG spectra. *Electroencephalography and clinical neurophysiology*, 55(1) :108–113, 1983.
- [McNemar47] Q. McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12 :153–157, 1947.
-

-
- [Mohanty09] M. Mohanty, A. Mishra, and A. Routray. A non-rigid motion estimation algorithm for yawn detection in human drivers. *International Journal of Computational Vision and Robotics*, 1(1) :89–109, 2009.
- [Morris96] T.C. Morris and J.C. Miller. Electrooculographic and performance indices of fatigue during simulated flight. *Biological Psychology*, 42 :343–360, 1996.
- [MPTSite] <http://mplab.ucsd.edu/grants/project1/free-software/mptwebsite/api/>.
- [Muzet03] A. Muzet, T. Pébayle, J. Langrognet, and S. Otmani. AWAKE pilot study no.2 : Testing steering grip sensor measures. Technical Report IST-2000-28062, CEPA, 2003.
- [Noguchi07] Y. Noguchi, R. Nopsuwanchai, M. Ohsuga, and Y. Kamakura. Classification of blink waveforms towards the assessment of driver’s arousal level - an approach for HMM based classification from blinking video sequence. In *Proc. of the Symposium on Biological and Physiological Engineering*, Harbin, China, 2007.
- [Ohsuga07] M. Ohsuga, Y. Kamakura, Y. Inoue, Y. Noguchi, and R. Nopsuwanchai. Classification of blink waveforms toward the assessment of driver’s arousal levels - an EOG approach and the correlation with physiological measures. In *Proc. of the Symposium on Biological and Physiological Engineering*, Harbin, China, 2007.
- [Ohzeki06] K. Ohzeki and B. Ryo. Video analysis for detecting eye blinking using a high-speed camera. In *Proc. of the 40th Signals, Systems and Computers conference*, pages 1081–1085, Pacific Grove, USA, 2006.
- [Oken06] B.S. Oken, M.C. Salinsky, and S.M. Elsas. Vigilance, alertness, or sustained attention : physiological basis and measurement. *Clinical Neurophysiology*, 117 :1885–1901, 2006.
- [Omi08] T. Omi, F. Nagai, and T. Komura. Driver drowsiness detection focused on eyelid behaviour. In *Proc. of the 34th Congress on Science and Technology of Thailand*, Bangkok, Thailand, 2008.
- [ONISR09] J. Chapelon. La sécurité routière en France - Bilan de l’année 2008. Technical report, ONISR, 2009.
- [Otmani05] S. Otmani, T. Pebayle, J. Rogé, and A. Muzet. Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. *Physiology and Behavior*, 84 :715 – 724, 2005.
- [Pal08] N. R. Pal, C.-Y. Chuang, L.-W. Ko, C.-F. Chao, T.-J. Jung, S.-F. Liang, and C.-T. Lin. EEG-based subject- and session-independent drowsiness detection : An unsupervised approach. *EURASIP Journal on Advances in Signal Processing*, 2008, 2008.
- [Papadelis06] C. Papadelis, C. Kourtidou-Papadeli, P. D. Bamidis, I. Chouvarda, D. Koufogiannis, E. Bekiaris, and N. Maglaveras. Indicators of sleepiness in an ambulatory EEG study of night driving. In *Proc. of the 28th IEEE EMBS conference*, New York City, USA, 2006.
-

-
- [Parmar02] N. Parmar. Drowsy driver detection system. Technical report, Department of Electrical and Computer Engineering, Ryerson University., 2002.
- [Peiris05] M. Peiris, R. Jones, P. Davidson, G. Carroll, T. Signal, P. Parkin, M. van den Berg, and P. Bones. Identification of vigilance lapses using EEG/EOG by expert human raters. In *Proc. of the 27th IEEE EMBS conference*, pages 5735–5738, Shanghai, China, 2005.
- [Peiris06] M. Peiris, R. Jones, P. Davidson, G. Carroll, and P. Bones. Frequent lapses of responsiveness during an extended visuomotor tracking task in non-sleep-deprived subjects. *Journal of Sleep Research*, 15 :291–300, 2006.
- [Peng05] K. Peng, L. Chen, S. Ruan, and G. Kukharev. A robust algorithm for eye detection on gray intensity face without spectacles. *Journal of Computer Science and Technology*, 5(3) :127–132, 2005.
- [Picot08] A. Picot, S. Charbonnier, and S. Caplier. On-line automatic detection of driver drowsiness using a single electroencephalographic channel. In *Proc. of the 30th IEEE EMBS conference*, pages 3864–3867, Vancouver, Canada, 2008.
- [Picot09] A. Picot, S. Charbonnier, and A. Caplier. *Recent Advances in Biomedical Engineering*, chapter Monitoring drowsiness on-line using a single encephalographic channel. IN-TECH, 2009, à paraître.
- [Picot09b] A. Picot, A. Caplier, and S. Charbonnier. Comparison between EOG and high frame rate camera for drowsiness detection. In *Proc. of the IEEE Workshop on Applications of Computer Vision*, Snowbird, USA, 2009.
- [Popieul00] J-C. Popieul, P.Simon, M-A. Dillies, and J-C. Angué. Car driver eye scanning behaviour : Study during a long duration highway travel in a simulator. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, Detroit, USA, 2000.
- [Quinlan86] J. Quinlan. Induction of decision trees. *Machine Learning*, 1 :81–106, 1986.
- [Ragot90] J. Ragot, M. Darouach, D. Maquin, and G. Bloch. *Validation de données et diagnostic* *Traité des nouvelles technologies*. Traité des nouvelles technologies, 1990.
- [Rechtschaffen68] A. Rechtschaffen and A. Kales. *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subject*. National Institute of Health Publication, Washington DC, us government printing office edition, 1968.
- [Renner97] G. Renner and S. Mehring. Lane departure and drowsiness - two major accident causes - one safety system. Technical report, Transport Research Laboratory, 1997.
- [Rimini01] M. Rimini-Doering, D. Manstetten, T. Altmueller, U. Ladstaetter, and M. Mahler. Monitoring driver drowsiness and stress in a driving simulator. In *Proc. Intern. Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Aspen, USA, 2001.
- [Rosipal07] R. Rosipal, B. Peters, G. Kecklund, T. Akerstedt, G. Gruber, M. Woertz, P. Anderer, and G. Dorffner. EEG-based drivers' drowsiness monitoring
-

- using a hierarchical gaussian mixture model. *Hum. Comp. Interaction*, 16 :294–303, 2007.
- [Santamaria87] J. Santamaria and K. H. Chiappa. the EEG of drowsiness in normal adults. *Journal of clinical Neurophysiology*, 4(4) :327–382, 1987.
- [Schleicher08] R. Schleicher, N. Galley, S. Briest, and L. Galley. Blinks and saccades as indicators of fatigue in sleepiness warnings : looking tired? *Ergonomics*, 51(7) :982–1010, 2008.
- [Sensation] <http://www.sensation-eu.org>.
- [Shafer76] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
- [Sharabaty08] H. Sharabaty, B. Jammes, and D. Esteve. EEG analysis using HHT : One step toward automatic drowsiness scoring. In *22nd Int. Conf. on Advances Information Networking and Applications*, 2008.
- [Sharabaty08b] H. Sharabaty, D. Esteve, and B. Jammes. Automatic EOG analysis : a first step toward automatic drowsiness scoring during wake sleep transitions. *Somnologie*, 12(3) :227–232, 2008.
- [Sharabaty08c] H. Sharabaty. *Détection de la somnolence d'un opérateur : analyse automatique de signaux physiologiques*. PhD thesis, Université de Toulouse, 2008.
- [Simon03] P. Simon, J-C. Popieul, P. Loslever, R. Leroux, and D. Sonnerat. Car drivers eyes scanning behaviour during a monotonous journey on highway. In *Proceedings of the IEEE Intelligent Vehicles Symposium*, Columbus, USA, 2003.
- [Singh99] S. Singh and N.P. Papanikolopoulos. Monitoring driver fatigue using facial analysis techniques. In *IEEE. Intelligent Transport System Proceedings*, Tokyo, Japan, 1999.
- [Sipina] <http://eric.univ-lyon2.fr/ricco/sipina.html>.
- [Smit04] A. Smit, P. Eling, and A. Coenen. Mental effort affects vigilance enduringly : after-effects in EEG and behavior. *International Journal of Psychophysiology*, 53 :239–243, 2004.
- [Smith03] P. Smith, M. Shah, and N. da Vitoria Lobo. Determining driver visual attention with one camera. *IEEE Transactions on intelligent transportation systems*, 4 :205–218, 2003.
- [SR08] Sécurité Routière. Les grandes données de l'accidentologie en 2007, 2008.
- [Stern00] E. J. Sirevaag and J. A. Stern. *Engineering psychophysiology*, chapter Ocular measures of fatigue and cognitive factors, pages 269–288. Lawrence Erlbaum Associates, 2000.
- [Stern94] J. A. Stern, D. Boyer, and D. Schroeder. Blink rate : a possible measure of fatigue. *Human Factors*, 36 :285–297, 1994.
- [Su07] H. Su and G. Zheng. A non-intrusive drowsiness related accident prediction model based on D-S evidence theory. In *Proc of the 1st International*
-

- Conference on Bioinformatics and Biomedical Engineering*, Wuhan, China, 2007.
- [Subasi05] A. Subasi. Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. *Expert systems with applications*, 28(4) :701–711, 2005.
- [Svensson04] U. Svensson. Blink behaviour based drowsiness detection ũ method development and validation. Master’s thesis, Linköping University, Dept. Biomedical Engineering, 2004.
- [Tova] <http://www.tovatest.com/>.
- [Trejo05] L. J. Trejo, R. Kochavi, K. Kubitz, L. D. Montgomery, R. Rosipal, and B. Matthews. EEG-based estimation of cognitive fatigue. In *Proceedings of SPIE : Bio-monitoring for physiological and cognitive performance during military operations.*, pages 105–115, Orlando, USA, 2005.
- [Ueno94] H. Ueno, M. Kaneda, and M. Tsukino. Development of drowsiness detection system. In *Proc. of the Vehicle Navigation and Information Systems Conference*, Yokohama, Japan, 1994.
- [Vigario97] R. N. Vigario. Extraction of ocular artefacts from EEG using independent component analysis. *Electroencephalography and Clinical Neurophysiology*, 103(3) :395–404, 1997.
- [Viola02] P. Viola and M. Jones. Robust real-time object detection. *International Journal of Computer Vision*, 57(2) :137–154, 2002.
- [Vion08] J. Vion-Dury and F. Blanquet. *Pratique de l’EEG - Bases neurophysiologiques - Principes d’interprétation et de prescription*. Broché, 2008.
- [Virkkala07] J. Virkkala, J. Hasan, A. Värri, S.-L. Himanen, and M. Härmä. The use of two-channel electro-oculography in automatic detection of unintentional sleep onset. *Journal of Neuroscience Methods*, 163 :137–144, 2007.
- [Vural07] E. Vural, M. Cetin, A. Ercil, G. Littlewort, M. Bartlett, and J. Movellan. Drowsy driver detection through facial movement analysis. *Human-Computer Interaction*, 4796 :6–18, 2007.
- [Welch67] P.D. Welch. The use of fast fourier transform for the estimation of power spectra : A method based on time averaging over short, modified periodograms. *IEEE Trans. Audio & Electroacoust.*, 15 :70–73, 1967.
- [Wierwille94] W.W. Wierwille, L.A. Ellworth, S.S. Wreggit, R.J. Fairbanks, and C.L. Kim. Research on vehicle-based driver status/performance monitoring : development, validation, and refinement of algorithms for detection of driver drowsiness. Technical report, National Highway Traffic Safety Administration, 1994.
- [Wilson00] B. J. Wilson and T. D. Bracewell. Alertness monitor using neural networks for EEG analysis. In *Proc. of the 2000 IEEE Signal Processing Society Workshop*, volume 2, pages 814–820, Sydney, Australia, 2000.
- [Wylie96] C.D. Wylie, T. Shultz, J.C. Miller, M.M. Milter, and R.R. Mackie. Commercial motor vehicle driver fatigue and alertness study. Technical report, Federal Highway Administration, U.S. Department of Transportation, 1996.
-

- [Yang02] M.H. Yang, D. Kriegman, and N. Ahuja. Detecting face in images : a survey. *IEEE Trans on PAMI*, 24 :34–58, 2002.
- [Yang05] G. Yang, Y. Lin, and P. Bhattacharya. A driver fatigue recognition model using fusion of multiple features. In *Proc. of the IEEE International Conference on Systems, Man and Cybernetics*, Hawaii, USA, 2005.
- [Zadeh65] L. Zadeh. Fussy sets. *Information and Control*, 8(3) :338–353, 1965.
- [Zhou04] Z.-H. Zhou and X. Geng. Projection functions for eye detection. *Pattern Recognition*, 37(5) :1049–1056, 2004.
-

Publications

Chapitres de Livre

- [Picot09] A. Picot, S. Charbonnier, et A. Caplier. Recent Advances in Biomedical Engineering, chapter *Monitoring drowsiness on-line using a single encephalographic channel*. IN-TECH, 2009, à paraître.

Conférences internationales

- [Picot09b] A. Picot, A. Caplier, and S. Charbonier. Comparison between EOG and high frame rate camera for drowsiness detection. In *Proc. of the IEEE Workshop on Applications of Computer Vision*, Snowbird, USA, 2009.
- [Picot08] A. Picot, S. Charbonnier, et S. Caplier. On-line automatic detection of driver drowsiness using a single electroencephalographic channel. In *Proc. of the 30th IEEE EMBS conference*, pages 3864-3867, Vancouver, Canada, 2008.
-

