



**HAL**  
open science

# Nouvelle approche pour l'estimation du rythme respiratoire basée sur la photopléthysmographie sans contact

Duncan Luguern

► **To cite this version:**

Duncan Luguern. Nouvelle approche pour l'estimation du rythme respiratoire basée sur la photopléthysmographie sans contact. Traitement du signal et de l'image [eess.SP]. Université de Bourgogne Franche-Comté, 2021. Français. NNT: . tel-03169483v1

**HAL Id: tel-03169483**

**<https://u-bourgogne.hal.science/tel-03169483v1>**

Submitted on 15 Mar 2021 (v1), last revised 19 May 2021 (v2)

**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.



RÉGION  
BOURGOGNE  
FRANCHE  
COMTÉ



THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT UNIVERSITÉ  
BOURGOGNE FRANCHE-COMTÉ

PRÉPARÉE À L'UNIVERSITÉ DE BOURGOGNE

École doctorale n°37

Sciences pour l'Ingénieur et Microtechniques

Doctorat d'Informatique et Instrumentation de l'Image

par

DUNCAN LUGUERN

Nouvelle approche pour l'estimation du rythme respiratoire basée sur la  
photopléthysmographie sans contact

Thèse présentée à Dijon, le 5 janvier 2021

Composition du Jury :

BERTRAND GRANADO	Professeur à Sorbonne Université	Rapporteur
OLIVIER MESTE	Professeur à Université Côte d'Azur	Rapporteur
FRÉDÉRIC BOUSEFSAF	Maître de Conférences à Université de Lorraine	Examineur
FAN YANG	Professeure à Université de Bourgogne	Présidente du Jury
YANNICK BENEZETH	Maître de conférences HDR à Université de Bourgogne	Co-directeur de thèse
JULIEN DUBOIS	Maître de conférences HDR à Université de Bourgogne	Directeur de thèse



# Sommaire

<b>Remerciements</b>	<b>7</b>
<b>Notations mathématiques</b>	<b>9</b>
<b>I Contexte et motivations</b>	<b>11</b>
<b>1 Introduction</b>	<b>13</b>
1 La photopléthysmographie et la photopléthysmographie sans contact . . . . .	14
2 Applications possibles de la photopléthysmographie sans contact . . . . .	16
2.1 Détection de maladies . . . . .	16
2.2 Suivi de l'état de santé d'un patient . . . . .	18
2.3 Suivi de personnes au quotidien . . . . .	20
2.4 Applications non médicales . . . . .	21
3 Objectifs et contributions de la thèse et plan du manuscrit . . . . .	23
<b>2 Mesure de paramètres physiologiques : état de l'art et analyse</b>	<b>25</b>
1 Les appareils et les méthodes de mesure de paramètres physiologiques . . . . .	26
1.1 Les méthodes utilisant un contact . . . . .	27
1.2 Les méthodes sans contact . . . . .	36
2 Algorithmes et dispositifs pour la mesure par la rPPG . . . . .	40
2.1 Configurations expérimentales . . . . .	40
2.2 Traitement vidéo . . . . .	43
2.3 Pré-traitements . . . . .	45
2.4 Combinaison des signaux RGB . . . . .	46
2.5 Estimation de paramètres physiologiques . . . . .	56
2.6 Estimation de la qualité du signal . . . . .	61
2.7 Le cas de l'apprentissage profond . . . . .	62



<b>II Contributions</b>	<b>65</b>
<b>3 Analyse du lien entre vecteur de combinaison et qualité du signal : Méthode Bruteforce</b>	<b>67</b>
1 Étude préliminaire . . . . .	68
2 Méthode de combinaison des signaux couleur . . . . .	74
<b>4 Méthodes d'extraction de la modulation de la ligne de base</b>	<b>79</b>
1 EVM-FFT . . . . .	81
2 EVM-filtre . . . . .	83
3 WVM . . . . .	86
4 DWVM . . . . .	91
<b>5 Protocole expérimental et comparaison des méthodes</b>	<b>95</b>
1 Base de validation . . . . .	96
2 Métriques . . . . .	98
3 Graphiques . . . . .	99
4 Méthodes de l'état de l'art implémentées . . . . .	100
5 Résumé des traitements . . . . .	101
6 Comparaison des méthodes . . . . .	102
6.1 Temps de calcul . . . . .	102
6.2 Métriques et graphiques . . . . .	103
<b>6 Fusion des modulations</b>	<b>111</b>
1 Méthode . . . . .	112
2 Résultats obtenus . . . . .	115
<b>7 Conclusion et projets</b>	<b>119</b>
1 Conclusion . . . . .	120
2 Discussion et ouvertures . . . . .	122
<b>Glossaire médical</b>	<b>123</b>
<b>Table des abréviations</b>	<b>127</b>
<b>References</b>	<b>131</b>
<b>Publications personnelles</b>	<b>131</b>

**Titre** : Résumé

**Mots-clés** : rPPG, GEVD, Respiration

**Description** : Le rythme respiratoire est une information importante dans le contexte médical puisqu'elle permet de prédire plusieurs complications potentiellement mortelles. Malgré cela, elle est souvent négligée par le personnel médical faute de temps ou de bien comprendre les enjeux associés. Dans ce contexte, les méthodes de mesure automatisées permettent d'améliorer le statu quo en fournissant en continu une mesure du rythme respiratoire. La plupart des méthodes actuelles comme la ceinture respiratoire ou l'ECG nécessitent un contact avec la personne pour pouvoir mesurer efficacement le rythme respiratoire. Malheureusement, cela introduit des problèmes qui peuvent empêcher la mesure dans certains cas ou la rendre contraignante lors d'une mesure en continu et au quotidien, là où il serait souhaitable que la mesure soit la plus discrète possible. Afin de pallier à ces problèmes, plusieurs méthodes de mesure du rythme respiratoire sans contact sont actuellement en développement. Parmi celles-ci, la photopléthysmographie sans contact utilise la variation de la couleur de la peau en fonction du volume sanguin présent dans les capillaires afin de trouver un signal cardiaque et respiratoire. Dans la thèse présentée, nous nous attachons à améliorer la qualité de la mesure du rythme respiratoire à l'aide de la photopléthysmographie sans contact en développant des méthodes dont le but est de combiner efficacement les signaux couleur extraits à partir d'une vidéo de manière à obtenir un seul signal maximisant l'information respiratoire. Dans un deuxième temps, une chaîne de traitement est mise en place de façon à utiliser ces méthodes de combinaison pour déterminer le rythme respiratoire en utilisant toutes les informations pouvant être extraites du signal photopléthysmographique.

**Title** : Abstract

**Keywords** : rPPG, GEVD, Breathing

**Description** : Respiratory rhythm is important information in medical context. Its assessment allows to predict some medical complications that could lead to death. However, it is often neglected by the medical staff due to a bad comprehension of its importance, or a lack of time. Automated measurement methods allow to improve this by continuously giving respiratory rate. Most of these methods needs a contact with the patient to efficiently measure the breathing rate. Unfortunately it leads to some issues which could forbid measurement or make it uncomfortable for continuous monitoring. The continuous, every-day monitoring especially needs to be as discrete as possible to be forgotten by the patient. To deal with these drawbacks, several non contact respiratory rate assessment methods are currently developed. In these methods, the remote photoplethysmography uses the color variation of the skin due to blood volume in capillaries to obtain a cardiac and respiratory correlated signal. In this thesis, we focused on improving the respiratory rate measurement using remote photoplethysmography. To do this, we developed some methods whose goals are to efficiently combine color signals that were extracted from the video to obtain a single signal that maximizes the breathing information. In a second part, we developed a processing pipeline to assess the breathing rate from all informations that could be extracted from remote photoplethysmography signals. The processing pipeline was tested with state of the art combination methods and the methods that were developed during the thesis.



---

## REMERCIEMENTS

La thèse a pu être réalisée à l'aide d'un cofinancement du Conseil Régional de Bourgogne Franche-Comté et du laboratoire ImViA (anciennement LE2i).

Je remercie le laboratoire ImViA qui m'a accueilli durant ces trois années afin de réaliser mes recherches et me former au mieux. Au sein du laboratoire, je pense particulièrement à mes encadrants, M. Julien Dubois et M. Yannick Benezeth, qui m'ont beaucoup aidé en me guidant dans mes recherches, en corrigeant de nombreuses fois les écrits que j'ai pu produire et en résolvant les différentes difficultés que j'ai pu rencontrer concernant ma recherche comme l'administratif. Je remercie également M. Serge Bobbia, M. Mustapha Bouderbane et M. Ali Douiyek avec lesquels j'ai travaillé notamment durant les stages ingénieurs et qui m'ont amené à m'orienter vers la photopléthysmographie sans contact. Enfin, mes remerciements vont également auprès de M. Richard Macwan dont les travaux et l'expertise ont été une source d'informations importante pour le développement des méthodes présentées, ainsi qu'auprès de M. Simon Perche qui m'a aidé pour la mise en place de la plateforme de validation dans le cadre de son stage de Master.

La thèse a également eu lieu en partie dans le laboratoire du CSEM de Neuchâtel qui m'a accueilli pendant deux stages et financé les défraiements correspondants. Au sein du CSEM, je remercie Mme Virginie Moser et Mme Andrea L. Dunbar qui m'ont encadré ainsi que Mme Laura Bujouves qui a également travaillé sur le suivi de nouveaux nés à l'aide de la photopléthysmographie sans contact.



---

## NOTATIONS MATHÉMATIQUES

Afin de faciliter la lecture des équations présentées dans le manuscrit, les notations ont été harmonisées autant que possible. Les scalaires sont présentés par des symboles en minuscule, les vecteurs sont représentés par des lettres minuscules en gras et les matrices par des lettres capitales en gras. Les signaux correspondent à des signaux temporels de taille  $[N \times M]$  avec  $N$  la durée du signal et  $M$  le nombre de canaux. Ils sont présentés avec une lettre minuscule grasse (pour les signaux unidimensionnels) ou majuscule grasse (pour les signaux multidimensionnels) avec un indice temporel entre parenthèses (par exemple  $\mathbf{s}(t)$ ). Les matrices sont dimensionnées avec un échantillon (par exemple temporel) par ligne et une variable (par exemple la couleur) par colonne. Les vecteurs reprennent cette notation, les vecteurs de combinaison des signaux couleur (détaillés par la suite) sont donc des vecteurs lignes de taille  $[1 \times 3]$ .

Les multiplications scalaire-scalaire ou scalaire-matrice sont représentées à l'aide de l'opérateur  $\times$  ou sans opérateur. Les multiplications matricielles sont notées avec l'opérateur  $\cdot$  et les multiplications de Hadamard (élément par élément) sont indiquées à l'aide de l'opérateur  $\odot$ .

Les indices temporels sont représentés entre parenthèses tandis que les autres types d'indices sont en sous-script. Les opérateurs appliqués sur des matrices ou des signaux multidimensionnels comme la moyenne ou la transformée de Fourier opèrent sur la matrice ou le signal indiqué entre accolades et sur la première dimension (temporelle). Les autres notations qui se retrouvent régulièrement dans le manuscrit sont présentées dans le tableau 1.

TABLE 1 – Table des notations courantes du manuscrit.

Symbole	Description
$\delta x$	Variation d'une variable $x$
$\mu\{x\}$	Moyenne de $x$ le long de la dimension 1 ( $x$ matrice, vecteur ou signal)
$\sigma\{x\}$	Écart-type de $x$ le long de la dimension 1 ( $x$ matrice, vecteur ou signal)
$\mathcal{F}\{\mathbf{s}(t)\}$	Transformée de Fourier de $\mathbf{s}(t)$ le long de la dimension 1
$\mathcal{E}\{\mathbf{s}(t)\}$	Énergie de $\mathbf{s}(t)$ le long de la dimension 1
$\boldsymbol{\omega}_i(t)$	Vecteur de combinaison (ou vecteur propre d'indice $i$ ) à l'instant $t$ [ $1 \times M$ ]
$\boldsymbol{\Omega}$	Matrice de rotation GEVD [ $M \times M$ ]
$\mathbf{P}_x(t)$	Matrice signal utilisée dans la GEVD [ $M \times M$ ]
$\mathbf{C}_x(t)$	Matrice bruit utilisée dans la GEVD [ $M \times M$ ]
$\mathbf{X}(t)$	Signal utile, matrice [ $N \times M$ ] avec $N$ échantillons temporels
$\mathbf{B}(t)$	Signal représentant le bruit, matrice [ $N \times M$ ] avec $N$ échantillons temporels
$\mathbf{x}_c(t)$	Signal utile après combinaison [ $N \times 1$ ]
$\mathbf{b}_c(t)$	Signal représentant le bruit après combinaison [ $N \times 1$ ]
$\mathbf{C}(t)$	Signaux couleur [ $N \times M$ ]
$\mathbf{C}_n(t \rightarrow t+l)$	Signaux couleur [ $N \times M$ ] normalisés et intégrés de $t$ à $t+l$ (normalisation / intégration)
$\mathbf{h}_{f_c}(f)$	Fenêtre de séparation du signal et du bruit centrée en $f_c$ et à la fréquence $f$
$\lambda_i(t)$	Valeur propre d'indice $i$ à l'échantillon $t$

---

Première partie

---

CONTEXTE ET MOTIVATIONS





# Chapitre 1

## Introduction

Ce chapitre présente le contexte général dans lequel la thèse se situe. Dans un premier temps, l'importance des mesures du rythme respiratoire et ses limitations actuelles sont présentées. La photopléthysmographie sans contact, qui est la méthode étudiée au cours de la thèse, présente des caractéristiques intéressantes afin de mesurer le rythme respiratoire. Ces principes généraux ainsi que ces applications possibles sont détaillées afin de positionner cette méthode par rapport à la mesure du rythme respiratoire et par rapport à ses autres champs d'application. Enfin, les différents objectifs de la thèse ainsi que le plan du manuscrit sont détaillés afin de comprendre le cheminement réalisé.

La médecine a fortement évolué au fil des siècles. Ses avancées l'ont rendue plus fiable et plus efficace dans la prise en charge des malades. En parallèle, celle-ci s'est démocratisée. Là où les soins étaient auparavant restreints aux personnes les plus aisées ou les plus puissantes, l'évolution des techniques et leur professionnalisation se sont faites tout en augmentant l'accessibilité des traitements au plus grand nombre.

Au fur et à mesure des avancées médicales, la quantité d'informations nécessaires à la prise de décision a augmenté de manière significative. Des besoins nouveaux en terme de fiabilité de la mesure et de coût notamment ont émergé. Le développement de méthodes de mesures de paramètres physiologiques est alors devenu de plus en plus critique. Ces méthodes présentent de nombreux objectifs, parfois antagonistes : ils doivent notamment être précis, non invasifs, économiques, hygiéniques, simples à utiliser et permettre une mesure continue. En effet, la capacité à étudier un paramètre physiologique dans le cadre d'une surveillance sur une longue durée permet de fournir plus d'informations sur l'état du patient [93, 107]. Dans certains cas, il n'est pas possible de reconnaître une pathologie sans utiliser une surveillance long terme, par exemple dans la recherche de battements ectopiques (battement de cœur qui s'intercale entre deux battements normaux). De plus, la récupération et le traitement de grandes quantités de données, aujourd'hui possible à l'aide de l'apprentissage profond, font entrevoir des possibilités nouvelles en terme de télémédecine. Le fait de déléguer au moins partiellement la prise de décision des médecins aux ordinateurs permet de libérer du temps au praticien tout en permettant de détecter plus tôt des problèmes de santé nécessitant une consultation. Pour permettre de faire une mesure sur une longue durée, il est nécessaire d'avoir une méthode de mesure aussi confortable que possible pour le patient. La question du coût et de la facilité d'utilisation des appareils de mesures est aussi intimement liée à la question de la mesure en continue. En effet, permettre la mesure de paramètres physiologiques à la maison limite l'encombrement des services hospitaliers et est plus confortable pour le patient. Pour cela, le dispositif de mesure doit être suffisamment simple à utiliser pour que le patient puisse effectuer la mesure lui-même. Au delà des questions de confort, la mesure d'un paramètre physiologique se doit d'être la moins ressentie possible afin de ne pas modifier ledit paramètre. Par exemple, faire une mesure douloureuse du rythme cardiaque risque sans doute de l'augmenter et de fausser les informations obtenues.

Dans ce contexte, le travail réalisé au cours de la thèse a consisté à concevoir une méthode de mesure du rythme respiratoire sans contact à l'aide de la rPPG. Tout d'abord, la suite de ce manuscrit présentera brièvement la photopléthysmographie et son extension en mesure sans contact dans le chapitre suivant. Par la suite, les objectifs et les contributions de la thèse seront donnés afin de pouvoir introduire l'état de l'art présenté dans le chapitre 2 et les différentes méthodes développées présentées dans la partie II.

## **1 La photopléthysmographie et la photopléthysmographie sans contact**

La mesure des paramètres physiologiques est un point crucial dans la prise en charge des patients. Dans ce cadre, de nombreuses technologies ont été développées afin de permettre d'alléger le travail des personnels de santé et de rendre plus fiables les décisions que celui-ci peut être amené à prendre. Par exemple, la photopléthysmographie (ou PPG) est une méthode permettant de mesurer une large variété de paramètres physiologiques pour un coût relativement faible [133]. Cette technologie utilise l'absorption de la lumière par le sang pour estimer notamment le rythme cardiaque d'une part, et la saturation en oxygène d'autre part [126]. Elle se retrouve principalement dans les oxymètres qui sont des appareils relativement économiques (environ 100€), non invasifs et assez précis pour exploiter les informations qui sont acquises. Les oxymètres utilisent l'absorption de la lumière par le sang afin de pouvoir extraire des paramètres physiologiques comme le rythme cardiaque, le rythme

respiratoire ou la saturation en oxygène. Son principe de fonctionnement est montré dans la figure 1.1. Les oxymètres présentent de nombreux avantages comme leur faible coût ou le nombre de paramètres pouvant être lus avec un seul appareil. Néanmoins, il est difficile de les utiliser dans le cadre d'une surveillance en continue, puisque la plupart des modèles sont fixés au doigt de la personne. En plus de cela, la plupart des appareils de mesure en contact avec les personnes comme l'oxymètre posent des problèmes d'hygiène. Ils doivent être nettoyés entre deux utilisations ce qui peut représenter un délai supplémentaire et l'utilisation de consommables. Le risque d'oubli ou d'erreur lors du nettoyage implique la possibilité d'une contamination. Enfin, dans certaines situations le patient peut être particulièrement sensible au risque de contamination. On trouve notamment cette situation pour les grands brûlés qui nécessitent donc l'utilisation de capteurs stériles [104]. Les méthodes basées sur un contact avec la peau peuvent également poser des problèmes lorsque celle-ci est particulièrement fragile. Ce problème se pose par exemple dans le cas des bébés prématurés dont la peau est très fragile et peut être irritée par les électrodes de l'électrocardiogramme (ECG) ou par les Diodes Électro-Luminescentes (DEL) de l'oxymètre [116].

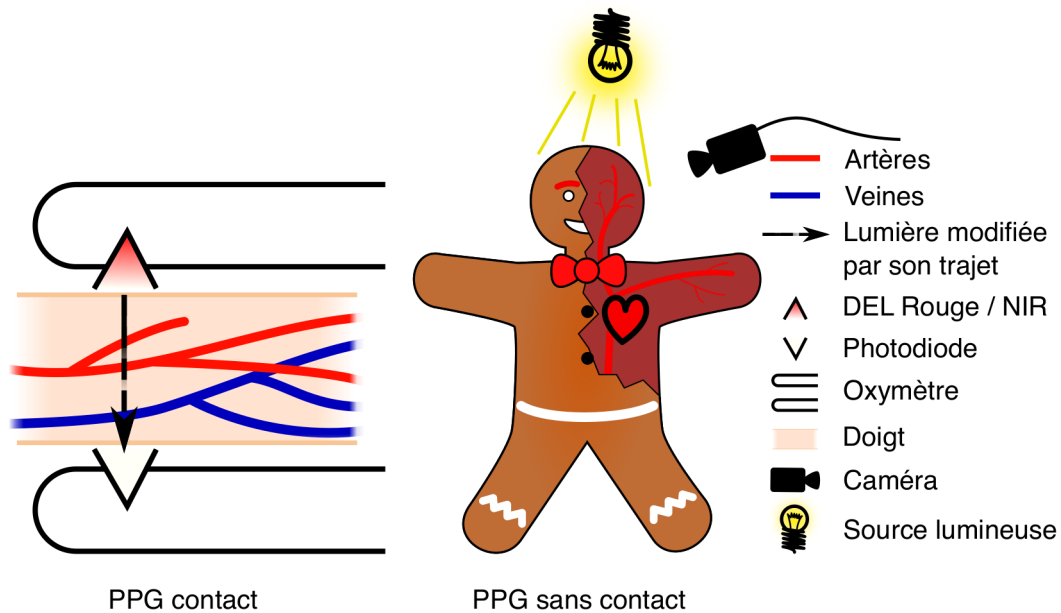


FIGURE 1.1 – Principes de fonctionnement de la photopléthysmographie contact et sans contact.

Afin de résoudre ces différents problèmes, des méthodes n'utilisant pas de contact avec la peau ont été développées ces dernières années. Ces méthodes présentent généralement quelques inconvénients comme un rapport signal sur bruit plus faible. Néanmoins, elles résolvent certains problèmes posés par les méthodes utilisant un contact. De plus, certaines mesures peuvent être faites à l'aide d'appareils du quotidien. C'est notamment le cas de la photopléthysmographie sans contact qui peut être utilisée sur des smartphones [115] ou des ordinateurs [139]. Cette technique a été étudiée au cours de cette thèse. Ses nombreux avantages lui donnent un avenir prometteur tandis que ses limitations actuelles posent de nombreux défis à relever dans la recherche afin de rendre cette technologie utilisable à la fois par le grand public et par le personnel médical. La méthode repose sur les mêmes principes que la photopléthysmographie contact, mais en employant une caméra et une source de lumière distante (éventuellement naturelle). Cela permet de ne pas nécessiter de contact avec la peau, mais en contrepartie réduit généralement la qualité du signal obtenu et donc la fiabilité de la mesure.

Une fois mature, la photopléthysmographie sans contact, ou rPPG (pour *remote photoplethysmography*), pourra être utilisée dans de nombreux scénarios afin d'améliorer la qualité de la prise en charge médicale. La possibilité de faire la mesure sans contact apporte un bénéfice dans le cas où la peau serait fragile comme pour les grands brûlés ou les prématurés. De plus, le matériel nécessaire à sa mise en place est relativement commun et permet donc un déploiement plus rapide et moins onéreux de la technique en comparaison des autres méthodes qui vont généralement nécessiter des dispositifs moins communs et plus cher (comme par exemple des caméras thermiques). D'autres opportunités sont également prévisibles comme l'utilisation de la photopléthysmographie sans contact afin de renforcer la sécurité des systèmes biométriques [81] ou de faciliter la détection des *deepfakes* [27]. Nous présenterons dans le détail la photopléthysmographie et sa variante sans contact dans la section 1 du chapitre 2 mais dans un premier temps, afin de mieux comprendre l'intérêt de la rPPG, un certain nombre d'applications potentielles seront présentées dans la suite de cette introduction.

## 2 Applications possibles de la photopléthysmographie sans contact

La photopléthysmographie sans contact ouvre des voies prometteuses dans plusieurs champs d'applications. Qu'il s'agisse de l'implémenter sur des dispositifs légers et quotidiens (smartphones) ou de l'utiliser dans un contexte médical avec un matériel dédié, cette méthode apporte beaucoup d'informations utiles tout en restant suffisamment simple dans sa mise en place.

Le fait d'utiliser une caméra plutôt que des capteurs spécialisés permet à la photopléthysmographie sans contact de capitaliser sur l'ensemble des évolutions technologiques dans le domaine de l'image, et donc de profiter de la miniaturisation du capteur, de sa baisse de coût, de ses performances plus élevées et de son omniprésence dans le quotidien.

Ainsi, on peut généralement classer les applications en fonction de leur accessibilité au grand public. La création de capteurs photopléthysmographiques pouvant se faire avec des caméras existantes (smartphones) ou dans des dispositifs dédiés. Les dispositifs dédiés pour la photopléthysmographie sans contact est utile par exemple aux utilisations dans un contexte de consultation en cabinet, avec la création de caméras et de sources de lumière dédiées. Les capteurs et sources de lumière pouvant alors tirer parti du matériel et bénéficier d'une meilleure qualité de signal en modifiant la fréquence d'échantillonnage, la résolution de la caméra, les longueurs d'onde mesurées ou en ajoutant des filtres et des optiques spéciales. La création d'un dispositif dédié permet également de pouvoir profiter d'une puissance de calcul plus adaptée avec éventuellement l'emploi de circuits pour la parallélisation du calcul via des GPU (*Graphical Processing Unit*), des FPGA (*Field Programmable Gate Array*) ou des ASIC (*Application Specific Integrated Circuit*).

Les différentes applications proposées par la suite correspondent à des besoins pour lesquels la photopléthysmographie sans contact peut apporter des solutions. La plupart de ces situations peuvent être gérées à l'aide d'autres capteurs ou algorithmes, néanmoins la photopléthysmographie sans contact apporte à ces problèmes des réponses qui sont généralement moins coûteuses et donc plus simples et plus rapides à déployer.

### 2.1 Détection de maladies

Dans la majorité des applications médicales de la photopléthysmographie sans contact, le but final est de pouvoir faciliter et améliorer la détection de pathologies. Dans un rapport

de la cour des comptes publié en 2019<sup>1</sup>, celle-ci estime qu'une consultation aux urgences sur cinq ne donne lieu à aucune intervention médicale. Le fait de pouvoir faire un prédiagnostic à la maison pourrait probablement réduire ce chiffre, pour peu que la méthode soit précise et simple à mettre en œuvre. De ce fait, les nouvelles technologies en général et la photopléthysmographie en particulier trouvent toute leur place pour répondre à ce nouveau besoin. En effectuant la mesure du signal rPPG et en l'analysant, il est possible de détecter des maladies, notamment cardiaques. Bien que ces détections ne puissent pas remplacer un avis médical, ils peuvent motiver ou démotiver un passage aux urgences ou en consultation. De plus, certaines anomalies cardiaques sont intermittentes. Leur détection se fait en appareillant le patient pour mesurer au quotidien les signaux physiologiques. À l'heure actuelle on trouve notamment l'utilisation d'un holter ECG pour faire la mesure [50]. Il s'agit d'un petit appareil porté sous les vêtements et qui effectue la mesure d'un électrocardiogramme avec moins de dérivations que lors d'un examen au cabinet. Cet appareil, bien que fiable, nécessite un contact et son remplacement par des dispositifs sans contact améliorerait le confort du patient.

De nombreuses maladies peuvent être analysées à l'aide des méthodes décrites précédemment comme par exemple les extrasystoles. Les extrasystoles correspondent à des battements supplémentaires dans le rythme cardiaque. Certaines maladies plus complexes peuvent être détectées en analysant l'activité électrique du cœur avec un ECG. On trouve par exemple la fibrillation qui correspond à une décharge anarchique et inefficace des foyers ectopiques (qui sont en quelque sorte des roues de secours du nœud sinusal). Il est également possible que le nœud sinusal cesse complètement son activité. Dans ce cadre là, les foyers ectopiques prennent le relais à un rythme plus faible pour maintenir une activité cardiaque. Selon le foyer ectopique qui prend le relais, le rythme de l'onde ainsi que sa forme peuvent être modifiés. Certaines maladies du système circulatoire peuvent également être évaluées à l'aide de ces méthodes. Par exemple, les pathologies vasculaires périphériques correspondent à des rétrécissements des artères ou des veines dans le système circulatoire. On peut les détecter à l'aide de la photopléthysmographie contact en regardant la forme des signaux en plusieurs sites de mesure [10]. Cela laisse penser qu'il serait possible d'effectuer également cette mesure en utilisant la rPPG, à condition d'avoir accès aux deux sites et de s'assurer qu'il n'y ait pas de variation des conditions, par exemple avec moins de luminosité ambiante dans l'un des deux sites de mesure.

Un autre exemple de pathologie qui pourrait être détectée à l'aide de la rPPG est le COVID19. Lors de l'épidémie de COVID19 en 2020, il a été mis en évidence un besoin accru de nombreux produits, notamment en masques et en tests virologiques. L'explosion du nombre de contaminations a provoqué des tensions fortes sur les circuits d'approvisionnement habituels et introduit un besoin à innover afin de réduire ces manques. Concernant les tests virologiques, un ensemble de méthodes permettant la détection de la maladie ont été essayés afin de permettre un diagnostic plus précoce et favoriser le respect strict des mesures barrières pour les personnes potentiellement infectées. On retrouve dans ces nouveautés notamment l'interrogatoire du patient [167] afin d'étudier les risques et l'analyse des images de radiographie par le biais d'algorithmes d'apprentissage profond [153]. Une des méthodes étudiées consistait à mesurer la variabilité du rythme cardiaque afin de pouvoir diagnostiquer les cas positifs au COVID19 [113]. Il s'est également avéré que la saturation en oxygène pouvait être non seulement un marqueur du COVID19 mais également permettre de prédire une détérioration de l'état de santé des patients [164]. Bien que cette mesure n'ait pas été testée avec la photopléthysmographie sans contact, la  $S_pO_2$  fait partie des paramètres physiologiques qui pourraient être mesurés à l'aide de la rPPG dans le futur. Le but de ces méthodes était de permettre d'effectuer ces tests en utilisant le moins de ressources possibles (examen médical, consommables) afin de pouvoir tester rapidement un grand nombre de personnes. À ce titre, l'interrogatoire du patient est une méthode pertinente mais elle repose sur les

---

1. <https://www.ccomptes.fr/system/files/2019-02/08-urgences-hospitalieres-Tome-2.pdf>

déclarations faites par les personnes qui peuvent éventuellement être sous-estimées. À l'inverse, la détection des cas COVID19 à l'aide de la variabilité du rythme cardiaque présente l'intérêt d'être assez robuste aux erreurs de manipulation de la part du patient. La question posée par cette mesure est de réussir à obtenir ce paramètre en utilisant du matériel présent au quotidien d'une part, et facilement utilisable d'autre part par le patient. C'est ainsi que des groupes de recherche se sont intéressés à la possibilité d'utiliser la photopléthysmographie sans contact en implémentant les algorithmes dans une application pour smartphone. Ce travail ne semble pas avoir fait l'objet de publications à l'heure actuelle. À partir de là, il devient possible de faire la mesure chez soi, seul et sans utiliser de matériel particulier. Ces trois critères permettent potentiellement de tester l'intégralité de la population quotidiennement sans problème de surcoût. Cette méthode se base néanmoins sur la mesure de la variabilité du rythme cardiaque, qui reste assez difficile à obtenir à l'aide de la rPPG mesurée via des dispositifs quotidiens (du fait de la faible fréquence d'échantillonnage).

## 2.2 Suivi de l'état de santé d'un patient

Bien que la détection de maladies soit sans doute l'application la plus importante de la rPPG, d'autres applications comme le suivi d'un patient sont également très intéressantes à mettre en œuvre. La possibilité de suivre des paramètres physiologiques présente bien entendu un intérêt pour le diagnostic. En plus de cela, il est également possible d'en tirer parti dans le cadre de suivi de lésions, de la surveillance de la perfusion d'un organe lors d'une opération chirurgicale ou de l'observation de l'état de santé de nourrissons.

L'observation de l'évolution des lésions comme les coupures se fait généralement de manière visuelle. Cette méthode a pu être légèrement améliorée avec l'informatique en permettant la sauvegarde des photos pour comparer, ou l'utilisation de l'intelligence artificielle [163]. Malheureusement, ces technologies passent outre les informations les plus importantes dans le cadre du suivi de la cicatrisation. En effet, l'important n'est pas tant l'aspect extérieur de la plaie mais l'état sous-cutané de la plaie. La présence d'une circulation et d'une oxygénation optimale par exemple sont très intéressants pour pouvoir prendre les meilleures décisions possibles. Afin de faire cela, la photopléthysmographie sans contact présente plusieurs caractéristiques qui la rendent intéressante dans ce contexte. En effet, celle-ci produit une image de la peau, ce qui permet de visualiser l'état de la plaie dans son ensemble et non pas en un point donné [16]. De plus, la photopléthysmographie permet la mesure de la saturation en oxygène et du rythme cardiaque qui peuvent être par la suite utilisés dans cette application. Enfin, cette méthode peut se baser sur des appareils relativement peu coûteux (voir même « gratuit » si implémenté sur un smartphone). Ce dernier point est important puisqu'il permet d'effectuer ce suivi sur un grand nombre de plaies sans trop se soucier du coût du suivi. En supposant que le patient ait la possibilité de faire lui-même la prise de vue, cela permet également d'éviter des trajets supplémentaires et de réduire le temps dépensé pour le patient et le médecin.

Dans un contexte hospitalier, la photopléthysmographie sans contact pourrait permettre d'améliorer la prise en charge dans le cadre d'opérations chirurgicales par exemple. L'étude de la perfusion d'un organe n'a pas encore été étudiée à l'aide de la photopléthysmographie sans contact. Néanmoins, plusieurs systèmes photopléthysmographiques au contact ont été développés dans le cadre de la recherche afin de vérifier la bonne perfusion et oxygénation d'organes au cours d'une opération chirurgicale. Les deux principaux cas d'études sont le foie et l'intestin. Dans le cas du foie, le principal but est le suivi post-greffe, considérant qu'un nombre assez conséquent de rejets se produisent dans les jours suivant la greffe. Le suivi post-opératoire ne permet pas à l'heure actuelle le suivi de la bonne perfusion et oxygénation de l'organe greffé, ce qui permettrait de réagir plus précocement à un éventuel rejet. Afin de permettre le suivi du foie, deux études *in vitro* ont porté sur la survie après transplant [7, 9]. Celles-ci ont respectivement étudié la perfusion sanguine au niveau de la

veine porte (la veine principale conduisant le sang vers le foie) et la forme du signal PPG. Le but de ces études était de valider la faisabilité d'un dispositif de suivi continu. Dans une autre étude [9] le dispositif est cette fois testé en utilisant des porcs comme cobayes. Dans cette étude le but est de mesurer la perfusion du foie et sa consommation en oxygène lors d'occlusions vasculaires (blocage d'un vaisseau). L'intestin fait également l'objet d'études sur la possibilité d'effectuer un suivi à l'aide de la photopléthysmographie [8, 59]. Dans ces travaux, l'objectif est de suivre l'oxygénation de l'intestin ainsi que sa perfusion. La première étude [8] porte sur l'oxygénation et la perfusion de l'intestin dans le cadre d'une opération chirurgicale sur un patient victime de trauma abdominal. Dans la deuxième étude [59] la perfusion de l'intestin est suivi au cours d'une opération chirurgicale pour un choc septique (surplus de bactéries). Bien que la technologie étudiée ne soit pas mature dans le cadre de la PPG contact (et donc encore plus lointaine pour la rPPG), elle pourrait être bénéfique dans les deux situations proposées. Pour le suivi post-transplantation le fait de pouvoir faire la mesure sans contact pourrait limiter le besoin en capteurs fixés sur le corps ce qui améliorerait grandement la qualité de vie du patient. Dans le cas de la chirurgie, l'utilisation d'un capteur sans contact peut limiter le risque d'infections et une éventuelle gêne pour le chirurgien.

Enfin, la photopléthysmographie sans contact peut également être utilisée dans le cadre du suivi de prématurés [17]. Les nouveaux-nés prématurés présentent généralement une santé particulièrement fragile et nécessitent une surveillance continue. Dans ce contexte, la détection d'apnée est très importante dans le sens où ce genre d'événement peut provoquer rapidement des problèmes neurologiques voire la mort du nourrisson. Cette surveillance se fait soit régulièrement par une mesure manuelle, soit en continu à l'aide d'appareils reliés comme un ECG ou un capteur PPG contact. La mesure manuelle n'est pas faite fréquemment, d'une part car elle nécessite la présence d'un personnel soignant et d'autre part car il est possible que la mesure soit négligée par celui-ci, de part un manque de temps et une mauvaise compréhension de l'importance de la mesure [39]. La mesure en continu par un appareil reste la meilleure méthode pour détecter une apnée, en effet même en réglant les questions posées par la mesure manuelle il est peu réaliste de considérer que celle-ci soit faite juste au moment d'un épisode d'apnée. Il s'agit donc d'un point critique pour permettre au personnel soignant de réagir vite à ce genre de complications et donc de sauver la vie des bébés concernés. Ces méthodes présentent néanmoins des inconvénients à l'heure actuelle. Tout d'abord, les appareils de mesure nécessitent un contact avec le nourrisson. Ce contact peut engendrer une lésion de la peau puisque celle-ci est très fragile, notamment pour les prématurés. La peau peut être brûlée par une mesure en continue avec un capteur PPG (LED infrarouge), ou être irritée par les électrodes de l'ECG (collées à la peau). Un autre problème associé à ces méthodes de mesure est leur fiabilité. À l'heure actuelle, une majorité des alarmes provoquées par les appareils de mesure sont des fausses alertes [61]. Bien qu'il soit préférable de faire des fausses alertes plutôt que des négligences, celles-ci sont trop fréquentes ce qui provoque une désensibilisation du personnel médical. Dans ce contexte, un grand nombre de méthodes sont en cours de développement afin de pouvoir d'une part améliorer le confort des patients mais également de limiter le nombre de fausses alertes afin de contrer la lassitude du personnel médical. La photopléthysmographie sans contact fait partie des méthodes qui ont été développées dans ce but. Pour cela, une caméra fonctionnant dans le proche infrarouge est nécessaire afin de permettre la surveillance des nouveaux-nés de nuit en utilisant la lumière dans le proche infrarouge. La lumière infrarouge n'est pas visible et permet une mesure de la rPPG suffisamment précise. Le fait que la caméra récupère une image du patient permet dans une certaine mesure de corriger le bruit associé aux mouvements du nourrisson (qui est un problème majeur dans cette situation). De plus, la caméra peut-être placée sans problème à l'extérieur de la couveuse et ne pose donc pas de problème lié à l'hygiène. La seule limitation associée à ce dernier point est la présence de reflets sur la paroi qui peut brouiller la qualité du signal obtenu.



## 2.3 Suivi de personnes au quotidien

Au delà du domaine médical, de nombreuses applications de la rPPG sont également à l'étude afin d'apporter un plus dans le suivi de personnes dans un contexte non médical. Cela va correspondre principalement à l'aide au maintien à domicile de personnes âgées afin de pouvoir réagir plus rapidement à des problèmes de santé tout en limitant la présence d'auxiliaires, le suivi de la fatigue des conducteurs qui est très importante pour limiter les accidents routiers et la mesure des paramètres physiologiques lors d'exercices sportifs.

Les personnes âgées font généralement face à une perte d'autonomie qui limite leur possibilité de rester chez eux. Le risque de chute ou de problème cardiaque est assez fort et nécessite donc a minima d'avoir une personne présente pour pouvoir alerter les secours en cas de besoin. Généralement, les personnes âgées présentant le plus de difficultés sont placées dans des établissements spécialisés comme des EHPAD. Néanmoins, le coût de ceux-ci et leur capacité d'accueil est un problème sur lequel butent beaucoup de personnes âgées. De plus, le fait de pouvoir rester chez eux représente un plus pour leur bien être et leur autonomie. Afin de permettre leur maintien à domicile, plusieurs dispositifs ont été développés de façon à pouvoir alerter les secours en cas de souci. Il existe par exemple des systèmes de bracelets qui permettent d'alerter facilement par l'appui sur un bouton, ou lors d'une détection de chute (à l'aide d'accéléromètre par exemple) [98]. Dans ce contexte, l'utilisation d'algorithmes d'analyse d'image a son rôle à jouer. En plaçant des caméras dans la maison, il est possible d'utiliser des algorithmes de détection de chute ou des algorithmes de mesure des paramètres physiologiques afin de pouvoir détecter les problèmes et de pouvoir alerter les secours [34]. Un autre intérêt de cette méthode est la possibilité de faire une première mesure des paramètres physiologiques à l'aide de la rPPG de façon à pouvoir envoyer un bilan plus complet aux secours.

La photopléthysmographie sans contact peut également être employée pour suivre des conducteurs afin de limiter le risque d'accident. En effet, la fatigue est une des principales causes des accidents de la route. D'après le site de la sécurité routière<sup>2</sup>, la somnolence multiplie les risques d'accidents par 8. De ce fait il est très important de pouvoir déterminer les premiers signes de fatigue le plus tôt possible afin de pouvoir limiter les risques d'accidents. Pour cela, des appareils ont été développés afin de détecter la fatigue en se basant sur l'ouverture des yeux ou la bascule de la tête. La photopléthysmographie sans contact peut également apporter sa pierre à l'édifice en surveillant les paramètres physiologiques du conducteur afin de pouvoir détecter des problèmes, notamment la baisse de vigilance. Bien qu'il s'agisse d'un point moins traité dans la littérature, le suivi de la fatigue présente de nombreuses caractéristiques qui permettent de l'implémenter simplement. Tout d'abord, il ne s'agit pas d'un dispositif médical au sens strict du terme, ce qui implique qu'il puisse être mis à disposition du public en nécessitant moins de validations et de normes que les dispositifs servant par exemple au suivi de nourrissons. La situation est plus simple à gérer car on a peu de mouvements de la part du conducteur et l'angle de la tête change moins. De ce fait le dispositif peut employer des traitements plus simples pour détecter la peau par rapport au cas notamment de l'aide au maintien à domicile. La principale difficulté concernant ce dispositif sont les changements dans les conditions d'illumination. En effet, à l'intérieur de l'habitacle d'une voiture, la luminosité est naturelle (soleil) ou artificielle (phares, éclairage intérieur) et dépend de l'environnement. Elle peut donc fortement varier au fil du temps du fait par exemple des ombres. Ces variations peuvent s'effectuer dans les hautes fréquences, là où l'éclairage naturel va habituellement présenter des variations plus lentes (changement de l'heure du jour et nuages). Les scénarios en extérieur sont assez peu présents dans les études sur la photopléthysmographie. Un exemple de travail portant sur ce type de situation a consisté à améliorer le filtrage des variations de luminosité en se basant sur un réseau de neurones [161]. Dans cette étude un ensemble de réseaux de

2. <https://www.securite-routiere.gouv.fr/dangers-de-la-route/la-fatigue-et-la-conduite>

neurones sont entraînés dans différentes conditions d'éclairage puis un réseau est sélectionné automatiquement en fonction de la situation de façon à améliorer la qualité du signal extrait. Un autre problème de cette application correspond à l'étude du signal rPPG obtenu. En effet, le but final est de pouvoir détecter une baisse de vigilance afin de prévenir le conducteur pour qu'il s'arrête. Paracchini *et al.* ont par exemple publié une étude portant sur cette application [105].

Enfin, le suivi des personnes peut également être mis en place dans des contextes sportifs. Ces dernières années ont vu émerger auprès du grand public des capteurs visant à mesurer les paramètres physiologiques d'une personne durant son activité physique. Le but de ces capteurs va être de détecter les changements d'état durant la journée (sommeil, éveil, sport), mais également de mieux conseiller son utilisateur lors d'une activité physique. De cette façon, le capteur peut transmettre des recommandations à l'utilisateur en le conseillant, en fonction de son rythme cardiaque, sur l'augmentation ou la diminution de ses efforts physiques. De plus, la mesure de ces paramètres permet d'améliorer le calcul des calories consommées. À l'heure actuelle, les dispositifs de mesure sont généralement au contact. Il peut s'agir de bracelets mesurant l'activité cardiaque à l'aide de la photopléthysmographie par exemple. Ce capteur présentant l'intérêt d'être porté en continu et donc de pouvoir avoir une plus grande connaissance de l'utilisateur. Il existe également des ECG dissimulés dans des machines de fitness. Les électrodes sont placées dans les poignées que l'utilisateur agrippe durant l'effort, ce qui va permettre de mesurer un signal cardiaque. Cette méthode est intéressante dans les salles de sport mais la mesure n'est effectuée que lors de l'effort et nécessite que la personne tienne les poignées pour pouvoir faire la mesure. Dans ce dernier cas, il peut y avoir un intérêt à pouvoir mesurer les paramètres physiologiques d'une personne sans contact. En effet certains mouvements vont être difficiles à réaliser en tenant les poignées. La principale difficulté associée à la mesure est la présence d'un mouvement périodique lié à l'effort. Dans le cas d'un appareil de course ou de vélo, le mouvement aura une fréquence d'environ 1 Hz, ce qui va entrer en conflit avec le rythme cardiaque. Des algorithmes spécifiques sont nécessaires pour pouvoir filtrer cette composante et pouvoir conserver le signal cardiaque [158].

## 2.4 Applications non médicales

Au delà des applications présentées précédemment qui s'attachent à mesurer l'état de santé des patients, il existe également des applications qui sont totalement détachées du domaine médical. En effet, la photopléthysmographie sans contact peut être employée pour renforcer des algorithmes comme ceux dédiés à l'analyse d'états émotionnels, à la biométrie ou à la détection de *deepfakes*.

L'analyse des états émotionnels d'une personne peut être effectuée à l'aide des appareils de mesures cardiaque. Les émotions d'une personne influent en effet sur le rythme cardiaque, de manière similaire à la glycémie ou la respiration. À partir de cette variation du rythme cardiaque, il est possible de remonter à une estimation de l'état émotionnel de la personne. À l'heure actuelle, cette méthode nécessite une pré-calibration du système afin d'identifier efficacement l'état émotionnel d'une personne [21]. Cela pose une limitation dans le sens où on ne peut pas filmer des inconnus et déterminer leur état émotionnel. Les méthodes actuelles nécessitent de connaître les variabilités du rythme cardiaque associées à chaque état émotionnel pour pouvoir faire une détection fiable. Cette détection du stress ou de l'état émotionnel de la personne présente des intérêts dans le domaine de la sécurité (protéger un aéroport), dans le domaine publicitaire (connaître les réactions à une campagne), dans le domaine de la recherche (dans le cadre d'expérimentations psychologiques) et dans le domaine du jeu vidéo (améliorer le vécu en adaptant le scénario).

La photopléthysmographie sans contact peut également être employée dans le cadre de la sécurité biométrique. La démocratisation de la biométrie entraîne également de nombreux

problèmes de sécurité. Les algorithmes de vérifications les plus simples sont facilement induits en erreur en utilisant de simples photographies du doigt ou du visage de la personne à authentifier. Afin d'améliorer l'efficacité de la biométrie, de nombreux algorithmes dits d'*antispoofing* sont actuellement en développement [122]. On a notamment l'utilisation du relief afin de pouvoir différencier une image d'un doigt ou d'un visage réel. La photopléthysmographie sans contact de son côté permet de valider la présence d'une vascularisation et donc d'ajouter une étape de validation supplémentaire [81]. Cette validation sera probablement trompée par la suite, en utilisant des variations de luminosité par exemple, mais permet néanmoins de compliquer la tâche du pirate et donc de renforcer la sécurité du système.

Enfin, la rPPG permet d'améliorer légèrement les algorithmes de détection de *deepfakes*. Les *deepfakes* sont des images générées par ordinateur qui vont remplacer le visage d'une personne par celui d'une autre personne. Le principe repose sur l'utilisation du *deep learning* et permet de générer des images voire des vidéos de plus en plus réalistes. Bien qu'à l'heure actuelle, les *deepfakes* soient principalement utilisés dans le cadre de parodies ou de publicités, il existe un risque de voir émerger dans le futur des *deepfakes* dont le but serait d'usurper l'identité d'une personne, de modifier le résultat d'un scrutin électoral ou de participer à une vague de cyberharcèlement. En effet il est techniquement possible de réaliser des *deepfakes* de façon à créer des scandales (en mettant par exemple une personne en scène dans un film à caractère pornographique ou en lui faisant faire de fausses déclarations). Parmi les exemples existants de *deepfakes* visant des politiques, on trouve notamment la parodie de Donald Trump annonçant la fin du SIDA dans le cadre d'une publicité<sup>3</sup> ainsi qu'une fausse déclaration faite par l'ambassadeur Français au Cameroun<sup>4</sup>. L'autre risque associé au développement des *deepfakes* est non pas la création de fausses vidéos mais la mise en doute de la véracité de vidéos existantes. Ainsi, la mort de Georges Floyd survenue aux États-Unis<sup>5</sup> a conduit à beaucoup de rumeurs complotistes à son sujet. Le décès de Georges Floyd s'est produite en 2020 suite à une interpellation policière et a conduit à des mouvements sociaux d'ampleur<sup>6</sup>. De ce fait sa mort est devenu un événement important dans la vie politique. Cela a amené de nombreuses théories visant à limiter la responsabilité des policiers voire à simplement nier l'existence de l'interpellation. Ainsi, un site internet conçu spécialement sur le sujet visait à présenter la vidéo de l'arrestation comme étant un *deepfake* de façon à discréditer les différents mouvements sociaux s'étant produits par la suite<sup>7</sup>. Bien que les *deepfakes* ne soient actuellement pas parfaits, ils sont déjà très réalistes et leur qualité augmente de jour en jour. La détection de *deepfakes* est alors un enjeu majeur notamment pour les GAFAM (les géants d'internet comme Google ou Facebook) [134]. Selon les utilisations des *deepfakes* les plateformes risquent dans le futur de se voir reprocher leur diffusion. Dans cette situation, il leur est nécessaire de pouvoir différencier les vraies vidéos des montages de façon à pouvoir supprimer le contenu ou a minima indiquer le fait qu'il s'agisse d'une parodie. Les algorithmes de classifications utilisent actuellement de l'intelligence artificielle. Bien que les algorithmes rPPG puissent difficilement rivaliser avec les algorithmes basés sur l'apprentissage machine, ceux-ci peuvent apporter leur pierre à l'édifice en proposant une validation relativement simple. En effet il est peu probable que le signal rPPG soit présent dans une vidéo truquée, en considérant que les algorithmes de *deep learning* s'entraînent sur une base d'images qui n'ont pas de cohésion dans le temps. De plus, le signal rPPG est extrêmement faible dans les vidéos et n'est donc pas beaucoup pris en compte dans le calcul des fonctions d'erreur. De ce fait, les *deepfakes* ne présenteront pas de signal rPPG alors que les vidéos réelles en présenteront un. Cette détection reste limitée puisqu'il est possible

3. <https://www.brut.media/us/news/why-an-aids-charity-made-a-deepfake-of-trump-a489852b-889e-4ce1-985d-3618044774f6>

4. <https://observers.france24.com/fr/20200629-attention-video-ambassadeur-francais-cameroun-guillhou-biya-deepfake>

5. [https://www.lemonde.fr/international/article/2020/05/30/george-floyd-le-doux-geant-tue-par-la-police-americaine\\_6041287\\_3210.html](https://www.lemonde.fr/international/article/2020/05/30/george-floyd-le-doux-geant-tue-par-la-police-americaine_6041287_3210.html)

6. <https://www.lefigaro.fr/flash-actu/decès-de-george-floyd-les-manifestations-se-multiplient-20200530>

7. <http://investigatedeepfakefloyd.com/>

d'ajouter un signal rPPG à un *deepfake* en rajoutant une étape de post-traitement. De plus, les vidéos en mouvement ou compressées avec pertes ne présenteront pas de signal rPPG sans qu'on puisse en conclure qu'il s'agisse d'un *deepfake*.

Cette section a présenté un panorama des applications possibles de la rPPG. Certaines de ces applications ont des chances de pouvoir bénéficier à court terme de la photopléthysmographie sans contact, tandis que d'autres ne sont à l'heure actuelle que des hypothèses basées sur les bons résultats déjà obtenus avec la photopléthysmographie au contact. Dans tous les cas, ces différentes applications montrent le grand intérêt qu'il y a à améliorer la qualité de la mesure sans contact afin de pouvoir utiliser la rPPG dans le futur et bénéficier de ses différents avantages par rapport aux autres méthodes. Dans ce cadre, la thèse présentée a pour but d'apporter sa contribution en proposant de nouveaux algorithmes dont le but serait d'obtenir des signaux de meilleure qualité. En particulier, la thèse porte sur l'extraction de la respiration à l'aide de la photopléthysmographie sans contact. Le rythme respiratoire a été choisi afin de pouvoir proposer de nouveaux algorithmes sur un paramètre physiologique qui fait l'objet d'assez peu de travaux dédiés. En effet, l'essentiel des travaux portant sur le rythme respiratoire se basent sur l'extraction du rythme cardiaque, le rythme respiratoire étant extrait généralement avec les mêmes méthodes. De plus, le rythme respiratoire est un paramètre physiologique suffisamment accessible pour pouvoir faire avancer l'état des connaissances. Les paramètres plus complexes comme la saturation en oxygène ou la tension artérielle sont plus complexes à mesurer et à contrôler, ce qui rend plus difficile l'étude de ces paramètres sans mettre en place des expériences poussées.

### 3 Objectifs et contributions de la thèse et plan du manuscrit

L'objectif principal de la thèse est d'utiliser la photopléthysmographie sans contact de manière à extraire le rythme respiratoire à partir d'une vidéo d'une personne. Pour ce faire, des algorithmes dédiés à l'amélioration de la mesure ont été développés. La photopléthysmographie sans contact consiste généralement en un ensemble d'étapes de traitements qui seront détaillées par la suite. En résumé, une vidéo est analysée pour extraire des pixels de peau (par exemple avec une détection de visage). La valeur moyenne de ces pixels est calculée en chaque instant pour obtenir des signaux couleur qui seront ensuite combinés afin de ne garder qu'un seul signal. Le signal obtenu est finalement analysé afin d'extraire le ou les paramètres physiologiques souhaités.

À la suite d'une première analyse de la littérature, la thèse s'est focalisée sur la partie du traitement consistant à combiner les signaux couleur RGB en un seul signal de qualité optimale. Cette étape consiste à transformer un ensemble de signaux généralement bruités et de les combiner de façon à pouvoir obtenir un seul signal de meilleure qualité. Généralement, les méthodes de combinaison pour le rythme respiratoire consistent à reprendre une combinaison dédiée au rythme cardiaque puis à l'appliquer sur des signaux couleur qui seront filtrés différemment [151]. Cela suppose que le signal respiratoire et le signal cardiaque sont tous les deux présents sur une même combinaison qui sera obtenue avec une des méthodes de la littérature. L'hypothèse de travail proposée est que la qualité du signal respiratoire en sortie peut être améliorée en estimant le signal respiratoire directement à l'aide d'une combinaison supposée distincte de celle servant à obtenir le signal cardiaque.

Dans un premier temps, les vecteurs de combinaison obtenus avec les méthodes existantes sont présentées graphiquement afin de déterminer des schémas ou des tendances dans la façon dont ils se répartissent sur l'ensemble des solutions possibles. Cela permet d'étudier la possibilité d'utiliser une nouvelle méthode de combinaison pour améliorer la qualité de la mesure. Cet algorithme de combinaison représentant la fondation de la thèse est détaillée et

ses performances sont estimées dans le chapitre 3. Cette première méthode appelée Brute-force correspond plus à un travail préliminaire étant donné qu'elle nécessite des quantités de calculs beaucoup trop importantes pour être utilisée dans des applications concrètes. Néanmoins, elle permet de démontrer des améliorations possibles du côté de la combinaison des signaux RGB. De ce fait, la grande majorité des algorithmes réalisés durant la thèse vont correspondre à des améliorations de ce premier algorithme avec pour objectif principal de rendre la méthode plus rapide sans perdre en fiabilité.

Dans un deuxième temps, le chapitre 4 détaille les différents algorithmes issus de ce travail préliminaire. Ceux-ci se basent sur un algorithme nommé GEVD (*Generalized Eigen Value Decomposition*) qui permet d'obtenir une combinaison similaire à celle calculée via la méthode Brute-force mais en réduisant fortement le temps de calcul. Ces méthodes vont chacune améliorer la précédente de façon à devenir petit à petit suffisamment fiable et rapide pour une utilisation en temps réel.

Dans un troisième temps, les différents algorithmes de l'état de l'art avec lesquels nous nous comparons, ainsi que les méthodes utilisées pour la comparaison (métriques, graphiques, données d'entrée) sont présentés dans le chapitre 5. La base de vidéos utilisée pour la comparaison, les métriques de qualité, les graphiques et les méthodes de l'état de l'art sont précisés dans ce chapitre. Ces différents éléments seront utilisés par la suite pour pouvoir comparer l'efficacité de chaque méthode.

En plus de ces travaux, une autre étude décrite dans le chapitre 6 a été menée et a consisté à réutiliser une chaîne de traitements « complète » qui extrait le rythme respiratoire à l'aide de la photopléthysmographie contact. Cette chaîne de traitement décompose la détection en plusieurs étapes, à savoir l'extraction du signal respiratoire, l'estimation des rythmes respiratoires et éventuellement la fusion de plusieurs résultats pour affiner la qualité de la mesure. Le but de ce travail est tout d'abord d'insérer les différents algorithmes développés dans une chaîne de traitement plus complète de façon à pouvoir valider de manière plus fiable les algorithmes, notamment en tirant parti des trois modulations. Un autre intérêt était d'évaluer si les algorithmes développés pour la PPG contact pouvaient être réutilisés dans une configuration sans contact, et si les grandes tendances restaient les mêmes en terme de types d'algorithmes à privilégier ou à éviter. Enfin, cette chaîne de traitements permettait de détecter les rythmes respiratoires sur deux types de signaux. En considérant que les algorithmes de combinaison pouvaient obtenir soit un signal cardiaque, soit un signal respiratoire, la chaîne de traitements pouvait être utilisée avec l'un ou l'autre de ces types de signaux. En faisant cela, il est possible de déterminer s'il est préférable d'extraire directement le signal respiratoire à l'aide de la combinaison des signaux couleur ou s'il vaut mieux extraire le signal respiratoire depuis un signal cardiaque.

## Chapitre 2

# Mesure de paramètres physiologiques : état de l'art et analyse

Dans ce chapitre, un état de l'art est présenté afin d'avoir une vision plus complète de la situation. Les dispositifs actuellement utilisés ou utilisables pour la mesure des paramètres physiologiques sont présentés et groupés en deux types : les méthodes nécessitant un contact avec la peau et les méthodes sans contact. Puisque la photopléthysmographie sans contact est étudiée au cours de la thèse, une partie de l'état de l'art s'attache à présenter les différentes méthodes déjà existantes pour mesurer des paramètres physiologiques à l'aide de la photopléthysmographie sans contact. Ces algorithmes sont replacés sur une chaîne de traitement habituellement utilisée dans le cadre de la photopléthysmographie sans contact.

De nombreuses méthodes de mesure des paramètres physiologiques ont été développées dans l'histoire de la médecine. Le but de celles-ci est de permettre la récupération d'informations utiles au diagnostic et à la prise en charge. Ces méthodes de mesure se basent sur un phénomène lié au paramètre physiologique comme la déformation de la cage thoracique lors de la respiration. Ce phénomène est mesuré par un capteur pour produire un signal qui sera par la suite éventuellement traité à l'aide d'algorithmes afin de mettre en évidence l'information utile (rythme cardiaque, respiratoire, saturation en oxygène...).

Afin de pouvoir positionner au mieux l'objectif de la thèse et l'impact des travaux réalisés, un état de l'art présentant les connaissances préexistantes dans le domaine est réalisé dans ce chapitre. Dans un premier temps, nous présentons un aperçu des méthodes et capteurs les plus utilisés pour mesurer les paramètres physiologiques liés au système cardiorespiratoire. Le but de cette section est de placer la photopléthysmographie sans contact dans le contexte plus général des méthodes de mesure afin de mieux en comprendre les avantages et les inconvénients. Les paramètres physiologiques sont également précisés dans cette partie pour en comprendre l'intérêt d'une part et pour connaître les phénomènes qui permettent de les mesurer. Le phénomène mesuré va avoir un impact sur la qualité de la mesure et sa simplicité (nécessité d'utiliser des capteurs coûteux pour mesurer certains phénomènes). Nous nous focaliserons dans un deuxième temps sur les méthodes utilisées par la photopléthysmographie sans contact qui sont au cœur de cette thèse. La thèse portant sur le développement d'algorithmes pour la mesure de la respiration, une attention toute particulière sera portée sur les méthodes existantes permettant de faire cette mesure.

Détailler l'intégralité des méthodes de mesure n'a pas de sens au vu de leur grand nombre. Afin de garder un panorama compréhensible et représentatif de ces méthodes, le choix des paramètres physiologiques s'est limité aux paramètres cardiorespiratoires, à savoir le rythme cardiaque, le rythme respiratoire, la saturation en oxygène et la tension artérielle. Certaines informations supplémentaires associées à ces derniers sont également évoquées comme la détection d'apnées (utile dans l'étude du sommeil [97] et la surveillance des nouveaux nés [76]) ou de pathologies cardiaques (troubles du rythme notamment).

## 1 Les appareils et les méthodes de mesure de paramètres physiologiques

Il existe de nombreux paramètres physiologiques dont le suivi permet une meilleure prise en charge. Ces paramètres apportent une information importante concernant l'état de santé de la personne, et leur mesure doit donc être aussi précise et aussi peu coûteuse que possible. Dans ce but, plusieurs méthodes de mesure ont été développées au fil du temps afin de répondre à ce besoin. De manière générale, les méthodes les plus utilisées actuellement sont les méthodes basées sur un capteur au contact de la peau de la personne. Ce type de dispositif permet une mesure plus précise car le contact limite l'introduction de bruits associés à l'environnement et permet de mieux contrôler les paramètres. Néanmoins, ces méthodes présentent certains inconvénients. Notamment, ils sont plus inconfortables à porter sur du long terme et posent des problèmes d'hygiène. Afin de réduire les inconvénients des méthodes de mesure au contact, d'autres méthodes fonctionnant à distance ont été développées au cours du temps. Bien qu'elles ne soient pas à l'heure actuelle suffisamment matures pour être utilisées dans le milieu médical, elles présentent de nombreux avantages qui pourraient les rendre incontournables dans le futur.

Cette section présente une sélection de ces méthodes en séparant les méthodes contact et les méthodes sans contact. Les algorithmes utilisés par la suite pour obtenir ces mesures à l'aide de la rPPG sont présentés dans le chapitre 2.

## 1.1 Les méthodes utilisant un contact

Dans cette partie, nous détaillons les différentes méthodes utilisées pour la mesure des paramètres physiologiques utilisant un contact. Ces méthodes sont très utilisées actuellement afin de mesurer et suivre des paramètres physiologiques dans le temps. Le contact avec la peau permet d'obtenir un signal de meilleure qualité mais pose des problèmes d'hygiène ou de confort dans le cadre d'une surveillance long terme.

### 1.1.1 Le capnographe, le pneumotachographe, le sphygmomanomètre

Le capnographe, le pneumotachographe et le spiromètre sont des appareils de mesure basés sur l'étude de l'air inspiré ou expiré. Ils sont constitués d'un masque dans lequel le patient respire. Le capnographe mesure la concentration de  $\text{CO}_2$  dans l'air expiré. Le spiromètre mesure soit le volume, soit le débit d'air inspiré et expiré. Un exemple de spiromètre portable est présenté sur la figure 2.1.



FIGURE 2.1 – Exemple de spiromètre portatif (Crédit Cosmed, CC-BY-SA).

Le pneumotachographe mesure le débit instantané d'air. Il a remplacé le spiromètre car il fonctionne en circuit ouvert et n'a donc pas besoin d'être compensé en oxygène et en  $\text{CO}_2$ . Ces méthodes permettent de mesurer le rythme respiratoire ainsi que certaines mesures de la capacité pulmonaire. Le principal problème de ces méthodes est le besoin d'un appareil fixé sur la bouche, ce qui limite la possibilité de mesurer en continu la qualité de la respiration.

### 1.1.2 La ceinture respiratoire

La ceinture respiratoire est soit une ceinture gonflable, soit une ceinture élastique. Celle-ci mesure le rythme respiratoire en se basant sur les déformations du ventre et du thorax lors de la respiration. Dans le cas de la ceinture gonflable, un capteur de pression mesure les variations. Dans le cas de la ceinture élastique, un capteur présent sur la ceinture mesure la variation de la résistance électrique liée à sa déformation. Ces appareils permettent principalement de mesurer le rythme respiratoire mais ont pour inconvénient principal de gêner la respiration de la personne.

### 1.1.3 Les accéléromètres

Les battements cardiaques produisent une onde de pression qui se propage dans le corps. Cela produit des déformations et des mouvements qui peuvent être captés en utilisant un accéléromètre [14]. Cette méthode de mesure s'appelle la ballistocardiographie (BCG). Les



mouvements mesurés sont provoqués par les déplacements de l'onde de pression dans les grosses artères comme l'aorte.

En plus des battements cardiaques, les accéléromètres peuvent être employés pour mesurer la respiration. Les mouvements de la cage thoracique liés à la respiration sont captés par l'accéléromètre et permettent d'estimer un signal respiratoire.

Cette méthode peut par exemple être mise en place en utilisant une balance qui mesure les petites variations de poids (et donc les mouvements) de la personne [52]. On peut également utiliser les accéléromètres dans des tissus intelligents [127] ou des montres connectées [136] afin de suivre au quotidien la personne. Les accéléromètres permettent de détecter des pathologies cardiaques [160], mais ils peuvent difficilement être mis en place pour de la surveillance en continu car ils sont très sensibles aux mouvements.

#### 1.1.4 Le tensiomètre

La principale méthode de mesure de la tension est par le biais d'un tensiomètre. Bien qu'il ne s'agisse pas de la méthode standard, son utilisation est non-invasive, peu coûteuse et suffisamment simple pour être faite par le patient avec un modèle automatique et quelques explications du praticien.

Le tensiomètre manuel est constitué d'un brassard gonflable et d'une poire avec un manomètre. La mesure ajoute également un stéthoscope pour écouter les bruits cardiaques dans l'artère du bras. Sa version automatique comprend un brassard gonflable et un boîtier électronique qui gère le gonflage, le dégonflage, la mesure de la pression et l'écoute des bruits. Le principe du tensiomètre repose sur les bruits de Korotkoff [13] provoqués par la pression du brassard. En règle générale, le flux sanguin ne produit pas de sons lors de sa circulation au niveau des artères car il s'écoule de façon fluide. L'ajout d'une compression à l'aide d'un brassard gonflable provoque des turbulences dans le flux sanguin et donc un bruit audible à l'aide du stéthoscope. Ces turbulences se produisent lorsque la pression du brassard excède la pression diastolique (au repos) du système circulatoire. Au delà d'une certaine pression, la circulation est totalement coupée ce qui stoppe les bruits de Korotkoff. Cette deuxième valeur est la pression systolique et correspond à la pression dans le système circulatoire durant le battement cardiaque.

La tension artérielle est un paramètre physiologique dont la connaissance présente un intérêt pour les diagnostics. Elle est utilisée notamment pour qualifier l'hypertension lorsqu'elle est trop élevée. Une diminution forte de la pression artérielle par rapport à ses valeurs habituelles peut se retrouver dans les états de chocs (choc anaphylactique, hémorragie...).

#### 1.1.5 L'électrocardiographie

L'électrocardiographie est une technique mesurant le champ électrique lié à l'activité cardiaque. Ce champ électrique est mesuré sur un ensemble de directions possibles afin d'avoir une vue d'ensemble de la structure anatomique et du fonctionnement du cœur. Pour cela, on utilise un électrocardiographe. Celui-ci est constitué d'un appareil d'acquisition et d'un ensemble d'électrodes placées sur le corps du patient. Certaines électrodes sont placées sur les poignets et la cheville et évaluent les signaux électriques prenant place dans le plan frontal. Les autres électrodes sont placées sur la poitrine et dans le dos afin d'évaluer l'activité électrique dans le plan longitudinal. Un exemple d'ECG est présenté sur la figure 2.2.

Le cœur est dirigé par le nœud sinusal qui est un ensemble de cellules particulières présent au niveau de l'oreillette droite. Celui-ci crée des impulsions électriques qui se propagent dans le cœur et provoque sa contraction. Le schéma de la propagation de l'impulsion électrique est présenté sur la figure 2.3. Dans un premier temps, l'impulsion électrique se propage dans les

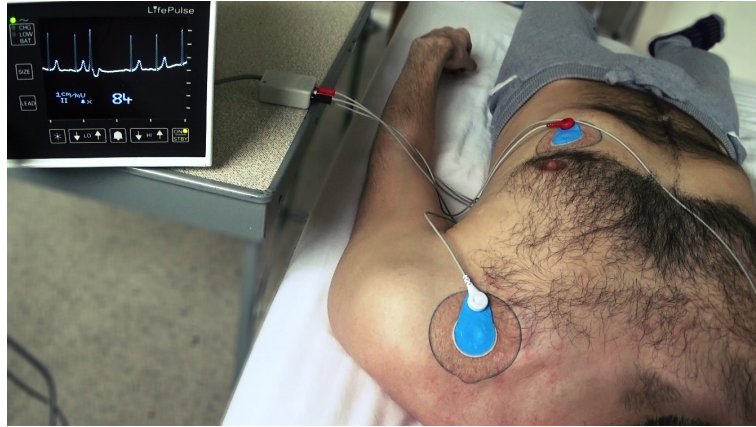


FIGURE 2.2 – Un exemple de mesure ECG (Source : speedoglyn1 on Flickr CC0).

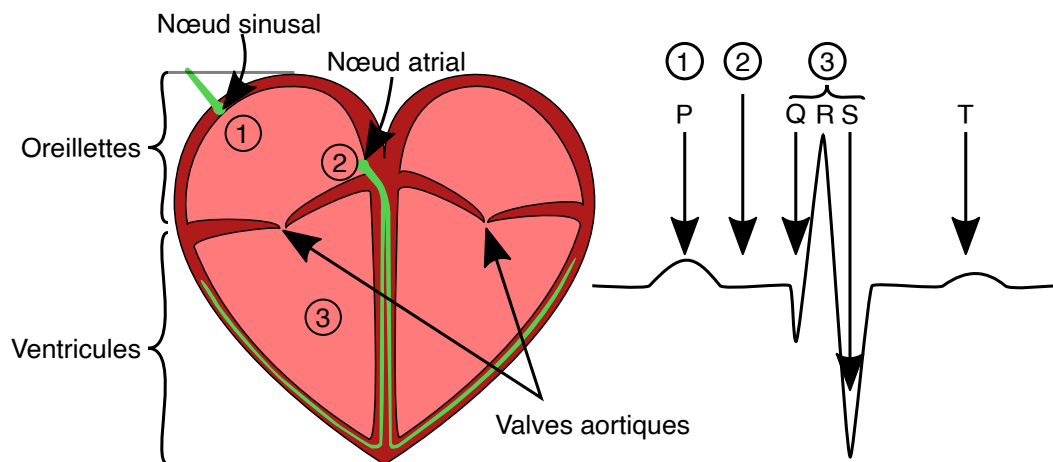


FIGURE 2.3 – Les étapes du battement cardiaque.

oreillettes et provoque leur contraction. L'impulsion est bloquée par les valves cardiaques et ne se transmet dans les ventricules que par le biais d'un goulot d'étranglement. Ce goulot, nommé nœud atrial (ou atrioventriculaire), introduit un délai du fait de sa composition chimique. Une fois ce délai passé, l'impulsion traverse rapidement le ventricule et provoque sa contraction. Cet enchaînement provoque des ondes électriques qui peuvent être mesurées par l'électrocardiographe et qui sont généralement présentes sur le signal sous la forme du complexe PQRST. Ce complexe correspond à l'enchaînement de l'onde P, du complexe QRS et de l'onde T et est montré sur la figure 2.3.

Le complexe QRS correspond à la systole, c'est-à-dire la compression des ventricules pour envoyer le sang dans les artères. L'onde P précédant le complexe QRS, correspond à la compression des oreillettes. L'onde T marque la phase rapide de la repolarisation du cœur pour se préparer à la prochaine contraction. Le motif PQRST est présent sur une personne saine dans toutes les dérivations de l'ECG. L'analyse de sa répartition dans les dérivations permet de déterminer des informations importantes comme l'orientation du cœur ou une hypertrophie d'un ventricule. De plus, l'analyse du signal en lui-même permet d'obtenir un rythme cardiaque et un rythme respiratoire. Un grand nombre de maladies cardiaques vont modifier la forme du signal et peuvent être détectées par un ECG [117, 36].

Bien que l'ECG n'ait pas vocation à mesurer la respiration de prime abord, on peut la retrouver dans le tracé de l'ECG. L'information respiratoire est présente sous la forme de trois modulations présentées sur la figure 2.4.

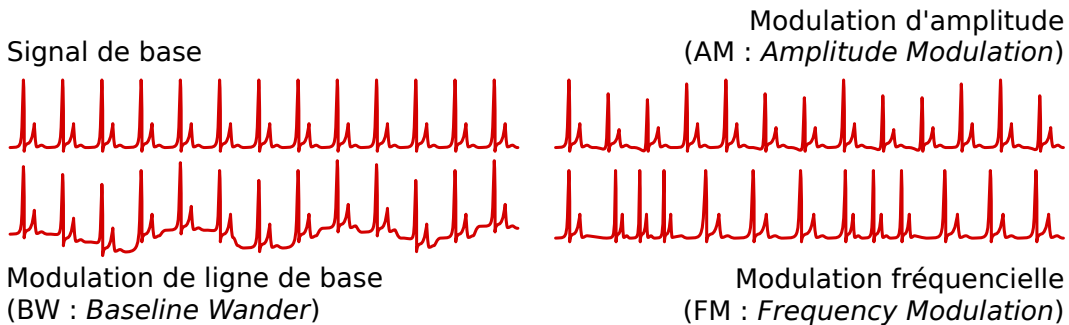


FIGURE 2.4 – Présentation des modulations du signal ECG par la respiration (Source : Charlton *et al.* [22]).

Sur cette figure, on voit une première modulation qui s'ajoute au signal (modulation de la ligne de base). Cette modulation est provoquée par la déformation de la cage thoracique. Cette déformation va écarter et rapprocher les électrodes et donc décaler le signal. La modulation de l'amplitude des pics est provoqué par le changement de l'axe électrique (lui aussi dû à la déformation de la cage thoracique). La dernière modulation (fréquentielle) se trouve au niveau du rythme cardiaque et correspond à une réponse physiologique du cœur. Ses principes sont résumés sur la figure 2.5. Lors de l'inspiration, le cœur reçoit une quantité plus importante de sang à cause des variations de pression liées aux modulations de la ligne de base et d'amplitude. Cette augmentation du volume est détecté par des capteurs de pression (barorécepteurs) présents notamment dans l'aorte et les carotides. Ces capteurs « avertissent » le nœud sinusal de l'augmentation de la pression artérielle en inhibant le système nerveux parasympathique. Le système nerveux parasympathique diminue la fréquence cardiaque, en l'inhibant cette fréquence est donc augmentée. Cette augmentation du rythme cardiaque s'appelle l'arythmie sinusale respiratoire. Elle permet de gérer le surplus de travail à faire d'une part, et d'envoyer plus de sang aux poumons d'autre part lors de l'inspiration. Ce surplus de sang aux poumons va optimiser les échanges gazeux qui ont plus d'intérêt lorsque les poumons sont remplis d'air.

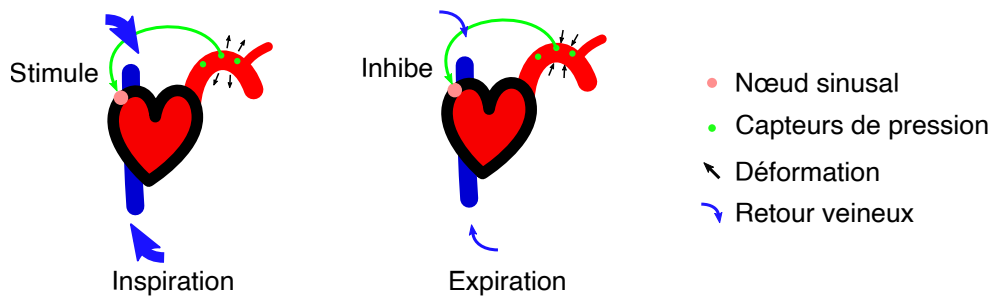


FIGURE 2.5 – Principes physiologiques de la modulation en fréquence.

Les différentes informations qui peuvent être obtenues font de l'ECG une méthode de surveillance très importante dans le cadre hospitalier. Cette méthode présente l'avantage d'être très précise et de donner un grand nombre d'informations sur le cœur à l'aide d'une mesure peu coûteuse. De plus, il est possible de fabriquer des ECG simplifiés afin de mesurer l'activité cardiaque en continu, en ne gardant qu'une dérivation et en l'implémentant dans un vêtement intelligent par exemple [168]. Néanmoins, dans le cadre d'une mesure à plusieurs dérivations l'ECG présente l'inconvénient d'être relativement complexe à positionner sur le corps du patient. Cela rend cette technique difficile à mettre en œuvre sans l'aide d'un praticien. De plus, l'ECG nécessite un contact avec la peau, souvent en plusieurs points, et possède généralement des composants métalliques qui peuvent interférer avec d'autres appareils comme l'IRM.

### 1.1.6 La photopléthysmographie au contact

La photopléthysmographie contact (PPG) consiste à mesurer la variation du volume sanguin en un point du corps [10]. Pour cela, une lumière dédiée éclaire la peau. Cette lumière traverse les différents tissus et liquides et est altérée par ceux-ci. En considérant que le flux sanguin est le seul à varier dans le temps, la variation de lumière mesurée dépend principalement de ce flux sanguin. Ainsi, en ne conservant que la partie variable du signal, on peut connaître la variation du volume de sang artériel dans les capillaires. Cette variation est corrélée au rythme cardiaque, mais également au rythme respiratoire. Certaines études ont par ailleurs démontré la possibilité d'employer le capteur PPG afin de mesurer la tension en mesurant le décalage entre le signal ECG et le signal PPG [135] ou en utilisant des réseaux de neurones [119]. D'autres articles ont étudié la détection de maladies données, comme les battements ectopiques [37], les asystolies [106] ou la sténose artérielle de la jambe [162].

Le rythme cardiaque est présent dans le signal PPG du fait de la variation de volume qui se produit à chaque battement. Cette variation de volume va modifier la quantité de lumière absorbée. La manière dont la lumière interagit avec les tissus et le sang est présentée sur la figure 2.6.

On peut reconstruire un signal cardiaque à l'aide de cette variation de la lumière transmise dans la peau (ou réfléchi depuis la peau). La pulsation PPG obtenue pour chaque battement cardiaque est généralement composée de deux parties : la systole et la diastole. La systole correspond au sang envoyé par le battement de cœur. Elle est représentée par le pic le plus haut et arrivant le plus tôt dans le signal. Durant la descente du pic systolique, un second pic plus faible est présent sur l'onde PPG. Il s'agit du pic diastolique, qui correspond au retour de l'onde dans les veines vers les ventricules. Cette onde est plus faible que la première car elle a parcouru plus de trajet depuis qu'elle a quitté le cœur (le trajet représentant le cœur vers les extrémités puis le retour vers le point de mesure). Il lui arrive dans certains cas de disparaître, notamment chez les personnes âgées. De manière générale, on ne divise

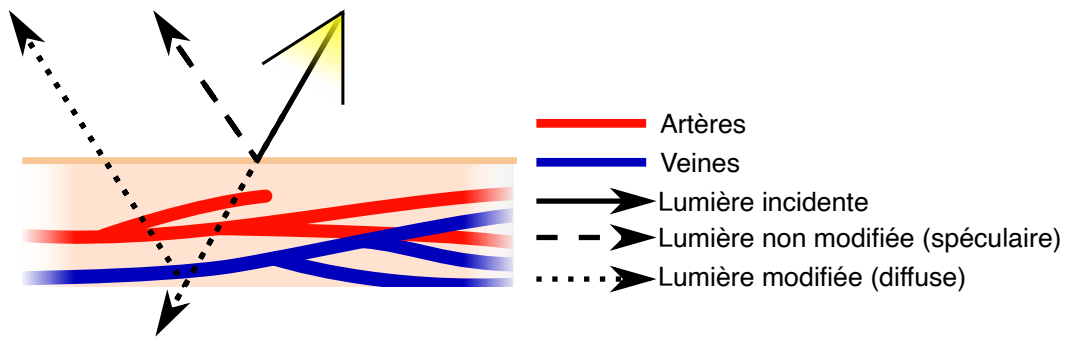


FIGURE 2.6 – Principes de l'interaction lumière-tissu.

pas la pulsation PPG en deux ondes mais en considérant le « creux » provoqué par l'onde diastolique que l'on nomme l'entaille dichrotique. Ces points sont représentés sur le schéma 2.7.

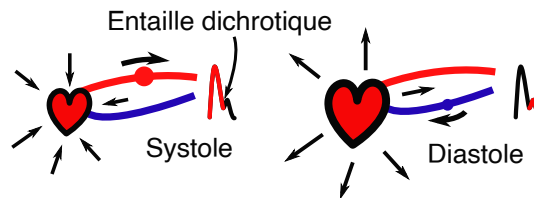


FIGURE 2.7 – Le battement PPG.

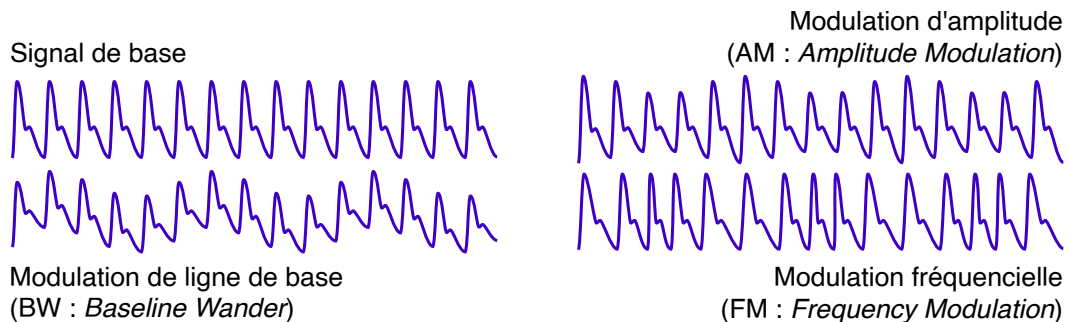


FIGURE 2.8 – Présentation des modulations du signal PPG par la respiration (Source : Charlton *et al.* [22]).

Le signal photopléthysmographique contient également des informations liées à la respiration de la personne. Ces modulations, très similaires à ceux présents sur l'ECG (figure 2.4) sont représentées sur la figure 2.8. Lors de l'inspiration, une diminution de la pression dans la cage thoracique est provoquée par son augmentation de volume. Cette diminution de la pression permet à l'air d'entrer dans les poumons : la pression intérieure diminue ce qui provoque un débit d'air entrant. Cependant cette diminution de la pression est également relative au reste du corps (tête et membres). Cela provoque de manière similaire une « aspiration » du sang depuis ces zones vers le thorax. De ce fait, l'inspiration de la personne provoque une diminution du volume sanguin dans les artères des membres et du visage, ce qui est visible sur un signal PPG. L'aspiration du volume sanguin provoque une surcharge des ventricules du cœur et donc une augmentation du volume expulsé par battement. On retrouve donc une modification de l'amplitude des pulsations PPG qui augmentent durant l'inspiration.

Enfin, comme indiqué dans le paragraphe sur l'ECG (figure 2.5), une réponse physiologique à ces changements provoque une augmentation du rythme cardiaque durant l'inspiration. Cette augmentation permet d'optimiser les échanges gazeux au niveau du poumons lorsque ceux-ci sont pleins. Cette réponse est provoquée par la surcharge des ventricules qui étire les cavités. De petits capteurs de pression sont situés sur la paroi cardiaque et informent le nœud sinusal du travail supplémentaire à effectuer. En réponse à cela, le nœud sinusal augmente son rythme d'impulsion pour gérer la charge de travail. On peut comparer ce phénomène à un coup de feu dans un restaurant.

Les trois modulations présentées sont visibles sur l'ECG et sur le signal PPG. Néanmoins, à part la modulation fréquentielle, les modulations d'amplitude et de ligne de base sont liées à des phénomènes différents. De manière générale, les études publiées ont démontré l'utilité de fusionner les estimations de chaque modulation de manière à avoir un résultat aussi fiable que possible [22]. Il faut également noter que certaines réponses physiologiques dépendent de l'état de santé du patient. Ainsi, la modulation de fréquence liée à la respiration est moins visible sur des personnes âgées [30]. Par la suite, la plupart des méthodes développées se sont basées sur la variation de la ligne de base afin d'estimer le rythme respiratoire car celle-ci ne nécessite pas d'algorithmes complexes pour la démodulation. Les principes physiologiques associés à la respiration sont résumés sur la figure 2.9.

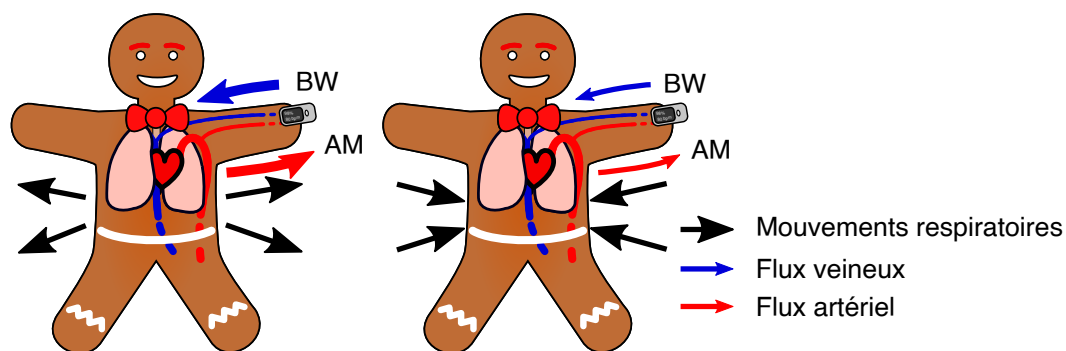


FIGURE 2.9 – Les effets de la respiration sur l'ECG et l'oxymètre.



FIGURE 2.10 – Un exemple d'oxymètre utilisé pour la mesure de la saturation en oxygène (Source : Rama on Wikimedia CeCILL).

Enfin, la principale mesure effectuée à l'aide d'un capteur PPG au contact est la saturation pulsée en oxygène ( $S_pO_2$ ) [12]. La saturation en oxygène est la concentration relative en sang oxygéné (oxyhémoglobine) par rapport au sang total. Il existe plusieurs variantes de celle-ci ( $S_aO_2$  pour la mesure faite par prise de sang,  $S_pO_2$  pour la mesure faite via un

oxymètre) qui dépendent principalement de la méthode de mesure. La  $S_pO_2$  se mesure à l'aide d'un oxymètre de pouls (voir figure 2.10). Pour obtenir la concentration relative, on utilise les différences dans le spectre d'absorption du sang oxygéné et du sang désoxygéné. Ces spectres sont présentés sur la figure 2.11. La mesure de la  $S_pO_2$  nécessite d'une part d'isoler le sang des différents tissus présents entre la source de lumière et le capteur, et d'autre part de différencier le sang oxygéné du sang désoxygéné. La comparaison du sang oxygéné et du sang désoxygéné se fait à l'aide de deux signaux PPG mesurés à deux longueurs d'onde différentes (dans le rouge et le proche infrarouge). Sur la figure 2.11, la longueur d'onde rouge est représentée avec PPG R tandis que la longueur d'onde du proche infrarouge est montrée avec PPG NIR. Ces absorptions sont situées à proximité d'un point de bascule, appelé point isobestique. Sur des longueurs d'onde inférieures à ce point de bascule, l'oxyhémoglobine est moins absorbée que la désoxyhémoglobine (d'où la couleur rouge du sang artériel). Sur les longueurs d'onde plus grandes, les courbes d'absorption s'inversent et l'oxyhémoglobine absorbe plus la lumière que la désoxyhémoglobine.

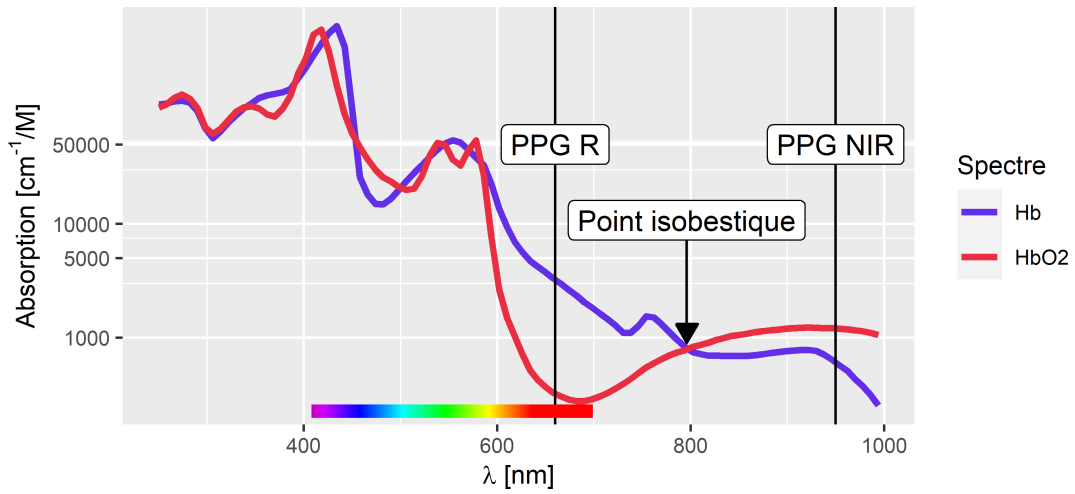


FIGURE 2.11 – Spectres d'absorption de l'oxyhémoglobine et de la désoxyhémoglobine.

Ces valeurs d'absorption affectent la lumière reçue selon la loi de Beer-Lambert [140]. Cette loi est présentée dans l'équation (2.1) :

$$A(\lambda) = \epsilon(\lambda) \times L \times C \quad (2.1)$$

avec  $\lambda$  la longueur d'onde de la lumière,  $A(\lambda)$  l'absorbance de la lumière,  $\epsilon(\lambda)$  le coefficient d'extinction molaire qui correspond à une valeur du spectre d'absorption.  $L$  est la longueur du trajet parcouru par la lumière dans le milieu et  $C$  est la concentration du milieu. Cette loi exprime la façon dont la lumière est atténuée par son parcours dans un milieu (par exemple le sang), en fonction de la longueur de traversée et de la concentration en un produit (les hémoglobines). Dans notre cas, cette loi permet de retrouver la  $S_pO_2$  en faisant un ratio entre les valeurs obtenues par le capteur. La valeur de la lumière reçue par le capteur vaut alors :

$$I(\lambda) = I_e(\lambda) \times 10^{-A(\lambda)} \quad (2.2)$$

avec  $I$  l'intensité reçue et  $I_e$  l'intensité émise par la lumière. Il est possible de décomposer l'équation (2.2) de façon à montrer la contribution du sang oxygéné, du sang désoxygéné et des autres tissus dans l'intensité reçue par l'oxymètre :

$$I(\lambda) = I_e(\lambda) \times 10^{-(A_{tissus}(\lambda) + A_{HbO_2}(\lambda) + A_{Hb}(\lambda))}. \quad (2.3)$$

Dans cette équation, les absorptions du sang oxygéné et désoxygéné dépendent de la même longueur de traversée  $L_{sang}$ . Lors du battement cardiaque, cette longueur de traversée est

augmentée de  $\delta L_{sang}$ . Comme la longueur de traversée des tissus reste la même lors du battement cardiaque, le fait de diviser l'intensité lumineuse reçue lors d'un battement par celle reçue au repos permet de supprimer la contribution des tissus :

$$\log_{10} \left( \frac{I_{r,pic}(\lambda)}{I_{r,creux}(\lambda)} \right) = \epsilon_{Hb}(\lambda) c_{Hb} \delta L_{sang} + \epsilon_{HbO_2}(\lambda) c_{HbO_2} \delta L_{sang}. \quad (2.4)$$

En considérant le ratio de l'équation (2.4) sous deux longueurs d'onde différentes, il est possible de supprimer la dépendance à  $\delta L_{sang}$  :

$$R = \frac{\log_{10} (I_{r,pic}(\lambda_1)/I_{r,creux}(\lambda_1))}{\log_{10} (I_{r,pic}(\lambda_2)/I_{r,creux}(\lambda_2))} = \frac{\epsilon_{Hb}(\lambda_1) c_{Hb} + \epsilon_{HbO_2}(\lambda_1) c_{HbO_2}}{\epsilon_{Hb}(\lambda_2) c_{Hb} + \epsilon_{HbO_2}(\lambda_2) c_{HbO_2}}. \quad (2.5)$$

R est généralement appelé dans la littérature le « ratio des ratios » du fait de l'équation (2.5). De manière générale, les longueurs d'onde utilisées sont prises dans le rouge et le proche infrarouge pour améliorer la qualité de la mesure, du fait de l'inversion des courbes d'absorption de l'oxyhémoglobine et de la désoxyhémoglobine (comme montré sur la figure 2.11). Le ratio des ratios de l'équation (2.5) permet de retrouver théoriquement la  $S_pO_2$  via l'équation (2.6) :

$$S_pO_2 = \frac{c_{HbO_2}}{c_{HbO_2} + c_{Hb}} = \frac{\epsilon_{Hb}(R) - R \epsilon_{Hb}(R)}{(\epsilon_{Hb}(R) - \epsilon_{HbO_2}(R)) - R (\epsilon_{Hb}(NIR) - \epsilon_{HbO_2}(NIR))}. \quad (2.6)$$

Cette équation permet d'obtenir une valeur correspondante à la  $S_pO_2$ . Plusieurs problèmes compliquent ce modèle. Tout d'abord le modèle de Beer-Lambert ne prend pas en compte la diffusion de la lumière dans les tissus. Afin de régler ce problème, un étalonnage est effectué pour corriger ces valeurs. Cela permet d'avoir des valeurs fiables sur une plage physiologique allant de 80% à 100% de saturation. Comme la courbe réelle n'est pas linéaire, les valeurs en dessous de 80% s'écartent de la saturation réelle. L'autre problème posé par ce calcul est le besoin de pouvoir déterminer efficacement les pics et les creux du signal rPPG de façon à pouvoir faire le ratio de l'équation (2.4). Dans le cas où le signal serait bruité à cause de mouvements du capteur, la détection de pics et de creux serait moins efficace et ne permettrait pas de donner une saturation fiable. Par exemple, il est parfois nécessaire d'évaluer une personne non consentante lorsque son état est grave et que celle-ci n'est pas en capacité d'exprimer un avis éclairé (personne mineure, sous tutelle, sous stupéfiants ou présentant un trouble psychologique comme le délire paranoïaque ou l'épisode psychotique qui peuvent être associés à une idée de persécution). Une autre limitation liée à ce ratio est que la composante variante du signal doit être peu bruitée pour avoir une mesure précise. L'utilisation de l'oxymètre est donc proscrite sur des doigts vernis ou lorsque la personne est en hypothermie (constriction des vaisseaux). Enfin, une dernière limitation de cette technique est qu'elle ne peut pas distinguer l'oxyhémoglobine de la carboxyhémoglobine. Cette dernière correspond à de l'hémoglobine chargée en monoxyde de carbone. En effet, le monoxyde de carbone se fixe à l'hémoglobine pour former du carboxyhémoglobine en prenant la place de l'oxygène. La carboxyhémoglobine n'est généralement pas présente en grande quantité dans l'organisme, sauf en cas d'intoxication au monoxyde de carbone. Elle présente un spectre similaire à l'oxyhémoglobine sur les longueurs d'onde étudiées. De ce fait, la mesure obtenue considère la carboxyhémoglobine comme de l'oxyhémoglobine, alors que la première ne joue pas son rôle dans l'oxygénation des organes.

Malgré ces inconvénients, cette méthode est très peu coûteuse, rapide à mettre en œuvre et non-invasive. Elle peut-être utilisée afin de mesurer l'état de santé d'un patient, mais également pour surveiller le bon fonctionnement d'un organe (en étudiant la consommation d'oxygène par celui-ci) [99]. Elle est également présente dans des appareils portables comme des montres connectées afin de mesurer le rythme cardiaque notamment [142]. Néanmoins, elle nécessite un contact avec la peau et perd de sa fiabilité dans un certain nombre de cas d'utilisation.



## 1.2 Les méthodes sans contact

Les différentes méthodes présentées précédemment sont suffisamment fiables pour être employées dans un cadre clinique. Néanmoins, plusieurs limitations de ces méthodes encouragent le développement de méthodes sans contact. Notamment, les méthodes contact sont inconfortables dans le cadre d'une surveillance long terme, pose des problèmes d'hygiène dans des situations à risques et peut modifier la mesure physiologique. Afin de pallier à ces problèmes, des méthodes basées sur des capteurs sans contact sont en cours de développement. Ces méthodes sont généralement moins fiables que les mesures contact puisque des bruits supplémentaires peuvent être ajoutés par l'environnement. De plus, certaines mesures sont impossibles dans le cadre d'une mesure sans contact, comme la mesure du volume d'air expiré. Bien que ces méthodes présentent des inconvénients, il existe des situations dans lesquelles elles prennent toute leur importance. Afin de mieux comprendre les avantages et les inconvénients de ces méthodes, la prochaine section présente les principes de fonctionnement qui y sont liées et les paramètres physiologiques qu'elles permettent de mesurer. Une revue de la littérature sur la mesure sans contact a également été réalisée par [95], en présentant notamment les méthodes basées sur les caméras thermiques, l'analyse d'image et la photopléthysmographie sans contact.

### 1.2.1 Les méthodes basées sur les variations de température

Certaines méthodes basées sur des caméras infrarouge mesurent le rythme respiratoire en se basant sur l'air expiré [40], [67]. En effet, l'air expiré est plus chaud que l'air ambiant, ce qui permet de « voir » les expirations de la personne. L'inconvénient principal de cette méthode est la nécessité d'employer une caméra thermique, qui peut être assez coûteuse. De plus, cette méthode nécessite de voir le visage de la personne ce qui n'est pas toujours possible, notamment pour une mesure continue. Un exemple d'imagerie thermique utilisée pour la détection de la respiration est présenté sur la figure 2.12. Celle-ci est extraite du travail de Al-Khalidi *et al.* [67].

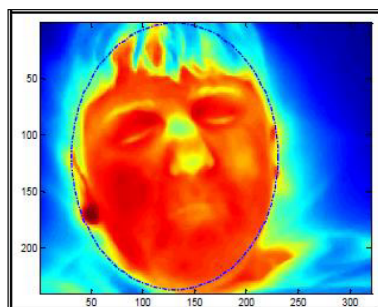


FIGURE 2.12 – Image infrarouge utilisée pour la mesure de la respiration (extraite du travail de Al-Khalidi *et al.* [67]).

### 1.2.2 Les méthodes basées sur le mouvement

Une autre possibilité est de mesurer le mouvement de la personne. Lorsque celle-ci respire, la cage thoracique se déforme. Cela induit un mouvement de la personne qui peut être mesuré. Un autre mouvement, lié à l'activité cardiaque, est également produit. Il a déjà été développé dans la section 1.1.3 du chapitre 2. Dans le cas de la mesure sans contact, on s'intéresse plus particulièrement aux mouvements de la tête. En effet, la tête est très mobile et est irriguée par de grosses artères : les artères carotidiennes (situées de part et d'autre du cou). Les

ondes de pression qui parcourent les artères carotidiennes font bouger la tête de haut en bas, ce qui peut être mesuré à l'aide d'une caméra.

Il est possible de mesurer à distance le mouvement de la personne en se basant sur une analyse de l'image [78] ou sur des méthodes radio [70]. L'analyse de l'image consiste par exemple à filmer une personne et à en extraire des points d'intérêt. Ces points sont suivis dans le temps et leurs déplacements peuvent permettre d'estimer le mouvement respiratoire de la personne.

Les méthodes radio consistent elles à envoyer des ondes radio sur la personne et à mesurer la fréquence de l'écho de l'onde retour. La fréquence varie en fonction du déplacement de la surface sur laquelle l'onde a rebondi. Cela est dû à l'effet Doppler, qui est utilisé par exemple pour les radars routiers. L'effet Doppler produit une légère variation de la fréquence lorsque l'objet sur lequel l'onde est produite (ou rebondit) est en mouvement. Cet effet peut être noté lors de déplacement de véhicules d'urgence. Leurs sirènes ont une fréquence différente une fois que le véhicule dépasse l'observateur. Les principales technologies permettant de mesurer la respiration et le rythme cardiaque à l'aide des ondes radio sont les lasers Doppler, les radars Doppler à micro-ondes et les radars à ultra large bande. Un exemple de dispositif de mesure basé sur les ondes radio est présenté dans la figure 2.13. Ce dispositif permet d'estimer les mouvements respiratoires de la personne et donc son rythme respiratoire.

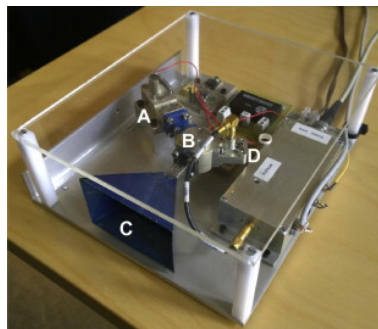


FIGURE 2.13 – Exemple de montage permettant la mesure de la respiration à partir des ondes radio (extrait de Kuutti et al. [70]).

La méthode radio est plus robuste que l'analyse de l'image et sans contact contrairement à l'accéléromètre. Néanmoins elle nécessite un matériel coûteux. De plus, les méthodes radio et basées sur l'image sont extrêmement sensibles aux mouvements de la personne. Cela rend ces appareils inefficaces notamment lorsque la personne étudiée fait du sport.

### 1.2.3 Les méthodes basées sur les variations de couleur : la photopléthysmographie sans contact

La photopléthysmographie sans contact ou rPPG est une extension de la photopléthysmographie contact. Elle reprend les mêmes principes mais emploie une caméra et une source de lumière distante. La caméra peut être simple (tel qu'une webcam [139] ou le capteur vidéo d'un smartphone [115]) afin de rendre la technologie plus accessible, ou une caméra plus complexe comme une caméra multispectrale ou une caméra avec une fréquence d'échantillonnage élevée [166]. Les caméras les plus évoluées sont plus coûteuses et ne se trouvent pas dans l'environnement quotidien. De ce fait il n'est pas possible de les utiliser pour un déploiement à très grande échelle. Néanmoins, ces dispositifs restent intéressants afin de mesurer des paramètres physiologiques particuliers ou pour obtenir une meilleure qualité du signal. En ce qui concerne la source de lumière, l'utilisation de la photopléthysmographie

peut employer de nombreuses sources spécialisées ou non. Ainsi, on retrouve l'éclairage naturel, l'éclairage artificiel habituel ou les sources de lumière contrôlées. L'éclairage naturel est le plus simple en terme de mise en place, néanmoins il n'y a aucun contrôle concernant son intensité qui peut facilement varier au fil du temps (par exemple avec le passage d'un nuage). L'éclairage artificiel habituel correspond aux lampes électriques. Leur luminosité est plus stable dans le temps et on a un contrôle plus important sur leur façon d'éclairer (la direction et la puissance notamment). La principale limitation de ces éclairages est la présence d'un bruit à 50 Hz, correspondant aux variations électriques, et qui se retrouvent dans le signal rPPG sous la forme d'une onde basse fréquence. La dernière possibilité en terme d'éclairage est d'utiliser une source de lumière dédiée, conçue pour rester stable dans le temps et avec un spectre connu. Il s'agit du dispositif idéal mais il est beaucoup moins facile à mettre en place contrairement aux deux autres qui se retrouvent dans le quotidien. Ces éclairages peuvent éventuellement être complétés en utilisant des filtres polariseurs [148] ou une lumière structurée. Ces méthodes seront présentées dans la section 2.1 du chapitre 2.

Dans certaines études, la rPPG est combinée avec d'autres méthodes sans contact afin d'améliorer la qualité de l'estimation [66]. On trouve par exemple des systèmes combinant un système rPPG avec une caméra infrarouge pour mesurer la respiration comme présenté à la section 1.2.1 du chapitre 2. La rPPG peut également être utilisée avec des algorithmes d'analyse d'images présentés dans la section 1.2.2 du chapitre 2 afin de mesurer la respiration d'après le mouvement. Cette combinaison de méthodes est plus simple à mettre en œuvre puisqu'elle ne nécessite pas de capteurs supplémentaires.

La photopléthysmographie sans contact consiste typiquement à filmer la peau d'une personne (généralement son visage), d'en extraire les pixels de peau et de faire une moyenne spatiale pour chaque image afin de récupérer des signaux (un par canal). Ces signaux sont combinés entre eux afin d'obtenir un signal rPPG similaire à un signal PPG mais généralement avec une qualité dégradée. Le signal rPPG peut enfin être traité de manière à obtenir le paramètre physiologique recherché. Dans certaines situations il est possible de mesurer ce paramètre sur différentes petites zones de façon à créer une carte du paramètre physiologique [171]. Un choix doit alors être fait sur la résolution de l'image, sachant que des zones trop petites ne seront pas exploitables du fait de la présence de bruit d'échantillonnage.

En considérant le modèle dichromatique, l'image perçue par la caméra sera constituée de différentes composantes présentées sur la figure 2.14. Ce modèle utilisé pour représenter la lumière décompose la lumière en deux composantes et sert de fondement à la photopléthysmographie sans contact. On a tout d'abord la composante spéculaire qui est la partie de la lumière qui est réfléchiée directement sans être altérée par la matière. Cette composante se retrouve notamment sur les reflets blancs. Puisque la composante spéculaire ne pénètre pas dans la matière, celle-ci ne porte que l'information de mouvement et n'a généralement pas d'intérêt dans la rPPG. De son côté, la composante diffuse traverse la matière et est altérée par celle-ci. C'est cette composante qui va notamment prendre la couleur de l'objet de part cette altération. Habituellement, comme l'objet est inerte la composante diffuse est continue dans le temps. Dans le cas de la peau, la lumière traverse également le système circulatoire et est donc modulée par celui-ci. Les différentes variations de volume sanguin pourront donc être estimées à l'aide de la composante diffuse.

La rPPG va généralement présenter plus de bruit lié au mouvement puisque le capteur n'est pas situé au contact de la peau de la personne. Cependant, le fait de récupérer une vidéo permet d'utiliser une analyse de celle-ci afin de réduire ce bruit lié au mouvement. Cela donne la possibilité de développer des algorithmes rendant la rPPG plus robuste au mouvement [156]. De plus, certaines analyses peuvent s'intéresser à une cartographie de la peau plutôt qu'à un signal 1D [170]. De ce fait, il est théoriquement possible de faire une cartographie de la  $S_pO_2$  à l'aide de la rPPG. Cela peut être utile par exemple afin d'étudier la gravité d'une lésion sur la peau [132]. La rPPG peut dans certains cas être implémentée sur

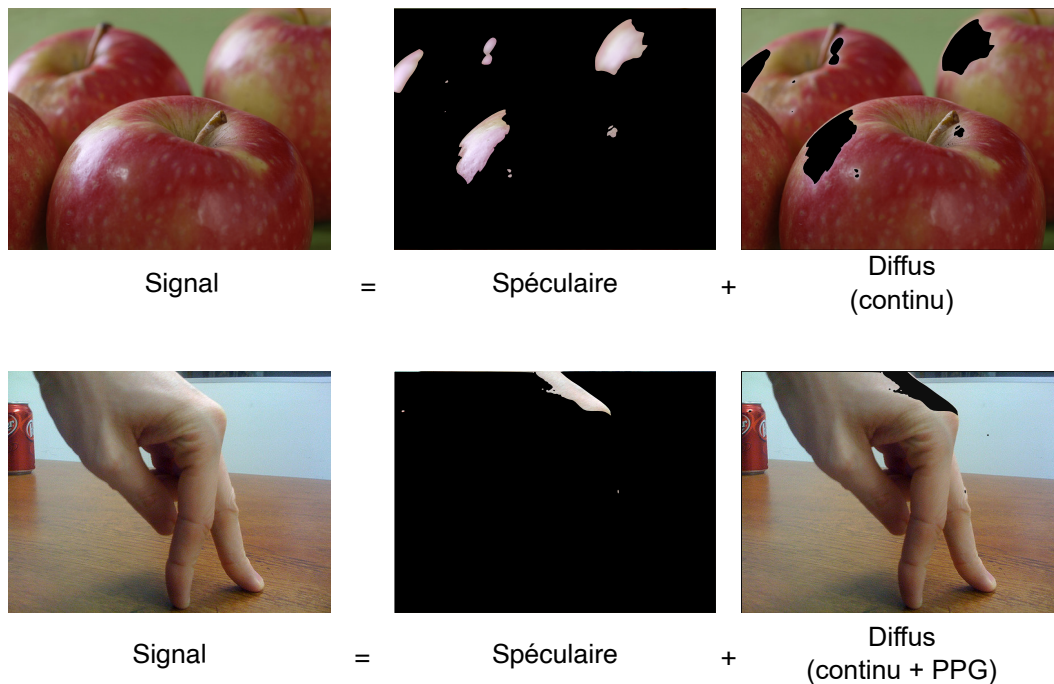


FIGURE 2.14 – Les différentes composantes de la lumière.

des appareils du quotidien comme des ordinateurs avec webcam ou des smartphones, ce qui les rend particulièrement intéressants pour la télémédecine par exemple. Enfin, l'utilisation d'une caméra au lieu d'un dispositif au contact permet également de faire plusieurs mesures simultanément. Cela permet notamment de faire des mesures de masse, par exemple dans le but de suivre une épidémie. En effectuant la mesure dans des zones particulières comme des aéroports ou des hôpitaux, il est possible soit de détecter les personnes à risque pour compléter le diagnostic par la suite, soit d'évaluer la quantité de personnes contaminées afin de prendre les décisions sanitaires qui s'imposent. La possibilité de réaliser des mesures multiples et sans contact dans des lieux publics est non seulement utile dans ces contextes, mais permet également de fluidifier les flux de personnes en évitant de créer des points de contrôle. Néanmoins, ces méthodes génèrent massivement des informations médicales personnelles ce qui pose le problème de la sécurisation de ces données sensibles et du consentement des personnes.

La rPPG peut théoriquement faire toutes les mesures pouvant être faites à l'aide de la PPG contact, étant donné que ses principes sous-jacents sont strictement les mêmes. Certains paramètres supplémentaires pourraient éventuellement être mesurés, comme par exemple la tension artérielle qui se base généralement sur une mesure en deux points du corps. En effet, la mesure de la tension à l'aide de la PPG se base généralement sur la mesure de la vitesse de propagation de la pulsation. Dans le cas des oxymètres contact, le délai est généralement obtenu en comparant les temps des pulsations sur l'ECG et sur l'oxymètre. Dans le cas de la rPPG, le fait d'utiliser des images en données de départ peut éventuellement permettre de faire cette mesure sans nécessiter d'ECG. L'estimation des paramètres physiologiques est généralement moins précise et empêche en pratique la récupération d'un certain nombre d'informations importantes. En effet, le développement de la rPPG sur des appareils moins coûteux implique d'utiliser des résolutions et des fréquences d'échantillonnage plus faibles, et des bruits d'échantillonnage plus élevés. De plus, l'environnement est moins contrôlé que dans le cas de la PPG contact et introduit plus de bruits liés aux mouvements ou à la

variation de la luminosité. En particulier, mesurer la PPG avec une lumière non contrôlée (comme la lumière naturelle) peut diminuer la qualité de la mesure.

Les méthodes de mesure ont été présentées dans la section précédente. Le but de celles-ci est de mesurer un ou plusieurs paramètres physiologiques. Pour cela, ces méthodes se basent sur un capteur mesurant un signal. Ce signal doit ensuite être traité à l'aide d'algorithmes de façon à en extraire l'information tout en restant le plus robuste aux artefacts qui peuvent avoir été ajoutés. Toutes les méthodes proposées dans cette thèse pour estimer le rythme respiratoire sans contact se basent directement sur le signal rPPG, il semble donc opportun de détailler dans la section suivante les algorithmes employés pour la mesure du signal rPPG et la mesure des rythmes cardiaques et respiratoires à partir de ce signal.

Dans la suite de l'état de l'art, un panorama des différents algorithmes utilisés pour la rPPG est présenté en regroupant les méthodes par étape de traitement. Certaines méthodes ne suivent pas le schéma habituel et sont alors placés dans la section la plus pertinente possible.

## 2 Algorithmes et dispositifs pour la mesure par la rPPG

La photopléthysmographie sans contact est une méthode particulièrement flexible dans son implémentation, ce qui lui permet d'être améliorée de différentes façons, que ce soit en modifiant la configuration de l'environnement (source de lumière, caméra...) ou les algorithmes utilisés. Les grandes étapes permettant l'obtention du signal rPPG sont présentées sur la figure 2.15. Tout d'abord, la vidéo est enregistrée à l'aide d'une caméra, les pixels de peau sont récupérés en utilisant un ou plusieurs algorithmes puis la moyenne spatiale des pixels est calculée par image. Au fil du temps, cette moyenne devient un ensemble de signaux temporels (typiquement un par canal de couleur). Ces signaux sont combinés de façon à éliminer le maximum de bruit tout en conservant le signal rPPG. Enfin, les informations physiologiques sont extraites du signal rPPG à l'aide d'une analyse du signal.

Cet ensemble de traitements peut éventuellement être adapté en fonction du paramètre à mesurer. Un exemple de cette adaptation présentant l'extraction des rythmes respiratoires depuis un signal (r)PPG a été réalisé dans le cadre de la thèse et sera présenté dans le chapitre 6.

On trouve également des méthodes plus récentes qui s'affranchissent de cette structure en utilisant notamment les méthodes basées sur l'approche *deep learning* [69, 25, 131, 83]. Ces algorithmes prennent la vidéo en entrée et sortent directement le paramètre physiologique ce qui limite les problèmes de choix d'algorithme mais nécessite d'importants jeux de données pour l'entraînement.

### 2.1 Configurations expérimentales

Bien qu'une grande partie des travaux portants sur la rPPG correspondent à des choix de traitements logiciels, certaines études ont été effectuées en utilisant un matériel ou une configuration plus spécifique pour la mesure. Par exemple, certaines études se sont attachées à étudier l'effet de la direction de la lumière sur la qualité du signal rPPG. Dans l'article de Moço *et al.* [91], le signal rPPG est dérivé en faisant varier l'angle de la source lumineuse par rapport à la peau de la personne. Le but était d'étudier la proportion de bruit associé à la ballistocardiographie en fonction des conditions d'illumination. Il découle de ce travail que la source de lumière doit être si possible tangente à la peau de façon à maximiser la proportion de signal rPPG et minimiser le signal BCG. Comme montré avec cet exemple, le choix de la configuration ou du matériel permet d'améliorer la qualité de la mesure. Cela peut permettre

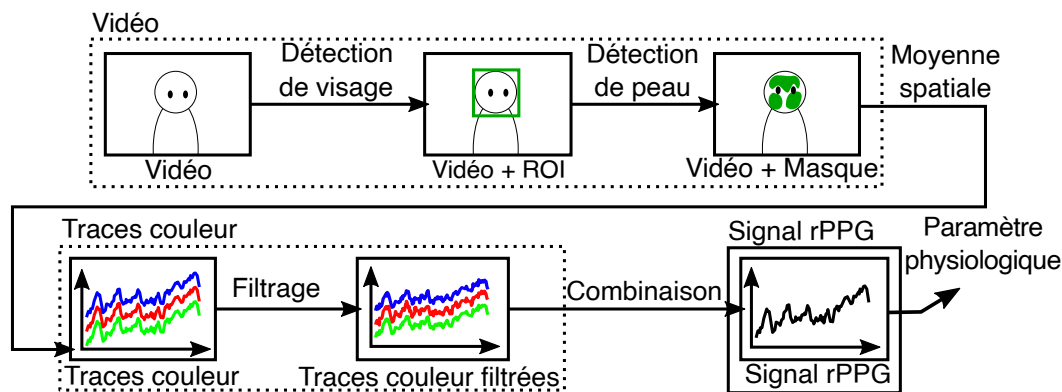


FIGURE 2.15 – Étapes pour l'extraction d'un signal rPPG.

également d'extraire des paramètres physiologiques impossibles à obtenir en utilisant une simple caméra RGB et la lumière ambiante comme source lumineuse. L'inconvénient majeur lié à ces choix est la nécessité d'utiliser des conditions d'illumination ou des dispositifs matériels (caméra) spécifiques. Cela réduit la possibilité de l'implémenter à grande échelle comme on peut l'espérer avec des algorithmes fonctionnant via une caméra RGB. Dans la suite de cette section nous décrirons brièvement quelques choix de configurations pouvant être mises en place afin d'améliorer la qualité de la mesure.

### 2.1.1 Filtre polarisateur et lumière structurée

Une des limitations de la photopléthysmographie sans contact est la faible amplitude de la composante pulsatile. Le signal pulsatile est présent dans la composante diffuse (traversant les tissus et portant l'information rPPG). La composante spéculaire est directement réfléchiée et n'est pas altérée par les tissus. Bien que la composante spéculaire puisse porter une partie de l'information de part les mouvements de la peau associés aux battements cardiaques ou à la respiration, elle porte également le bruit associé aux mouvements. De ce fait, un grand nombre de travaux cherchent à supprimer la composante spéculaire afin de rendre la mesure plus robuste aux mouvements (notamment périodiques). Afin de supprimer la composante, il est possible d'effectuer des traitements sur les images et les signaux extraits mais aussi d'utiliser un matériel spécifique afin de supprimer le bruit associé au mouvement le plus tôt possible dans la chaîne de traitements.

Par exemple, certains travaux se sont attachés à la suppression de la composante spéculaire en modifiant l'environnement de la mesure afin d'améliorer la qualité du signal sans utiliser de traitements supplémentaires. Pour cela, les deux méthodes les plus présentes dans l'état de l'art consistent à modifier l'éclairage et la caméra. Le premier exemple de modification utilise un filtre de polarisation afin de supprimer la composante spéculaire [128, 148]. La lumière émise est polarisée selon un angle et éclaire la peau de la personne. La composante spéculaire, qui n'est pas modifiée par la peau, va conserver son spectre et sa polarité. À l'inverse, la lumière qui traverse les tissus (composante diffuse) subit un ensemble de changements dans les tissus et perd sa polarisation. La lumière qui arrive dans la caméra est alors constituée d'une composante spéculaire polarisée et d'une composante diffuse non polarisée. Un nouveau filtre polariseur, orthogonal au premier est placé sur la caméra et supprime la composante spéculaire. Cette méthode permet de ne garder que la composante diffuse de la lumière sans avoir de traitements à faire par la suite. L'inconvénient principal est le fait de perdre environ la moitié de l'intensité de la composante diffuse. Le schéma de la configuration expérimentale est présenté sur la figure 2.16.

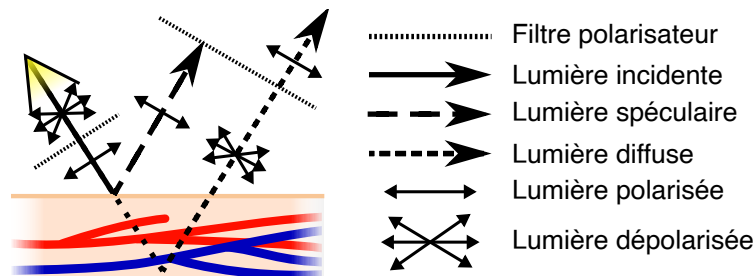


FIGURE 2.16 – Utilisation de filtres polariseurs afin de supprimer la composante spéculaire.

Une autre méthode proposée consiste à employer une lumière structurée [108]. Le concept de la lumière structurée est d'utiliser un éclairage qui va projeter un « motif » particulier. L'image qui est enregistrée va correspondre au motif d'illumination modifié par plusieurs facteurs. Par exemple, éclairer un objet avec une ligne lumineuse va permettre d'obtenir une image avec les lignes déformées en fonction de la forme de l'objet. Connaissant le motif de départ, il est possible d'appliquer des traitements afin de reconstruire l'objet en 3D. Dans le cadre de la rPPG, l'intérêt de la méthode est de supprimer la composante spéculaire en considérant que cette composante respecte strictement le motif tandis que la composante diffuse va être « floutée » du fait de son passage dans les tissus. En comparant les pixels éclairés et non éclairés par le motif, on peut retrouver l'intensité de la composante diffuse. L'intérêt de cette méthode par rapport aux filtres polariseurs est de ne pas réduire l'intensité lumineuse en entrée de la caméra. Son inconvénient principal est le besoin de traitements spécifiques au niveau de l'image, ce qui n'était pas nécessaire avec les filtres. Le principe de la lumière structurée est présenté sur la figure 2.17.

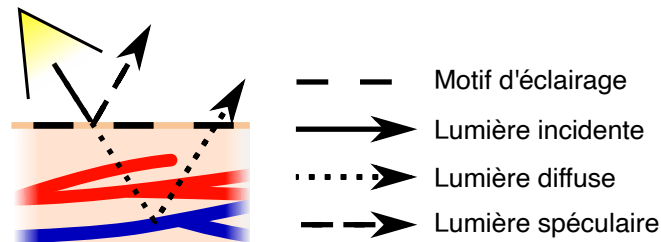


FIGURE 2.17 – Principes de l'utilisation de la lumière structurée pour supprimer la composante spéculaire. La composante spéculaire ne se retrouve que dans les rayures éclairées (en rose) tandis que la composante diffuse se retrouve partout de part sa diffusion dans les tissus.

### 2.1.2 Caméra multispectrale

Bien qu'une part importante de la photopléthysmographie sans contact soit faite sur des caméras RGB traditionnelles, il arrive dans certaines situations que l'utilisation de caméras multispectrales soit privilégiée. Les caméras RGB sont plus accessibles ce qui permet de plus facilement déployer la technologie dans des dispositifs du quotidien. D'un autre côté, l'utilisation de longueurs d'onde particulières peut améliorer la qualité de la mesure ou faire un suivi des paramètres sans utiliser d'illumination visible (avec une caméra NIR). Par exemple, la mesure de la saturation en oxygène se base sur un principe proche de la spectrométrie. En considérant l'absorption effectuée sur deux longueurs d'onde précises il est possible d'obtenir la proportion de sang oxygéné par rapport au sang désoxygéné. Bien qu'il soit possible de faire la mesure avec une caméra RGB [157], le fait de sélectionner

finement les longueurs d'onde permet d'avoir une meilleure qualité de mesure. Pour cela les longueurs d'onde doivent notamment présenter des absorptions différentes dans le spectre de l'oxyhémoglobine et de la désoxyhémoglobine.

Dans le cadre de la mesure rPPG, toutes les informations contenues dans les longueurs d'onde ne se valent pas. Un choix approprié de la longueur d'onde permet donc de potentiellement améliorer la qualité de la mesure. Ce choix va dépendre de nombreux facteurs : il faut que la longueur d'onde choisie soit peu absorbée par la mélanine et les autres composants présents dans la peau mais fortement absorbée par le sang. De manière générale, on considère les longueurs d'onde de la couleur verte comme des longueurs d'onde pertinentes pour la mesure de la photopléthysmographie. À l'inverse, les longueurs d'onde de la couleur bleue ne traversent pas suffisamment les tissus pour être modulées par le sang, elles ne portent donc que les informations associées aux déformations des tissus liées au flux sanguin [152]. Dans les travaux de McDuff *et al.* [88], l'utilisation d'une caméra multispectrale avec 5 canaux est employée de façon à améliorer significativement la qualité de la mesure de paramètres physiologiques comme le rythme cardiaque ou respiratoire.

Le suivi d'une personne sur du long terme peut nécessiter l'emploi de caméra infrarouge afin de pouvoir faire la mesure dans le noir. On retrouve par exemple ce genre d'usage dans [109] et dans [150]. L'utilisation de la photopléthysmographie sans contact dans le noir permet de faire le suivi d'un patient sur du long terme ou d'analyser la qualité du sommeil d'une personne sans utiliser de capteurs au contact de la peau par exemple.

Il est également possible d'utiliser les caméras multispectrales en combinaison avec une technologie rPPG afin de rendre plus fiables certaines opérations. On utilise notamment cette combinaison pour l'excision des tissus brûlés [145]. Le travail réalisé dans l'article consiste à créer une image de la microcirculation de la peau à l'aide de la rPPG et de rajouter une image multispectrale afin d'améliorer l'estimation. Le tissu brûlé comprend une microcirculation réduite qui peut être mesurée à l'aide de la rPPG.

## 2.2 Traitement vidéo

Comme présenté dans la figure 2.15, la première étape logicielle de la photopléthysmographie sans contact est de convertir le flux vidéo en un ensemble de signaux. Le but de cette étape est de réduire la quantité de données à traiter et de réduire le bruit associé à la quantification de la caméra. Idéalement l'étape de traitement de la vidéo est une occasion de supprimer des bruits qui ne pourront plus l'être une fois les données converties en signaux 1D.

Pour ce faire, une première approche consiste à sélectionner la région d'intérêt qui ne doit contenir que les pixels de peau. Bien que les premiers travaux concernant la rPPG se basaient sur des définitions manuelles de la région d'intérêt, de nombreux algorithmes existent aujourd'hui pour améliorer la qualité de la détection. On trouve notamment la détection de visage [71], la détection de points spécifiques (*facial landmarks*) [81] ou la détection de pixels de peau [155]. Le fait d'avoir une segmentation plus fine des pixels de peau permet d'améliorer la qualité du signal en sortie en limitant le bruit associé aux changements d'illumination ou aux mouvements de la personne. La détection de visage ou de points spécifiques est très souvent utilisée dans les travaux car le visage est une partie rarement couverte, assez peu mobile et facile à extraire de la vidéo. La région d'intérêt est généralement suivie dans le temps en utilisant des algorithmes qui estiment le mouvement de la tête plutôt que de refaire une détection. En effet, les détections sont généralement assez chronophages et ne prennent pas en compte les images et les détections précédentes.

Les algorithmes de détection de peau considèrent les valeurs RGB des pixels et estiment s'ils correspondent à de la peau ou non. Ces algorithmes peuvent être basés sur des seuillages (valeurs dans des bornes prédéfinies) [46, 130] ou sur de l'apprentissage machine avec par



exemple des algorithmes de *random forest* ou des classifieurs bayésiens [68]. Certains pré-traitements permettent d'améliorer la qualité de la détection en normalisant les couleurs. Ainsi, il n'est pas rare d'effectuer une détection de pixels de peau sur des données que l'on a normalisé (à l'aide d'un algorithme d'« *histogram matching* » par exemple), ou sur lesquelles on a modifié l'espace colorimétrique (de façon à porter un maximum d'information utile sur un seul canal).

Les méthodes de segmentation peuvent être utilisées en cascade avec par exemple une détection de visage puis une détection de pixels sur la zone découpée. L'intérêt dans ces cas est de réduire les données à traiter par le deuxième algorithme, de réduire les faux positifs et de grouper les pixels de peau dans le cas de l'étude d'un ensemble de personnes.

Il arrive dans certaines études que la région d'intérêt soit redimensionnée de façon à retirer des parties du visage qui sont peu vascularisées ou génératrices de bruit [41]. Par exemple, la barbe ne porte pas d'informations rPPG tandis que les lèvres en portent plus que la peau du fait de la finesse de la muqueuse. En revanche, la récupération du signal rPPG depuis les lèvres peut poser problème car celles-ci bougent beaucoup lors d'expressions faciales ou de discussions. Le front est considéré comme une zone particulièrement favorable pour réaliser la mesure du signal rPPG. En effet cette région bouge relativement peu et est fortement vascularisé.

Certains travaux appliquent des traitements plus spécialisés sur la région d'intérêt de façon à améliorer la qualité du signal. Par exemple, la région d'intérêt peut être redécoupée en sous-régions de façon à évaluer la qualité de chacune d'entre elles et de pouvoir filtrer celles qui n'ont pas une qualité de signal satisfaisante. Les régions d'intérêt peuvent être analysées en utilisant une matrice auto-adaptative comme présentée dans les travaux de Tulyakov *et al.* [149]. Ces sous-régions sont utilisées pour créer une matrice dont la réduction va permettre d'améliorer la qualité du signal et de connaître les sous-régions qui contiennent beaucoup de signal et peu de bruit. Un autre exemple de traitement de ces sous-régions est par l'intermédiaire d'un algorithme de *K-means* [44]. L'algorithme est utilisé pour grouper les régions d'intérêt qui ont des métriques de qualité similaires. Le meilleur groupe est conservé pour les analyses futures.

La segmentation des pixels de peau peut se faire à l'aide d'algorithmes de segmentation en superpixels. Ces algorithmes définissent un ensemble de zones regroupant des pixels proches et de couleur similaires. Cela permet de récupérer un ensemble restreint de zones qui pourront être filtrées et suivies dans le temps afin d'obtenir plusieurs signaux. L'un des principaux avantages de cette segmentation est de ne pas faire d'hypothèses sur la vidéo et donc d'être totalement indépendant de la zone à traiter et donc de pouvoir être employé sur d'autres parties du corps que le visage de la personne. Le principal algorithme employé pour la segmentation en superpixels est SLIC (*Simple Linear Iterative Clustering*) [5]. Une variante de cet algorithme, nommée IBIS (*Iterative Boundaries implicit Identification for Segmentation*) [2], a été récemment présentée dans le but de réduire le temps de calcul des superpixels et de permettre l'implémentation de l'algorithme de manière totalement parallèle. Un exemple de segmentation en superpixels est présenté sur la figure 2.18.

Enfin, le traitement de la vidéo permet également de supprimer des bruits associés aux conditions d'illumination en récupérant un signal non-rPPG à utiliser comme « antimodèle » (par exemple, [143] filtre le bruit provoqué par la variation à 50 Hz de la lumière artificielle et retrouvé dans les plus basses fréquences en prenant un signal non rPPG et en appliquant une annulation de pôles sur son modèle auto-régressif). Ce filtrage peut par exemple se faire à l'aide de suppression de pôles sur un modèle auto-régressif ou en implémentant un filtrage adaptatif prenant comme modèle le signal extrait de l'arrière plan.

Une fois la zone d'intérêt obtenue, une moyenne spatiale des pixels de peau est effectuée sur chaque image pour obtenir un signal temporel par canal (trois pour les caméras RGB). Dans le cas des superpixels, comme évoqué précédemment, les signaux sont plus nombreux,

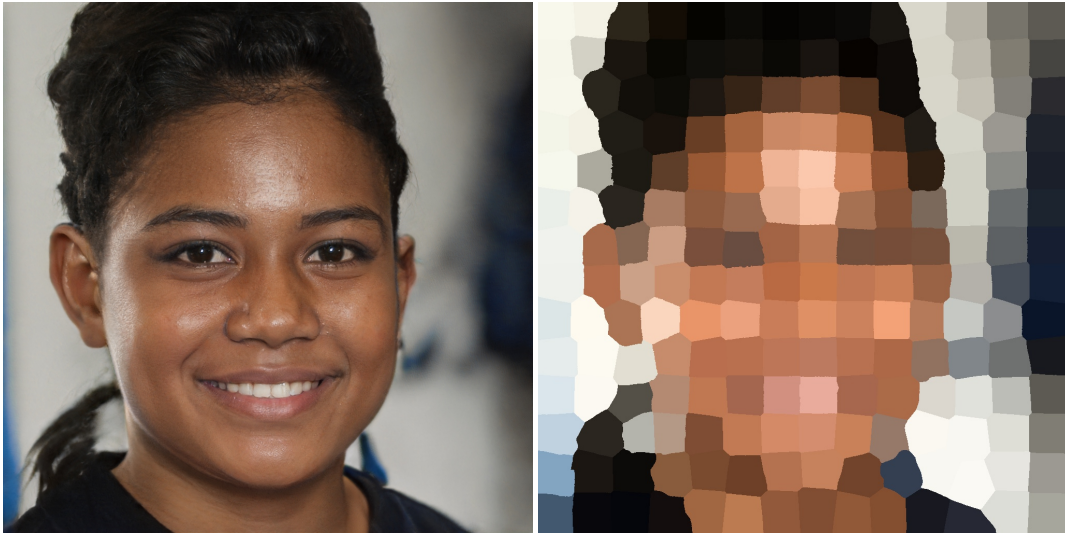


FIGURE 2.18 – Exemple de segmentation superpixels avec SLIC (image tirée de [thispersondoesnotexist.com](http://thispersondoesnotexist.com)).

mais les principes restent les mêmes. Ces signaux doivent ensuite être combinés de manière à obtenir un signal d'une seule dimension. Ce signal correspond au signal rPPG et permet d'extraire les informations physiologiques.

### 2.3 Pré-traitements

Avant d'effectuer la combinaison des signaux couleur, il est possible de filtrer les signaux obtenus. Le but est d'améliorer la qualité de la combinaison ou de tirer profit d'informations qui seront supprimées au cours de la combinaison, comme par exemple l'amplitude relative de deux canaux.

Parmi les filtrages existants, il existe notamment les filtrages habituels comme les filtrages fréquentiels ou la normalisation par la moyenne et l'écart type. Des filtrages fréquentiels plus avancés comme le filtrage en ondelettes ou l'annulation de pôles sur un modèle auto-régressif se retrouvent parfois dans la littérature afin de mieux filtrer les signaux. L'emploi d'un filtrage adaptatif est parfois retrouvé en utilisant comme modèle un signal portant du bruit, par exemple en utilisant le résultat de la détection du mouvement ou la variation de la luminosité extraite d'une région d'intérêt en arrière-plan [58, 156]. Un autre exemple de filtrage présenté par [75] consiste à estimer une courbe de la variation d'illumination de façon à retirer sa contribution du signal rPPG obtenu. Le but étant par exemple de pouvoir faire fonctionner la rPPG dans des scénarios où la variation de la luminosité est très forte, comme lorsqu'un individu regarde la télévision dans le noir.

Dans certaines combinaisons des signaux [32, 157], une méthode de normalisation temporelle est mise en place afin de supprimer les tendances basse-fréquence des signaux. La méthode considère le modèle dichromatique introduit brièvement dans la section 1.2.3 du chapitre 2 de façon à supprimer de manière analytique les bruits associés à la composante spéculaire de la lumière. Dans le cadre de l'article de De Haan *et al.* [31], une normalisation de chaque signal est effectuée par une moyenne glissante incluant au minimum une pulsation (afin de ne pas le déformer). En supprimant les variations de la ligne de base, les signaux en sortie ne contiennent plus les variations basses fréquences qui correspondent principalement aux variations des conditions d'illumination, c'est-à-dire des changements de la luminosité (dans

le cas notamment d'un éclairage naturel) et des changements dans la distance lumière-peau-caméra. Les signaux normalisés sont obtenus à l'aide de l'équation (2.7) :

$$\mathbf{C}_n(t) = \frac{1}{\mu_L\{\mathbf{C}(t)\}} \mathbf{C}(t) - 1 \quad (2.7)$$

avec  $\mathbf{C}(t)$  un des signaux couleur et  $\mathbf{C}_n(t)$  sa version normalisée.  $\mu_L\{\mathbf{C}(t)\}$  représente la moyenne temporelle calculée sur une fenêtre glissante de taille  $L$ . Dans le cas de l'article de De Haan [32],  $L$  est considéré comme supérieur à 32 échantillons afin de garantir un filtrage passe-bande de bonne qualité. Dans l'article de Wang [157],  $L$  est fixé à 32 échantillons ce qui reste cohérent d'un article à l'autre.

Un autre pré-filtrage plus intéressant a été décrit dans l'article de Van Gastel *et al.* [151]. La normalisation présentée dans l'article consiste à effectuer une première étape de normalisation des signaux couleur puis une intégration des signaux précédemment normalisés. Dans un premier temps, la normalisation des signaux produit la dérivée normalisée dans le temps. Le but de la normalisation est de supprimer des bruits associés aux mouvements de la région d'intérêt. Dans le cadre de l'article les signaux couleur étaient définis pour un ensemble de sous-régions d'intérêt qui se déplacent d'une image à l'autre selon un vecteur déplacement  $(\delta x, \delta y)$ . Le pré-traitement consiste à effectuer une première étape de normalisation temporelle sur les signaux RGB via :

$$\mathbf{C}(t \rightarrow t+1) = \frac{\mathbf{C}(t+1) - \mathbf{C}(t)}{\mathbf{C}(t+1) + \mathbf{C}(t)} \quad (2.8)$$

avec  $\mathbf{C}(t \rightarrow t+1)$  les signaux RGB normalisés à l'échantillon  $t$  et  $\mathbf{C}(t)$  les signaux RGB avant normalisation à l'échantillon  $t$ . Cette première étape de normalisation introduit une dérivation de par le numérateur. Afin de supprimer cette dérivée, une seconde opération d'intégration est réalisée avec :

$$\mathbf{C}_n(t \rightarrow t+l) = \sum_{k=0}^{l-1} \mathbf{C}(t+k \rightarrow t+k+1) \quad (2.9)$$

avec  $l$  la taille de la fenêtre d'intégration et  $\mathbf{C}_n(t \rightarrow t+l)$  les signaux RGB à l'échantillon  $t$  après le pré-traitement. Cette méthode sera utilisée par la suite dans la comparaison des méthodes de l'état de l'art et des méthodes développées.

Enfin, la méthode SSA (*Singular Spectrum Analysis*) permet d'effectuer une décomposition en sous-signaux sur le même principe que la transformée de Fourier (qui décompose le signal en sinusoïdes de fréquences croissantes). Dans [165], SSA est utilisé afin de construire une matrice en prenant comme lignes le signal avec différents délais. Cette matrice va permettre d'étudier les relations entre chaque segment décalé dans le temps afin de mettre en évidence les composantes périodiques. La matrice est analysée en utilisant une décomposition en valeurs singulières. Cette décomposition va définir une matrice de rotation qui extrait de la matrice SSA les différents sous-signaux qui la compose. Les signaux qui sont obtenus sont supposés séparer au mieux le signal cardiaque du bruit de mouvement, en considérant que le signal cardiaque est périodique et que le signal de mouvement est non périodique et défini sur une durée très courte. Dans le travail de Zhao *et al.* [169], la méthode SSA est reprise en rajoutant une deuxième itération afin d'affiner le signal en sortie. Pour cela, une première estimation du rythme cardiaque est effectuée et utilisée afin de guider la deuxième décomposition spectrale.

## 2.4 Combinaison des signaux RGB

Une fois les pré-traitements des signaux couleur effectués, on cherche à trouver une combinaison idéale dont le but est de maximiser le signal rPPG et de limiter le bruit (associé aux

mouvements ou aux changements des conditions d'illumination). Cette étape correspond à l'étape « Combinaison » de la figure 2.15. Le bruit associé au mouvement n'affecte plus les signaux de par le pré-traitement normalisation-intégration présenté dans la section 2.3 (chapitre 2). Le bruit qui persiste dans les signaux couleur correspond principalement aux variations des conditions d'illumination.

Afin de trouver la combinaison la plus pertinente dans le cadre de la rPPG, de nombreuses méthodes ont été développées au fil du temps. Celles-ci peuvent être regroupées en deux grandes catégories, à savoir les méthodes qui utilisent les propriétés statistiques des signaux pour maximiser la qualité du signal, et les méthodes qui emploient des principes physiques liés à l'interaction lumière-tissu, en considérant le bruit ou le signal porté sur un vecteur RGB particulier. Cette dernière catégorie est généralement plus robuste au mouvement, moins coûteuse en temps et nécessite une fenêtre d'analyse plus restreinte. Néanmoins, les méthodes statistiques présentent l'avantage d'être utilisables quelque soit la configuration matérielle utilisée. En effet, les méthodes basées sur l'interaction lumière-tissu se basent généralement sur des connaissances a priori que l'on obtient à l'aide d'une analyse du spectre de la lumière, de la peau ou du sang. De cette analyse ressort des vecteurs « paramètres » qui sont utilisables dans le cadre où ils ont été définis. Par exemple, la méthode CHROM (*Chrominance Based*) [31] présentée par la suite emploie un plan orthogonal à la composante spéculaire. Cette méthode est très robuste dans le cadre d'une illumination par lumière blanche et d'une caméra RGB, mais nécessite de recalculer le vecteur lorsqu'on souhaite utiliser la rPPG dans d'autres contextes comme par exemple avec des canaux dans le proche infrarouge (pour le suivi de nuit).

De manière générale, les différents algorithmes présentés considèrent la combinaison des signaux RGB sous la forme d'une équation  $\mathbf{s}(t) = \mathbf{C}(t) \cdot \boldsymbol{\omega}^T$  avec  $\mathbf{C}(t)$  les signaux RGB de taille  $[N \times 3]$ ,  $\mathbf{s}(t)$  le signal rPPG de taille  $[N \times 1]$  et  $\boldsymbol{\omega}$  un vecteur unitaire  $[1 \times 3]$  représentant la proportion de rouge, de vert et de bleu dans le signal de sortie. Les traces RGB sont construites via le modèle :

$$\mathbf{C}(t) = \mathbf{x}(t) \cdot \mathbf{v}_1 + \mathbf{b}(t) \cdot \mathbf{v}_2 \quad (2.10)$$

avec  $\mathbf{v}_1$  et  $\mathbf{v}_2$  deux vecteurs différents  $[1 \times M]$  qui donnent les répartitions respectives du signal  $\mathbf{x}(t)$   $[N \times 1]$  et du bruit  $\mathbf{b}(t)$   $[N \times 1]$  sur les différents canaux de couleur. Le but des combinaisons est alors de maximiser le ratio signal sur bruit, par exemple en déterminant une combinaison orthogonale au vecteur  $\mathbf{v}_2$ . Afin d'avoir une idée plus précise de la manière dont se répartissent les différents algorithmes de combinaison, les principales méthodes de combinaison sont associées à une image de sphère RGB. Cette sphère unitaire représente le ou les vecteurs  $\boldsymbol{\omega}$  utilisés par la combinaison en l'indiquant par une croix noire et obtenus sur des signaux RGB. Ces vecteurs sont obtenus en calculant les combinaisons idéales sur les signaux couleur découpés en segments de 30 secondes avec un décalage de 1 seconde par segment. La vidéo utilisée pour extraire les signaux analysés est une vidéo prise en interne mettant en scène une personne assise respirant lentement (environ 0,1 Hz) et éclairée à l'aide d'une lumière artificielle. La caméra employée réalise des vidéos RGB avec une résolution de  $1280 \times 720$  et une fréquence d'échantillonnage de 30 ips. La vérité terrain respiratoire est obtenue en utilisant une ceinture respiratoire. La représentation des vecteurs sur une sphère RGB permet d'obtenir une meilleure vue de la manière dont fonctionne les différents algorithmes. Néanmoins, il faut garder à l'esprit que les différents segments sont fortement corrélés ce qui peut limiter la répartition de l'ensemble des vecteurs. L'angle de vue place la composante (1 1 1) au centre de l'image. Cette composante correspond après l'étape de normalisation-intégration à la composante portant le bruit.

### 2.4.1 Méthodes de combinaison statistiques

Les méthodes de combinaison statistiques reposent sur la maximisation d'une valeur statistique du signal combiné. Ces méthodes peuvent être soit aveugles (BSS : *Blind Source Separation*) comme ICA (*Independent Component Analysis*) [112] et PCA (*Principal Component Analysis*) [77], ou guidées à l'aide de contraintes supplémentaires comme cICA (*constrained ICA*) [84].

**2.4.1.1 ICA** ICA [112] est une méthode permettant de déterminer une matrice de rotation qui combine les signaux pour maximiser leur non-gaussianité. Cette non-gaussianité permet d'exprimer l'indépendance statistique des signaux. Dans le cadre de la rPPG, on considère que les signaux couleur correspondent à un mélange de signal rPPG et de bruit, et on souhaite trouver une matrice de rotation qui sépare au maximum ces deux composantes qui sont par nature indépendantes.

L'avantage de ICA est d'employer les informations statistiques associées au signal pour choisir la meilleure combinaison possible. L'inconvénient est que la méthode nécessite une grande quantité de signal et donc plus de calculs afin d'avoir suffisamment d'informations statistiques. On considère généralement 30 secondes comme une durée correcte pour ce genre de combinaison, ce qui implique une grande latence. Un autre inconvénient d'ICA est le fait qu'on ne connaisse pas a priori le signal rPPG parmi les signaux de sortie. On a donc besoin d'une étape supplémentaire pour évaluer le meilleur candidat. Généralement cette étape consiste à sélectionner le signal avec la périodicité la plus forte, ou avec le meilleur SNR (Rapport Signal sur Bruit) dans la bande de fréquences physiologiques. Dans le cas de la figure 2.19, la composante conservée est celle avec le meilleur SNR.

Sur la sphère couleur, la combinaison ICA donne le résultat de la figure 2.19. On voit tout d'abord que le vecteur de combinaison va dépendre des données en entrée ce qui explique qu'il y ait plusieurs vecteurs différents. On remarque également que l'essentiel des vecteurs de combinaison sont aux environs du plan orthogonal au vecteur bruit normalisé  $(1 \ 1 \ 1)$ . Enfin, il semble qu'un vecteur de combinaison se retrouve sur ce vecteur de bruit en particulier. Cela est très probablement dû à un problème dans le choix de la composante à garder.

L'un des problèmes posés par l'utilisation d'ICA est celui de la sélection de la composante à conserver. Dans [92], une méthode de sélection basée sur de l'apprentissage machine permet de sélectionner la composante en fonction d'un ensemble de descripteurs. Les algorithmes utilisés sont soit la régression linéaire, soit KNN. Les descripteurs utilisés sont estimés d'après le spectre de densité de puissance du signal.

**2.4.1.2 ICA contraint** La méthode ICA présentée dans la section précédente ne prend pas en compte les spécificités du signal rPPG dans sa combinaison. On peut supposer que guider ICA en utilisant ces informations peut réduire l'espace de recherche (et donc le temps de calcul), rendre l'algorithme plus robuste et obtenir directement le signal rPPG sans avoir à sélectionner la composante selon la périodicité des signaux en sortie.

Macwan *et al.* [84] proposent l'utilisation d'ICA afin d'extraire le signal rPPG des signaux couleur. Afin de le rendre plus robuste, ils intègrent dans le calcul deux contraintes associées au signal attendu. Ces contraintes font qu'ICA va ressortir les signaux en fonction des propriétés expliquées précédemment mais en n'effectuant la recherche que dans l'ensemble des combinaisons possibles qui respectent les deux contraintes. La première contrainte est obtenue à l'aide d'une autocorrélation du signal sur lui-même. Le but de cette contrainte est de ne considérer que les combinaisons qui présentent une périodicité provoquée par les battements cardiaques. L'autre contrainte est une contrainte de teinte. Le but de cette contrainte est de réduire l'espace de recherche en prenant en compte les spécificités de la

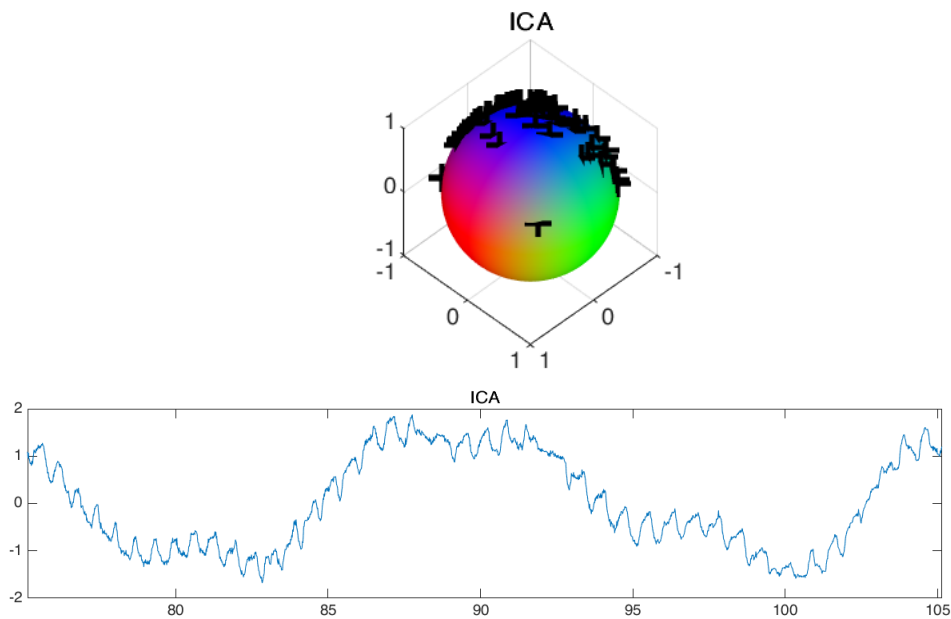


FIGURE 2.19 – Sphère RGB représentant le vecteur en sortie d’ICA et signal correspondant.

rPPG. On retrouve cette idée par exemple dans CHROM [31] et PBV [32] (*Pulse Blood Vector*) qui seront détaillés dans les sections 2.4.2.3 et 2.4.2.4 du chapitre 2.

**2.4.1.3 PCA** La méthode PCA [77] est une méthode similaire à ICA dans le sens où elle maximise une propriété statistique des signaux en sortie. Dans le cas d’ICA il s’agissait de l’indépendance (exprimée par exemple via la non-gaussianité) des signaux, dans le cas de PCA il s’agit de la variance du signal de sortie. Sur un signal, l’analyse PCA permet de faire une combinaison qui maximise l’énergie portée par le signal.

Les avantages et les inconvénients de PCA sont similaires à ceux d’ICA. Les deux méthodes sont basées sur la statistique des signaux et nécessitent donc une taille de fenêtre d’analyse longue. Un autre problème est le manque de robustesse aux bruits. Enfin, les signaux ne sont pas connectés d’une fenêtre à l’autre, ce qui nécessite une analyse supplémentaire pour sélectionner la composante souhaitée en sortie.

Sur la figure 2.20, on remarque que la combinaison de PCA donne des projections proches du vecteur  $(1 \ 1 \ 1)$ . Cela est très probablement lié à la prépondérance du mouvement dans la vidéo qui se retrouve sur ce vecteur. Dans cette situation, PCA n’est pas en capacité de reconnaître ces variations comme du bruit pour les minimiser. Afin de réussir cela, il faudrait soit utiliser un filtrage capable de supprimer au mieux cette composante, soit de rechercher la combinaison d’intérêt sur les autres composantes (qui seront orthogonales à la composante spéculaire).

**2.4.1.4 PVM** Les méthodes basées sur la physiologie consistent généralement à supprimer un bruit associé au mouvement. Cette approche a du sens car maximiser simplement l’énergie du signal rPPG ne permet pas d’avoir le signal de la meilleure qualité. En effet on considère la qualité du signal en se basant sur le rapport entre l’énergie du signal et celle du bruit. En considérant que le bruit est porté par une seule combinaison proche de la combinaison maximisant le signal, alors utiliser la combinaison maximisant le signal (comme pour PCA chapitre 2 section 2.4.1.3) va également augmenter le bruit et limiter le SNR. À

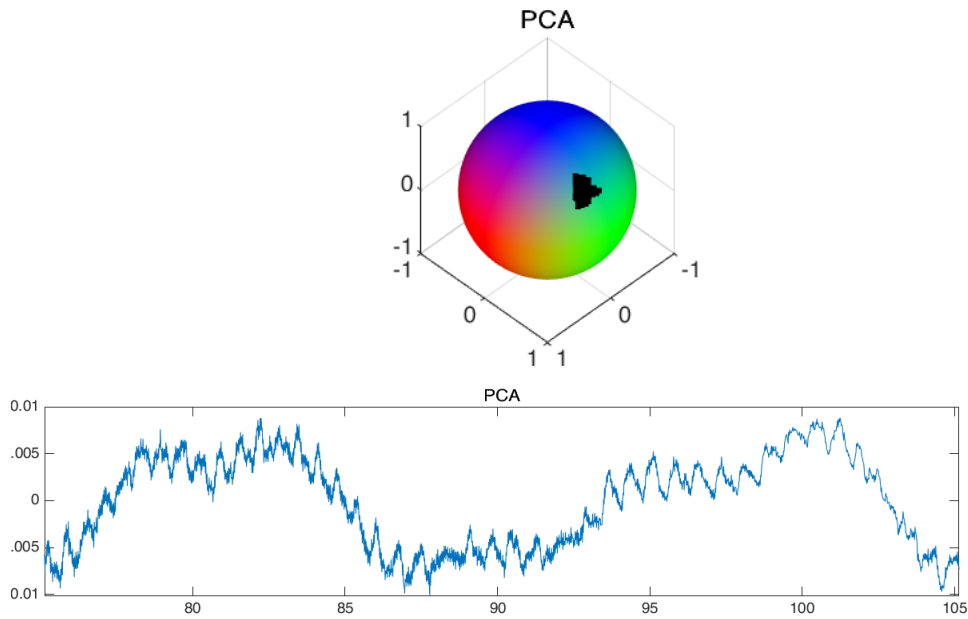


FIGURE 2.20 – Sphère RGB représentant le vecteur en sortie de PCA et signal correspondant.

l'inverse, prendre une combinaison orthogonale à la combinaison du bruit va limiter le bruit et donc maximiser le SNR. Cela suppose que l'essentiel du bruit soit porté sur un vecteur, ce qui est cohérent avec la philosophie de CHROM [31] qui considère le bruit principalement porté par la composante spéculaire sur le vecteur unité (après normalisation).

L'idée de maximiser l'énergie du signal est présentée par PCA et consiste à diagonaliser la matrice de covariance des signaux. En diagonalisant la matrice, on peut maximiser l'énergie portée par le signal combiné. PVM (*Periodic Variance Maximization*) [85] reprend une idée similaire à PCA, mais en utilisant la décomposition généralisée en valeurs propres (GEVD pour *Generalized Eigen Value Decomposition*). Cette décomposition prend en entrée deux matrices de covariance du signal et du bruit (respectivement  $\mathbf{P}_x$  et  $\mathbf{C}_x$ ) et retourne le vecteur de combinaison optimal. En fonction de la manière de créer les matrices de covariance, on peut obtenir une combinaison qui met en évidence une propriété particulière du signal. Dans le cas de PVM, cette propriété est la périodicité du signal de sortie, elle est embarquée dans la matrice  $\mathbf{P}_x$  en calculant la covariance des signaux couleur avec leur version décalée dans le temps.

Sur la sphère couleur (figure 2.21), la combinaison PVM se retrouve généralement sur la composante rouge ou à son opposée. Par ailleurs, les vecteurs de combinaison se répartissent parfaitement sur le plan orthogonal au vecteur bruit normalisé  $(1 \ 1 \ 1)$ . Cela reste cohérent avec les hypothèses faites précédemment, à savoir que la composante  $(1 \ 1 \ 1)$  porte principalement le bruit et que pour obtenir un meilleur SNR il faut trouver un plan orthogonal à la composante bruit et non orthogonal à la composante signal (donc aux environs du plan orthogonal à  $(1 \ 1 \ 1)$ ).

#### 2.4.2 Méthodes de combinaison basées sur l'interaction lumière-tissu

Les méthodes de combinaison basées sur l'interaction entre la lumière et les tissus permettent d'améliorer la qualité de l'extraction sans nécessiter de ressources de calculs élevées. Le cas de la méthode *Green* par exemple se base sur la meilleure absorption de la couleur verte

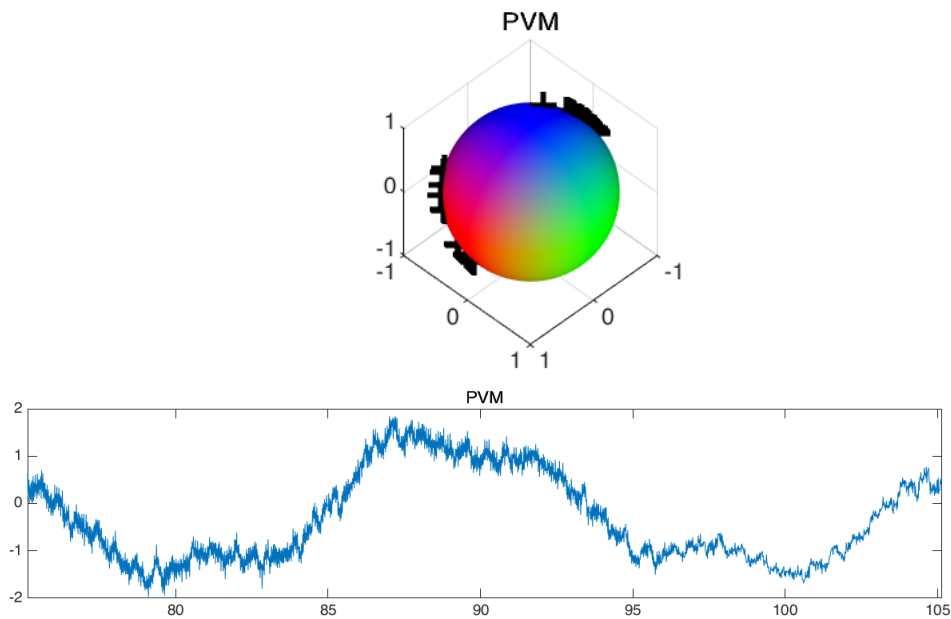


FIGURE 2.21 – Sphère RGB représentant le vecteur en sortie de PVM et signal correspondant.

par le sang. L'exemple de CHROM supprime le bruit lié au mouvement en le considérant comme fortement lié à la composante spéculaire.

**2.4.2.1 Green** La première méthode proposée est de conserver uniquement le canal vert. Cette méthode est la moins coûteuse en temps de calcul et prend en compte la physiologie de la personne. En effet, le sang absorbe beaucoup plus les longueurs d'onde de la bande verte que celles des autres couleurs [63]. Cette plus grande absorption se traduit par plus d'amplitude dans le signal rPPG. Certaines interrogations demeurent concernant l'utilisation de la bande verte [63]. En effet, la longueur d'onde du vert ne peut théoriquement pas traverser les tissus de l'épiderme et du derme. Plusieurs chercheurs supposent donc que l'information provenant du canal vert ne serait pas due à l'absorption de la lumière par le sang mais par les mouvements associés aux battements cardiaques.

La combinaison *Green* ne dépend pas du signal et prend donc toujours la même position sur la sphère couleur (figure 2.22).

**2.4.2.2 Green-Red** La deuxième méthode la plus simple est l'utilisation des canaux vert et rouge pour améliorer la qualité du signal [45]. Cette méthode est robuste et peu coûteuse en temps de calcul, néanmoins elle reste perfectible pour supprimer le bruit ou améliorer le signal rPPG. Le vecteur de combinaison de cette méthode est  $(-1 \ 1 \ 0)$ . Il est orthogonal à la composante  $(1 \ 1 \ 1)$  et supprime donc les bruits associés à cette composante, contrairement à la méthode *Green* présentée à la section 2.4.2.1 du chapitre 2. Sur la sphère couleur (figure 2.23), la combinaison reste toujours fixe car elle ne dépend pas du signal d'entrée.

**2.4.2.3 CHROM** La méthode CHROM a été présentée par De Haan [31]. Elle sert à supprimer l'effet de la composante spéculaire de la lumière en effectuant une combinaison linéaire qui est orthogonale à cette composante. Afin de maximiser le signal dans cette



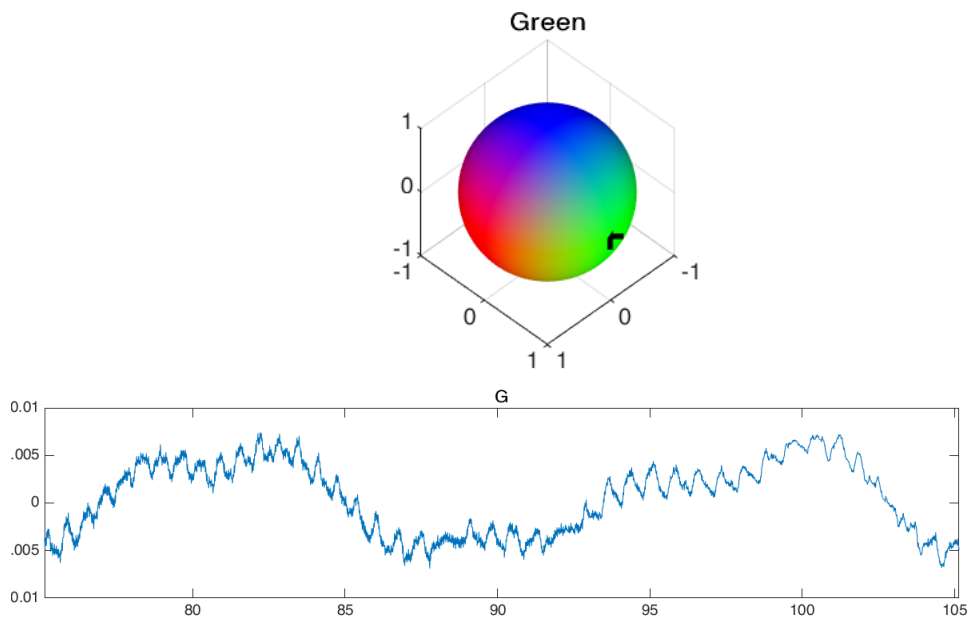


FIGURE 2.22 – Sphère RGB représentant le vecteur en sortie de *Green* et signal correspondant.

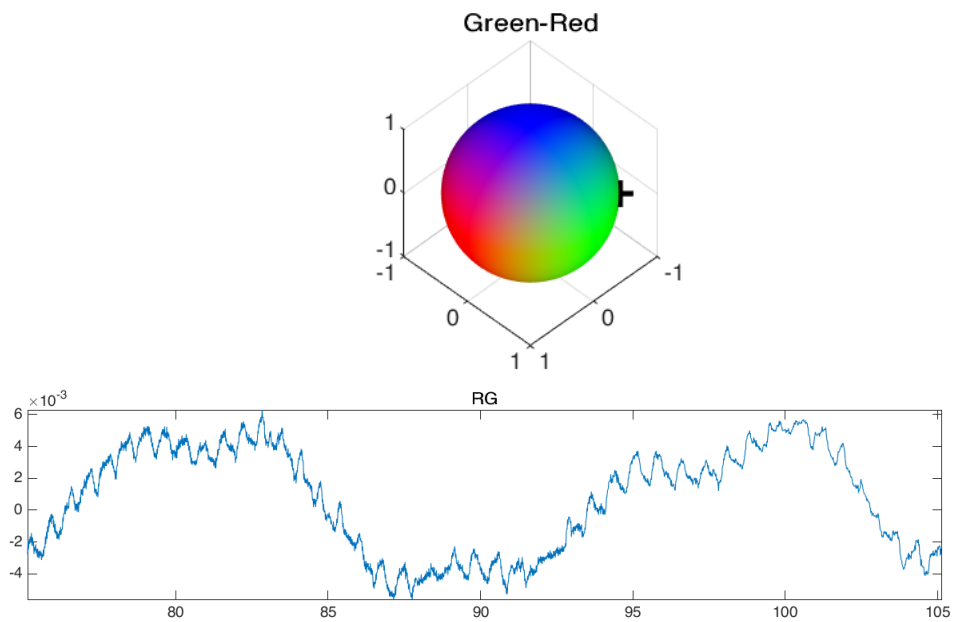


FIGURE 2.23 – Sphère RGB représentant le vecteur en sortie de *Green-Red* et signal correspondant.

combinaison, on considère une combinaison type qui va correspondre à une « teinte de peau normalisée ». Cette teinte de peau est supposée être la même pour toutes les personnes lors d'une mesure sous lumière blanche. Le fait d'utiliser ce modèle normalisé permet de fonctionner sur des vidéos n'utilisant pas la balance automatique des blancs [32]. Bien que la plupart des peaux soient suffisamment proches de cette peau normalisée, une dernière étape permet d'ajouter de l'adaptabilité en introduisant un coefficient  $\alpha$ . Ce coefficient est le ratio des écarts types de deux signaux intermédiaires  $\mathbf{x}(t)$  et  $\mathbf{y}(t)$ . Les deux signaux intermédiaires sont orthogonaux à la composante spéculaire et contiennent des proportions fixes de chaque composante. Les deux signaux sont ensuite recombinaison pour obtenir l'équation (2.11) :

$$\mathbf{s}(t) = \mathbf{x}(t) - \frac{\sigma\{\mathbf{x}(t)\}}{\sigma\{\mathbf{y}(t)\}}\mathbf{y}(t). \quad (2.11)$$

L'équation (2.11) correspond à une procédure d'*alpha-tuning* (avec  $\alpha = \sigma\{\mathbf{x}(t)\}/\sigma\{\mathbf{y}(t)\}$ ). Cette procédure permet de normaliser les signaux  $\mathbf{x}(t)$  et  $\mathbf{y}(t)$  de façon à ce que leurs composantes alternatives aient la même énergie. Si l'énergie portée par les deux composantes correspond à du mouvement, alors l'*alpha-tuning* va supprimer la composante du mouvement (puisque celle-ci est en phase dans les deux signaux intermédiaires). À l'inverse, comme les signaux rPPG sont en opposition de phase dans les deux signaux intermédiaires, si la composante rPPG représente l'essentiel de l'énergie des deux signaux alors la procédure d'*alpha-tuning* va simplement améliorer la qualité du signal en sortie.

Sur la sphère couleur (figure 2.24), la combinaison CHROM se situe comme prévu dans le plan orthogonal au vecteur bruit normalisé (1 1 1). Les deux points présentés en gris représentent les combinaisons intermédiaires  $\mathbf{x}(t)$  et  $\mathbf{y}(t)$ . Afin de simplifier la représentation, la composante Y a été inversée sur le schéma. Comme la formule finale est  $\mathbf{s}(t) = \mathbf{x}(t) - \alpha \times \mathbf{y}(t)$  et que  $\alpha$  est toujours positif, alors en inversant  $\mathbf{y}(t)$  on montre sur la sphère couleur l'ensemble des combinaisons possibles de CHROM qui se situent sur l'arc  $\mathbf{x}(t) \rightarrow \mathbf{y}(t)$ . Comme prévu, les différentes combinaisons sont sur le plan orthogonal à (1 1 1), dans l'arc  $\mathbf{x}(t) \rightarrow \mathbf{y}(t)$  et relativement proches les unes des autres (considérant que toutes les combinaisons sont faites sur des signaux de la même vidéo, donc assez proches).

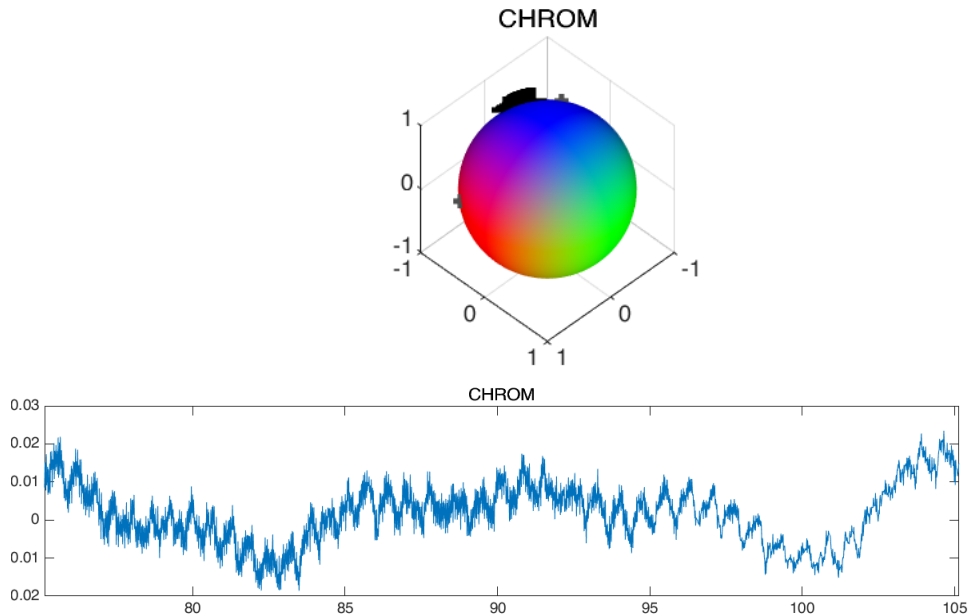


FIGURE 2.24 – Sphère RGB représentant le vecteur en sortie de CHROM et signal correspondant.

**2.4.2.4 PBV** La méthode PBV [32] repose sur des principes similaires à CHROM mais utilise un autre vecteur pour guider la combinaison. Un vecteur « guide », dépendant des conditions de la mesure, est utilisé comme paramètre pour la combinaison. Ce vecteur est ajusté en fonction des signaux couleur en entrée afin d'obtenir le vecteur de combinaison final.

Alors que CHROM [31] supprime le bruit associé aux variations locales d'intensité et à la réflexion spéculaire de la lumière, PBV utilise la différence entre les spectres d'absorption des tissus et du sang. Cette différence implique que les variations se font le long d'un vecteur spécifique. PBV propose le meilleur vecteur possible pour la combinaison. Malheureusement il est probable que des signaux ayant peu de distorsions rendent plus difficile l'estimation du vecteur de combinaison avec cette méthode. Cela provoquerait alors une diminution de la qualité de la mesure. De plus, ce vecteur existe à condition de se placer dans un espace couleur normalisé. PBV nécessite donc de connaître le spectre lumineux de la lumière et les caractéristiques caméra. Une variation même faible de la configuration de l'expérience va diminuer la qualité de la combinaison.

La combinaison PBV représentée sur la figure 2.25 montre un exemple de combinaison obtenue. La croix en gris représente le vecteur PBV qui a guidé la combinaison (c'est-à-dire (0,33 0,78 0,53)). Les différents vecteurs de combinaison sont relativement proches à la structure obtenue avec CHROM.

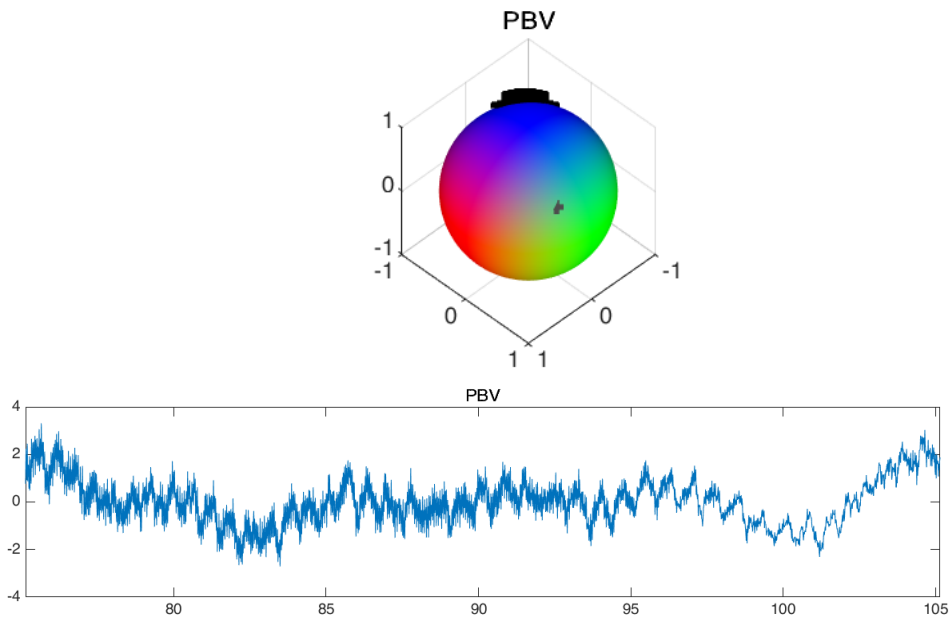


FIGURE 2.25 – Sphère RGB représentant le vecteur en sortie de PBV et signal correspondant.

**2.4.2.5 POS** POS (*Plane Orthogonal to Skin*) [157] est une méthode comparable à CHROM dans ses principes de base. Elle reprend l'idée de projeter les signaux sur un plan (orthogonal à une composante indésirable) puis de recombinaison les deux signaux intermédiaires en employant une méthode d'*alpha-tuning*. Dans le cas de CHROM, le plan intermédiaire supprimait la composante associée à la réflexion spéculaire de la lumière. La procédure d'*alpha-tuning* servait alors à affiner en maximisant le signal pulsatile et en minimisant la composante « variation d'intensité ». À l'inverse, POS supprime d'abord la contribution associée à la variation d'intensité puis combine les deux signaux intermédiaires afin de minimiser la contribution de la réflexion spéculaire. Dans le cadre de CHROM, la

procédure d'*alpha-tuning* faisait une différence entre les deux signaux. POS pour sa part fait une somme des deux signaux, en considérant que le signal pulsatile est en-phase avec les deux signaux intermédiaires et que les composantes spéculaires sont anti-phase. De ce fait, additionner les deux signaux permet d'augmenter l'énergie pulsatile et de diminuer celle portée par la réflexion spéculaire. La procédure d'*alpha-tuning* permet de calculer l'énergie portée par chaque signal de manière à les « normaliser » l'un l'autre. Ainsi, si la composante pulsatile est majoritaire dans les signaux, celle-ci est encore plus améliorée. Dans le cas où le bruit associé à la variation d'intensité est majoritaire, la normalisation permet aux deux signaux d'avoir le même niveau de bruit et donc de l'annuler via la somme des deux. Le problème associé à la procédure d'*alpha-tuning* est la même que celle proposée par CHROM, à savoir que celle-ci est moins efficace lorsque les niveaux de bruit et de signal pulsatile sont similaires. Dans ce cas là, la procédure ne permet pas de supprimer le bruit de manière optimale. La méthode POS permet d'être plus robuste que CHROM aux variations de specularité puisque la phase adaptative (*alpha-tuning*) gère ce bruit. Néanmoins, CHROM est plus robuste que POS aux variations de configurations expérimentales. En effet, POS projette sur un plan orthogonal à la peau normalisée. De ce fait, un changement des conditions d'illumination peut limiter l'efficacité de la méthode.

Sur la sphère couleur (figure 2.26), on voit les combinaisons POS obtenues. Il semble d'après cette sphère que la combinaison soit différente de PBV et CHROM mais qu'elle présente une plus grande stabilité.

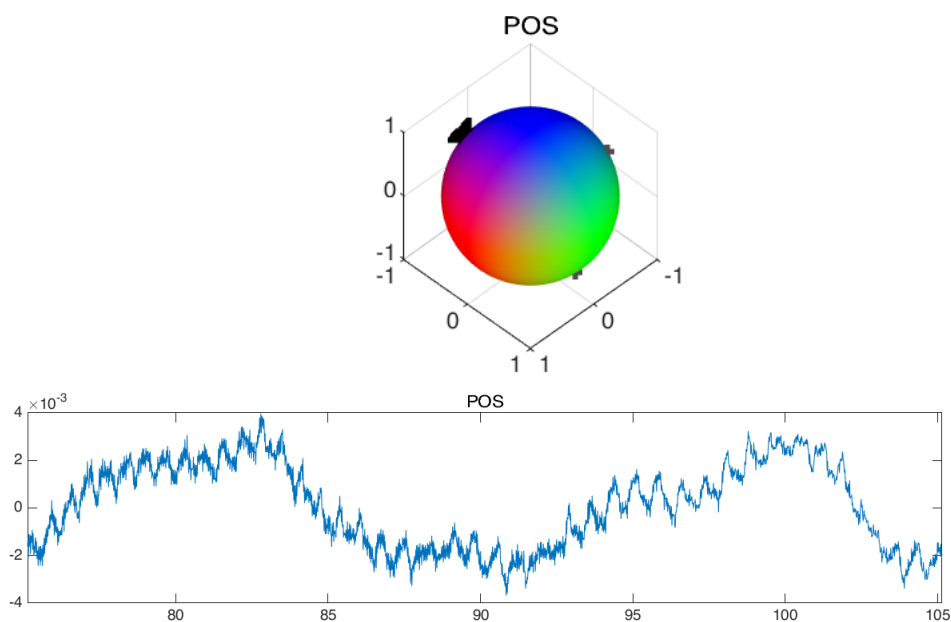


FIGURE 2.26 – Sphère RGB représentant le vecteur en sortie de POS et signal correspondant.

**2.4.2.6 SSR** Bien que les algorithmes présentés précédemment respectent tous le schéma présenté dans la figure 2.15, il existe des méthodes qui vont plus ou moins transgresser ce schéma par exemple en utilisant des données provenant d'autres étapes. C'est par exemple le cas de la méthode SSR (*Spatial Subspace Rotation*) [155] qui se base sur les valeurs des pixels de peau pour déterminer le vecteur de combinaison idéal. Pour cela, les pixels de peau sur une image sont représentés sur un espace colorimétrique et résumés sous la forme d'une matrice  $\mathbf{S}$  telle que :

$$\mathbf{S} = \frac{\mathbf{V}^T \cdot \mathbf{V}}{n} \quad (2.12)$$

avec  $n$  le nombre de pixels de peau et  $\mathbf{V}$  est une matrice de taille  $[n \times 3]$  contenant les valeurs RGB des pixels de peau. Cette matrice est décomposée via un algorithme QR pour obtenir la décomposition suivante :

$$\mathbf{S} \cdot \mathbf{U} = \mathbf{\Lambda} \cdot \mathbf{U} \quad (2.13)$$

avec  $\mathbf{U}$  la matrice des vecteurs propres et  $\mathbf{\Lambda}$  la matrice des valeurs propres. Le premier vecteur propre de la décomposition  $\omega_1$  correspond au vecteur modèle de la peau. Il représente la teinte de couleur sur laquelle on maximise la variation de la luminosité (cette variation étant représentée par la valeur propre).

On peut raisonnablement s'attendre à ce que la variation d'intensité sur ce vecteur soit due à des facteurs comme le mouvement de la personne ou les changements dans les conditions d'illumination. Il est moins probable en revanche que ces facteurs entraînent une variation de la teinte de la peau. Cette variation peut en revanche se produire du fait de la variation de la quantité de sang dans les capillaires et donc du signal photopléthysmographique. En considérant la rotation du vecteur  $\omega_1$  entre une image modèle  $\tau$  et une nouvelle image  $t$  telle que  $t \leq l$  (avec  $l$  le délai maximal avant de recalculer l'image modèle), on obtient une matrice de rotation  $R$  qui représente ce changement de teinte :

$$R = \omega_1(t)^T \cdot \omega_1(\tau). \quad (2.14)$$

Dans un deuxième temps, la mesure est raffinée en ajoutant la variation des valeurs propres. Enfin, la mesure est projetée vers l'espace RGB afin de faciliter les combinaisons des deux signaux. Pour cela, une procédure d'*alpha-tuning* est mise en place pour combiner les deux signaux.

Cette méthode permet d'avoir de meilleurs résultats que CHROM et PBV [155], notamment avec l'ensemble des couleurs de peau possibles. Le principal inconvénient de cette méthode est le besoin d'un ensemble de pixels représentant la peau. Cette méthode n'est donc applicable qu'avec un système rPPG sans contact. De plus, il est nécessaire d'avoir un algorithme fiable pour la détection des pixels de peau qui n'a pas de faux positifs.

## 2.5 Estimation de paramètres physiologiques

Une fois le signal (r)PPG extrait, le rythme cardiaque ou respiratoire est estimé à l'aide de méthodes assez simples. On trouve parmi les algorithmes employés des méthodes d'analyse fréquentielles comme la transformée de Fourier, les modèles autorégressifs [141] ou les transformées en ondelettes [20]. Le but de ces méthodes est de représenter le signal dans un domaine fréquentiel à partir duquel il est facile de déterminer la fréquence d'une oscillation (dans notre cas cardiaque ou respiratoire). Les autres méthodes typiquement utilisées pour déterminer les rythmes cardiaques et respiratoires effectuent une analyse temporelle du signal afin de détecter les événements d'intérêt (battement cardiaque, respiration). Ces méthodes sont souvent un peu moins robustes puisqu'elles ne prennent pas en compte l'ensemble du signal, néanmoins elles peuvent traiter correctement des signaux non-stationnaires et permettent également de détecter des événements particuliers comme une apnée ou une extrasystole (ce qui est plus difficile dans le domaine fréquentiel). Les méthodes d'analyse temporelle vont généralement se baser sur une détection des pics ou des creux dans le signal. Ces pics et ces creux peuvent ensuite être affinés selon des critères d'amplitude, de distance ou de largeur.

Dans le cas de la respiration, les méthodes temporelles peuvent être utiles dès l'instant où l'on cherche à extraire un signal respiratoire depuis les modulations d'amplitude ou de fréquence d'un signal cardiaque. En effet, ces deux méthodes nécessitent de récupérer des modulations qui ne seront pas forcément très visibles dans le domaine fréquentiel (deux

pics éloignés pour la modulation d’amplitude et un ensemble de pics pour la modulation fréquentielle).

Bien que de manière générale la mesure de paramètres physiologiques puisse être effectuée de manière générique avec les algorithmes présentés précédemment, certains travaux se sont attachés à mesurer un seul paramètre physiologique en faisant intervenir des algorithmes spécialisés dans ce but. Ces algorithmes peuvent ou non respecter le schéma présenté précédemment dans la figure 2.15. Les paramètres physiologiques présentés dans la suite de cette section sont le rythme cardiaque et le rythme respiratoire. Ces paramètres ont été sélectionnés principalement de part leur intérêt dans un contexte médical et leur littérature existante concernant la rPPG. À ces paramètres s’ajoutent la détection d’apnées qui est liée à la mesure du rythme respiratoire et pouvait donc difficilement être ignoré dans cette section. Enfin, les méthodes d’apprentissage profond seront présentées par la suite dans une section à part du fait de leurs spécificités. En particulier, les méthodes d’apprentissage profond ne s’appuient pas sur le pipeline présenté dans la figure 2.15 et implémentent une sorte de « boîte noire » qui transforme directement la vidéo en une mesure physiologique. Pour des questions d’exhaustivité et du fait des similarités entre les signaux PPG contact et les signaux rPPG, les algorithmes utilisés pour les deux types de signaux ont été rassemblés.

### 2.5.1 Méthodes dédiées à la mesure du rythme cardiaque

Le rythme cardiaque a été le premier paramètre physiologique étudié dans le cadre de la photopléthysmographie sans contact. Tout d’abord, il s’agit d’un paramètre relativement simple à extraire, contrairement à la saturation en oxygène par exemple. Ensuite, il est directement lié au signal rPPG puisque celui-ci représente les variations du volume sanguin dans les artères. Malgré cette apparente simplicité, plusieurs travaux ont essayé de fournir des méthodes d’extraction plus fiables du rythme cardiaque. Par exemple, le travail présenté dans [57] considère les combinaisons obtenues à l’aide d’ICA et de CHROM et combine leurs caractéristiques à l’aide d’un algorithme d’apprentissage machine. Pour cela, les signaux RGB sont découpés en segments et chaque segment est analysé à l’aide d’une transformée de Fourier. La rotation obtenue par ICA permet de créer trois segments fréquentiels supplémentaires, tandis que CHROM en donne encore un autre. Les segments d’ICA sont tous conservés car ICA ne définit pas d’ordre de préférence, contrairement à CHROM qui a pour but de retourner un et un seul segment de qualité maximale. L’ensemble de ces 7 transformées est ensuite normalisé en divisant par la moyenne puis utilisé dans un algorithme de type *Support Vector Regression* (SVR) afin d’extraire un rythme cardiaque. Afin de considérer la connexion des différents segments (qui se chevauchent), l’article propose soit de fusionner les différentes détections en effectuant une moyenne (fusion retardée) ou d’effectuer la fusion en entrée de l’algorithme SVR (fusion antérieure).

La modélisation du signal rPPG peut se faire sous la forme d’un processus de Markov d’ordre 1. En considérant le signal sous cette forme, [55] considère la possibilité de filtrer efficacement le bruit provoqué par le mouvement de la personne. Pour cela, un ensemble de signaux rPPG sont extraits depuis la région d’intérêt découpée en 6 sous-régions. En considérant que les signaux présentent un signal cardiaque corrélé et un signal bruit (mouvement et illumination) décorrélé, le signal est représenté sous la forme d’un processus Markovien. Une fonction de coût est dérivée de ce processus en considérant un terme pour représenter le bruit et un terme pour représenter le signal cardiaque. En ajustant les deux termes précédents il est possible de filtrer efficacement le bruit associé au mouvement.

Certaines méthodes se basent sur des décompositions du signal en modes. Le principal exemple de ce type de décomposition est EMD (*Empirical Mode Decomposition*) utilisé par exemple dans [137]. Dans cet article, les signaux PPG contact sont analysés avec une décomposition EMD afin de séparer le signal cardiaque d’éventuels bruits associés aux mouvements de la personne. Concernant la photopléthysmographie sans contact, on retrouve notamment

l'utilisation de la décomposition en modes non-linéaires (NMD) [35]. Cette décomposition est faite en 5 modes de façon à séparer la composante cardiaque, la composante respiratoire, les mouvements de la tête, les changements dans les conditions d'illumination, les déformations du visage (sourire, froncements de sourcils. . .). La méthode permet donc de séparer les différentes composantes de façon à utiliser une analyse de Fourier sur un signal ne contenant que les informations cardiaques.

### 2.5.2 Méthodes dédiées à la mesure du rythme respiratoire

Le rythme respiratoire est un paramètre physiologique plus difficile à extraire que le rythme cardiaque. Comme la fréquence de la respiration est plus basse, le risque de confondre une fréquence respiratoire avec une variation très basse fréquence (comme les variations de luminosité, les mouvements de la caméra ou les ondes de Mayer) est assez élevé. Les ondes de Mayer sont particulièrement problématiques puisqu'elles sont embarquées dans le signal rPPG et varient à des fréquences de 6 à 10 oscillations par minute [62]. Malgré ces difficultés, la mesure de ce paramètre physiologique est très importante pour une grande partie des diagnostics. En effet, le rythme respiratoire peut être utilisé afin de prévoir diverses situations à risques comme les septicémies [129], les arrêts cardiaques [24], les accidents ischémiques [124] ou les états de choc [114]. Malgré cela, le rythme respiratoire est très souvent négligé par le personnel médical faute de bien comprendre son intérêt et à cause du temps nécessaire à la mesure. Dans ce contexte, le développement des méthodes de mesure automatiques du rythme respiratoire est très important pour pouvoir régler ces différents problèmes. Malgré tout, la mesure du rythme respiratoire à l'aide de la rPPG reste assez peu développé dans la littérature. Un état de l'art portant sur l'extraction de la respiration à partir de la photopléthysmographie contact a été réalisé par Charlton *et al.* [23, 22]. Bien qu'il s'agisse de la version contact et que la qualité du signal est de ce fait bien meilleure, l'article présente un grand nombre d'algorithmes qui peuvent être réutilisés dans le cadre de la rPPG. De manière générale, l'extraction de la respiration depuis la rPPG se fait à l'aide de trois modulations différentes (présentées précédemment dans la figure 2.8).

**2.5.2.1 Modulation additive (Baseline Wander)** La modulation additive correspond à la variation du retour veineux en fonction des changements de pression dans le thorax. Celle-ci se traduit par une variation de la ligne de base dans le signal rPPG. La mesure de la respiration depuis la modulation additive peut facilement être bruitée du fait de la présence de composantes très basse fréquence (ondes de Mayer, mouvements, variation des conditions d'illumination. . .). Cela rend important le choix des algorithmes afin de ne pas intégrer de bruit dans le signal final ni filtrer le signal respiratoire.

Un exemple de mesure respiratoire depuis la modulation additive est présenté dans Sun *et al.* [138], l'analyse du signal est faite en employant une pseudo distribution de Wigner Ville. La principale contribution de l'article est la compensation du mouvement à l'aide d'un algorithme de séparation de source aveugle.

Dans Tarassenko *et al.* [143], la mesure du rythme respiratoire est obtenue via une analyse fréquentielle du signal respiratoire en se basant sur un modèle auto-régressif avec des fenêtres de 30 secondes. Une seconde partie du travail a consisté à réaliser une carte représentant le rythme respiratoire en fonction de la localisation sur la peau, de façon à pouvoir analyser la qualité de la mesure selon les localisations.

Enfin, dans [151], la respiration est obtenue également en utilisant une modulation de ligne de base. Pour combiner les signaux RGB, le vecteur de combinaison est d'abord estimé avec CHROM ou PBV sur des signaux RGB filtrés pour ne contenir que la composante cardiaque. Ce vecteur est ensuite réutilisé sur les signaux RGB filtrés pour ne garder que la composante respiratoire.

**2.5.2.2 Modulation d’amplitude (Amplitude Modulation)** La modulation d’amplitude correspond à une variation de la hauteur des pulsations PPG liée à la respiration. Il ne semble pas exister de travaux sur la rPPG étudiant l’extraction de la respiration depuis la modulation d’amplitude. Il est probable que la modulation de la ligne de base soit préférée pour des raisons de simplicité, et que la modulation de fréquence soit de son côté privilégiée pour sa robustesse aux variations d’intensité lumineuse. Malgré ce manque, on retrouve des algorithmes de mesure de la respiration dans les travaux sur la photopléthysmographie au contact. Ainsi, dans les articles de Karlen *et al.* [65] et Pimentel *et al.* [111], on retrouve l’utilisation de la modulation d’amplitude. Cette modulation est extraite de façon relativement simple (différence d’amplitude entre le haut de la pulsation PPG et de son creux). De plus, elle est fusionnée dans le cas de ces articles avec les autres modulations et les estimations sont filtrées en fonction d’une métrique de qualité. Tout cela laisse supposer que l’estimation du signal respiratoire en utilisant la distance entre un pic et un creux n’est pas suffisante en soi pour les signaux PPG, et posera donc encore plus de soucis lors de son utilisation sur des signaux rPPG.

Afin d’améliorer la qualité de la mesure, on trouve également dans la littérature des travaux utilisant des méthodes d’analyse du signal plus avancées. L’utilisation de méthodes d’analyse temps-fréquence permet de prendre en compte l’ensemble de points de la pulsation et non plus un seul. Dans [96, 26], la transformée en ondelettes continue et la VFCDM (*Variable Frequency Complex Demodulation*) sont employées pour extraire le signal respiratoire soit par le biais des modulations de fréquences ou d’amplitude. Ces méthodes donnent des résultats intéressants car elles peuvent traiter des signaux non-stationnaires (ce qui est souvent le cas pour les signaux biologiques).

La transformée en ondelettes continue est également utilisée dans [6] comme entrée dans l’algorithme présenté par l’article. Le but de cet algorithme est d’extraire des informations cachées dans le signal sous la forme de modulation d’amplitude ou de fréquence, et est donc applicable dans le cas des signaux PPG. Une autre adaptation de la transformée en ondelettes se retrouve dans [33] par le biais d’une transformée SST (*SynchroSqueezing Transform*). Cette transformée améliore la transformée en ondelettes de façon à augmenter la résolution de la transformée et avoir une représentation de la fréquence respiratoire instantanée.

Une autre transformée nommé CSD (*Correntropy Spectral Density*) permet également de tirer parti de la modulation d’amplitude en représentant la fréquence de celle-ci directement dans sa transformée, contrairement à une transformée de Fourier qui placerait la fréquence de la modulation avant et après la fréquence porteuse (cardiaque). Cette transformée est utilisée dans [51] afin d’extraire à la fois le rythme cardiaque et respiratoire.

Deux autres algorithmes sont présentés dans [74] pour la démodulation d’amplitude. Le premier consiste à employer un opérateur de Teager afin d’effectuer la démodulation AM / FM. Cet opérateur permet de faire la démodulation et donc de récupérer le signal respiratoire par la suite à l’aide d’un filtrage simple. La deuxième méthode, qui a donné de meilleurs résultats, consiste à employer la méthode SSA (*Singular Spectrum Analysis*). Celle-ci construit une matrice en décalant un segment de signal dans le temps, puis la décompose de façon à en extraire les informations importantes. Cette méthode donne de meilleurs résultats que la méthode utilisant l’opérateur de Teager.

La modulation d’amplitude est un peu moins sensible aux variations de luminosité que la modulation de la ligne de base, néanmoins celle-ci nécessite des algorithmes plus avancés afin de pouvoir l’extraire correctement. Une modulation plus robuste encore est la modulation de fréquence. Celle-ci est totalement indépendante des variations de la luminosité mais nécessite des algorithmes plus avancés pour pouvoir l’extraire. La fréquence d’échantillonnage est un point crucial sur la modulation de fréquence, ce qui pose certains problèmes dans le cas de la rPPG.



**2.5.2.3 Modulation de fréquence (Frequency Modulation)** La modulation de fréquence, présentée précédemment sur la figure 2.5 est provoquée par une réaction du système nerveux autonome qui régule le rythme cardiaque en fonction de la respiration. La réponse du système nerveux autonome va être notamment dépendante de l'âge, celle-ci diminuant avec le vieillissement [43].

Les algorithmes d'extraction de la respiration qui se basent sur la modulation de fréquence vont être plus robuste que les autres modulations car la fréquence n'est pas dépendante des conditions d'illumination et du mouvement. Le problème majeur de la modulation de fréquence est le besoin d'une détection des pulsations qui soit très robuste et très précise dans les moments des pics. En effet, la variation de la fréquence reste très faible ce qui fait qu'une petite erreur va fortement modifier la qualité de la mesure.

Une première estimation du rythme respiratoire en utilisant l'*InterBeat Interval* (IBI) a été présenté par Poh *et al.* [112]. Dans cet article, les signaux RGB sont combinés en utilisant l'algorithme ICA présenté dans le chapitre 2, section 2.4.1.1. En détectant les battements sur le signal pulsatile, le signal IBI est extrait puis analysé à l'aide d'un périodogramme de Lomb.

Dans les travaux de Mirmohamadsadeghi *et al.* [90], une estimation du signal respiratoire est faite en utilisant l'IBI qui est globalement équivalent à la variabilité du rythme cardiaque. Le calcul de l'IBI est obtenu à l'aide d'un ensemble de filtres passe-bande et raffiné en utilisant une décomposition en valeurs singulières.

L'extraction de l'IBI peut également se faire à l'aide de méthodes d'analyse temps fréquence comme une décomposée en ondelettes. Ainsi, dans Bousefsaf *et al.* [20], le signal rPPG est extrait et analysé à l'aide d'une transformée en ondelettes continue. Cette transformée permet de filtrer le bruit plus finement d'une part, mais également d'extraire le signal respiratoire en détectant les coefficients d'ondelettes ayant la plus grande amplitude.

**2.5.2.4 Spectral Kurtosis** La *spectral kurtosis* est une transformée comparable à la transformée de Fourier. Contrairement à cette dernière, le *spectral kurtosis* permet de suivre les non-stationnarités du signal et la localisation de leurs changements. Là où la transformée de Fourier est dédiée à l'analyse de signaux stationnaires (ou presque stationnaires), le *spectral kurtosis* permet de prendre en compte les cas non-stationnaires, très présents dans les signaux médicaux. De plus, il est plus robuste au bruit que la transformée de Fourier. Une méthode utilisée pour le calcul du *spectral kurtosis* a été présenté [11]. Cette transformée a ensuite été réutilisée dans [38] comme une information supplémentaire liée à la respiration, au même titre que les trois modulations présentées précédemment.

**2.5.2.5 Fusion des modulations** D'après les travaux de Charlton *et al.* [22] portant sur les algorithmes utilisés sur la PPG contact pour l'estimation du rythme respiratoire, le fait d'effectuer une fusion de plusieurs des modulations présentées précédemment permet d'améliorer de manière significative la qualité de la mesure. Pour cela, on retrouve plusieurs méthodes de fusion, la principale étant la *smart fusion* introduite par Karlen *et al.* [65].

La *smart fusion* consiste à calculer la moyenne et la variance des estimations du rythme respiratoire faites sur chaque modulation. Si la variance est suffisamment faible (c'est-à-dire inférieure ou égale à 4 rpm), la moyenne des rythmes est conservée comme estimation finale. Dans le cas contraire on considère que le rythme respiratoire n'a pas pu être mesuré sur le segment concerné.

D'autres méthodes plus complexes, utilisant l'apprentissage machine, ont également vu le jour. Ainsi, les travaux de [38] présentent une nouvelle mesure corrélée à la respiration (le *spectral kurtosis* présenté dans la section 2.5.2.4 du chapitre 2). En plus de cette nouvelle

méthode pour mesurer le rythme respiratoire, une fusion de cette métrique avec les trois modulations habituelles a été mise en place en employant un algorithme d'*extreme learning machine*.

### 2.5.3 Méthodes dédiées à la détection d'apnées

Un paramètre secondaire à la respiration est la détection d'apnée au cours de la mesure. Cette information permet de confirmer certains diagnostics (apnée du sommeil) ou de gagner du temps sur des situations d'urgence (apnée d'un nourrisson pouvant amener à un arrêt cardiaque). La détection d'apnée ne semble pas être un sujet très étudié avec la photopléthysmographie sans contact. Néanmoins, les méthodes utilisées avec d'autres méthodes peuvent potentiellement être réutilisées avec la photopléthysmographie à condition de faire leur analyse sur un signal respiratoire. Par exemple, Lorato *et al.* [82] réalise une détection d'apnée en utilisant des vidéos de nouveaux nés. Pour cela, le travail proposé consiste à mesurer le signal respiratoire en se basant sur un algorithme de détection du mouvement puis à faire une détection des apnées en comparant l'écart type sur des fenêtres courtes et longues. Une forte baisse de l'écart type de la fenêtre courte par rapport à l'écart type de la fenêtre longue va alors correspondre à une apnée.

Les algorithmes de *machine learning* peuvent également être mis à contribution dans la détection d'apnées. Par exemple, les travaux de [144] ont consisté à extraire un signal respiratoire à l'aide de la rPPG et de détecter des apnées sur le signal respiratoire obtenu. Pour cela, les algorithmes SVM (*Support Vector Machine*), *K-nearest neighbors* et *XGBoost* ont été testés sur le signal respiratoire. L'article parle de rPPG mais la description de la méthode et notamment les figures laissent penser que la mesure du signal respiratoire se fait à l'aide des mouvements du torse de la personne. Néanmoins, comme la classification se fait sur un signal respiratoire on peut supposer que les résultats seraient similaires en utilisant des signaux extraits à l'aide de la rPPG traditionnelle.

## 2.6 Estimation de la qualité du signal

Malgré les améliorations de ces méthodes, il n'est pas possible de toujours pouvoir extraire un signal rPPG de qualité suffisante. Dans des situations réelles, la présence de mouvements, de variations de luminosité ou la qualité de la caméra font que le signal rPPG peut contenir des segments inutilisables. Dans ce contexte, un traitement important des signaux rPPG va être de pouvoir déterminer les segments qui n'ont pas une qualité suffisante pour pouvoir avoir une mesure fiable. Cette différenciation va généralement se faire en calculant une métrique de qualité et en effectuant un seuillage. Le choix du seuil va dépendre de la nécessité d'avoir des estimations fiables tout en ayant suffisamment de mesures. Le point principal des algorithmes de mesure de la qualité va être de fournir une valeur qui soit aussi proche que possible de la qualité réelle du signal. Pour cela, de nombreuses approches ont été mises en place afin de pouvoir fournir une valeur de la qualité la plus fiable possible.

Par exemple, dans [154], la qualité du signal est estimée en calculant deux métriques. La première métrique se calcule en amont des algorithmes rPPG et la deuxième métrique en aval, ce qui permet de les appliquer indépendamment de l'algorithme utilisé pour l'extraction du signal. La métrique en amont étudie les conditions d'illumination et les propriétés de la peau. Quatre aspects de la source de lumière sont considérés sur cette métrique : l'intensité lumineuse, le type de lumière (spot ou parallèle, spot étant à éviter), le spectre lumineux (qui affecte la composante spéculaire) et la direction de la lumière (qui détermine la proportion de la lumière modulée par le mouvement). À cela s'ajoute deux caractéristiques de la peau : la teinte de la peau (les peaux foncées présentant une qualité de signal plus faible du fait

de l'absorption par la mélanine) et le nombre de pixels de peau qui sont considérés (ceux-ci permettant de supprimer le bruit d'échantillonnage).

Une autre façon d'estimer la qualité de la mesure se fait à l'aide de modèles statistiques. Le signal est modélisé et une fois adaptés les paramètres obtenus permettent de fournir des renseignements sur la ressemblance du signal avec le modèle. Ainsi, [15] propose une méthode basée sur un modèle de Markov caché afin de modéliser la forme de la pulsation cardiaque. Le modèle est constitué de quatre états (signal croissant, maximum local, signal décroissant, minimums). En comparant le signal avec le modèle, il est possible de déterminer la qualité du signal en calculant le *Bayesian Information Criterion*.

## 2.7 Le cas de l'apprentissage profond

Le schéma présenté dans la figure 2.15 est très souvent respecté dans les travaux réalisés avec un cloisonnement fort de chaque algorithme. Néanmoins, il existe des situations dans lesquelles ce cloisonnement est réduit afin d'améliorer l'efficacité des algorithmes. Ce cas a notamment été discuté dans la section sur les combinaisons (chapitre 2, section 2.4) avec la méthode SSR qui réutilisait la détection de pixels de peau pour déterminer une combinaison idéale. D'autres travaux ignorent totalement ce pipeline en introduisant un seul algorithme qui regroupe plusieurs traitements en un seul algorithme, généralement de *deep learning*. Ces algorithmes sont entraînés par un jeu de données de très grande taille de façon à adapter automatiquement leurs paramètres et déterminer au mieux les mesures attendues. On trouve notamment ces algorithmes pour l'extraction du rythme cardiaque et de la variabilité du rythme cardiaque. Ces méthodes ont connu un essor à la suite du début de la thèse et représentent une nouvelle façon de considérer les méthodes de photopléthysmographie sans contact. Elles permettent de mieux prendre en compte les non-linéarités qui sont souvent retrouvées dans des signaux biologiques en ajustant des fonctions extrêmement complexes de manière automatisée. En retour, elles nécessitent de très grands jeux de données et il est difficile de garantir que leurs performances resteront les mêmes sur des données inconnues (à part en utilisant un autre jeu de données pour l'évaluation). Les méthodes d'apprentissage profond sont très éloignées des travaux réalisés au cours de la thèse. Néanmoins, les méthodes d'apprentissage profond représentent aujourd'hui une part très importante des nouvelles contributions à la photopléthysmographie sans contact, et il n'était donc pas possible de les ignorer complètement.

Parmi ces nouvelles contributions, on trouve le travail de Kopeliovich *et al.* [69] dans lequel une estimation du rythme cardiaque est obtenue en utilisant un réseau de neurones convolutionnel. Pour cela, on considère comme entrée les signaux RGB sur un ensemble de portions du visage. Ces signaux sont définis sur une durée donnée et fournis en entrée du réseau. Le réseau de neurones permet alors d'obtenir une estimation du rythme cardiaque sur la portion considérée. Dans l'article, l'estimation du rythme cardiaque est considérée soit comme un problème de régression ou de classification.

Dans l'article [25], un réseau de neurones convolutionnel est employé de façon à extraire un rythme cardiaque et respiratoire directement depuis une vidéo. L'intérêt de cette méthode est de mieux gérer les mouvements de la tête, néanmoins il y a moins de contrôle sur les différents paramètres à ajuster. Le système repose sur une représentation du mouvement basé sur la réflexion de la peau et sur un modèle d'attention. Ce modèle repose sur le concept de la vue humaine qui prend en compte à la fois un point de petite taille mais avec une grande résolution (focus) et une zone plus floue mais très grande. Cette attention dépend des caractéristiques de l'image et est utilisée de façon à représenter le mouvement de la personne.

L'article [131] utilise également un réseau de neurones convolutionnel mais le but de celui-ci n'est plus de traiter une vidéo de bout en bout mais d'utiliser une image générée par des

méthodes d'extraction traditionnelles, de façon à améliorer la qualité de la mesure du rythme cardiaque. Afin de créer l'image en entrée du réseau de neurones, trois régions d'intérêt sont considérées. Un signal pulsatile est obtenu pour chacune de ces régions d'intérêt et une matrice est construite à l'aide de décalages successifs du signal pulsatile. Les trois images en niveaux de gris obtenues sont ensuite regroupées dans une seule image couleur afin de servir de donnée d'entrée au réseau de neurones. Le réseau est d'abord entraîné à l'aide de données synthétiques, puis une étape de transfert est effectuée avec des données réelles de façon à améliorer la détection du rythme cardiaque.

Enfin, l'emploi de l'apprentissage profond a également fait l'objet d'études pour l'extraction du HRV [83]. Cette méthode utilise l'apprentissage profond pour pouvoir estimer le HRV directement depuis les vidéos de la personne. Elle remplace donc l'ensemble du pipeline de la figure 2.15.

L'état de l'art présenté avait pour objectif de brosser un portrait assez large des différents algorithmes et des différentes configurations pouvant être utilisées pour la photopléthysmographie sans contact. Celles-ci sont extrêmement nombreuses et peuvent se combiner les uns avec les autres. Dans ce cadre très large, il a été nécessaire d'effectuer des choix afin de concentrer ses efforts dessus et pouvoir produire un résultat utilisable. Dans ce contexte, le choix a été fait de mesurer le rythme respiratoire à l'aide d'une caméra RGB, en gardant comme objectif une implémentation en temps réel. Le choix du rythme respiratoire a été fait en tenant compte de sa popularité qui semble assez faible par rapport à l'importance de la mesure. En regardant la littérature il semble qu'une grande partie des méthodes utilisées pour la respiration se basent en fait sur l'extraction du rythme cardiaque. De ce fait, on suppose fortement que la mesure du rythme respiratoire présente un potentiel à être amélioré en réalisant des algorithmes dédiés à sa mesure. L'utilisation de caméras RGB a été faite afin de pouvoir rester dans un cadre assez classique et s'assurer que les algorithmes développés pourront être utilisés dans le plus grand nombre d'applications possibles. Enfin, l'objectif de temps réel retenu pour cette thèse était très important. En effet, la photopléthysmographie sans contact côtoyant fortement le domaine médical, un grand nombre d'applications de cette technologie ne présentent aucun intérêt si les informations obtenues le sont après un délai trop important. Par exemple, la détection d'apnées chez les bébés prématurés va permettre de prévoir une détresse respiratoire et réagir plus rapidement. De ce fait, il est nécessaire de détecter les apnées le plus rapidement possible afin que le personnel médical puisse réagir efficacement. Dans la suite du manuscrit, les différents travaux réalisés au cours de la thèse seront présentés en détail. Dans un premier temps, les différentes méthodes utilisées pour la comparaison des algorithmes sont présentés, avec notamment la base de données, les métriques de qualité et les graphiques utilisés. Dans un deuxième temps, les différents algorithmes sont présentés. Les principes généraux, le détail des algorithmes et les résultats obtenus avec les tests sont précisés afin de pouvoir comparer au mieux les différents algorithmes.



---

## Deuxième partie

---

### CONTRIBUTIONS

De nombreuses méthodes ont été développées au fil du temps afin d'évaluer au mieux l'état de santé d'une personne. Ces méthodes doivent répondre à de nombreux critères et mesurer plusieurs paramètres physiologiques, ce qui explique leur diversité. De plus, elles font l'objet au fil des années d'avancées et de découvertes scientifiques dont le but est de les rendre plus efficaces, plus portables et plus faciles à utiliser, de façon à accompagner les nouveaux usages associés à la médecine. Pour cela, les dispositifs de mesure évoluent non seulement du côté matériel, avec l'utilisation de nouveaux matériaux ou l'ajout de filtres, mais également au niveau logiciel avec des algorithmes plus précis et moins consommateurs de ressources.

Dans le cadre de mes stages ingénieurs, des travaux préliminaires à ceux réalisés durant la thèse avaient pour objectif principal d'implémenter sur des cibles embarquées une mesure basée sur la photopléthysmographie sans contact. Ces travaux se sont appuyés sur l'expérience acquise par l'équipe de recherche en photopléthysmographie sans contact [18, 19] mais également dans le domaine de l'implantation de traitement d'images embarquées [86, 123, 146]. En se basant sur cette expérience, l'intégration d'une chaîne complète de mesure à l'aide d'une caméra intelligente (potentiellement mobile) a été étudiée.

Ainsi, le stage portait sur une implémentation d'un algorithme rPPG sur une plateforme GPU. Cette implémentation comprenait notamment une segmentation en superpixels et la détection d'un rythme cardiaque. Par ailleurs, une implémentation de la chaîne complète de mesure a été réalisée sur une carte Raspberry Pi. L'intérêt d'implémenter des traitements au niveau de la caméra est de ne transmettre sur le réseau que des informations de haut niveau (par exemple la présence d'une personne) plutôt que le flux vidéo. Cela permet de réduire l'encombrement du réseau, de limiter les problèmes en cas de piratage et de mieux protéger la vie privée des utilisateurs.

Le but de l'algorithme implémenté était de mesurer les paramètres physiologiques afin de pouvoir détecter le malaise d'une personne dans le bâtiment. Cet élément s'insérerait dans un système plus complet où, après une première détection de la chute d'une personne, un robot était envoyé afin de faire une détection plus fiable d'une part et de mesurer les paramètres physiologiques d'autre part. La programmation du robot a été le sujet d'un autre stage et le but de celui-ci était de permettre son déplacement vers la victime afin de faire une confirmation. Pour cela, le robot était équipé d'une caméra qui donne une image RGB et une carte de profondeur.

Ces différents travaux réalisés ont permis d'avoir une première approche de la photopléthysmographie sans contact. Ceux-ci ont amené à la réalisation d'une chaîne de traitements complète, implémentée sur des plateformes matérielles dédiées et conçue dans un but précis (par exemple la détection de fatigue chez les conducteurs). Les algorithmes devaient être implémentés de façon à fonctionner en temps réel sur des plateformes assez limitées en terme de ressources. Ce problème était très important à résoudre car la mesure rPPG en temps réel est nécessaire dans un grand nombre d'applications. Ces stages ont permis de réaliser un survol de la photopléthysmographie sans contact dans son ensemble et de participer à une première publication [2].

Par la suite, la thèse présentée avait pour objet la mesure de paramètres physiologiques autres que le rythme cardiaque. Afin d'apporter sa pierre à l'édifice, les travaux réalisés se

sont attachés à améliorer la mesure du rythme respiratoire. Pour cela, la partie du traitement correspondant à la combinaison des signaux couleur a fait l'objet d'une attention toute particulière. Cela a donné lieu à un ensemble de méthodes détaillées dans la suite du manuscrit. Les méthodes ont été comparées à l'aide d'un protocole expérimental constant d'une méthode à l'autre afin de pouvoir comparer au mieux les résultats. Le protocole expérimental est détaillé dans le chapitre 5 avec notamment la base de vidéos utilisée, les méthodes de l'état de l'art, les métriques et les graphiques utilisés pour comparer la qualité des méthodes. Une fois le protocole expérimental présenté, une première hypothèse est formée pour améliorer l'état de l'art et une méthode de combinaison est mise en place afin de tester cette hypothèse. Cette première méthode permet de valider l'hypothèse mais est inutilisable pour des situations réelles. Pour corriger ce problème, plusieurs nouvelles méthodes utilisant la modulation de la ligne de base sont ensuite développées et présentées par la suite. Ces méthodes déterminent la combinaison permettant de maximiser le signal respiratoire directement en se basant sur cette modulation plutôt que les modulations d'amplitude et de fréquence, ce qui simplifie le problème dans un premier temps. Enfin, une dernière méthode fusionne les différentes modulations respiratoires afin d'améliorer la qualité de la mesure du rythme respiratoire. Chacune des méthodes est présentée en montrant un rapide contexte et l'idée générale de l'algorithme. Elles sont ensuite détaillées et testées afin de comparer la méthode avec l'état de l'art et déterminer les limitations de chaque algorithme et les améliorations qui feront l'objet de travaux futurs.

## Chapitre 3

# Analyse du lien entre vecteur de combinaison et qualité du signal : Méthode Bruteforce

Afin d'apporter une contribution à la littérature, une première analyse a été réalisée en cherchant à mieux comprendre le fonctionnement de l'étape de combinaison des signaux couleur. Dans celle-ci, l'ensemble des combinaisons possibles est réalisé et la qualité de chaque signal obtenu est calculée afin de déterminer quelles combinaisons donnent les meilleurs résultats. Cette analyse amène au développement d'une première méthode, Bruteforce, qui détermine la combinaison donnant la meilleure qualité de signal d'un point de vue du SNR.



Les travaux réalisés au cours de la thèse se sont principalement concentrés sur la meilleure façon possible de combiner les signaux couleur obtenus depuis la vidéo en signal 1D contenant le plus d'information PPG (respiratoire et cardiaque) et le moins de bruit possible. Dans la littérature, cette étape consiste dans la majorité des cas à réaliser une combinaison linéaire des signaux. Cela permet d'avoir des calculs plus simples et prend en compte le modèle dichromatique présenté précédemment dans la section 1.2.3 du chapitre 2. Ainsi, le problème se résume dans le cas des signaux RGB à déterminer un vecteur  $\omega \in \mathbb{R}^3$  idéal et à projeter les signaux sur ce vecteur. On retrouve cette idée dans l'état de l'art chapitre 2 section 2.4 dans laquelle les méthodes de combinaison courantes sont représentées à l'aide d'une sphère de couleur. Le fait de multiplier le vecteur  $\omega$  par un scalaire augmente l'énergie du signal et du bruit dans le signal 1D en sortie dans des proportions similaires. De ce fait, on peut réduire l'espace des solutions à une sphère unitaire. De manière plus générale, l'objectif de cette étape est de maximiser le ratio du signal sur le bruit dans le signal résultat.

## 1 Étude préliminaire

Lors d'une première étude, nous avons observé la répartition des SNR sur la sphère unitaire en calculant le SNR sur tous les signaux possibles, calculés à partir de tous les vecteurs  $\omega$  possibles. L'étude représente une fonction  $f(\omega, \mathbf{C}(t)) = SNR$  telle que  $\omega \in \mathbb{R}^3$  représente le vecteur de combinaison avec  $|\omega| = 1$  et  $\mathbf{C}(t)$  représente les signaux RGB à combiner,  $\mathbf{C}(t) \in \mathbb{R}^{N \times 3}$ . Cette méthode revient donc à calculer la projection des signaux RGB sur le vecteur  $\omega$  puis à calculer le SNR sur le signal 1D généré (équation (3.1)) :

$$f(\omega, \mathbf{C}(t)) = SNR(\mathbf{C}(t) \cdot \omega^T). \quad (3.1)$$

De manière générale, le SNR est considéré comme étant :

$$SNR = 10 \log_{10} \frac{P_s}{P_b} \quad (3.2)$$

en considérant  $P_s$  la puissance du signal et  $P_b$  la puissance du bruit. Afin de limiter le nombre de calculs à réaliser, le choix a été fait de calculer le SNR sous la forme :

$$SNR = 10 \log_{10} \frac{P_s}{P_\phi} \quad (3.3)$$

avec  $P_\phi$  la puissance du signal complet (signal et bruit). Cela permet par la suite de réduire les calculs en n'effectuant qu'un seul fenêtrage pour séparer le signal et le bruit, et en ne calculant qu'une seule matrice de covariance du bruit  $\mathbf{C}_x$  pour les méthodes EVM-filtre et EVM-FFT qui seront présentées dans les chapitres suivants. Une amélioration pouvant être testée par la suite serait de supprimer la contribution du signal dans la partie « bruit » en la soustrayant.

Pour pouvoir utiliser la formule (3.3), on doit démontrer que maximiser le SNR de l'équation (3.2) revient à maximiser le SNR de l'équation (3.3). Si le signal et le bruit sont indépendants,

on a  $P_\phi = P_s + P_b$ . Dans ce cas, déterminer le vecteur  $\omega$  maximisant le SNR de l'équation (3.2) revient à calculer :

$$\operatorname{argmax}_{\omega \in \mathbb{R}^3, |\omega|=1} 10 \log_{10} \frac{P_s}{P_b} = \operatorname{argmax}_{\omega \in \mathbb{R}^3, |\omega|=1} \frac{P_s}{P_b} \quad (3.4)$$

$$= \operatorname{argmin}_{\omega \in \mathbb{R}^3, |\omega|=1} \frac{P_b}{P_s} \quad (3.5)$$

$$= \operatorname{argmin}_{\omega \in \mathbb{R}^3, |\omega|=1} \frac{P_\phi - P_s}{P_s} \quad (3.6)$$

$$= \operatorname{argmin}_{\omega \in \mathbb{R}^3, |\omega|=1} \frac{P_\phi}{P_s} - 1 \quad (3.7)$$

$$= \operatorname{argmin}_{\omega \in \mathbb{R}^3, |\omega|=1} \frac{P_\phi}{P_s} \quad (3.8)$$

$$= \operatorname{argmax}_{\omega \in \mathbb{R}^3, |\omega|=1} \frac{P_s}{P_\phi} \quad (3.9)$$

$$= \operatorname{argmax}_{\omega \in \mathbb{R}^3, |\omega|=1} 10 \log_{10} \frac{P_s}{P_\phi}. \quad (3.10)$$

À partir du moment où les composantes signal et bruit sont indépendantes, on peut utiliser indifféremment les formules (3.2) ou (3.3) pour la maximisation. Une erreur légère peut être introduite dans le vecteur de combinaison final en utilisant cette formule, probablement à cause des approximations dans le calcul numérique.

Dans le cas de la méthode Bruteforce, le SNR est calculé avec :

$$SNR(\mathbf{C}(t), \mathbf{h}(f), f_c, f_{min}, f_{max}) = 10 \log_{10} \left( \frac{\int_{f_{min}}^{f_{max}} \mathbf{h}_{f_c}(f) |\mathcal{F}\{\mathbf{C}(t)\}(f)|^2 df}{\int_{f_{min}}^{f_{max}} |\mathcal{F}\{\mathbf{C}(t)\}(f)|^2 df} \right) \quad (3.11)$$

$\mathbf{C}(t)$  étant les signaux couleur,  $\mathbf{h}(f)$  la fonction utilisée pour séparer le signal du bruit,  $f_c$  la fréquence sur laquelle centrer la fenêtre  $\mathbf{h}(f)$ ,  $f_{min}$  et  $f_{max}$  les fréquences physiologiques minimale et maximale (0,1 Hz à 0,7 Hz pour la respiration). En considérant  $\mathbf{C}(t)$  comme étant un ensemble de signaux RGB représentatifs des signaux habituellement rencontrés et  $\mathbf{h}(f)$ ,  $f_c$ ,  $f_{min}$ ,  $f_{max}$  un ensemble de paramètres fixés empiriquement, on simplifie l'équation (3.1) pour donner l'équation (3.12) :

$$\begin{aligned} f &: \mathbb{R}^3 &\rightarrow \mathbb{R} \\ \omega &&\mapsto SNR. \end{aligned} \quad (3.12)$$

On simplifie enfin l'équation en considérant :

- que le vecteur  $\omega$  est un vecteur unitaire ;
- que le vecteur opposé  $-\omega$  donne le même SNR puisque celui-ci correspond au signal opposé.

Cela nous donne ainsi la possibilité de représenter l'espace des solutions sur une demi sphère. Cette simplification limite très fortement le nombre de valeurs à tester, vu que l'espace est limité et que l'une des valeurs du vecteur  $\omega$  sera liée aux autres (puisque le vecteur est unitaire). Le calcul du SNR est fait soit en considérant une fenêtre fréquentielle fixe comme étant le signal, soit en considérant une fenêtre fréquentielle mobile (centrée sur un pic dans la FFT qui peut varier selon la combinaison). En considérant la fenêtre comme fixe, on peut par exemple considérer toutes les fréquences physiologiques comme le signal et considérer

les autres fréquences comme du bruit. Cette première méthode sera très peu précise et donc moins intéressante. Une autre façon de positionner la fenêtre est de la centrer sur le pic dans la FFT de la vérité terrain. Enfin, une troisième possibilité est d'utiliser une fenêtre centrée sur le pic de la FFT du signal combiné  $C(t) \cdot \omega^T$ . L'avantage de l'utilisation de la vérité terrain est de calculer le SNR sur la fréquence respiratoire réelle d'une part, et de limiter le risque de changement brusque du SNR d'autre part (voir figure 3.1).

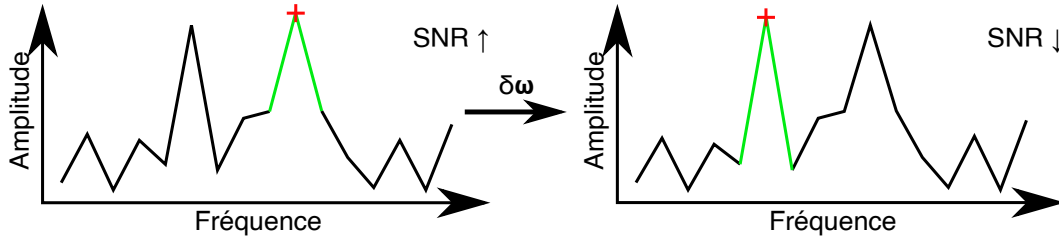


FIGURE 3.1 – Exemple de variation forte du SNR à partir d'une petite variation du vecteur de combinaison notée  $\delta\omega$  : le pic maximal est modifié par le changement ce qui change fortement la valeur du SNR.

Sur la figure 3.1, on voit que lorsque le pic respiratoire n'est pas clairement défini (à cause de la présence de bruit), plusieurs pics peuvent être proches du maximum et être des candidats valables. Dans le cas de la détection implémentée, seul le pic le plus haut est considéré comme pic signal. Un léger changement dans le vecteur de combinaison ( $\omega + \delta\omega$ ) va modifier la transformée de Fourier du signal combiné. Dans certaines situations, ce changement va modifier le pic considéré comme pic signal (puisqu'il y avait plusieurs candidats) et donc modifier brutalement le SNR. Ce changement introduit des irrégularités dans la sphère SNR qui ne représente alors plus le cas général mais un cas particulier associé à ces signaux RGB en particulier. Un exemple de sphère SNR est présenté dans la figure 3.2. On remarque sur celle-ci la présence d'irrégularités qui forment des zones bien définies de jaune, de rouge et de rouge foncé.

Afin de corriger ce problème, il est possible d'utiliser la vérité terrain pour déterminer la fréquence signal comme indiqué précédemment. L'avantage de ce choix est que la vérité terrain est plus fiable que les signaux couleur d'une part, et que celle-ci ne dépend pas du vecteur de combinaison d'autre part. L'utilisation de la vérité terrain permet de ne pas être soumis à ces irrégularités dans la sphère SNR et donne une image plus juste de la sphère SNR. Néanmoins, cette façon de faire est impossible dans un cas réel puisque la vérité terrain est une information inconnue dans ce cas. L'utilisation de cette méthode pour optimiser le SNR passe donc forcément par le calcul du pic sur les signaux combinés dans le cadre des tests et des applications. La fenêtre fixe, basée sur la vérité terrain, sert principalement à faire une validation de la méthode d'une part, et de voir si la structure de la sphère SNR présente des motifs particuliers qui pourront être exploités par d'autres méthodes.

Cette première version limite légèrement les irrégularités présents dans la sphère SNR. Il reste néanmoins d'autres problèmes qui peuvent éventuellement apparaître. En effet, la fenêtre rectangulaire utilisée sur la transformée de Fourier pour séparer le signal du bruit introduit une distinction stricte entre les fréquences du signal et les fréquences du bruit. Cette limite stricte n'est pas représentative de la réalité puisque l'énergie associée à la respiration se trouve principalement au centre de la fenêtre. Les bords de la fenêtre contiennent une proportion de signal plus faible, et devraient donc être moins pris en compte dans le calcul de l'énergie du signal. L'utilisation d'une fenêtre rectangulaire va également provoquer des variations brusque du SNR lors d'un déplacement de celle-ci. Si le pic signal se décale légèrement entre deux projections, il est possible que le SNR varie fortement dans le cas où il y aurait un pic présent au niveau du bord de la fenêtre. Idéalement la fenêtre devrait donc appliquer une pondération plus forte au centre de la fenêtre et plus faible en ses bords.

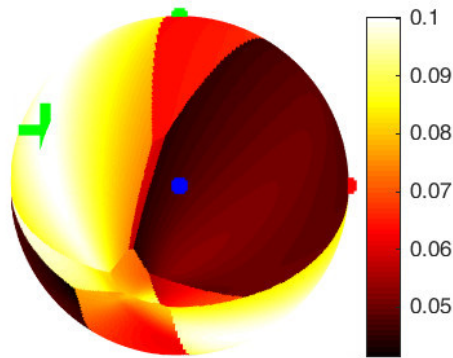


FIGURE 3.2 – Exemple de sphère présentant des irrégularités dues aux changements dans le pic signal.

Afin d'améliorer l'algorithme existant, la fenêtre est donc modifiée pour obtenir une fenêtre en cloche plutôt qu'une fenêtre rectangulaire. Cette modification supprime les variations provoquées par la sortie d'un pic de la fenêtre signal, comme montré dans la figure 3.3. En utilisant une fenêtre en cloche, telle qu'une gaussienne, le décalage d'un pic modifie sa contribution à l'énergie du signal ou du bruit de manière plus régulière. Sur la figure les tracés de gauche montrent une première transformée obtenue sur un vecteur de projection donné et les tracés de droite montrent une transformée obtenue sur un vecteur de projection légèrement différent. Le changement de vecteur provoque un léger décalage en fréquence du pic à gauche de la fenêtre. Dans le premier cas, ce pic va être considéré comme du signal, tandis que dans le deuxième cas le pic sort de la fenêtre et est donc considéré comme du bruit. Dans le cas de la fenêtre rectangulaire, ce pic contribue soit complètement au signal, soit complètement au bruit. Dans le cas de la fenêtre gaussienne, ce pic contribue un peu au signal et un peu au bruit. Plus la fréquence considérée est proche du bord et plus celle-ci contribue au bruit. De ce fait, un décalage du pic va simplement diminuer un peu le SNR mais de manière beaucoup plus lisse.

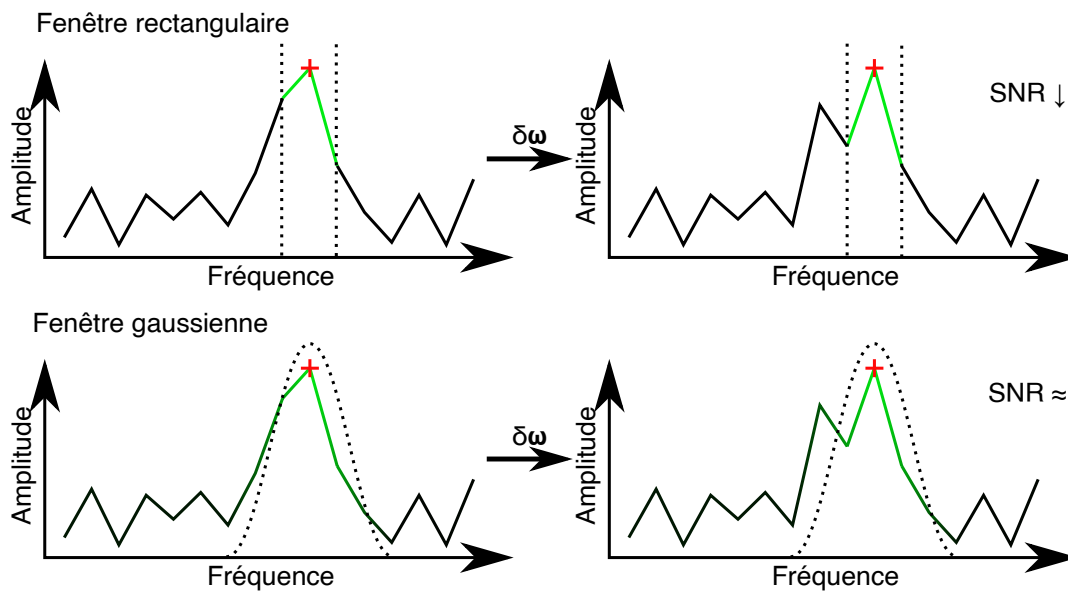


FIGURE 3.3 – Exemple de variation du SNR avec une fenêtre rectangulaire et une fenêtre gaussienne.

Il existe enfin des cas dans lesquels la manière de calculer le SNR n'est pas suffisamment représentative de la qualité du signal. En considérant un signal combiné, le SNR est actuellement calculé en appliquant une fenêtre sur la transformée pour obtenir l'énergie du signal, et en prenant le reste pour obtenir l'énergie du bruit. La fenêtre est centrée sur un pic présent dans la transformée de Fourier. Comme le SNR prend en compte une fenêtre et non pas la valeur de ce pic seul, cette solution n'est pas idéale et peut amener à déterminer une fréquence autre que celle avec le SNR le plus élevé (notamment lorsque le voisinage de ce pic a une amplitude faible). Pour choisir la position de la fenêtre qui va donner le meilleur SNR, le pic signal est recherché sur la convolution de la FFT avec la fenêtre plutôt que sur la FFT seule. Cela garantit d'une part de supprimer encore quelques irrégularités, mais également d'avoir le SNR le plus élevé possible sur la FFT. Cette dernière amélioration est représentée sur la figure 3.4.

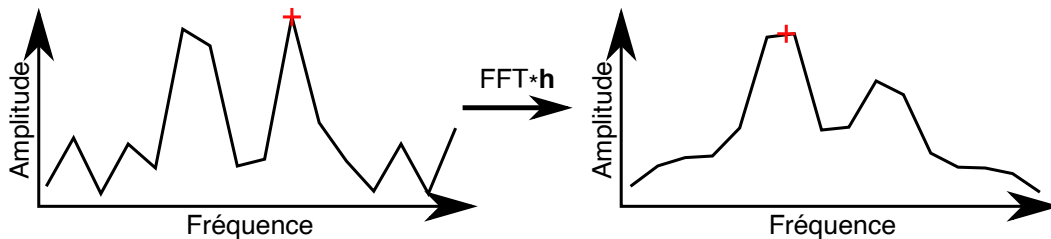


FIGURE 3.4 – Estimation de la fréquence du signal en se basant sur la convolution par  $h(f)$  plutôt que sur la FFT seule.

Sur la figure 3.4 on voit que sans la convolution on peut sélectionner un pic sans considérer son voisinage. Dans le cas de la convolution, le maximum va prendre en compte le voisinage et va donc être plus robuste à des pics isolés. Puisque le numérateur du SNR est calculé à l'aide de la même fenêtre, prendre le maximum sur la convolution de la transformée garantit d'avoir la fenêtre donnant le meilleur SNR. Ce point là est d'autant plus important que la fréquence respiratoire peut varier dans le temps et que lorsqu'on utilise une FFT sur une durée réduite il y a de fortes chances qu'une fréquence respiratoire se retrouve dispersée sur plusieurs points de la FFT (du fait de la diminution de la résolution fréquentielle).

Pour terminer, le SNR peut considérer le bruit comme l'opposé de la fenêtre signal sur l'ensemble des fréquences physiologiques ou sur l'ensemble des fréquences existantes. Le premier cas va prendre plus de sens puisqu'il est toujours possible de filtrer les fréquences restantes. Cependant, dans le cas où on ne souhaiterait pas ou peu filtrer, l'utilisation de la transformée entière est plus intéressante. Cette étude a été réalisée pour voir si la théorie proposée dans [151] peut être affinée. Dans cet article, la combinaison idéale est calculée en utilisant la bande de fréquences cardiaques. Cette combinaison est ensuite réutilisée pour extraire le signal respiratoire, en considérant que celui-ci se trouve globalement sur la même projection que le signal cardiaque. Cette approximation est suffisante pour extraire un signal respiratoire. Néanmoins, il est peut-être possible d'améliorer la qualité de la combinaison en la choisissant spécifiquement pour le signal respiratoire. Dans ce cas, on peut considérer le signal cardiaque comme du bruit de façon à améliorer la séparation entre les deux combinaisons.

En considérant une sphère SNR qui utilise une fenêtre gaussienne centrée sur la fréquence respiratoire obtenue via la vérité terrain, on obtient la figure présente à la figure 3.5. À la lecture de celle-ci, on remarque une structure en couronne aux environs du plan orthogonal au vecteur  $(1 \ 1 \ 1)$ . Cette couronne montre un passage rapide d'un SNR maximal à un SNR minimal. De plus, on remarque que les zones de SNR maximal pour le rythme respiratoire

et cardiaque ne sont pas identiques. Sur la figure 3.5, la différence d'angle entre les deux vecteurs  $\omega_c$  et  $\omega_r$  définis dans les équations (3.13) et (3.14) est de  $46,7^\circ$  :

$$\omega_c = \operatorname{argmax}_{\omega \in \mathbf{V}} SNR_c\{\omega\} \quad (3.13)$$

et

$$\omega_r = \operatorname{argmax}_{\omega \in \mathbf{V}} SNR_r\{\omega\}. \quad (3.14)$$

Dans ces équations,  $\mathbf{V}$  représente l'ensemble des vecteurs unitaires de  $\mathbb{R}^3$ , c'est-à-dire l'espace des solutions recherchées.  $SNR_c$  et  $SNR_r$  représentent respectivement les SNR calculés sur des fréquences cardiaques et respiratoires.

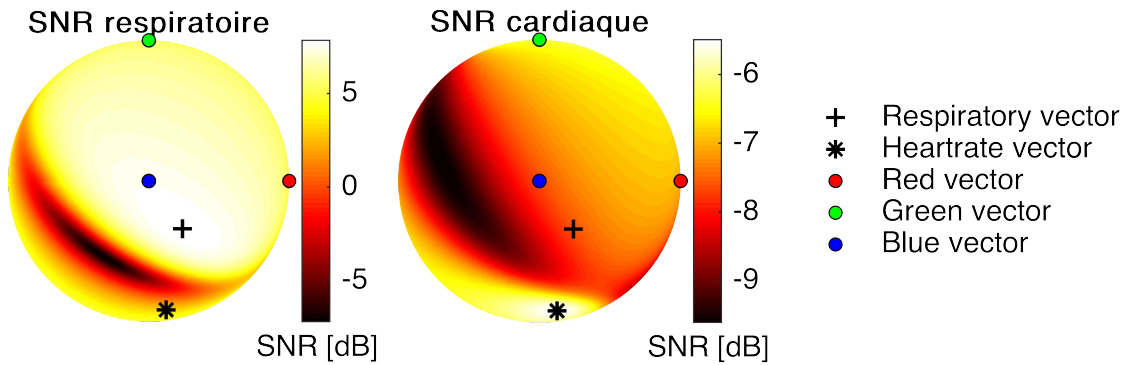


FIGURE 3.5 – Exemple de sphères SNR calculées sur une fréquence respiratoire (gauche) ou cardiaque (droite).

Les sphères SNR montrent également que l'essentiel du signal cardiaque se trouve sur la composante verte, tandis que le signal respiratoire se trouve sur la composante bleue. Cela reste cohérent avec la littérature et notamment [152] qui indiquait que « le signal cardiaque et le signal respiratoire sont plus présents respectivement dans les canaux vert et bleu ». Dans les algorithmes de l'état de l'art comme CHROM [31], on cherche généralement à trouver une combinaison orthogonale à  $(1 \ 1 \ 1)$  car ce vecteur (après normalisation) porte la composante spéculaire et donc le bruit associé au mouvement. Ce type de méthode est pertinent lorsqu'on cherche à récupérer un signal cardiaque (qui provoque peu de variations du spéculaire puisque la peau bouge peu à cause des battements cardiaques). Le signal respiratoire est généralement dérivé en réutilisant une méthode de combinaison dédiée au signal cardiaque [151], cependant d'après l'étude présentée il semble que cette approximation diminue légèrement la qualité du signal respiratoire en sortie, et donc que l'on peut améliorer la mesure en calculant un vecteur respiratoire avec des méthodes dédiées.

Un exemple d'extraction du signal respiratoire est présenté dans la figure 3.6. Les signaux proviennent de la même vidéo et ont tous les deux été combinés avec un vecteur qui maximise le SNR. Dans le premier cas, les vecteurs de combinaison est choisi en maximisant le SNR dans les fréquences cardiaques. Dans le deuxième cas, les vecteurs de combinaison sont calculés de façon à maximiser directement le signal respiratoire. On remarque que dans le premier cas, une partie du signal est de très mauvaise qualité et il n'est pas possible d'estimer le rythme respiratoire dessus.

L'exemple présenté sur la figure 3.5 a été réalisé sur une vidéo spécifiquement enregistrée pour l'étude avec un rythme respiratoire d'environ 0,1 Hz. La structure de la sphère reste similaire dans la majorité des exemples analysés à condition que les signaux ne soient pas trop bruités. Dans le cas où les signaux seraient trop bruités on obtient une sphère moins régulière et plus difficile à analyser.

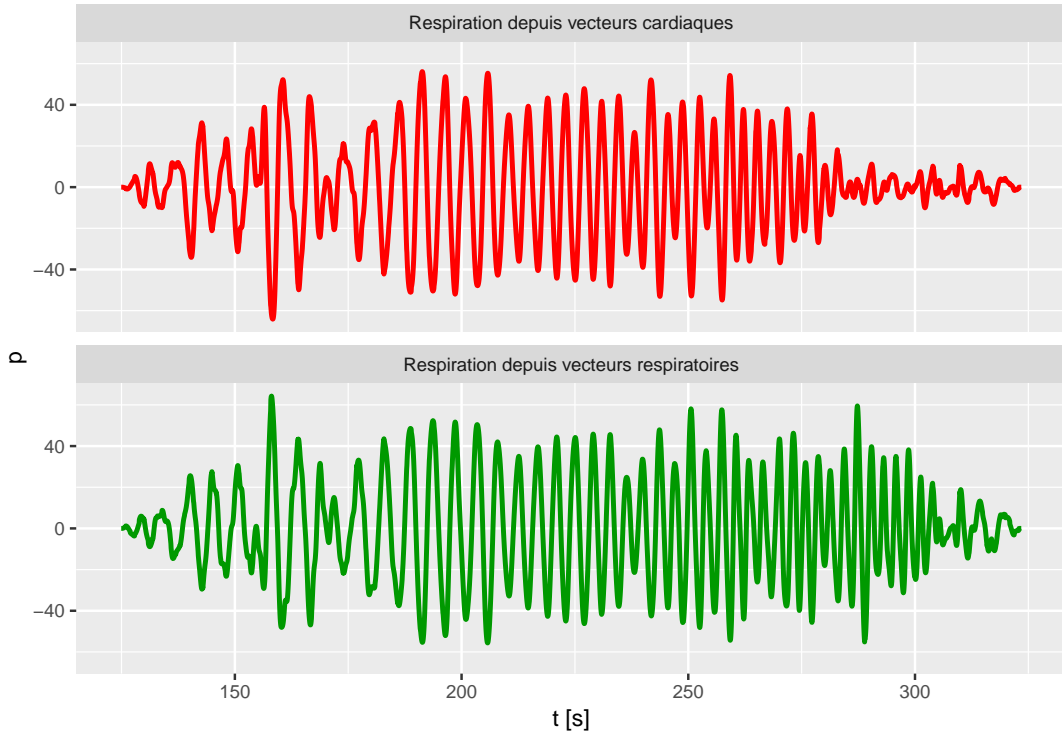


FIGURE 3.6 – Exemples de signaux respiratoires combinés afin de maximiser le SNR cardiaque ou respiratoire.

## 2 Méthode de combinaison des signaux couleur

À partir des résultats obtenus précédemment, une première méthode développée pour l'estimation du rythme respiratoire consiste à calculer l'ensemble des SNR possibles et de garder la combinaison présentant le meilleur résultat. Étant donné que le SNR est calculé sur la transformée de Fourier du signal, celui-ci est découpé en segments avec chevauchement pour pouvoir effectuer l'analyse sur ces segments. Chaque segment est analysé et combiné et le résultat est ensuite reconstruit via une méthode d'*overlap add*. Cette méthode est présentée dans la figure 3.7. Les signaux sont analysés à l'aide d'une fenêtre glissante qui découpe le signal en segments se chevauchant. La fenêtre glissante permet de tenir compte des variations de la fréquence respiratoire d'une part, et des variations du vecteur de combinaison optimal d'autre part. Pour chacun de ces segments, la combinaison est effectuée de façon à obtenir le signal 1D. Cette méthode donne donc en sortie un ensemble de segments combinés. La méthode d'*overlap add*, détaillée notamment dans [31] permet de reconstituer un seul signal 1D à partir de l'ensemble des fenêtres. La procédure est représentée dans la figure 3.8. Elle consiste à fenêtrer les différents segments à l'aide d'une fenêtre (par exemple gaussienne) puis à les additionner comme présenté dans l'équation (3.15) :

$$s(t) = \sum_n \begin{cases} \mathbf{h}(t - \tau \times n) \times \psi_n(t - \tau \times n) & \text{si } t \in \llbracket \tau \times n; N + \tau \times n \rrbracket \\ 0 & \text{sinon} \end{cases} \quad (3.15)$$

avec  $\mathbf{s}(t)$  le signal rPPG après combinaison et *overlap add*,  $\psi_n(t)$  le  $n$ -ième segment à l'instant  $t$ ,  $\mathbf{h}(t)$  une fonction de fenêtrage temporelle,  $\tau$  le décalage entre deux segments et  $N$  la taille d'un segment. Afin d'aligner correctement les différents segments, un retard de  $\tau \times t$  est ajouté dans l'équation. Le but est de prendre en compte les variations du vecteur idéal et

de la fréquence respiratoire tout en ayant une taille de fenêtre suffisante pour avoir une résolution fréquentielle élevée dans la FFT. Cette méthode ne fonctionne pas directement avec la méthode Bruteforce puisque les segments ne sont pas nécessairement en phase. Dans le cas de la recombinaison des segments obtenus à l'aide de la méthode de combinaison Bruteforce, les signaux obtenus peuvent être en opposition de phase d'une fenêtre à l'autre. Dans ce cas, la somme des deux fenêtres provoque une interférence destructive qui donne un signal de très mauvaise qualité (comme on peut le voir sur la ligne supérieure de la figure 3.10).

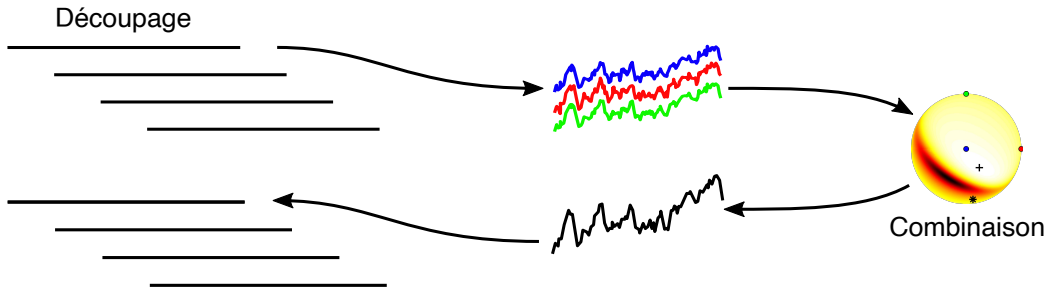


FIGURE 3.7 – Présentation de la combinaison par segments avec une méthode d'*overlap add*.

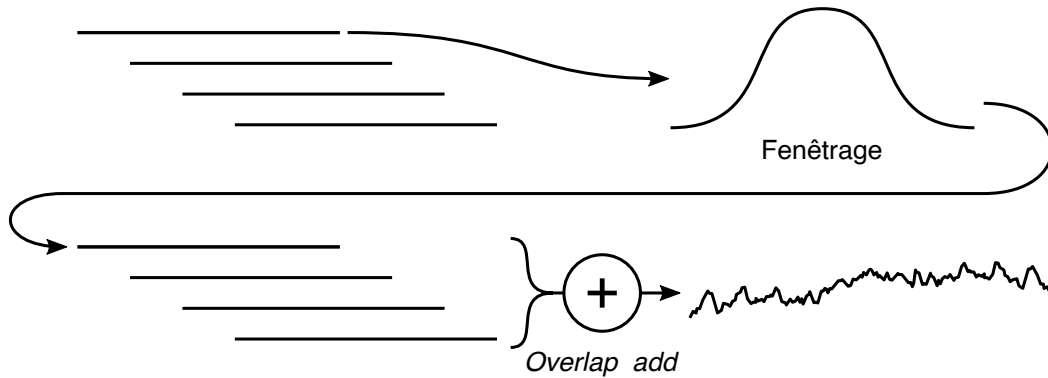


FIGURE 3.8 – L'*overlap add*.

Afin de corriger ce problème, plusieurs méthodes ont été testées. La première consiste à estimer un « plan de retournement ». L'idée est de considérer un plan dans l'espace des combinaisons qui représente le retournement du signal respiratoire. Ce plan est défini par un vecteur normal  $\mathbf{v}_n$  qui indique le sens de l'interférence. Ainsi, si deux fenêtres adjacentes et numérotées  $n$  et  $n + 1$  ont des combinaisons faites avec des vecteurs de projection  $\boldsymbol{\omega}(n)$  et  $\boldsymbol{\omega}(n + 1)$  tel que  $\boldsymbol{\omega}(n) \wedge \mathbf{v}_n$  est du même signe que  $\boldsymbol{\omega}(n + 1) \wedge \mathbf{v}_n$ , alors l'interférence est constructive. Ce cas correspond à celui avec les deux vecteurs de projection du même côté du plan de retournement. Dans l'autre cas, si les signes sont différents alors l'interférence est destructive. Cette méthode nécessite d'estimer le plan de retournement. Celui-ci correspond grossièrement à la « couronne » présentée sur la figure 3.5. Ce calcul nécessite de déterminer le plan de retournement en fonction de la sphère SNR et ne garantit donc pas que le plan soit suffisamment stable, notamment lorsque le signal est fortement bruité.

Afin de simplifier cette estimation, une autre possibilité consiste à estimer le vecteur normal  $\mathbf{v}_n$  par le biais d'une analyse en composantes principales. Cette analyse suppose que la structure de la sphère SNR soit similaire à celle présentée dans la figure 3.9. Sur cette coupe de la sphère, on considère que l'essentiel du signal est constitué d'un signal cardiaque, d'un signal respiratoire et de bruit lié au mouvement. Ces signaux sont répartis sur les signaux



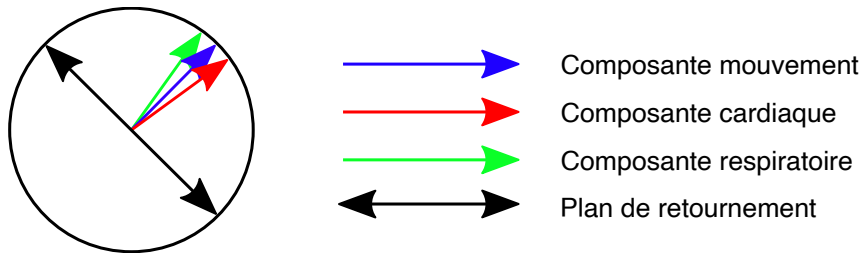


FIGURE 3.9 – Représentation de la sphère SNR.

RGB de façon relativement similaire. En effectuant une PCA sur les signaux RGB, le premier vecteur propre obtenu correspond à un vecteur proche de ces répartitions (puisque le signal cardiaque, respiratoire et le bruit associés au mouvement possèdent une énergie forte). Le plan orthogonal à ce vecteur est donc relativement proche du plan orthogonal au vecteur respiratoire et donc au plan de retournement.

Enfin, une troisième méthode testée consiste à faire la somme et la différence des deux signaux et de calculer le SNR sur chacune des deux combinaisons. Cette méthode est beaucoup plus simple et pragmatique que les deux précédentes. Le but n'est plus de comprendre comment les interférences constructives et destructives prennent place, mais simplement de voir si la somme des deux signaux améliore la qualité par rapport à leur différence. Cette méthode permet donc de choisir simplement s'il est nécessaire ou non de retourner un segment. Le principe de cette dernière méthode est présenté sur la figure 3.10.

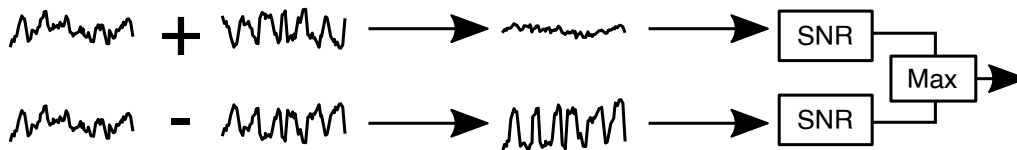


FIGURE 3.10 – Retournement des segments basé sur le SNR.

La comparaison de la méthode Bruteforce avec les autres méthodes sur la base de données du CSEM sera présentée par la suite dans la section 6 du chapitre 5. On remarque principalement que la méthode Bruteforce donne des résultats plus précis que les méthodes de l'état de l'art comme CHROM ou PBV. Du point de vue du temps de calcul, on peut voir que la méthode Bruteforce est extrêmement consommatrice de ressources et nécessite des temps de calcul élevés (environ 2800 fois plus lent que la méthode CHROM). Cela n'est pas un problème puisque cet algorithme n'avait pas pour objectif d'être rapide, néanmoins cette lenteur la rend inutilisable dans des situations réelles.

La méthode Bruteforce représente le point de départ des différentes méthodes qui seront présentées par la suite. Bien que celle-ci donne de très bons résultats dans des scénarios ne présentant pas de mouvements périodiques, elle présente trois problèmes qui devront être améliorés autant que possible.

Tout d'abord, la méthode ne fonctionne pas bien dans des situations où la personne bouge lors de la mesure. Cela est vraisemblablement dû à la détection de fréquence signal qui permet à la méthode Bruteforce de ne pas faire de suppositions particulières sur le modèle d'interaction lumière-tissu mais nécessite donc d'utiliser les données présentes dans les signaux couleur. Cette fréquence est détectée directement sur la transformée de Fourier des signaux ce qui veut dire que lorsque la vidéo contient un mouvement périodique il est possible que cette fréquence soit considérée comme une fréquence signal (du fait de son amplitude et de sa fréquence) et soit donc amplifiée par la combinaison au détriment du signal respiratoire

qui lui sera considéré comme du bruit. Certaines méthodes de l'état de l'art comme PVM ou PCA reposent également sur l'analyse des signaux et ne sont donc pas non plus robustes sur des signaux pollués par du mouvement. Les méthodes basées sur le modèle d'interaction lumière-tissu comme CHROM ou PBV seront en revanche plus robustes aux mouvements. Dans le cas où les signaux seraient bruités, une mauvaise fréquence est détectée et l'algorithme Bruteforce cherche alors à supprimer le signal respiratoire au lieu de le maximiser. Ce problème ne se posait pas dans le cadre de l'analyse préliminaire puisque la fréquence signal était estimée en se basant sur l'analyse de la vérité terrain. Dans le cas de la méthode Bruteforce, on cherche à réaliser un algorithme qui fonctionne sur des situations réelles dans lesquelles la vérité terrain n'existe pas, ce qui nous impose d'utiliser la combinaison des signaux RGB afin de déterminer la fréquence signal. Ce choix se fait néanmoins au détriment de la qualité de l'estimation puisque les signaux RGB peuvent contenir plus de bruit que la vérité terrain (notamment dans le cas où la vidéo contiendrait un mouvement périodique).

Un deuxième problème est la difficulté à ré-assembler les segments obtenus. Plusieurs méthodes pour tenter de ré-assembler les segments ont été évaluées, néanmoins elles nécessitent des calculs supplémentaires et peuvent éventuellement ne pas fonctionner parfaitement et réduire la qualité du signal. Pourtant, le fait de pouvoir ré-assembler les segments pour obtenir un signal unique est intéressant, notamment car cela permet de réduire les données en sortie de l'algorithme mais aussi car il est possible que l'on souhaite utiliser une taille de segment différente pour la combinaison et pour l'estimation des rythmes respiratoires.

Enfin, un troisième problème lié à la méthode Bruteforce est son fort besoin en calculs. La méthode Bruteforce nécessite de très grandes quantités de calculs par rapport aux méthodes habituelles. Pour chaque segment, la méthode calcule trois transformées de Fourier (une par couleur), un grand nombre de projections (dépendant de la résolution recherchée) et pour chacune de ces projections le SNR. Au delà de la qualité du signal en sortie, les performances de l'algorithme et éventuellement la possibilité de faire une implémentation parallélisée représentent deux critères importants pour toute nouvelle méthode. De ce fait, elle est inenvisageable dans une application réelle et tout particulièrement dans le cas où l'on souhaiterait obtenir les estimations du rythme respiratoire en temps réel (ce qui correspond à l'essentiel des cas). La méthode peut néanmoins être parallélisée puisque l'essentiel des calculs revient à calculer une fonction SNR sur un ensemble de combinaisons de très grande taille. Cette parallélisation pourrait améliorer la rapidité de l'algorithme mais nécessite une plateforme dédiée (par exemple un GPU) et donc augmente les coûts des appareils de mesure.

Afin de palier à ces problèmes, la méthode Bruteforce a fait l'objet d'améliorations de façon à pouvoir déterminer la combinaison idéale en utilisant le moins de calculs possible. Pour cela, les méthodes présentées par la suite reposent sur un algorithme nommé GEVD (*Generalized Eigen Value Decomposition*) qui était notamment utilisé par la méthode PVM et qui permet de déterminer le vecteur maximisant un ratio de deux propriétés. De par la nature de la GEVD, le calcul de la projection maximisant le SNR devient alors potentiellement beaucoup plus rapide.



# Chapitre 4

## Méthodes d'extraction de la modulation de la ligne de base

La méthode Bruteforce a permis d'améliorer la qualité du signal en sortie. Néanmoins, celle-ci est particulièrement lente ce qui la rend inutilisable pour mesurer le rythme respiratoire dans des conditions réelles. Afin de limiter le temps de calcul, une méthode de diagonalisation de matrices nommée GEVD est employée afin de remplacer l'étape de calcul de toutes les combinaisons possibles. Pour cela, deux matrices de covariance du signal et du bruit doivent être produites pour représenter la qualité du signal. En calculant ces matrices de différentes manières, plusieurs méthodes de combinaison ont pu être développées et sont présentées ici.

L'approche Bruteforce présentée précédemment permettait de trouver une combinaison linéaire des signaux RGB en reposant sur une recherche exhaustive afin de maximiser une métrique. Dans le cas de la méthode Bruteforce, la métrique maximisée était le SNR calculé sur la transformée de Fourier. Celle-ci était employée notamment afin de récupérer directement le signal respiratoire par le biais d'une de ses modulations : la modulation de la ligne de base (ou *baseline wander*). Les analyses ont été réalisées sur un seul jeu de données et dans une configuration précise où les personnes filmées ne font pas de mouvements parasites. Bien que cette méthode donne de bons résultats sur un scénario avec peu de mouvements, elle n'est pas applicable sur des cas concrets notamment à cause du temps de calcul conséquent. Afin de diminuer le temps de calcul nécessaire, nous allons dans la suite de ce manuscrit proposer plusieurs méthodes basées sur un algorithme nommé GEVD (*Generalized Eigen Value Decomposition*) qui permet de trouver la combinaison linéaire plus rapidement. Pour cela, cet algorithme repose sur la création de deux matrices carrées portant la covariance du signal et du bruit (respectivement nommées  $\mathbf{P}_x$  et  $\mathbf{C}_x$ ). À partir de ces matrices, l'algorithme calcule une matrice de rotation  $\mathbf{\Omega}$  qui va simultanément diagonaliser les matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$ . Cette matrice de rotation  $\mathbf{\Omega}$  est constituée des vecteurs propres généralisés  $\boldsymbol{\omega}_i$  (avec  $i$  l'indice du vecteur propre ou de la ligne de la matrice  $\mathbf{\Omega}^T$ ). Ces vecteurs propres vérifient la propriété suivante :

$$\mathbf{P}_x \cdot \boldsymbol{\omega}_i^T = \lambda_i \times \mathbf{C}_x \cdot \boldsymbol{\omega}_i^T \quad (4.1)$$

avec  $\mathbf{P}_x$  et  $\mathbf{C}_x$  les matrices définies précédemment et  $\lambda_i$  la valeur propre généralisée correspondante au vecteur propre  $\boldsymbol{\omega}_i$ . On peut alors réécrire l'équation (4.1) pour transformer les matrices dans l'équation en scalaires :

$$\boldsymbol{\omega}_i \cdot \mathbf{P}_x \cdot \boldsymbol{\omega}_i^T = \lambda_i \times \boldsymbol{\omega}_i \cdot \mathbf{C}_x \cdot \boldsymbol{\omega}_i^T \Leftrightarrow \lambda_i = \frac{\boldsymbol{\omega}_i \cdot \mathbf{P}_x \cdot \boldsymbol{\omega}_i^T}{\boldsymbol{\omega}_i \cdot \mathbf{C}_x \cdot \boldsymbol{\omega}_i^T}. \quad (4.2)$$

Cette dernière équation peut enfin être remaniée en divisant  $\mathbf{P}_x = \mathbf{X}(t)^T \cdot \mathbf{X}(t)$  et  $\mathbf{C}_x = \mathbf{B}(t)^T \cdot \mathbf{B}(t)$  ( $\mathbf{X}(t)$  et  $\mathbf{B}(t)$  sont respectivement les parties signal et bruit). Ces parties sont liées au signal combiné avec  $\mathbf{x}_{c,i}(t) = \mathbf{X}(t) \cdot \boldsymbol{\omega}_i(t)^T$  et  $\mathbf{b}_{c,i}(t) = \mathbf{B}(t) \cdot \boldsymbol{\omega}_i(t)^T$ . On obtient alors l'équation :

$$\lambda_i = \frac{\mathbf{x}_{c,i}(t)^T \cdot \mathbf{x}_{c,i}(t)}{\mathbf{b}_{c,i}(t)^T \cdot \mathbf{b}_{c,i}(t)}. \quad (4.3)$$

Cette équation correspond globalement à celle du SNR, à l'exception du logarithme qui ne figure pas dedans. Une propriété de la matrice de diagonalisation obtenue par la GEVD va être de contenir le meilleur SNR possible sur une de ses composantes. De ce fait, l'utilisation de la GEVD permet d'obtenir rapidement la projection qui maximise le SNR. La GEVD est un algorithme particulièrement flexible. Selon le choix de la construction de  $\mathbf{P}_x$  et  $\mathbf{C}_x$ , la GEVD permet de maximiser différents ratios de métriques. Ainsi, dans [120], la GEVD est utilisée en construisant des matrices de covariance à l'aide de propriétés comme l'énergie du signal (dans la FFT) ou la périodicité. L'idée générale reste dans tous les cas de filtrer les signaux de départ de façon à mettre en évidence soit le signal soit le bruit, puis à construire les matrices à l'aide des signaux filtrés. L'article [120] présente également un algorithme permettant de filtrer plus finement les signaux en utilisant une suite de décompositions / recompositions. Les signaux subissent une rotation vectorielle à l'aide de la GEVD pour mettre en avant la propriété d'intérêt. Les composantes sont ensuite filtrées selon leur position puis la rotation inverse est effectuée de façon à revenir aux signaux de départ. La décomposition par GEVD va alors être différente puisque les signaux auront été modifiés par le filtrage. Il est ainsi possible de réaliser une succession de décompositions / recompositions pour faire des filtrages aussi fins que possible. Il est également possible de modifier au fur et à mesure la propriété d'intérêt, le filtrage effectué ou les composantes sur lesquelles le filtrage est effectué. Ce traitement est présenté dans la figure 4.1.

Afin de permettre une implémentation simple et en-ligne de l'algorithme, cette partie du traitement n'est pas mise en place par la suite. Dans les sections suivantes, la GEVD est utilisée en modifiant les constructions des matrices de façon à maximiser différentes propriétés.

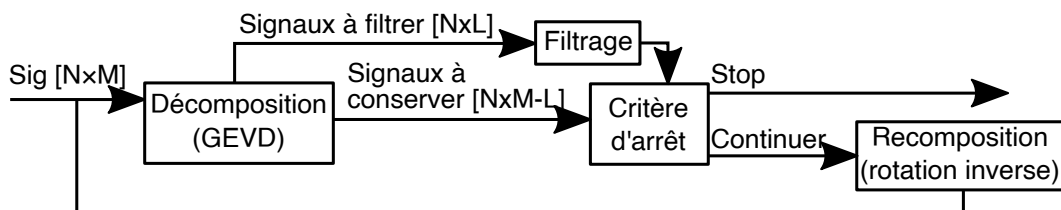


FIGURE 4.1 – Filtrage avancé à l'aide de la GEVD.

Le but final est d'obtenir une mesure qui soit aussi fiable que possible et la moins coûteuse en temps de calcul possible.

En résumant l'algorithme GEVD, on trouve :

$$\frac{\mathcal{E}\{\mathbf{X}(t) \cdot \boldsymbol{\omega}_1^T\}}{\mathcal{E}\{\mathbf{B}(t) \cdot \boldsymbol{\omega}_1^T\}} = \lambda_1 \quad (4.4)$$

avec  $\mathcal{E}\{\mathbf{X}(t)\}$  l'énergie du signal et  $\lambda_1$  la plus grande valeur pouvant être obtenue via une combinaison linéaire. Le vecteur  $\boldsymbol{\omega}_1$  est le vecteur de combinaison recherché et  $\mathbf{X}(t)$  et  $\mathbf{B}(t)$  sont respectivement la portion de signal et de bruit présent dans les signaux couleur. Les différentes variantes présentées par la suite consistant principalement à modifier les choix de  $\mathbf{X}(t)$  et de  $\mathbf{B}(t)$  afin de gagner en rapidité d'exécution ou de mieux séparer le signal respiratoire dans  $\mathbf{X}(t)$  et le bruit dans  $\mathbf{B}(t)$ .

## 1 EVM-FFT

Dans le chapitre précédent, une première méthode de combinaison des signaux RGB a été proposée. Cette méthode teste un grand nombre de combinaisons possibles et ne garde que celle qui maximise le SNR. Bien que cette méthode donne de meilleurs résultats dans le cas où le signal n'est pas trop pollué par le mouvement, elle est extrêmement gourmande en ressources ce qui la rend inutilisable dans la plupart des applications (qui dans la grande majorité des cas nécessite d'extraire les informations physiologiques en temps réel).

Afin de pouvoir prendre en compte la question de la vitesse de calcul, un second algorithme de combinaison a été mis en place. Celui-ci reprend les mêmes principes que la méthode Bruteforce (à savoir déterminer la combinaison qui assure le meilleur SNR), néanmoins, il utilise l'algorithme GEVD. Le but de la méthode EVM-FFT (*Energy Variance Maximization-Fast Fourier Transform*) est donc d'accélérer le calcul de la combinaison tout en gardant les résultats les plus proches possibles de ceux obtenus avec Bruteforce. Cela permet de vérifier que la méthode GEVD permet d'accélérer les calculs tout en gardant un vecteur de combinaison similaire à celui que l'on obtiendrait avec une méthode Bruteforce. Une fois ces points vérifiés il sera possible de modifier la propriété à maximiser afin d'améliorer la mesure ou d'accélérer le calcul de celles-ci.

Dans un premier temps, EVM-FFT doit utiliser GEVD de façon à maximiser le SNR. Comme indiqué dans le début du chapitre, l'algorithme GEVD prend en entrée deux matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$  représentant respectivement la covariance du signal et du bruit. Ces matrices sont diagonalisées par les vecteurs propres généralisés estimés grâce à la GEVD. Ces vecteurs propres peuvent être utilisés pour combiner les signaux RGB pour donner un signal combiné. En calculant la propriété d'intérêt (SNR) sur le signal combiné on obtient une valeur corrélée avec la valeur propre correspondant au vecteur de combinaison. Une propriété clef de la GEVD est que la valeur propre la plus grande correspond au coefficient de Rayleigh de  $\mathbf{P}_x$

et  $\mathbf{C}_x$  qui est la plus grande valeur pouvant être obtenue avec une combinaison linéaire. En résumé, les algorithmes basés sur la GEVD filtrent le signal et le bruit, calculent leurs matrices de covariance  $\mathbf{P}_x$  et  $\mathbf{C}_x$  et les fournit à la méthode GEVD. En retour GEVD donne une matrice de diagonalisation  $\mathbf{\Omega}$  et une matrice de valeurs propres généralisée  $\mathbf{\Lambda}$ . Le vecteur propre de  $\mathbf{\Omega}$  correspondant à la valeur propre la plus grande de  $\mathbf{\Lambda}$  est alors le vecteur de combinaison qui maximise la propriété SNR comme montré dans l'équation (4.2). Les algorithmes de combinaison basés sur la GEVD se différencient donc principalement par la façon de construire  $\mathbf{P}_x$  et  $\mathbf{C}_x$ . La construction de ces matrices peuvent utiliser des méthodes de filtrages qui dépendent du signal à analyser (signal stationnaire, périodique, avec un motif particulier, dans une bande de fréquences donnée...).

Dans le cas d'EVM-FFT, on cherche à maximiser le SNR défini dans l'équation (3.11). Dans cette équation, le signal est défini par  $\int_{f_{min}}^{f_{max}} \mathbf{h}_{f_c}(f) |\mathcal{F}\{\mathbf{C}(t)\}(f)|^2 \delta f$  et le bruit par  $\int_{f_{min}}^{f_{max}} (1 - \mathbf{h}_{f_c}(f)) |\mathcal{F}\{\mathbf{C}(t)\}(f)|^2 \delta f$ . Les matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$  sont alors construites via :

$$\mathbf{P}_x = (\mathcal{F}\{\mathbf{C}(t)\} \odot \mathbf{h}_{f_c}(f))^T \cdot (\mathcal{F}\{\mathbf{C}(t)\} \odot \mathbf{h}_{f_c}(f)) \quad (4.5)$$

et :

$$\mathbf{C}_x = (\mathcal{F}\{\mathbf{C}(t)\} \odot (1 - \mathbf{h}_{f_c}(f)))^T \cdot (\mathcal{F}\{\mathbf{C}(t)\} \odot (1 - \mathbf{h}_{f_c}(f))) \quad (4.6)$$

avec  $\mathcal{F}\{\mathbf{C}(t)\}$  la transformée de Fourier des signaux RGB (de taille  $N \times 3$ ),  $\mathbf{h}_{f_c}(f)$  la fenêtre signal centrée sur la fréquence  $f_c$  qui sert à séparer le signal et le bruit (le signal correspondant aux fréquences autorisées par  $\mathbf{h}_{f_c}(f)$  et le bruit correspondant aux fréquences autorisées par  $1 - \mathbf{h}_{f_c}(f)$ ),  $\odot$  le produit matriciel de Hadamard (élément par élément) et  $\cdot$  la multiplication matricielle. Ces matrices doivent ensuite être rendues symétriques et réelles de façon à obtenir par la suite des valeurs propres et des vecteurs propres réels. Les valeurs propres doivent appartenir à  $\mathbb{R}$  puisqu'ils représentent le SNR et le vecteur propre doit appartenir à  $\mathbb{R}^3$  car on ne souhaite pas déphaser les signaux. Afin de garantir ces propriétés, on somme les matrices avec leurs matrices adjointes :

$$\mathbf{P}_x := \mathbf{P}_x + \mathbf{P}_x^* \quad (4.7)$$

$$\mathbf{C}_x := \mathbf{C}_x + \mathbf{C}_x^* \quad (4.8)$$

avec  $:=$  une opération d'affectation.

De façon similaire à la méthode Bruteforce, le signal est découpé en un ensemble de segments et chaque segment est analysé via une transformée de Fourier. Dans le cas d'une caméra RGB, trois transformées sont obtenues, une par canal. Ces transformées sont ensuite utilisées pour la création des matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$ .

Afin de séparer le signal du bruit, il est nécessaire de déterminer une fréquence signal. La fréquence centrale de la fenêtre du signal est calculée avec une procédure similaire à celle de PVM section 2.4.1.4 du chapitre 2. Une succession de GEVD est effectuée en prenant plusieurs fréquences candidates et la fréquence avec le meilleur SNR, notée  $f_c$ , est conservée. Le signal est alors obtenu en ne gardant que les fréquences proches de cette fréquence centrale à l'aide d'une fenêtre gaussienne. Le bruit correspond de son côté aux autres fréquences.

Dans le cas présent, la propriété mise en valeur par la GEVD va être le ratio de l'énergie du signal dans la fenêtre  $\mathbf{h}_{f_c}(f)$  sur l'énergie du bruit hors de la fenêtre  $\mathbf{h}_{f_c}(f)$ . On a donc dans ce cas une définition de cette propriété assez proche de celle du SNR. La principale différence est l'absence de logarithme, néanmoins les maximums restent les mêmes dans les deux cas. Les différents traitements effectués dans cet algorithme sont présentés sur la figure 4.2.

La méthode EVM-FFT est comparée aux autres méthodes du point de vue de la précision des mesures et du temps de calcul. Les résultats et les graphiques obtenus et présentés dans la section 6 du chapitre 5 sont très proches de ceux obtenus précédemment via la méthode

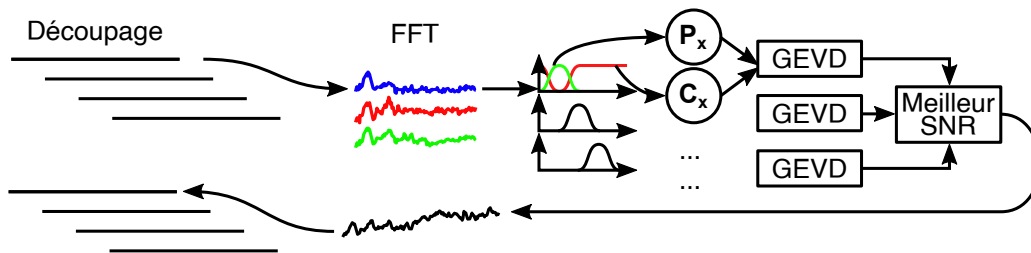


FIGURE 4.2 – Traitements utilisés par la méthode EVM-FFT.

Bruteforce. Cela était attendu puisque le but de la méthode était avant tout d'accélérer les traitements tout en gardant la même analyse. Les petites variations dans les rythmes obtenus sont probablement dûs à des différences dans l'implémentation, ou à l'utilisation de GEVD qui permet de s'affranchir de la question de la résolution à utiliser pour la sphère SNR. Au niveau des temps de calcul, on remarque que la méthode EVM-FFT est plus rapide que la méthode Bruteforce avec un gain de vitesse d'environ 30%. Bien que cela soit encore très insuffisant pour concurrencer les méthodes de l'état de l'art, il est clair que l'utilisation de GEVD permet d'accélérer les calculs de la combinaison idéale. La méthode peut être encore accélérée par la suite en jouant notamment sur la façon dont  $P_x$  et  $C_x$  sont déterminées.

La méthode EVM-FFT présentée permet de déterminer la combinaison permettant d'obtenir le meilleur SNR sans effectuer le même nombre de calculs que précédemment avec l'approche Bruteforce. Cette méthode présente l'inconvénient de faire une analyse de Fourier et plusieurs GEVD sur chaque segment. D'une part cela coûte un temps de calcul assez élevé, mais les différents segments ne sont pas aisément combinables entre eux (via une méthode d'*overlap add* présentée sur la figure 3.8 [31]). Le fait de réaliser l'analyse par segment empêche également de tirer parti des analyses précédentes, et augmente le risque d'avoir de grandes variations de fréquences entre deux segments adjacents lorsque ceux-ci sont trop bruités.

La méthode EVM-FFT est très proche de la méthode Bruteforce d'un point de vue algorithmique. Dans les deux cas, les méthodes cherchent une combinaison qui maximise le SNR du point de vue de la transformée de Fourier (équation (3.11)). Le but de la méthode EVM-FFT est surtout de limiter le nombre de calculs et de tester la possibilité d'utiliser la méthode GEVD pour obtenir la combinaison maximisant le SNR.

## 2 EVM-filtre

La méthode EVM-filtre [4] est un raffinement de la méthode EVM-FFT. Elle remplace la transformée de Fourier par une séparation du signal et du bruit en utilisant deux filtres de Butterworth. Le premier filtre correspondant au bruit va ne retenir que la bande de fréquences correspondant aux fréquences respiratoires possibles (c'est-à-dire de 0,1 à 0,7 Hz). Le second filtre, correspondant au signal, va également conserver les fréquences respiratoires mais avec une largeur de fenêtre plus fine (de  $\pm 0,1$  Hz). Ce filtre est placé au niveau d'une fréquence qui est considéré comme la fréquence respiratoire et est mis à jour régulièrement en utilisant une fréquence détectée sur le signal combiné. Les traitements effectués par l'algorithme sont présentés sur la figure 4.3.

L'essentiel de ce traitement reste similaire à celui utilisé pour la méthode EVM-FFT. Cependant, plusieurs modifications sont mises en place pour améliorer la qualité de la mesure ou sa rapidité. Tout d'abord, le fait d'employer des filtres va permettre d'effectuer la séparation du signal et du bruit sans utiliser de transformée de Fourier et donc sans introduire de délais. La méthode EVM-FFT de son côté nécessitait la récupération d'un segment complet avant



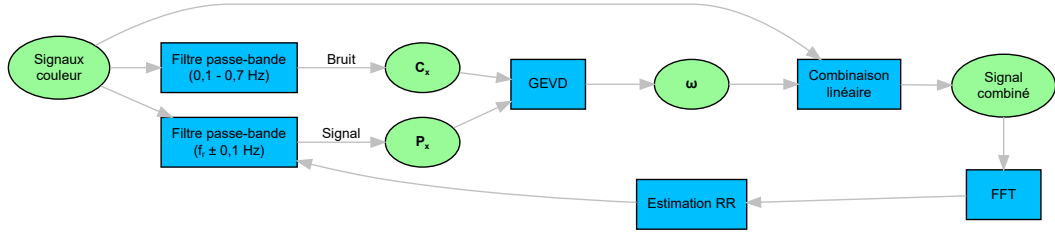


FIGURE 4.3 – Étapes de traitement de la méthode EVM-filtre.

de lancer une transformée de Fourier. Les échantillons en sortie de filtres sont utilisés de façon à construire les matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$  qui elles-mêmes seront employées comme entrées de l'algorithme GEVD. Pour terminer, on obtient grâce à la GEVD le vecteur de projection  $\boldsymbol{\omega}$  de taille  $[1 \times 3]$  qui permet de maximiser le signal respiratoire. Ce vecteur permet d'obtenir le signal de la meilleure qualité possible via une combinaison linéaire, puisqu'on considère que le signal respiratoire est réparti sur les signaux couleur de façon linéaire (comme présenté à l'équation (2.10)).

Les matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$  sont mises à jour pour chaque nouvel échantillon. Pour cela, elles sont représentées de façon récursive :

$$\mathbf{P}_x(t+1) = \gamma \mathbf{P}_x(t) + \mathbf{X}(t+1) \cdot \mathbf{X}(t+1)^T \quad (4.9)$$

et :

$$\mathbf{C}_x(t+1) = \beta \mathbf{C}_x(t) + \mathbf{B}(t+1) \cdot \mathbf{B}(t+1)^T \quad (4.10)$$

en notant  $\mathbf{X}(t)$  et  $\mathbf{B}(t)$  les échantillons de signal et de bruit en sortie de filtres à l'instant  $t$  (de tailles  $[1 \times 3]$ ),  $\gamma$  et  $\beta$  des facteurs d'oubli ( $\gamma \in [0, 1]$ ,  $\beta \in [0, 1]$ ). Ces facteurs vont permettre de remplacer le système de segments précédemment utilisé. Les échantillons les plus anciens sont multipliés de nombreuses fois par ces facteurs d'oubli et vont donc devenir négligeables par rapport aux nouveaux échantillons. Avec cette formule, on peut considérer que la matrice est l'approximation d'une matrice construite sur un segment. Le choix du facteur va permettre de choisir la taille du pseudo segment recherchée, en gardant une valeur proche de 1 le segment sera très grand et pourra prendre en compte des fréquences basses. En prenant des valeurs plus faibles, la matrice deviendra moins stable et moins utile pour capter les informations utiles.

Le fait d'écrire les matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$  en utilisant une formule récursive permet de la mettre à jour pour chaque nouvel échantillon. Là encore, cette implémentation peut être faite en-ligne et est donc moins consommatrice de ressources. Un autre intérêt non négligeable de cette écriture est la possibilité d'utiliser une version modifiée de la GEVD pour économiser du temps de calcul. En effet, comme présenté dans [42], l'algorithme GEVD peut être employé de façon à déterminer un vecteur de projection et à l'adapter pour chaque nouvel échantillon. Pour cela, les matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$  doivent être construites en utilisant une formule similaire à celle des équations (4.9) et (4.10). La matrice de rotation  $\boldsymbol{\Omega}$  déterminée à l'aide de la GEVD est mise à jour pour chaque nouvel échantillon via :

$$\boldsymbol{\omega}_i(t+1) = \frac{\boldsymbol{\omega}_i(t) \cdot \mathbf{C}_x(t+1) \cdot \boldsymbol{\omega}_i(t)^T}{\boldsymbol{\omega}_i(t) \cdot \mathbf{P}_x(t+1) \cdot \boldsymbol{\omega}_i(t)^T} \times \mathbf{C}_x^{-1}(t+1) \cdot \mathbf{P}_x(t+1) \cdot \boldsymbol{\omega}_i(t)^T. \quad (4.11)$$

L'équation (4.11) nécessite de connaître la matrice inverse  $\mathbf{C}_x^{-1}$ . De manière générale il est coûteux de calculer l'inverse d'une matrice directement. Cependant, dans le cas des matrices définies par les équations (4.9) et (4.10), il est possible d'utiliser le lemme d'inversion de Woodbury de façon à calculer plus rapidement l'inverse de la matrice. Celui-ci définit la formule suivante :

$$(\mathbf{A} + \mathbf{U} \cdot \mathbf{C} \cdot \mathbf{V})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \cdot \mathbf{U} \cdot (\mathbf{C}^{-1} + \mathbf{V} \cdot \mathbf{A}^{-1} \cdot \mathbf{U})^{-1} \cdot \mathbf{V} \cdot \mathbf{A}^{-1} \quad (4.12)$$

avec  $\mathbf{A}$ ,  $\mathbf{U}$ ,  $\mathbf{C}$ ,  $\mathbf{V}$  des matrices de tailles  $[n \times n]$ ,  $[n \times k]$ ,  $[k \times k]$ ,  $[k \times n]$  respectivement. Dans le cas de la rPPG sur une image RGB,  $\mathbf{A}$ ,  $\mathbf{U}$  et  $\mathbf{V}$  ont des tailles respectives de  $[3 \times 3]$ ,  $[3 \times 1]$  et  $[1 \times 3]$ , et  $\mathbf{C} = 1$ . Les temps de calcul nécessaires pour inverser une matrice  $[3 \times 3]$  directement ou avec le lemme de Woodbury semblent relativement similaires. Néanmoins, la possibilité d'implémenter ce lemme permet de gagner du temps lors de l'utilisation de matrices de plus grandes dimensions, et donc de pouvoir inclure dans le calcul de la combinaison d'autres données, par exemple lors de l'analyse de vidéos multispectrales ou pour combiner les informations présentes dans des superpixels.

Pour terminer, le filtre du signal nécessite une mise à jour de sa fréquence afin de pouvoir s'adapter aux changements de rythme respiratoire. Pour cela, la nouvelle fréquence est détectée à intervalles réguliers puis les coefficients du filtre sont modifiés en fonction de cette nouvelle fréquence. Les valeurs internes au filtre sont conservées pour assurer une modification plus fluide du filtrage. Cette méthode de filtrage simple suppose que les modifications de la fréquence respiratoire sont relativement faibles pour éviter de modifier trop fortement les coefficients du filtre. Afin de détecter la fréquence respiratoire, de nombreuses méthodes peuvent être mises en place.

Dans le cas du travail effectué, la nouvelle fréquence respiratoire est détectée en utilisant une transformée de Fourier sur le signal en sortie avec une taille de fenêtre de 10 secondes et un décalage de 5 secondes entre les fenêtres. On réalise de cette façon 15 fois moins de transformées de Fourier par rapport à la méthode EVM-FFT avec une transformée par canal et un décalage d'une seconde. De plus la taille de la fenêtre est réduite ce qui permet encore de réduire la durée du calcul. Cette diminution du temps de calcul s'obtient néanmoins au détriment de la qualité du suivi de la fréquence, 10 secondes ne représentant dans le pire des cas qu'une seule respiration complète à 0,1 Hz. Ce problème ne paraît néanmoins pas insurmontable : la fréquence utilisée pour adapter le filtre correspond à celle qui est recherchée en sortie du système. De ce fait, on peut considérer l'utilisation d'algorithmes plus avancés pour la détection de la fréquence vu que celle-ci sera juste « recyclée » par la suite.

Une autre possibilité plus complexe consisterait à effectuer plusieurs GEVD en parallèle afin de déterminer celle qui présente la meilleure qualité. Cela revient au système présenté dans la section 1 du chapitre 4. Du fait de l'implémentation en-ligne de l'algorithme, il est envisageable que les calculs se fassent tout le long de l'exécution de l'algorithme. La mise à jour de la fréquence et du signal en sortie pourrait donc être faite au fur et à mesure des changements de la qualité du signal en sortie des GEVD. Cette implémentation est en revanche plus complexe et nécessite probablement d'être implémentée sur une architecture dédiée afin de paralléliser complètement les différentes GEVD. Un exemple de cette chaîne de traitement est présentée dans la figure 4.4. La stabilité du filtre de Butterworth est également un point qui nécessite une validation. En effet, la méthode de la mise à jour des coefficients du filtre n'est pas un algorithme habituel, et bien qu'il fonctionne empiriquement, il serait souhaitable en perspective de ce travail de démontrer de manière formelle que cette mise à jour ne peut pas rendre le filtre instable.

Lors de la comparaison de la méthode EVM-filtre avec les autres méthodes (section 6, chapitre 5), on remarque principalement que les estimations des rythmes respiratoires sont moins précises avec la méthode EVM-filtre qu'avec les méthodes Bruteforce et EVM-FFT présentées précédemment. Bien que l'erreur systématique soit plus faible, on remarque que les erreurs présentées sur le graphique de Bland-Altman sont plus éparpillées, ce qui augmente le MAE et le RMSE tout en diminuant la précision. Ces erreurs sont probablement dues au suivi de la fréquence respiratoire qu'il est difficile d'estimer rapidement et précisément du fait de la non-stationnarité du signal respiratoire. En revanche, la méthode EVM-filtre est nettement plus rapide que les méthodes EVM-FFT et Bruteforce ce qui lui permet même de dépasser d'autres méthodes de l'état de l'art comme PVM. La vitesse d'exécution d'EVM-

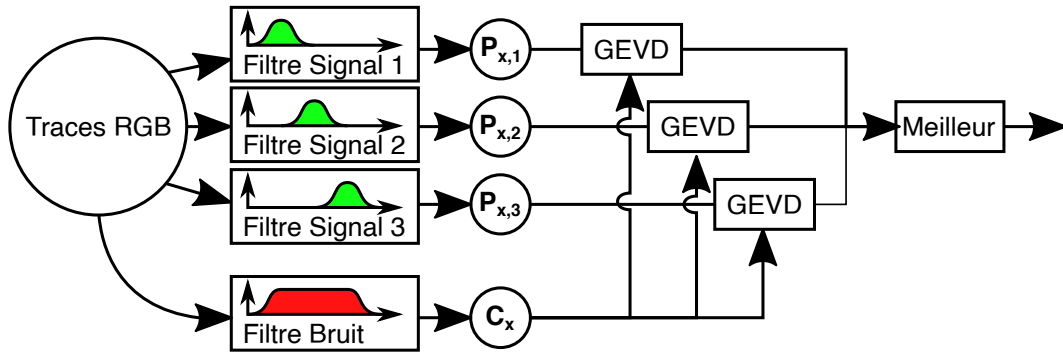


FIGURE 4.4 – Chaîne de traitement EVM-filtre utilisant une cascade de GEVDs.

filtre est proche de CHROM et de PBV ce qui en fait un bon candidat pour pouvoir être utilisée dans des applications réelles.

La méthode EVM-filtre présentée dans cette section est une amélioration des algorithmes étudiés précédemment. Contrairement aux précédents, le temps de calcul nécessaire pour le fonctionnement de l'algorithme le rend utilisable sur des cas réels. De plus, la structure de l'algorithme le rend plus facilement parallélisable sur des cibles matérielles dédiées. Plusieurs points restent à améliorer. Tout d'abord, la méthode nécessite de suivre de manière robuste et précise la fréquence respiratoire. Une erreur dans l'estimation de la fréquence ferait courir le risque d'avoir un filtre qui maximise le bruit et minimise le signal. Par ailleurs, la présence de mouvement périodique dans la vidéo risque fortement de perturber l'estimation de la fréquence respiratoire et donc de limiter la qualité de l'estimation. Les méthodes habituellement présentes dans l'état de l'art cherchent à pouvoir retirer la contribution du mouvement dans le signal sans se baser sur les données d'entrée et seront vraisemblablement plus robustes dans ces situations. Une possibilité d'amélioration pour gérer la présence de mouvement serait soit de sélectionner l'algorithme en fonction du type de vidéo ou de modifier EVM-filtre pour ne fonctionner que sur le plan orthogonal à la composante mouvement (spéculaire). Pour finir l'utilisation du filtre passe-bande ne permet pas d'analyser les signaux non-stationnaires, comme le signal respiratoire, de manière efficace. Afin de pouvoir mieux prendre en compte la non-stationnarité du signal respiratoire, la section suivante présente une nouvelle évolution de l'algorithme dont le but est d'employer une transformée différente de la transformée de Fourier : la transformée en ondelettes.

### 3 WVM

La méthode EVM-filtre présente plusieurs inconvénients qui peuvent limiter son utilisation. Tout d'abord, la fréquence d'intérêt est difficile à estimer. Elle est obtenue dans ce cas par une FFT faite sur un buffer, ce qui pose des problèmes liés à la stationnarité du signal respiratoire d'une part, et au temps de calcul d'autre part. Il paraît important d'utiliser une méthode efficace pour l'analyse des signaux non-stationnaires de façon à pouvoir mieux séparer le signal respiratoire et le bruit.

Afin de corriger ces problèmes, il est possible d'utiliser une méthode d'analyse différente comme par exemple la transformée en ondelettes [54]. Dans le cadre de la plupart des analyses fréquentielles, l'estimation de l'énergie portée par une fréquence se fait en calculant le produit scalaire du signal avec un noyau. Pour la FFT le noyau est une sinusoïde pure dont la durée couvre tout le signal analysé. Cela permet d'avoir une résolution fréquentielle élevée (connaître les fréquences portées dans le signal avec une bonne précision) mais une résolution temporelle basse (connaître où se situent les fréquences dans le signal). Pour améliorer la

situation, les STFT (*Short Term Fourier Transform*) emploient un noyau correspondant à une sinusoïde pure, mais n'existant que sur une durée plus faible. Avec ce noyau, la résolution fréquentielle est plus faible mais la résolution temporelle augmente. Le compromis entre résolution fréquentielle et résolution temporelle se fait en modifiant la durée du noyau. Une autre amélioration de ce système consiste à fenêtrer le noyau à l'aide d'une fenêtre gaussienne, c'est-à-dire utiliser la transformée de Gabor. L'intérêt principal est de supprimer certaines fréquences parasites qui apparaissent à cause des effets de bord de la transformée.

Dans la continuité de ces méthodes, la transformée en ondelettes est apparue. Contrairement aux transformées traditionnelles, les transformées en ondelettes se basent sur un noyau qui n'est plus une sinusoïde pure mais une fonction définie sur une durée courte. Cela permet de mieux analyser les signaux non-stationnaires. De plus, l'ondelette n'est plus de taille fixe mais de taille variable selon la fréquence considérée. Au lieu de garder une taille de noyau fixe (et d'augmenter le nombre de périodes avec la fréquence), la transformée en ondelettes utilise des noyaux de tailles variables et avec un nombre d'oscillations fixe. Le fait d'utiliser un noyau de taille variable permet d'avoir une meilleure résolution fréquentielle sur les basses fréquences (au détriment de la résolution temporelle), et une meilleure résolution temporelle sur les hautes fréquences (au détriment de la résolution fréquentielle). Contrairement aux analyses précédentes, les ondelettes ne sont plus des sinusoïdes et ne se caractérisent plus par leur fréquence mais par leur échelle, l'échelle étant le coefficient d'étirement de l'ondelette. Plus le coefficient est élevé, plus l'ondelette s'étire dans le temps et plus elle représente une fréquence basse. Pour relier l'échelle à la fréquence, on utilise une mesure intermédiaire, la fréquence principale qui est propre à chaque ondelette (voir figure 4.5). Sur cette figure, l'ondelette de type Bump<sup>1</sup> est représentée en rouge, et la sinusoïde correspondant à la fréquence principale de l'ondelette est représentée en gris. La sinusoïde représente grossièrement la fréquence portée par l'ondelette mais il y a des variations dans la fréquence ou dans la forme qui font de cette fréquence une approximation.

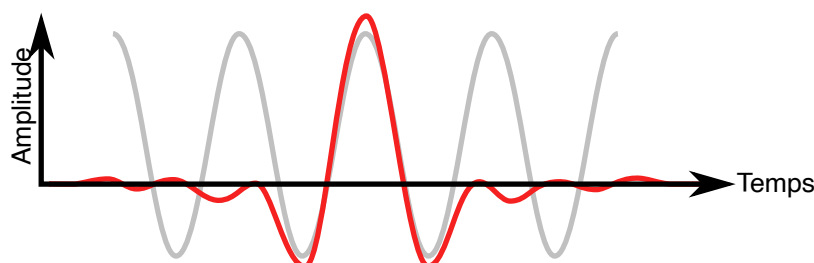


FIGURE 4.5 – Représentation de la fréquence principale sur une ondelette de type Bump.

Le choix de l'ondelette et de la méthode d'analyse peut être fait en tenant compte de nombreux facteurs comme le type de signal analyser, les événements que l'on souhaite mettre en évidence ou les temps de calcul. Du point de vue des ondelettes, on trouve plusieurs catégories d'ondelettes. Les ondelettes analytiques [79] possèdent une composante imaginaire ce qui permet d'obtenir une transformée de Fourier de l'ondelette définie uniquement sur des fréquences positives. Elles permettent d'obtenir des coefficients qui embarquent une information d'amplitude et de phase. Elles peuvent être intéressantes notamment pour l'analyse de signaux modulés en amplitude et en fréquence.

Les ondelettes orthogonales [89] sont des ondelettes qui permettent de conserver la même énergie dans la transformée et dans le signal de départ. Elles peuvent avoir un support compact dans le domaine temporel (faible durée) mais sont généralement non symétriques dans ce cas, à l'exception des ondelettes de Haar [56]. Le fait qu'elles soient non symétriques induit une phase non linéaire et donc une déformation du signal ce qui peut poser des problèmes notamment lorsque les ondelettes sont utilisées pour traiter des images.

1. <https://fr.mathworks.com/help/wavelet/ref/cwtft.html#buu64ch>

Les ondelettes biorthogonales [28] de leur côté sont très utilisées pour la compression. Elles sont symétriques ce qui garantit une phase linéaire et donc sont très utiles pour l'analyse d'images. De plus, elles permettent de réaliser des compressions de signaux puisqu'elles peuvent représenter un signal avec peu de coefficients, notamment lorsqu'elles possèdent un grand nombre d'oscillations.

Un autre point à déterminer pour le choix des ondelettes est leur nombre d'oscillations. En prenant des ondelettes avec peu d'oscillations il est possible de mettre en évidence des variations rapides du signal et d'avoir une bonne résolution temporelle. À l'inverse, une ondelette avec beaucoup d'oscillations permet de mieux représenter des phénomènes oscillatoires (transients) avec une bonne résolution fréquentielle. Cependant le nombre plus élevé d'oscillations implique d'avoir une ondelette plus large à une fréquence principale donnée et donc une résolution temporelle plus faible et des effets de bords plus présents.

Du point de vue de la transformée, il est possible de faire une analyse en ondelettes de différentes façons. La transformée en ondelettes continue [53] calcule les coefficients pour chaque échantillon et pour un grand nombre d'échelles. L'échelle d'une ondelette est échantillonnée sur des puissances de 2 ( $2^j$ ) afin de représenter l'énergie des ondelettes par octave. Dans le cas d'une transformée en ondelettes continue, plusieurs analyses peuvent être faites sur une même octave en considérant les facteurs d'échelles sous la forme  $2^{j/\nu}$  (avec  $\nu \in \mathbb{N} | \nu > 1$  et  $j \in \mathbb{N}^*$ ). Cela permet d'introduire une redondance qui facilite l'analyse du signal, un peu à la façon d'une transformée de Fourier avec ajout de zéros. La transformée en ondelettes continue est très utile pour effectuer une analyse du signal et peut employer des ondelettes analytiques. De plus, elle représente la transformée sous la forme d'un rectangle, à la manière d'une transformée de Fourier à court terme, ce qui est beaucoup plus simple pour effectuer les analyses.

La transformée en ondelettes discrète sans sous-échantillonnage [110] calcule les coefficients sur un nombre d'échelles plus réduit que la transformée en ondelettes continue ( $2^j$  avec  $j \in \mathbb{N}^*$ ). Cela permet d'avoir un coefficient par octave et par échantillon. À partir de cette décomposition, on obtient une analyse qui n'a plus de redondance sur l'axe des échelles mais qui en a toujours sur l'axe temporel. En effet, l'analyse est faite pour chaque échantillon ce qui implique d'analyser plusieurs fois avec une fenêtre glissante.

La transformée en ondelettes discrète avec sous-échantillonnage [110] ne contient aucune redondance. Elle reprend l'idée de la transformée en ondelettes sans sous-échantillonnage du point de vue du choix des échelles ( $2^j$  avec  $j \in \mathbb{N}^*$ ). En plus de ce choix d'échelles réduit, la transformée en ondelettes discrète avec sous-échantillonnage modifie la translation de l'ondelette pour ne plus avoir de redondance sur l'axe temporel. Cela permet d'obtenir un nombre minimal de coefficients et donc d'utiliser moins de ressources et de permettre certaines applications comme la compression de signaux ou d'images. En contrepartie, il est plus difficile de faire une analyse du signal avec ce type de méthodes. Les transformées en ondelettes discrètes avec sous-échantillonnage peuvent être implémentées à l'aide d'une cascade de filtres.

Enfin, une dernière transformée en ondelettes reprend les principes de la transformée en ondelettes discrète avec sous-échantillonnage mais en divisant l'approximation et les détails de manière à obtenir à la fin des signaux avec des bandes de fréquences de même largeur. C'est la transformée en paquets d'ondelettes [110] qui est utilisée par la méthode DWVM et sera donc détaillée plus tard.

Dans un premier temps, le choix a été fait d'utiliser la transformée en ondelettes continue pour simplifier les calculs en gérant une structure de données facile à représenter (rectangulaire avec beaucoup de redondance) [1]. La transformée en ondelettes continue correspond aux algorithmes précédents mais introduit des redondances dans le calcul des coefficients pour donner une transformée de taille rectangulaire (temps vs. fréquence) plus facile à utiliser pour les analyses. Les données utilisées dans le cadre de la méthode WVM (*Wavelet*

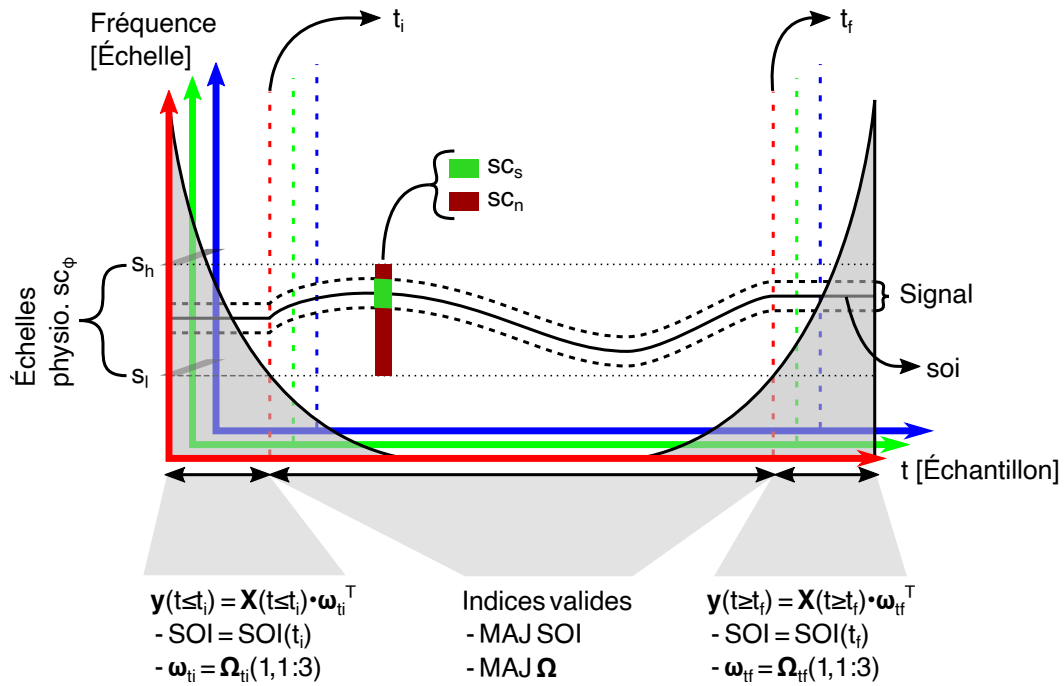


FIGURE 4.6 – Traitements et données de la méthode WVM.

*Variance Maximization*) sont représentées dans la figure 4.6. Le rectangle défini par les axes rouge, vert et bleu représente le résultat de la transformée en ondelettes continue sur les trois canaux RGB. L'axe des  $x$  correspond au temps, l'axe des  $y$  correspond à l'échelle utilisée et l'axe des  $z$  correspond aux canaux. À chaque point  $(x, y, z)$  correspond un coefficient d'ondelette complexe qui représente le produit scalaire de l'ondelette avec le signal du canal  $z$ , l'ondelette étant translatée d'un temps  $x$  et étirée par l'échelle  $y$ . Ce coefficient permet donc de représenter « l'énergie » sur une fréquence et en un temps donné. On obtient donc une analyse temps-fréquence et ces coefficients peuvent être utilisés pour construire les matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$  nécessaires à la GEVD. L'idée générale de la méthode WVM est de définir une fenêtre signal  $sc_s$ , centrée sur une échelle d'intérêt (SOI pour *Scale Of Interest*). La recherche du SOI se fait sur les bornes physiologiques définies par  $sc_\phi = [s_l, s_h]$  avec  $s_l$  l'échelle correspondant aux fréquences respiratoires les plus basses et  $s_h$  l'échelle correspondant aux fréquences respiratoires les plus hautes. L'énergie du bruit est calculée sur l'ensemble de  $sc_\phi$  en considérant que la maximisation de  $\mathcal{E}\{\mathbf{s}(sc_s)\}/\mathcal{E}\{\mathbf{s}(sc_n)\}$  est équivalente à maximiser  $\mathcal{E}\{\mathbf{s}(sc_s)\}/\mathcal{E}\{\mathbf{s}(sc_\phi)\}$ .

Dans un deuxième temps, on présente trois grandes étapes de l'analyse de la combinaison, représentées par les trois zones grises en bas du schéma. Ces zones correspondent respectivement à la phase d'initialisation, de calcul et de terminaison. Elles sont définies par la zone de validité. Cette zone en U est provoquée par un effet de bord des transformées en ondelettes. En effet, les ondelettes de plus basses fréquences ont une durée plus élevée que les ondelettes représentant les hautes fréquences. Pour avoir un coefficient valide, il faut que l'intégralité de l'ondelette soit apposée sur le signal. Ainsi, plus l'ondelette est large et plus le délai d'obtention d'un coefficient valide devient important. Ce problème se pose moins dans le cadre d'une surveillance long terme, étant donné que cette zone de validité n'existe qu'en début et fin du signal. Néanmoins, dans le cadre de l'évaluation de la méthode, cette zone représente une part importante du signal.

Afin de permettre une analyse suffisamment robuste de la méthode, les zones d'initialisation et de terminaison sont évaluées en tenant compte de leurs spécificités. La zone d'initialisation

commence au début du signal et se termine lorsque la fréquence physiologique basse ( $s_l$ ) voit ses coefficients devenir invalides. La zone de terminaison commence lorsque les coefficients de  $s_l$  deviennent invalides. Sur ces zones, les signaux sont combinés à l'aide de vecteurs de rotation fixes  $\omega(t_i)$  et  $\omega(t_f)$  de tailles  $[1 \times 3]$ . En effet, on considère qu'il n'est pas possible de déterminer les vecteurs de combinaison lorsque tous les coefficients de la bande physiologique ne sont pas valides. Ces vecteurs sont donc définis à l'aide des coefficients aux temps  $t_i$  et  $t_f$ . La fenêtre signal est centrée sur les SOI également définis sur ces temps  $t_i$  et  $t_f$ . Ainsi, les zones invalides sont évaluées mais sans utiliser la mise à jour des matrices de combinaison et de l'échelle d'intérêt. L'initialisation du SOI est faite à partir de la somme des trois transformées. Cette dernière n'est pas idéale puisque cela est équivalent à analyser un signal projeté sur le vecteur  $(1 \ 1 \ 1)$ , néanmoins il s'agit de la méthode ayant donné les meilleurs résultats de manière empirique par rapport aux autres méthodes testées dans la thèse, comme par exemple le fait d'effectuer un ensemble de GEVD sur un modèle similaire à PVM.

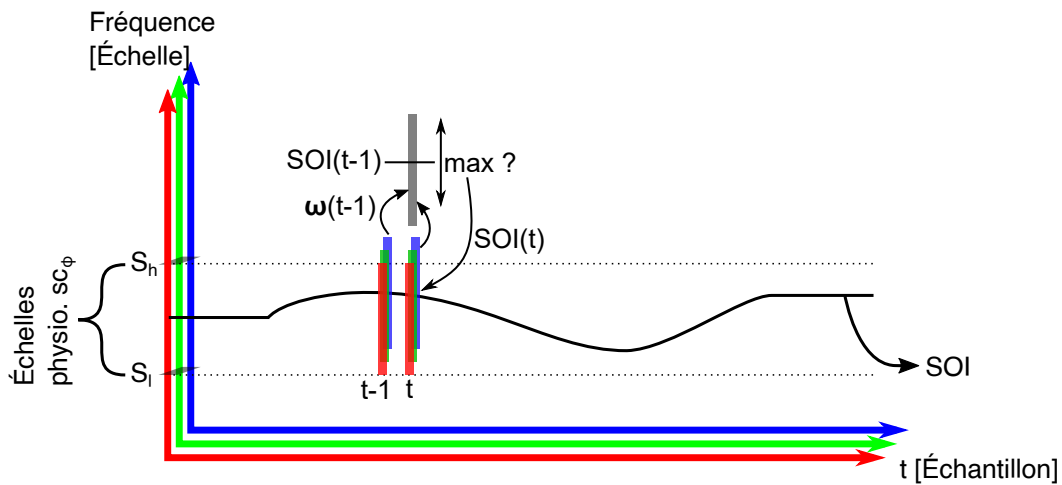


FIGURE 4.7 – Méthode de mise à jour du SOI et de  $\omega$ .

Dans la zone valide allant de  $t_i$  à  $t_f$ , l'échelle d'intérêt est mise à jour pour chaque nouvel échantillon. Pour cela, on utilise une méthode détaillée dans la figure 4.7. On considère la combinaison calculée à  $t - 1$  et on l'utilise pour combiner les coefficients à l'échantillon  $t$ . En considérant que la combinaison précédente était la meilleure, et que le vecteur de combinaison varie peu d'un échantillon à l'autre (délai de 1/20 s), cela permet d'obtenir le SOI de manière robuste. Une fois le nouvel SOI obtenu, les matrices de covariance  $P_x$  et  $C_x$  sont mises à jour de façon à utiliser les nouveaux coefficients. Le vecteur de combinaison  $\omega$  est alors recalculé à l'aide d'une nouvelle GEVD.

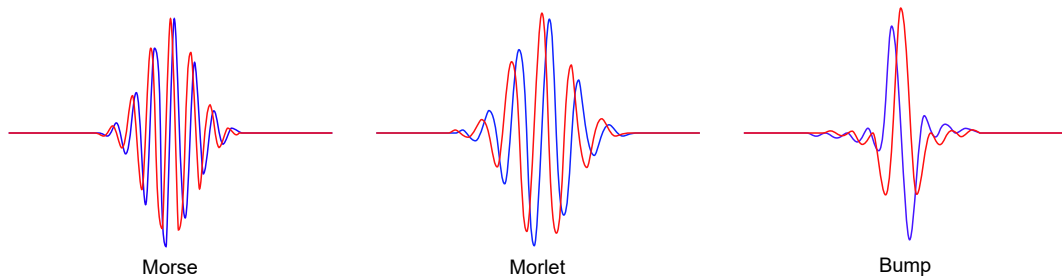


FIGURE 4.8 – Ondelettes de Morse, Morlet et Bump.

La méthode WVM a été testée avec plusieurs ondelettes analytiques différentes implémentées

sous Matlab 2018b [87], à savoir Bump<sup>2</sup>, Morse [101] et Morlet [94]. Les ondelettes testées sont présentées sur la figure 4.8. Au cours des tests, il s'est avéré que l'ondelette de Bump donne de meilleurs résultats pour l'application recherchée et avec la base de vidéos utilisée. On peut supposer que cela est dû au fait que l'ondelette soit plus compacte que les autres et présente moins d'oscillations. De ce fait, elle est plus robuste aux changements rapides de fréquence d'une part, et prend en compte plus de fréquences ce qui limite le risque en cas d'erreur dans l'estimation du SOI. Néanmoins, la recherche de la bonne ondelette a été faite de manière empirique en testant les trois ondelettes présentées. Le très grand nombre d'ondelettes existantes et la possibilité de modifier les paramètres de celles-ci laisse supposer qu'il est possible de déterminer des paramètres de manière plus rigoureuse pour améliorer encore les résultats.

Les transformées en ondelettes sont plus robustes pour l'analyse des signaux non-stationnaires puisque les ondelettes sont dimensionnées en fixant leur nombre d'impulsions plutôt que leur durée. En revanche, la transformée en ondelettes continue nécessite plus de temps de calcul puisqu'elle introduit une redondance dans le calcul des coefficients. Son utilisation vise principalement des signaux de durées finies ce qui limite leur application à des cas continus (plus intéressants pour la surveillance de patients). Enfin, l'introduction d'un nouvel échantillon modifie l'intégralité du buffer sur lequel  $\mathbf{P}_x$  et  $\mathbf{C}_x$  sont construits, ce qui empêche d'appliquer le lemme de Woodbury présenté auparavant dans l'équation (4.12). De façon à améliorer cet algorithme, une nouvelle méthode de combinaison est mise en place en utilisant non pas une transformée en ondelettes continue mais une transformée en ondelettes discrète implémentée à l'aide de filtres.

## 4 DWVM

La méthode WVM présentée dans la section précédente donne des résultats tout à fait satisfaisants dans des scénarios où la personne bouge peu. Néanmoins, sa conception la rend difficile à utiliser en pratique, notamment à cause de son coût en calcul élevé et son fonctionnement par segments. Une autre évolution de l'algorithme consiste à réutiliser le concept des transformées en ondelettes mais en l'implémentant sous la forme d'un traitement en-ligne. Pour cela, il existe des variantes de la transformée en ondelettes qui permettent de l'implémenter sous la forme de filtres traditionnels. Cette nouvelle implémentation, nommée DWVM (*Discrete Wavelet Variance Maximization*) est présentée dans cette section afin de vérifier s'il est possible de tirer parti des ondelettes tout en conservant une rapidité de calcul satisfaisante.

La méthode DWVM est un raffinement de la méthode WVM. Contrairement à celle-ci, elle n'effectue pas une transformée en ondelettes continue mais une transformée en paquets d'ondelettes [110]. La transformée en paquets d'ondelettes ne présente plus autant de redondances que dans la transformée continue. Elle permet en même temps de séparer le signal en bandes de fréquences de mêmes largeurs, contrairement à la transformée en ondelettes discrètes avec sous-échantillonnage. Cela permet de réaliser une comparaison plus juste des énergies des signaux en sortie de la transformée.

Cette méthode présente l'intérêt par rapport à WVM de pouvoir être implémentée avec des filtres traditionnels. Si on représente une ondelette à l'aide d'un filtre qui possède la même composition spectrale, alors le coefficient de l'ondelette va correspondre à l'énergie du signal en sortie du filtre. L'analyse sur plusieurs échelles se fait alors à l'aide d'opérations de sous-échantillonnage et de filtrage supplémentaires. La transformée en paquets d'ondelettes basée sur les filtres est représentée sur la figure 4.9. À chaque étage, le signal est divisé en « détails » et « approximation », obtenus en gardant respectivement les hautes et les basses

2. <https://fr.mathworks.com/help/wavelet/ref/cwtft.html#buu64ch>



fréquences. Les deux signaux obtenus sont ensuite sous-échantillonnés afin de réduire la quantité de données et de modifier le spectre des signaux. Les filtres suivants sont identiques aux précédents mais ne prennent qu'un échantillon sur deux ce qui modifie leur réponse en fréquence. La transformée en paquets d'ondelettes peut se faire sur un nombre arbitraire d'étages afin d'obtenir des résolutions fréquentielles très élevées. La différence principale entre la transformée en ondelettes discrètes avec sous-échantillonnage et la transformée en paquets d'ondelettes est la sous-division du signal « détails » (comprenant les hautes fréquences) aux étapes suivantes. Cela introduit des traitements supplémentaires mais permet en retour d'avoir des signaux qui représentent chacun une bande de fréquences de même largeur.

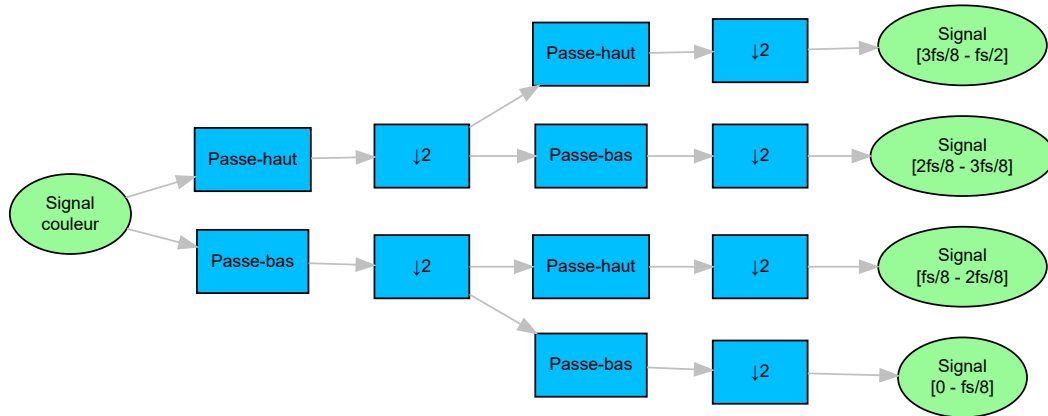


FIGURE 4.9 – Traitements permettant une transformée en paquets d'ondelettes à l'aide de filtres.

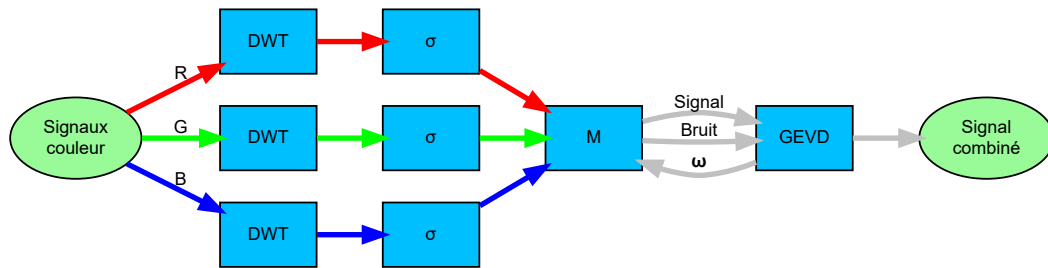


FIGURE 4.10 – Traitements effectués avec DWVM.

Dans le cas de DWVM, ce traitement est utilisé afin de permettre la sélection de la fréquence d'intérêt et le calcul des matrices  $\mathbf{P}_x$  et  $\mathbf{C}_x$  puis de la GEVD en elle-même. Ce traitement est présenté sur la figure 4.10. Dans ce schéma, on effectue une transformée en paquets d'ondelettes sur chaque canal afin de séparer les composantes fréquentielles. En sortie de cette étape, chaque canal est décomposé en un ensemble de signaux chacun contenant des bandes fréquentielles précises. L'énergie est calculée en mettant les signaux au carré et en les intégrant sur une durée donnée. Les signaux séparés sont ensuite passés dans un séparateur (M) qui doit déterminer la bande fréquentielle la plus à même de représenter le signal, en cherchant le signal (c'est-à-dire la bande fréquentielle) qui présente le plus d'énergie :

$$SOI = \operatorname{argmax}_{f \in F} \sum_{k=0}^{l-1} x_{c,f}^2(t-k) \quad (4.13)$$

avec  $SOI$  la bande fréquentielle sélectionnée par le séparateur et  $F$  l'espace de recherche. Les signaux sont ensuite recomposés en « signal » et « bruit » et utilisés pour mettre

à jour la GEVD. La partie « signal » correspond à la somme des signaux qui ont une bande fréquentielle proche de celle calculée par le séparateur, tandis que la partie « bruit » correspond à la somme des autres signaux. La GEVD ressort un signal de sortie d'une part, et un vecteur de combinaison qui peut être utilisé afin de mettre à jour la fréquence recherchée. Ainsi, le séparateur  $M$  peut déterminer la fréquence signal en effectuant un ensemble de GEVD au début (mode d'initialisation habituel) mais également mettre à jour cette fréquence en choisissant la combinaison retournée par la GEVD. Ainsi le temps de calcul requis diminue tout en préservant la cohérence du signal de sortie.

Le choix du nombre de sous-échantillonnages (ou d'« étages ») est un compromis à faire entre la résolution spectrale choisie et le délai nécessaire pour démarrer la combinaison. En effet, comme chaque étage effectue un sous-échantillonnage par deux, avec un filtre d'ordre  $o$  et une analyse passant par  $e$  étages donnera un délai initial de  $o \times 2^e$ . Ce délai correspond globalement aux zones de validité présentées dans la section 3 du chapitre 4. De manière générale il ne semble pas utile d'avoir un grand nombre d'étages, en considérant que le signal respiratoire n'est pas stationnaire.

L'ondelette utilisée pour le test de cette méthode est une ondelette de Daubechies d'ordre 2 [29]. D'après l'analyse des résultats présentée plus tard dans la section 6, on peut considérer que la méthode est suffisamment performante pour être utilisée dans des applications réelles. En plus de cela, elle est conçue de façon à pouvoir être aisément implémentable en parallèle avec par exemple un FPGA ou un GPU. Cela permettrait notamment d'augmenter le nombre d'étages et donc de déterminer plus finement la fréquence respiratoire. L'implémentation en parallèle permet également de faciliter l'utilisation de l'algorithme dans un contexte de temps réel, qui est nécessaire dans un grand nombre d'applications. Par exemple, dans le cas du suivi de nouveau nés, détecter un changement dans le rythme respiratoire plusieurs minutes après l'événement présente beaucoup moins d'intérêt, en particulier si ce changement est annonciateur d'une détresse vitale. Un autre intérêt est probablement le fait d'avoir une estimation de la fréquence respiratoire faite directement à l'étape de la combinaison. Cette estimation n'a pas été évaluée par rapport aux rythmes réels, néanmoins on peut supposer que celle-ci soit suffisamment précise pour pouvoir être utilisée directement de façon à sortir directement un rythme respiratoire plutôt qu'un signal combiné. Cela permettrait de réduire le nombre d'étapes de traitements et donc le coût en calcul.



# Chapitre 5

## Protocole expérimental et comparaison des méthodes

Chacune des méthodes présentées précédemment avaient pour but d'améliorer la précédente. Cette amélioration pouvait se faire du point de vue de la vitesse d'exécution de la méthode, qui est un point important à prendre en compte, ou du point de vue de la fiabilité des rythmes obtenus, qui doit être meilleure que les méthodes de l'état de l'art. Afin de vérifier l'intérêt qu'il y a à utiliser ces méthodes pour obtenir des rythmes respiratoires plus proches de la vérité terrain, une évaluation des méthodes a été réalisée et les différents détails la concernant ont été présentés dans ce chapitre.

L'évaluation des différents algorithmes se fait à l'aide d'une même base de données. Ces données sont traitées par les différents algorithmes qui fournissent ensuite des estimations du rythme respiratoire. Ces estimations sont alors comparées avec les valeurs données dans la vérité terrain ce qui permet d'obtenir des paires estimation / réalité. Ces paires sont enfin utilisées soit pour calculer des métriques de qualité, soit pour les afficher graphiquement afin de comprendre la répartition des erreurs. Les données et les méthodes de comparaison, présentées par la suite, ont été choisies de façon à permettre une comparaison équitable de chaque méthode. Dans ce chapitre, l'ensemble de ces points est décrit puis les résultats obtenus par leur biais et pour chaque méthode de combinaison sont présentés et analysés.

## 1 Base de validation

Plusieurs méthodes ont été développées au cours de la thèse. Celles-ci sont généralement des évolutions des précédentes faites dans le but d'améliorer les performances ou de permettre une implémentation plus rapide ou plus simple. De plus, comme nous avons pu voir dans le chapitre 2, section 2.5.2, il existe déjà des méthodes de mesure du rythme respiratoire basées sur la rPPG. À partir de là, un enjeu particulier dans la présentation de ces méthodes est de pouvoir évaluer l'efficacité de chacune par rapport aux précédentes et à l'état de l'art. En plus de cela, il est nécessaire de pouvoir étudier le comportement de chaque méthode en fonction de scénarios spécifiques afin de vérifier que les méthodes fonctionnent dans tous les cas de figure (respiration lente ou rapide, mouvements de la personne. . .). Afin de rester cohérent tout au long des analyses, la même base de vidéos est employée pour tester tous les algorithmes présentés. Cette base de données a été créée par le CSEM de Neuchâtel en Suisse. Une collaboration active avec le CSEM et plusieurs séjours de recherche au sein de cet institut, ont permis d'exploiter cette base. Elle regroupe trois scénarios différents. Dans le premier scénario (*Respiration*), le sujet effectue une apnée puis respire à des rythmes croissants allant de 5 à 30 rpm (respirations par minute). Dans le deuxième scénario (*Movement*), le sujet bouge la tête dans toutes les directions afin de valider la capacité de l'algorithme à détecter le rythme cardiaque ou respiratoire dans des situations où la personne bouge. Enfin, dans le dernier scénario (*HandGrip*), le sujet doit utiliser un *gripper* (un appareil que l'on serre dans sa main) afin de modifier son rythme cardiaque au cours de la vidéo. Chacun de ces scénarios regroupe un ensemble de 12 sujets différents de façon à pouvoir tester les algorithmes sur plusieurs types de personnes. Dans ces vidéos, le rythme cardiaque et le rythme respiratoire sont mesurés à l'aide d'un électrocardiogramme, d'un oxymètre et d'une ceinture respiratoire afin d'avoir une vérité terrain. Les différentes mesures sont gérées à l'aide d'un Système Biopac TSD210 ®. Les vidéos présentent un sujet allongé sur un lit d'hôpital avec des appareils de mesure fixés. Le cadrage montre le visage de la personne ainsi que le haut du torse. La vidéo est faite à l'aide d'une caméra RGB et une lumière artificielle. Elle utilise une taille d'image de  $1280 \times 1024$  pixels avec une fréquence de 20 ips. Les images ne sont pas compressées afin de ne pas perdre les informations de la photopléthysmographie sans contact. Suite à l'analyse des signaux, ceux correspondant aux vidéos 6 et 7 du scénario *HandGrip* ont semblé incohérentes avec notamment des rythmes respiratoires très élevés. De ce fait ces vidéos ont été retirées de l'analyse. Les analyses publiées ont été effectuées sur le scénario *Respiration* en retirant le segment d'apnée. Le retrait du segment d'apnée a été décidé car les algorithmes employés pour la détection du rythme respiratoire ne prenaient pas en compte ce cas d'utilisation et retourneraient des valeurs aberrantes. Une fois la partie d'apnée retirée, la durée moyenne d'une vidéo du scénario *Respiration* est de 3 minutes et 20 secondes, la durée moyenne d'une vidéo du scénario *Movement* est de 2 minutes et celle des vidéos du scénario *HandGrip* est de 4 minutes environ.

Afin de pouvoir comparer les algorithmes de combinaison, ces vidéos ont été utilisées de façon à extraire les signaux RGB (avec les étapes « Vidéo » de la figure 2.15). Les signaux sont extraits de la vidéo en effectuant une détection de pixels de peau sur l'ensemble de l'image à

l'aide de [100], la détection de visage ayant posé quelques problèmes de fiabilité dans le cas de ces vidéos, probablement du aux mouvements de la tête ou aux appareils de mesure. Cette détection de pixels de peau est obtenue en utilisant un modèle afin de calculer la probabilité de chaque pixel d'être un pixel de peau. Afin de raffiner le résultat de la détection, une opération d'ouverture est également effectuée. Cette opération était nécessaire à cause de petits artefacts présents dans la détection de peau au niveau de l'arrière plan. Un exemple d'image tiré de la vidéo, associé à sa vérité terrain et aux signaux RGB extraits est présenté dans la figure 5.3. Pour des questions de confidentialité et de licence, le visage du patient a été masqué. Le premier signal (PPG) représente la vérité terrain obtenue à l'aide de l'oxymètre et contenant notamment le rythme cardiaque de la personne. Le deuxième signal (*Respiration*) est le signal respiratoire obtenu à l'aide de la ceinture. Les signaux présents sur la troisième partie représentent un exemple de signaux RGB obtenus à partir de la vidéo. On peut déjà voir sur ce segment que le signal respiratoire est présent dans les signaux RGB puisqu'on peut observer une variation de la ligne de base à une fréquence proche des fréquences respiratoires sur la seconde moitié du signal. La première partie du signal est plus bruitée et il n'est pas possible de voir la composante respiratoire dans les signaux RGB. Pour obtenir le rythme respiratoire à partir des signaux RGB bruts, les signaux sont filtrés et normalisés selon la méthode employée. Les signaux sont ensuite combinés puis le résultat est utilisé pour obtenir les rythmes respiratoires. Le calcul du rythme respiratoire est fait sur des segments de 30 secondes, avec un décalage de 1 seconde entre chaque segment (voir figure 5.1). L'estimation des rythmes respiratoires est faite en se basant sur une détection des pics dans le signal respiratoire. Un exemple de cette estimation est présentée sur la figure 5.2 (RR est le Rythme Respiratoire). Les pics sont détectés sur le segment de 30 secondes et le délai total (entre le premier et le dernier pic) est obtenu. Ce délai est ensuite divisé par le nombre de pics dans la fenêtre afin d'obtenir un délai moyen entre deux pics. Enfin, ce délai moyen correspond à une période et donc à l'inverse du rythme respiratoire.

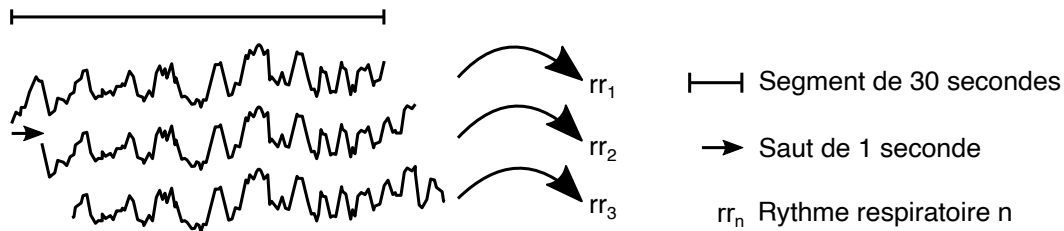


FIGURE 5.1 – Découpage des signaux en segments.

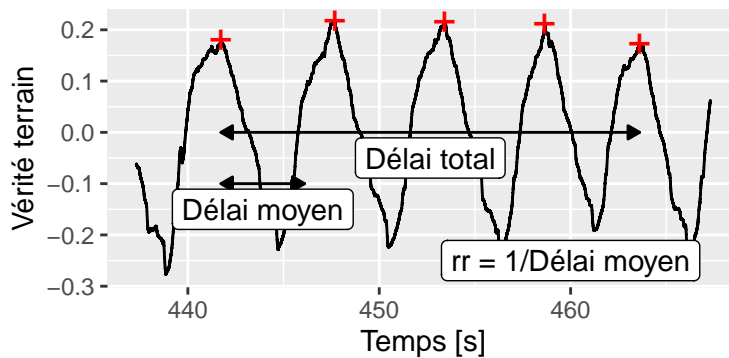


FIGURE 5.2 – Exemple d'estimation des rythmes respiratoires sur la vérité terrain.

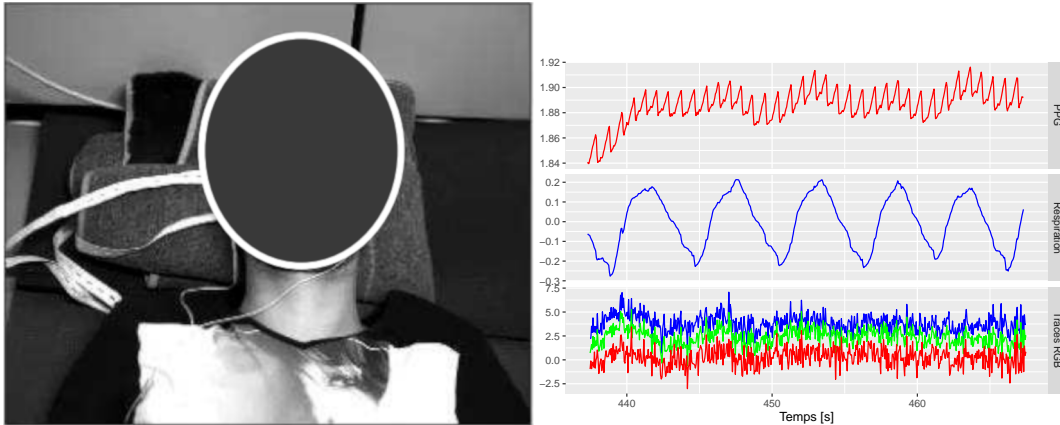


FIGURE 5.3 – Exemple tiré de la base de vidéos.

## 2 Métriques

Six métriques sont définies afin d'évaluer la qualité des algorithmes. Tout d'abord, le coefficient de Pearson, noté  $r$ , correspond à la qualité de la régression linéaire entre les rythmes respiratoires réels et estimés. Lorsque deux variables  $x$  et  $y$  sont comparées, le coefficient de Pearson vaut 1 si on a une relation de la forme  $y = ax + b$  avec  $a > 0$ . Celui-ci vaut -1 lorsque  $a < 0$  et lorsque la relation n'est pas parfaitement linéaire le coefficient de Pearson varie entre -1 et 1. Lorsque les deux variables sont totalement indépendantes, le coefficient de Pearson vaut 0. Cette métrique doit idéalement tendre vers 1. Le coefficient de Pearson ne donne aucune information sur la pente  $a$ .

Deux autres métriques utilisées correspondent à la précision de la mesure. La précision correspond au pourcentage des détections qui ont une erreur inférieure à une valeur donnée. Dans le cas des détections des rythmes respiratoires, les seuils sont définis à 1 rpm et 2 rpm. Ces deux valeurs doivent être aussi proches que possible de 1.

À cela s'ajoute les métriques MAE et RMSE. Le MAE (*Mean Absolute Error*) est l'erreur absolue moyenne. Elle doit être la plus petite possible. Le RMSE (*Root Mean Square Error*) est similaire au MAE mais utilise une erreur quadratique plutôt qu'une erreur absolue. De ce fait, le RMSE va augmenter plus fortement en présence de grandes erreurs. Enfin, le biais représente l'erreur systématique (ou moyenne) sans valeur absolue. Cette métrique représente une tendance des algorithmes à surévaluer ou sous-évaluer les rythmes respiratoires. Les métriques présentées sont résumées sur la figure 5.4.

Il est nécessaire de calculer plusieurs métriques de façon à pouvoir évaluer la qualité des algorithmes même dans des cas particuliers où l'une ou l'autre des métriques serait inappropriée. Par exemple, le coefficient de Pearson ne prend pas en compte la pente dans la relation entre les rythmes détectés et les rythmes réels. Ainsi, si la relation présente une pente différente de 1, le coefficient de Pearson peut rester idéal sans que l'algorithme considéré donne le rythme respiratoire réel. De la même façon, la précision présente quelques limites, comme le fait de devoir fixer le seuil de manière empirique. En choisissant soigneusement le seuil, il est probablement possible d'améliorer la mesure de la qualité d'un algorithme sur un autre. En plus de cela, la précision ne prend en compte l'erreur entre le rythme détecté et réel que de façon seuillée. Une erreur de 1,1 rpm et une erreur de 10 rpm donnera le même résultat dans la métrique. Le MAE et le RMSE de leur côté sont fortement impactés par les *outliers* contrairement à la précision par exemple. Cela peut limiter la qualité des algorithmes puisqu'il n'y a pas d'étape de filtrage des mauvais segments.

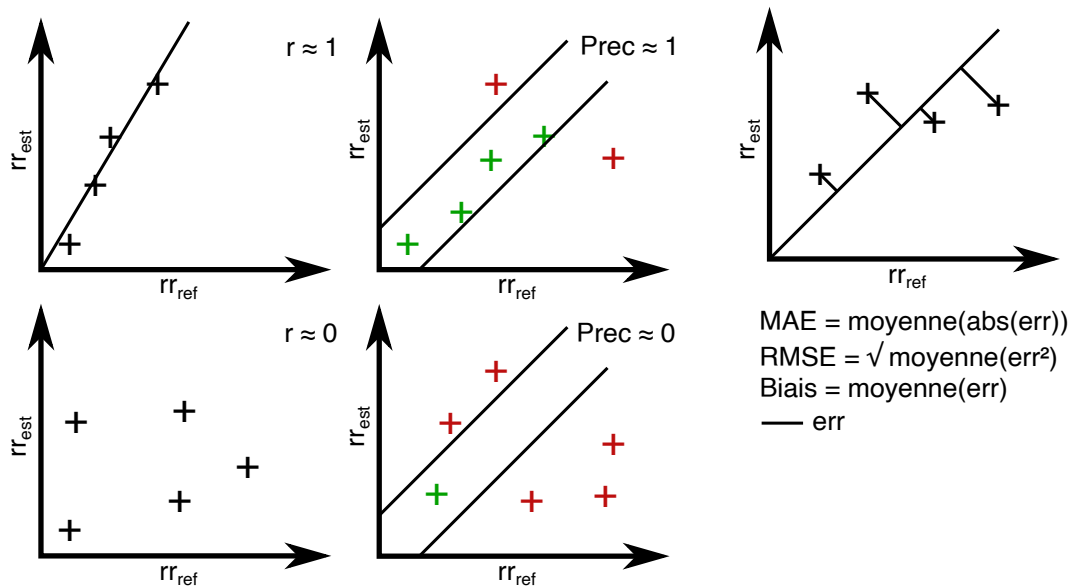


FIGURE 5.4 – Présentation des métriques de qualité.

En plus de ces métriques, les intervalles de confiance sont également estimés. Les données utilisées pour faire les évaluations sont relativement petites et de ce fait il est nécessaire de voir si les valeurs obtenues sont robustes statistiquement ou si elles dépendent uniquement d'un bruit statistique. Pour cela, il est nécessaire de regarder les intervalles de confiance pour voir si ceux-ci sont suffisamment petits pour considérer la métrique comme robuste. L'intervalle de confiance de  $r$  est calculé en utilisant le résultat de la méthode `corr.test` de R<sup>1</sup>. Un autre intervalle de confiance est calculé en prenant l'écart-type des erreurs. Cet intervalle de confiance est proche des limites d'agrément mais facilite la comparaison de la répartition des erreurs. Enfin, sur le graphique de corrélation, l'intervalle de confiance de la pente est obtenu via la méthode `confint` sur le modèle linéaire précédemment obtenu à l'aide de `lm`.

### 3 Graphiques

Les métriques présentées précédemment permettent de synthétiser au maximum la qualité des algorithmes. Le fait de les résumer autant ne permet pas de faire beaucoup d'observations sur comment les erreurs se répartissent. Il est intéressant de pouvoir évaluer par exemple si l'algorithme présente une erreur systématique (c'est-à-dire une tendance à sur ou sous-évaluer le rythme respiratoire), quelques erreurs fortes, ou des erreurs à des rythmes particuliers. Afin de pouvoir décrire les erreurs existantes, deux graphiques sont générés pour chaque algorithme. Le premier graphique est un graphe de Bland–Altman. Celui-ci représente sur l'axe des  $x$  le rythme respiratoire moyen et sur l'axe des  $y$  la différence entre le rythme respiratoire détecté et réel. L'erreur moyenne est représentée sous la forme d'une ligne horizontale bleue et l'intervalle de confiance (correspondant à  $\mu\{\text{erreurs}\} \pm 1,96\sigma\{\text{erreurs}\}$ ,  $\mu\{\text{erreurs}\}$  l'erreur moyenne et  $\sigma\{\text{erreurs}\}$  l'écart type des erreurs) est représenté à l'aide de deux lignes horizontales rouges. Afin de rendre le graphique plus lisible, deux zones rouges sont ajoutées par rapport au graphe habituel. La première zone, placée sur le côté droit du graphique, correspond à l'histogramme des erreurs. Celui-ci doit être aussi fin et aussi proche

1. <https://www.rdocumentation.org/packages/psych/versions/1.9.12.31/topics/corr.test>



de 0 que possible. Un décalage du pic dans l’histogramme correspond à une erreur systématique que l’on retrouve également sur la ligne de l’erreur moyenne. La deuxième zone rouge correspond à un MAE calculé localement afin de montrer la distribution des erreurs en fonction du rythme moyen. Pour cela, seuls les rythmes proches du rythme considéré sont utilisés afin de calculer l’erreur moyenne. Ce MAE est donc représenté par une fonction  $MAE_{local}(rr) = \mu\{rr_{ref, v} - rr_{est, v}\}$  avec  $MAE_{local}(rr)$  le MAE calculé au voisinage de  $rr$ ,  $rr_{ref, v} \in \{RR : |rr_{ref, v} - rr| < 1 \text{ rpm}\}$  et  $rr_{est, v}$  l’estimation du rythme respiratoire correspondant. Le deuxième graphique est un graphe de corrélation. Il représente sur l’axe des  $x$  le rythme respiratoire de référence ( $rr_{ref}$ ) et sur l’axe des  $y$  le rythme respiratoire estimé ( $rr_{est}$ ). Une zone rouge représentant un MAE local est également précisé. Celui-ci représente le décalage  $rr_{est}$  moyen calculé au voisinage d’un rythme  $rr_{ref}$  et permet donc d’estimer l’erreur absolue en fonction du rythme de référence.

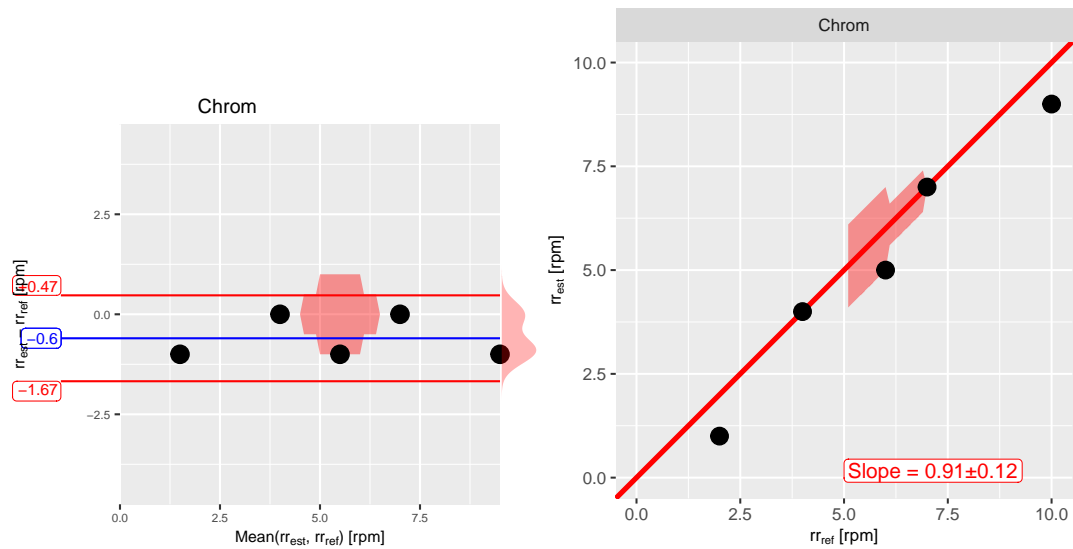


FIGURE 5.5 – Exemples de graphiques de Bland–Altman et de corrélations.

Des exemples de graphiques sont présentés sur la figure 5.5 afin de récapituler les différents éléments. Sur cette figure, on représente quelques rythmes afin de montrer le comportement des différents éléments. Sur le graphique de Bland–Altman, la zone rouge sur la droite du graphique indique l’histogramme des erreurs. Dans l’exemple, on a deux rythmes avec une erreur de 0 et trois rythmes avec une erreur de -1. Cela fait un histogramme avec deux bosses de tailles différentes. La zone rouge dans le plot indique quand à elle le MAE local. Celui-ci n’est pas calculé sur les premiers et les derniers rythmes pour éviter les effets de bord. Sur le graphique, on voit que l’erreur augmente vers le point situé en  $rr_{ref} = 6$  rpm. Sur le graphique de corrélation, la ligne rouge indique la relation  $rr_{ref} = rr_{est}$  qui est la courbe sur laquelle tous les points doivent idéalement se situer. L’étiquette *Slope* indique la pente de la régression linéaire effectuée sur les rythmes. Idéalement la valeur de *Slope* doit être de 1.

## 4 Méthodes de l’état de l’art implémentées

Les différents graphiques et métriques présentés dans les sections précédentes ont pour but de pouvoir analyser au mieux les différents algorithmes de l’état de l’art ainsi que ceux développés au cours de la thèse. Pour cela, la même base de données sera utilisée et décomposée en

trois scénarios distincts afin d’avoir une vision aussi globale que possible du fonctionnement des algorithmes. Dans les sections suivantes, nous comparerons les estimations obtenues par les algorithmes développés durant la thèse avec les estimations obtenues à l’aide de quelques algorithmes de l’état de l’art. Pour pouvoir conserver une certaine concision dans la quantité de graphiques à analyser, trois méthodes de l’état de l’art ont été conservées. La première est la méthode CHROM (présentée dans la section 2.4.2.3 du chapitre 2). Celle-ci se base sur l’utilisation d’un vecteur prédéfini représentant le bruit et guidant la combinaison (le vecteur de combinaison étant alors forcément orthogonal à ce vecteur de bruit). Il s’agit d’une méthode robuste notamment du fait de l’élimination des bruits associés aux mouvements. Pour cela, elle se base sur un modèle de l’interaction lumière-tissu. Une autre méthode utilisée pour la comparaison, nommée PBV (présentée dans le chapitre 2 section 2.4.2.4) reprend des principes assez similaires avec ceux de CHROM. Enfin, la méthode PVM (présentée dans la section 2.4.1.4 du chapitre 2 se base sur les propriétés statistiques des signaux couleur afin de trouver la combinaison idéale. Pour cela, la méthode cherche à déterminer le signal combiné qui présentera la périodicité la plus forte dans les bornes physiologiques. Contrairement à CHROM et PBV, cette méthode devrait être moins robuste lors d’une utilisation sur des vidéos polluées par un mouvement périodique (par exemple sur des vidéos de fitness), le signal respiratoire risquant d’être confondu avec le mouvement. Le choix de ces méthodes est principalement fait en fonction de leur bons résultats et afin d’avoir les deux grandes catégories d’algorithmes présentées précédemment dans l’état de l’art (à savoir les méthodes statistiques et les méthodes basées sur le modèle d’interaction lumière-tissu). La méthode PVM a été retenue plutôt que les méthodes PCA et ICA notamment car celle-ci repose sur des principes similaires à la plupart des algorithmes développés durant la thèse. De ce fait, cette méthode était incontournable pour effectuer l’évaluation de ces méthodes. Les pré-traitements sont différents d’un algorithme de combinaison à l’autre, pour les algorithmes PVM, PBV et CHROM il s’agit de la procédure de normalisation / intégration présentée dans [151] avec une taille de fenêtre de 8 secondes. Ce pré-traitement a été présenté dans les équations (2.8) et (2.9) de la section 2.3 du chapitre 2. Il a été présenté avec la méthode CHROM en utilisant une fenêtre de 8 secondes afin d’extraire le rythme respiratoire. De ce fait ce pré-traitement a été employé avec la méthode CHROM et par extension avec les méthodes PBV et PVM.

## 5 Résumé des traitements

L’ensemble des traitements effectués est présenté dans la figure 5.6. L’extraction des signaux couleur convertit une vidéo en un ensemble de signaux RGB. Ces signaux sont combinés en un seul signal à l’étape combinaison. Ce signal est découpé en segments qui sont employés pour estimer un rythme respiratoire. La même opération de découpage / estimation des rythmes est effectuée sur la vérité terrain. Enfin, chaque paire de rythme (estimation / vérité terrain) est utilisée soit pour être affichée sur un graphique soit pour calculer des métriques de qualité qui représentent les performances des algorithmes. Puisqu’on ne change que le pré-traitement et l’algorithme de combinaison, la comparaison est réalisée sur ces deux parties de la chaîne de traitement.

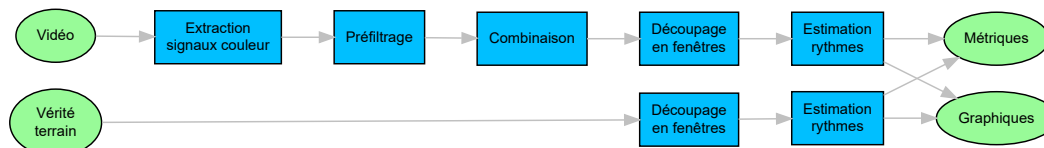


FIGURE 5.6 – Chaîne de traitements utilisée pour la comparaison des algorithmes de combinaison.

TABLE 5.1 – Temps d’exécution moyen de chaque méthode en secondes.

Algorithme	Temps
Chrom	0.16
EVM-filtre	0.17
PBV	0.19
PVM	0.73
WVM	1.09
DWVM	1.59
EVM-FFT	312.58
Bruteforce	459.45

## 6 Comparaison des méthodes

Dans cette section, les méthodes de l’état de l’art et les méthodes présentées dans la thèse ont été évaluées en utilisant les données et les méthodes de comparaisons présentées précédemment. Les vidéos utilisées pour la comparaison sont regroupées en trois scénarios distincts : *Respiration*, *HandGrip* et *Movement*. Les signaux RGB sont extraits de ces vidéos puis pré-traités et combinés avec les différentes méthodes évaluées. Les signaux obtenus sont ensuite découpés en segments et chaque segment donne un rythme respiratoire. L’ensemble des rythmes respiratoires sont comparés aux rythmes respiratoires de la vérité terrain afin de calculer des métriques de performances ou de dessiner des graphiques comparant l’estimation à la réalité.

Les temps de calcul utilisés pour chaque combinaison sont présentés dans un premier temps puis les métriques et les graphiques de résultats sont présentés afin de discuter de la capacité de chaque algorithme à fournir des rythmes proches de la réalité sur les différents scénarios.

### 6.1 Temps de calcul

En exécutant les différents algorithmes, les temps de calcul ont été mesurés afin de comparer les temps de calcul de chacune des méthodes. On obtient le tableau 5.1. Dans celui-ci, on voit que les méthodes Bruteforce et EVM-FFT prennent trop de temps pour être utilisables. Les méthodes WVM et DWVM prennent environ 10 fois plus de temps que les méthodes de l’état de l’art ce qui veut dire qu’elles peuvent ne pas être idéales lorsque l’on cherche à utiliser des systèmes avec des ressources de calcul limitées dans un contexte en temps réel. La méthode EVM-filtre de son côté semble suffisamment rapide pour être proche de CHROM et PBV. La méthode DWVM est moins rapide que les méthodes de l’état de l’art mais celle-ci peut facilement être parallélisée afin d’accélérer les traitements. Le temps de calcul obtenu dans le tableau est fait sans parallélisation, on peut donc tout à fait supposer que la méthode sera beaucoup plus rapide une fois implémentée en parallèle.

De manière plus globale, on constate que les différentes méthodes présentées dans la thèse ne sont pas plus rapides que les méthodes de l’état de l’art. Cela reste un handicap puisque la combinaison est une étape relativement simple qui ne peut pas s’accaparer les ressources de calcul au détriment des autres étapes comme l’extraction des signaux RGB ou l’estimation du rythme respiratoire. Néanmoins, certains algorithmes présentés comme EVM-filtre ou DWVM restent suffisamment rapides pour pouvoir être pris en compte dans des situations où la rapidité ou la puissance de calcul n’est pas un critère important. De plus, ces deux méthodes peuvent être implémentées en parallèle ce qui pourrait accélérer les différents traitement et potentiellement les rendre plus rapide que les méthodes de l’état de l’art.

Suite à cela, il est nécessaire de vérifier que les méthodes présentées dans la thèse donnent de meilleurs résultats que les méthodes de l'état de l'art afin de pouvoir justifier leur temps d'exécution plus élevé.

## 6.2 Métriques et graphiques

L'exécution des méthodes de combinaison donnent un signal respiratoire qui est ensuite analysé pour récupérer les rythmes respiratoires. Ces rythmes sont comparés à la vérité terrain et en fonction de leur écart un ensemble de métriques est obtenu. Ces métriques sont présentées dans le tableau 5.2. Elles sont calculées sur chaque scénario et de manière globale sur l'ensemble des scénarios. Ces métriques seront ensuite affichées dans des graphiques (figure 5.8 et 5.7) de façon à pouvoir comparer de manière intuitive les différentes méthodes. En plus des différentes métriques, les rythmes respiratoires obtenus sur le scénario *Respiration* ont été affichés sur des graphiques de Bland–Altman et de corrélation.

### 6.2.1 Comparaison par scénario

On remarque que dans le scénario *Respiration* les méthodes Bruteforce, EVM-FFT et DWVM donnent les meilleurs résultats. De manière générale, les méthodes présentées dans le manuscrit fonctionnent mieux sur ce scénario que les méthodes de l'état de l'art.

Sur les graphiques de Bland–Altman on remarque que les méthodes Bruteforce, DWVM, EVM-FFT et WVM estiment mieux le rythme respiratoire. Les graphiques de Bruteforce et EVM-FFT sont très proches ce qui est normal puisque les deux méthodes reposent sur les mêmes principes algorithmiques. Ce point était déjà visible dans les métriques et est confirmé par la répartition des erreurs sur le graphique. Les erreurs semblent réparties de façon homogène contrairement à CHROM, PBV et PVM qui présentent une erreur plus élevée aux alentours de 15 rpm. Par ailleurs, la méthode DWVM est celle qui présente l'erreur systématique la plus faible. Sur les graphes de corrélation on remarque que les erreurs sur DWVM et WVM sont bien réparties tout le long de la plage des valeurs possibles, là où des méthodes comme CHROM ou Bruteforce voient leur erreur augmenter avec le rythme. Il faut néanmoins garder à l'esprit que le nombre de rythmes respiratoires élevés est plus faible qu'aux environs de 10-15 rpm ce qui peut rendre la courbe du MAE local instable en début et fin de graphique (puisque l'on calcule l'erreur moyenne sur peu de rythmes).

Concernant le scénario *HandGrip* on remarque que la plupart des méthodes donnent des résultats similaires, à l'exception de la méthode EVM-filtre qui donne de moins bons résultats. Là encore, il est probable que la détection de la fréquence signal cause cette baisse de la qualité de la combinaison. La méthode Bruteforce fonctionne mieux du point de vue des précisions et du MAE et EVM-FFT semble relativement proche de Bruteforce. Néanmoins on remarque également que les méthodes de l'état de l'art fonctionnent légèrement mieux que les méthodes DWVM et WVM. Cela peut être dû à la présence de mouvement dans la vidéo du fait des gestes effectués ou à la plus grande proportion de signal cardiaque par rapport au signal respiratoire dans ces vidéos.

Les algorithmes donnent tous de mauvais résultats avec la méthode *Movement*. On remarque que PVM et EVM-filtre donnent des résultats moins bons que les autres méthodes. Les méthodes CHROM et PBV qui sont supposées être robustes au mouvement ne semblent pas donner de meilleurs résultats que les autres. Il est possible que l'extraction n'ait pas donné de signaux RGB avec une qualité suffisamment élevée ou que la vérité terrain ne soit pas fiable sur ces vidéos ce qui expliquerait qu'aucune méthode ne fonctionne correctement.

### 6.2.2 Comparaison par classe de méthode

Comme présenté dans l'état de l'art, les méthodes de combinaison peuvent être regroupées en deux grandes catégories : les méthodes basées sur une propriété statistique du signal (ICA, PCA, PVM) et les méthodes basées sur un modèle d'interaction lumière-tissu (CHROM, PBV, POS). En regroupant les différentes méthodes selon ces classes (nommée Statistique et Modèle), on représente les résultats du tableau 5.2 dans la figure 5.7.

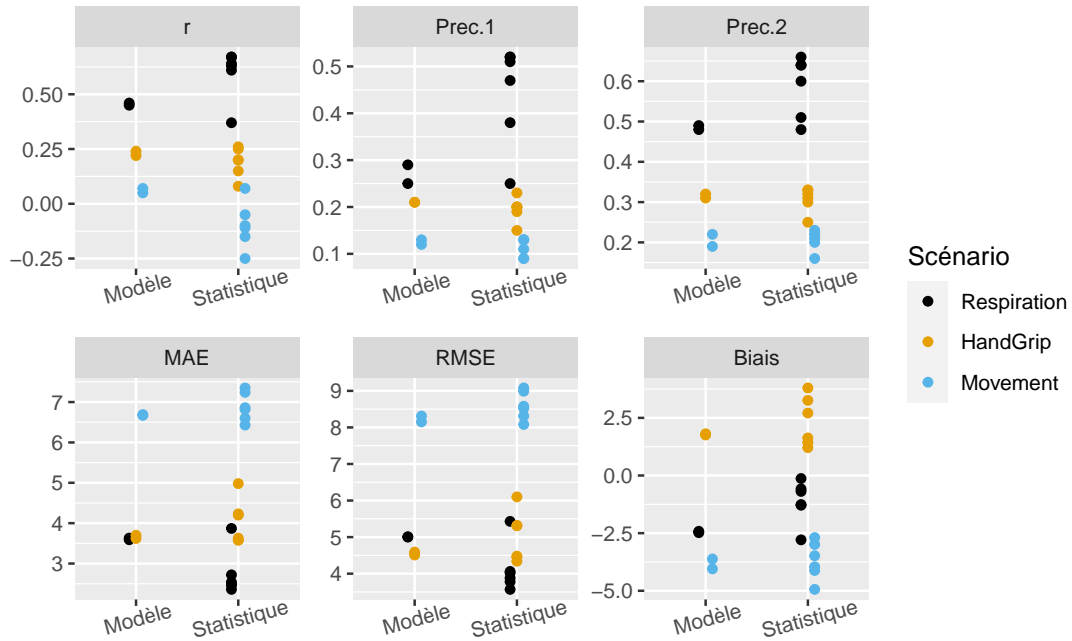


FIGURE 5.7 – Comparaison des méthodes regroupées par classe.

On remarque à la lecture de ces graphiques que les méthodes basées sur le modèle d'interaction lumière-tissu sont plus stable d'un scénario à l'autre que les méthodes Statistiques. Les méthodes Modèles semblent donner des résultats corrects pour la base de données *Movement* et *HandGrip* bien que dans certains cas les méthodes Statistiques peuvent donner de meilleurs résultats. À l'inverse, certaines méthodes Statistiques donnent de meilleurs résultats sur le scénario *Respiration* et améliorent donc la qualité de la mesure. Cela confirme les hypothèses précédemment évoquées, à savoir que les méthodes Statistiques s'adaptent aux données ce qui peut leur permettre de donner de meilleurs résultats dans certaines situations, mais les rendent inopérantes dans des scénarios où un mouvement périodique est présent. Il serait néanmoins intéressant de refaire cette analyse avec une base de vidéos plus conséquente et d'autres algorithmes Modèles, puisqu'à l'heure actuelle seuls deux algorithmes Modèles sont étudiés dans la thèse.

### 6.2.3 Apport des méthodes présentées dans la thèse

En comparant les méthodes présentées dans la thèse avec celles de l'état de l'art, on remarque principalement que certains résultats sont meilleurs avec ces méthodes plutôt qu'avec les méthodes existantes.

Les méthodes Bruteforce et EVM-FFT donnent des résultats relativement similaires, ce qui était attendu en considérant le fait que celles-ci reposent sur les mêmes types d'analyse (à savoir la transformée de Fourier). De son côté, EVM-filtre est moins robuste ce qui est probablement dû à la sélection de la fréquence d'intérêt qui n'est pas faite de manière

efficente. Le but d'EVM-filtre était avant tout d'accélérer les calculs de la combinaison ce qui se fait au détriment de la qualité. Pour finir, les méthodes WVM et DWVM avaient pour objectifs respectifs de ramener la qualité de la combinaison au niveau de Bruteforce et de réaccélérer les calculs en les vendant parallélisables. Cette dernière amélioration permet de revenir à une mesure aussi fiable que Bruteforce en utilisant moins de calculs et en facilitant la parallélisation.

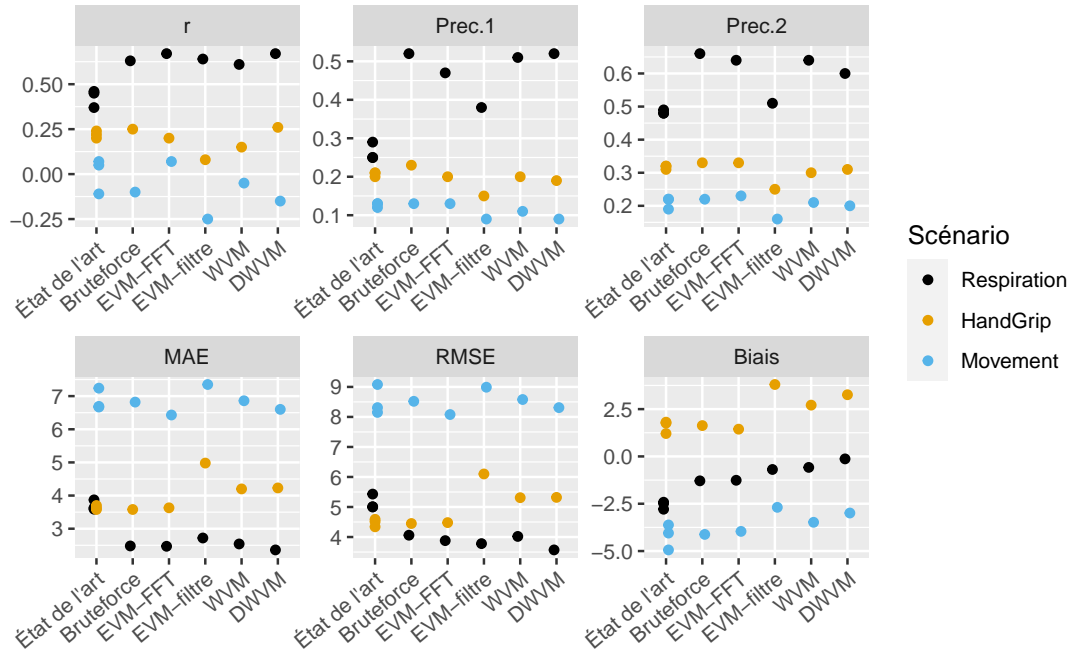


FIGURE 5.8 – Comparaison des méthodes regroupées entre état de l'art et nouvelles méthodes.

En représentant les métriques du tableau 5.2 sur la figure 5.8, on remarque que toutes les métriques sont meilleures sur la base *Respiration* avec les nouvelles méthodes. De manière générale, il semble qu'il y ait une amélioration des métriques au fur et à mesure des modifications, notamment en ce qui concerne le  $r$ , le MAE, le RMSE et le biais. On note également une chute dans la précision avec la méthode EVM-filtre qui reste néanmoins meilleure que les méthodes de l'état de l'art. Sur le scénario *HandGrip* on remarque que le  $r$  et les précisions varient peu sur les différentes méthodes. En revanche, le MAE, le RMSE et le biais augmentent de manière visible, ce qui laisse à penser que les méthodes de l'état de l'art donnent de meilleurs résultats sur ce scénario. Enfin, le scénario *Movement* semble plus difficile à interpréter que les autres, probablement du fait de la mauvaise qualité des signaux en entrée. De manière plus globale, il semble que les algorithmes de l'état de l'art sont plus efficaces pour gérer des scénarios qui comprennent un mouvement périodique léger comme dans le scénario *HandGrip*. Le scénario *Respiration* de son côté est mieux analysé à l'aide des méthodes Statistiques présentées dans la thèse. Une hypothèse possible est que les signaux RGB du scénario *Respiration* contiennent beaucoup de signal respiratoire et peu de signal cardiaque. On suppose que la présence de ce signal respiratoire visible est dû au fait que la respiration soit contrôlée dans le cas de ce scénario, ce qui modifie la manière dont la personne respire et augmente le signal dans les traces RGB. Du fait de la faible présence du signal cardiaque dans les signaux RGB, les méthodes spécialisées dans l'extraction d'un signal cardiaque sont moins efficaces sur ce type de scénario. Par exemple, PVM maximise la périodicité du signal combiné ce qui n'est pas nécessairement possible dans le cas d'un signal respiratoire non-stationnaire.

En regardant les différentes métriques, on remarque que certaines peuvent être maximisées dans certains scénarios alors que d'autres restent à un niveau moyen voire faible. Cela a déjà été vu dans la section 2, néanmoins les résultats présentés ici permettent de révéifier les explications précédentes. En observant le graphique de Bland–Altman (figure 5.9), on remarque que la méthode Bruteforce présente un grand nombre de rythmes avec une erreur faible. Cela lui permet d'avoir sur les métriques une précision plus élevée que les autres méthodes, en considérant que tous ces rythmes seront considérés comme bons. En revanche, on remarque également que certains rythmes ont une erreur exagérément élevée (environ 20 rpm) comparé à d'autres méthodes comme DWVM. Sur les métriques, cela se traduit par un MAE, un RMSE et un  $\sigma_{erreur}$  qui vont être assez proches de la méthode Bruteforce. Le coefficient de Pearson va dépendre de la présence d'une « ligne » dans le graphique de Bland–Altman ou du graphique de corrélation. Ainsi, si on compare les graphiques (figures 5.9 et 5.10) des méthodes EVM-FFT et EVM-filtre, on remarque que les rythmes affichés sur EVM-filtre ne forment pas de ligne visible mais plutôt un ensemble de rythmes répartis de façon homogène. À l'inverse, la méthode EVM-FFT présente une ligne assez visible au niveau de  $rr_{est} - rr_{ref} = 0$ . Cela indique qu'il y a une corrélation entre les rythmes estimés et les rythmes réels.

Plusieurs algorithmes de combinaison ont été développés dans le cadre de cette thèse. Ceux-ci reposent sur les mêmes principes généraux et extraient le rythme respiratoire en ne tirant parti que d'une seule modulation sur les trois existantes. Ainsi, en estimant les rythmes respiratoires à l'aide de la modulation de la ligne de base (précédemment étudiée) mais également des deux autres modulations (modulation d'amplitude et modulation de fréquence), il devrait être possible d'améliorer la robustesse et la précision des estimations. Dans le chapitre suivant, un dernier travail a consisté à utiliser les différents algorithmes développés précédemment en effectuant l'estimation des rythmes respiratoires sur les trois modulations, de façon à pouvoir améliorer encore la qualité de la mesure et vérifier que les conclusions tirées précédemment étaient fiables.

TABLE 5.2 – Métriques de performances algorithmes de combinaison.

Scenario	Algorithme	r	Prec.1	Prec.2	MAE	RMSE	Biais	$\sigma_{erreur}$
Global	CHROM	0,34	0,22	0,37	4,23	5,63	-1,07	5,53
Global	PBV	0,31	0,21	0,36	4,28	5,70	-1,16	5,58
Global	PVM	0,24	0,21	0,37	4,44	6,01	-1,70	5,77
Global	Bruteforce	0,41	<b>0,33</b>	<b>0,44</b>	3,78	5,39	-0,75	5,34
Global	EVM-filtre	0,21	0,23	0,34	4,52	6,04	0,62	6,01
Global	EVM-FFT	<b>0,46</b>	0,30	<b>0,44</b>	<b>3,72</b>	<b>5,21</b>	-0,78	<b>5,15</b>
Global	WVM	0,34	0,31	0,42	4,04	5,68	<b>0,10</b>	5,68
Global	DWVM	0,37	0,31	0,41	3,93	5,48	0,58	5,45
Respiration	CHROM	0,45	0,29	0,49	3,59	5,01	-2,42	4,39
Respiration	PBV	0,46	0,25	0,48	3,63	5,00	-2,48	4,35
Respiration	PVM	0,37	0,25	0,48	3,87	5,43	-2,79	4,65
Respiration	Bruteforce	0,63	<b>0,52</b>	<b>0,66</b>	2,48	4,06	-1,29	3,85
Respiration	EVM-filtre	0,64	0,38	0,51	2,72	3,78	-0,69	3,72
Respiration	EVM-FFT	<b>0,67</b>	0,47	0,64	2,47	3,88	-1,26	3,68
Respiration	WVM	0,61	0,51	0,64	2,54	4,02	-0,58	3,98
Respiration	DWVM	<b>0,67</b>	<b>0,52</b>	0,60	<b>2,36</b>	<b>3,57</b>	-0,13	<b>3,57</b>
HandGrip	CHROM	0,24	0,21	0,32	3,62	4,51	1,76	4,16
HandGrip	PBV	0,22	0,21	0,31	3,70	4,59	1,81	4,22
HandGrip	PVM	0,20	0,20	0,32	<b>3,58</b>	<b>4,34</b>	<b>1,21</b>	4,17
HandGrip	Bruteforce	0,25	<b>0,23</b>	<b>0,33</b>	<b>3,58</b>	4,45	1,63	<b>4,15</b>
HandGrip	EVM-filtre	0,08	0,15	0,25	4,98	6,10	3,80	4,77
HandGrip	EVM-FFT	0,20	0,20	<b>0,33</b>	3,63	4,48	1,44	4,24
HandGrip	WVM	0,15	0,20	0,30	4,20	5,31	2,71	4,56
HandGrip	DWVM	<b>0,26</b>	0,19	0,31	4,23	5,32	3,26	4,21
Movement	CHROM	<b>0,07</b>	0,12	0,19	6,67	8,15	-3,62	7,31
Movement	PBV	0,05	<b>0,13</b>	0,22	6,69	8,31	-4,05	7,26
Movement	PVM	-0,11	<b>0,13</b>	0,22	7,24	9,08	-4,94	7,63
Movement	Bruteforce	-0,10	<b>0,13</b>	0,22	6,82	8,52	-4,12	7,46
Movement	EVM-filtre	-0,25	0,09	0,16	7,35	8,99	-2,69	8,58
Movement	EVM-FFT	<b>0,07</b>	<b>0,13</b>	<b>0,23</b>	<b>6,43</b>	<b>8,08</b>	-3,96	<b>7,05</b>
Movement	WVM	-0,05	0,11	0,21	6,86	8,58	-3,48	7,85
Movement	DWVM	-0,15	0,09	0,20	6,60	8,31	-2,99	7,75



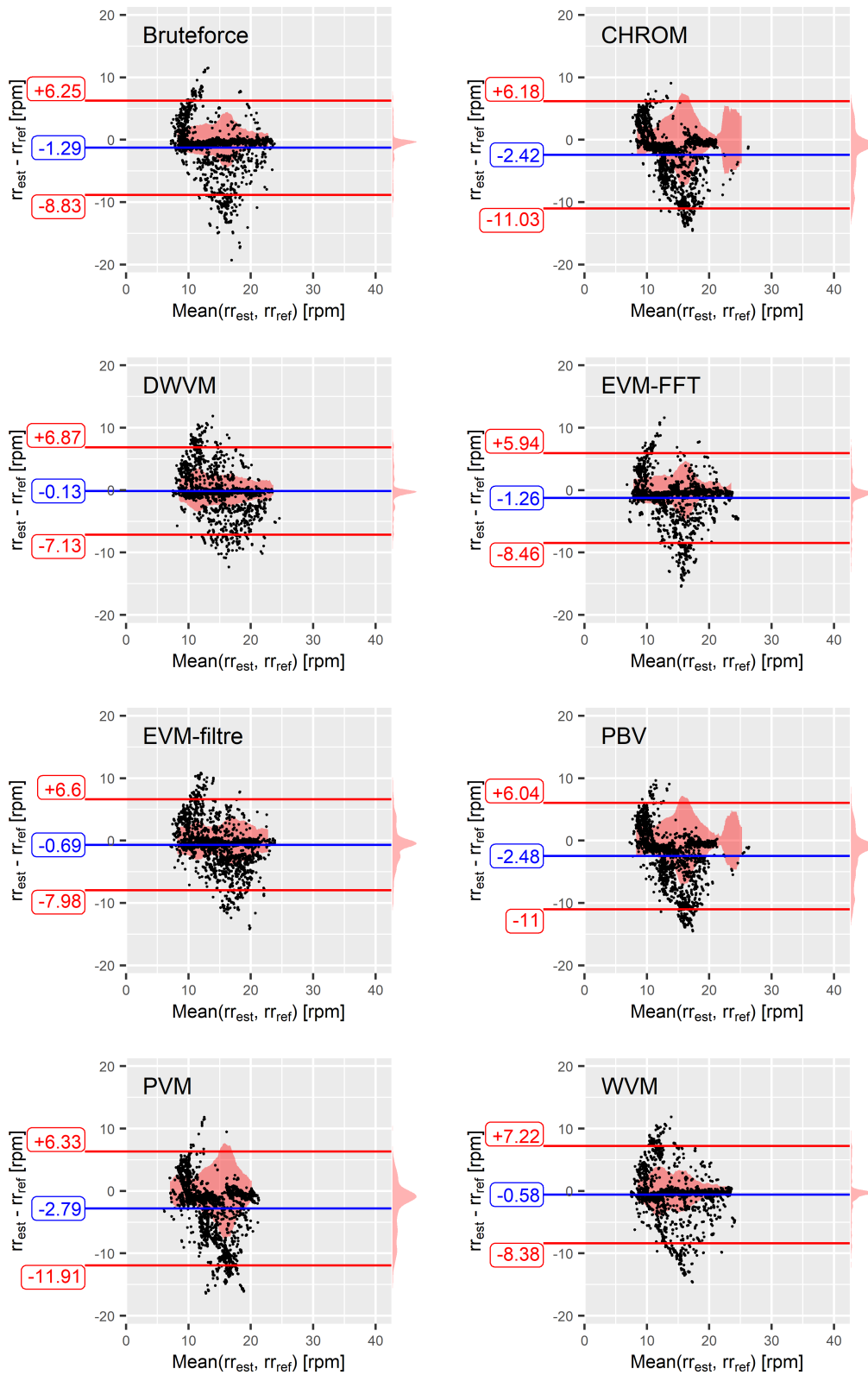


FIGURE 5.9 – Graphe de Bland–Altman représentant les rythmes détectés sur la base *Respiration*.

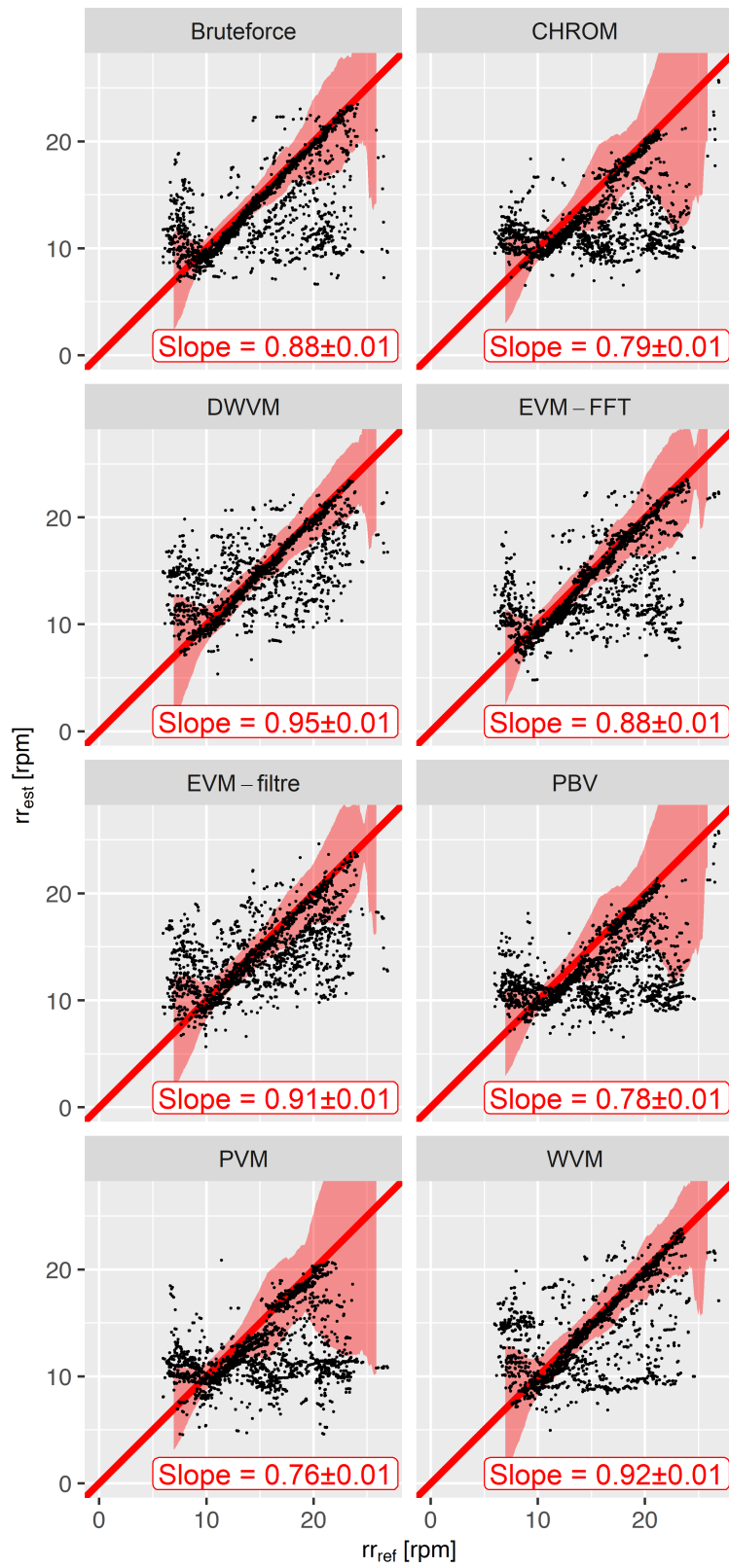


FIGURE 5.10 – Graphe de corrélation représentant les rythmes détectés sur la base *Respiration*.



# Chapitre 6

## Fusion des modulations

Comme précisé dans l'état de l'art, le rythme respiratoire peut être obtenu par le biais de trois modulations différentes du signal rPPG. Les analyses réalisées précédemment ne tenaient compte que d'une de ces modulations. Cela laisse à penser qu'il est possible d'améliorer la qualité de la mesure en intégrant les deux autres modulations. Pour cela, une chaîne de traitements a été réalisée et utilise chacune des modulations voire les trois en même temps, par le biais d'une fusion des rythmes obtenus. Les différents détails concernant cette chaîne de traitements sont présentés dans ce chapitre.

# 1 Méthode

Les algorithmes présentés précédemment avaient pour but d'améliorer l'estimation du rythme respiratoire en se focalisant sur l'estimation du vecteur de combinaison des signaux RGB  $\omega$ , avec  $\mathbf{s}(t) = \mathbf{C}(t) \cdot \omega^T$ . L'idée était de déterminer la combinaison linéaire des signaux couleur  $\mathbf{C}(t)$  tel que le SNR de  $\mathbf{s}(t)$ , le signal respiratoire estimé, soit maximal. Dans ce but une première chaîne de traitements a été mise en place pour évaluer la qualité des algorithmes de combinaison. Cette chaîne (présentée dans la figure 5.6) extrait des signaux couleur à partir d'une vidéo, les combine et estime les rythmes respiratoires à partir du signal obtenu. La chaîne de traitements ne prend en compte que la modulation de la ligne de base afin d'extraire la respiration, ce qui semble améliorable puisque la respiration influe sur le signal rPPG de trois façons différentes.

Afin de prendre en compte l'ensemble des modulations et de tester la qualité des algorithmes proposés de manière plus robuste, il était nécessaire de les implémenter sur une chaîne de traitements qui serait la plus complète possible [3]. Cette chaîne de traitements considère deux types de signaux pouvant être obtenus en sortie de la combinaison. En effet, l'étude présentée dans le chapitre 3 avait fait ressortir le fait que les combinaisons RGB pouvaient extraire soit du signal rPPG cardiaque soit du signal rPPG respiratoire. À partir de là, on peut supposer que le signal rPPG cardiaque va contenir les informations respiratoires principalement par le biais des modulations d'amplitude et de fréquence. À l'inverse, le signal rPPG respiratoire va contenir les informations respiratoires principalement sous la forme d'une modulation de la ligne de base.

Il est possible que les informations cardiaques et respiratoires soient mélangées dans les deux signaux, cependant comme la combinaison met en valeur soit le signal cardiaque soit le signal respiratoire, il serait sous-optimisé de récupérer le signal respiratoire à partir d'une combinaison cardiaque ou de récupérer un signal cardiaque à partir d'une combinaison respiratoire. Cette situation est vraie sauf si les vecteurs de combinaison cardiaque et respiratoire sont équivalents (ce qui donne de toutes façons le même signal) ou si l'estimation d'un des vecteurs de combinaison n'est pas efficace. Ce dernier cas justifiait la réutilisation du vecteur de combinaison cardiaque pour obtenir le signal respiratoire, puisque l'on considèrerait les deux vecteurs suffisamment proches. De plus, la combinaison cardiaque est un sujet étudié de longue date ce qui implique que des méthodes robustes et rapides ont vu le jour pour cette tâche. Néanmoins, les études préliminaires faites au cours de la thèse (présentées dans la section 1 du chapitre 3) laissent à penser que la qualité de la combinaison respiratoire peut être améliorée avec un vecteur dédié. De plus, on peut raisonnablement s'attendre à ce que des méthodes efficaces dédiées à la combinaison respiratoire voient le jour dans le futur au fil des recherches.

Comme on extrait d'une part un signal cardiaque et d'autre part un signal respiratoire, il est nécessaire d'ajouter une étape de traitement supplémentaire afin de convertir le signal cardiaque en un signal respiratoire. Cette étape va tenir compte soit de la modulation d'amplitude (présentée dans le chapitre 2, section 2.5.2.2), soit de la modulation de fréquence (présentée dans le chapitre 2, section 2.5.2.3) afin d'obtenir un signal respiratoire similaire au signal respiratoire directement obtenu par la combinaison. Pour cela, il est nécessaire d'introduire les signaux à deux étapes différentes de la chaîne de traitements.

Cette chaîne consiste à extraire les traces RGB depuis une vidéo et de les combiner pour obtenir soit un signal cardiaque soit un signal respiratoire. Le signal cardiaque est ensuite démodulé afin d'extraire un signal respiratoire via la modulation fréquentielle ou d'amplitude. Ces signaux respiratoires sont ensuite analysés pour estimer le rythme respiratoire. Afin d'affiner encore un peu plus le résultat, les rythmes respiratoires obtenus à l'aide des différentes modulations sont éventuellement fusionnés, à l'aide d'un lissage temporel et / ou une fusion des rythmes. Ces différentes étapes sont présentées sur la figure 6.1.

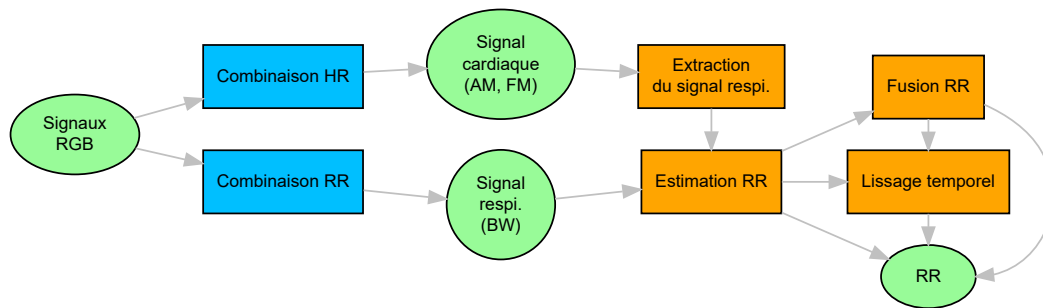


FIGURE 6.1 – Étapes pour l’estimation de rythmes respiratoires depuis des signaux RGB. Les signaux sont combinés pour extraire le signal cardiaque ou respiratoire. Le signal cardiaque permet d’obtenir une deuxième version du signal respiratoire. Les rythmes respiratoires sont ensuite estimés via l’ensemble des signaux respiratoires puis affinés avec des étapes de lissage ou de fusion facultatives.

Les algorithmes de combinaison (correspondant aux étapes en bleu) sont ceux basés sur la GEVD ainsi que les algorithmes de l’état de l’art. Afin d’obtenir soit un signal respiratoire soit un signal cardiaque, les fréquences physiologiques sont choisies différemment pour les algorithmes GEVD tandis que pour les méthodes de l’état de l’art seuls les filtrages sont modifiés. Pour chacune des étapes présentes en orange, un ensemble d’algorithmes est utilisé dans cette étude. Les rythmes respiratoires sont obtenus en testant l’ensemble des combinaisons d’algorithmes possibles.

L’idée derrière ce travail est d’avoir une comparaison plus fiable et plus exhaustive des différents algorithmes d’une part, et d’améliorer la qualité de la mesure en intégrant les autres modulations et une étape de fusion d’autre part. Les différentes publications qui ont pu être faites sur l’extraction du rythme respiratoire vont généralement utiliser chacune leur propre jeu de données et leur propre chaîne de traitements. De ce fait un nouvel algorithme d’estimation du rythme respiratoire sera testé avec un seul algorithme d’extraction et éventuellement un seul algorithme de fusion. Il n’est alors pas possible de savoir si les résultats n’auraient pas été meilleurs en utilisant un autre algorithme d’extraction par exemple. Dans le cas du travail réalisé précédemment, la qualité des méthodes est estimée avec un seul algorithme d’estimation (présenté dans la figure 5.2) et n’estime les rythmes respiratoires que par le biais de la modulation de ligne de base. Cela reste améliorable, en considérant qu’un autre algorithme d’extraction des signaux RGB ou d’estimation des rythmes respiratoires pourrait fortement améliorer les résultats. Pour tirer les conclusions les plus fiables concernant les algorithmes implémentés, il semble alors logique de tester toutes les combinaisons d’algorithmes possibles sur la même base de vidéos. De plus, cela facilite la comparaison de nouveaux algorithmes qui peuvent être ajoutés par la suite.

Ce travail avait déjà été réalisé par Charlton *et al.* dans [22] à partir de signaux ECG et PPG contact. Il est assez similaire au travail réalisé dans ce chapitre et les algorithmes implémentés par Charlton ont donc été réutilisés afin de pouvoir évaluer si ceux-ci peuvent être transposés à une utilisation sur des signaux rPPG. Afin de faire sa comparaison, Charlton a proposé les algorithmes qui sont présentés dans le tableau 6.1. On retrouve dans ce tableau les différentes classes d’algorithmes présentées en orange dans la figure 6.1. Tout d’abord, les signaux cardiaques sont traités afin d’extraire un signal respiratoire. Ce traitement peut être réalisé à l’aide d’un filtrage ou une analyse temporelle. Les filtrages vont consister en des filtres passe-bande, des transformées en ondelettes continues ou une méthode basée sur une *Centered-Correntropy Function*. Les méthodes par analyse temporelle détectent des points d’intérêts dans les battements cardiaques comme les maximums ou les creux, et en extraient des valeurs corrélées à la respiration (la hauteur du point, le délai entre deux points etc.). Ces

TABLE 6.1 – Algorithmes utilisés pour l’estimation du rythme respiratoire depuis un signal PPG (extraits de Charlton *et al.* [22]).

Algorithme	Type	Description
Wam	Extraction par filtrage	Amplitude maximale de la transformée en ondelettes continue [6]
Wfm	Extraction par filtrage	Fréquence à l’amplitude maximale de la transformée en ondelettes continue [6]
BFi	Extraction par filtrage	Filtre passe-bande [80]
CCF	Extraction par filtrage	<i>Centered-Correntropy Function</i> [51]
pulW	Extraction par analyse temporelle	Largeur de la pulsation PPG [73]
PK	Extraction par analyse temporelle	Amplitude des pics [65]
AM	Extraction par analyse temporelle	Différence d’amplitude entre les creux et les pics [65]
PCA	Extraction par analyse temporelle	<i>Kernel PCA</i> [159]
ON	Extraction par analyse temporelle	Amplitude des creux [118]
FM	Extraction par analyse temporelle	Délai entre deux pics [102, 65]
BWM	Extraction par analyse temporelle	Valeur moyenne du signal entre deux creux consécutifs [118]
BW	Extraction par analyse temporelle	Amplitude moyenne des creux et des pics
FTS	Estimation rythme (analyses fréquentielles)	Transformée de Fourier [65]
ARS	Estimation rythme (analyses fréquentielles)	Modèle auto-régressif d’ordre 8 [147, 102]
ARM	Estimation rythme (analyses fréquentielles)	Modèle auto-régressif sur plusieurs ordres et sélection par la médiane [125]
ARP	Estimation rythme (analyses fréquentielles)	Modèle auto-régressif <i>all-pole</i> [48]
ARPz	Estimation rythme (analyses fréquentielles)	Modèle auto-régressif (fréquence la plus basse) [49]
ACF	Estimation rythme (analyses fréquentielles)	Fonction d’autocorrélation [121]
WCH	Estimation rythme (analyses temporelles)	Périodogramme de Welch [73]
PKS	Estimation rythme (analyses temporelles)	Détection de pics [125]
ZeX	Estimation rythme (analyses temporelles)	<i>Positive gradient zero-crossing detection</i> [60]
PZX	Estimation rythme (analyses temporelles)	<i>Zero-crossing</i> et détection de pics [47]
CtO	Estimation rythme (analyses temporelles)	<i>Count-orig</i> [121]
CtA	Estimation rythme (analyses temporelles)	<i>Count-adv</i> [121]
PMC	Fusion rythmes	<i>Pole Magnitude Criterion</i> [102]
PRC	Fusion rythmes	<i>Pole Ranking Criterion</i> [103]
SFu	Fusion rythmes	<i>Smart Fusion</i> [65]
SPA	Fusion rythmes	<i>Spectral Peak Averaging</i> [73]
TFm	Lissage temporel	Lissage par filtre médian (sur 10 échantillons)
TFu	Lissage temporel	Lissage par filtre temporel $(0, 8 \times RR(n-1) + 0, 2 \times RR(n))$ [72]

différents points d’intérêt sont présentés dans la figure 6.2 avec la nomenclature présentée dans le tableau 6.1. Parmi les différentes valeurs extraites, certaines sont normalisées et ne correspondent donc pas exactement au schéma présenté, comme par exemple BW dont les valeurs sont divisées par la moyenne des amplitudes de pics).

Une fois les signaux respiratoires obtenus, que ce soit directement par la combinaison ou à la suite d’une étape supplémentaire d’extraction, les rythmes respiratoires sont estimés à l’aide d’analyses fréquentielles ou temporelles. Les analyses fréquentielles transforment le signal pour le représenter dans le domaine spectral. On trouve parmi ces méthodes la FFT, la fonction d’autocorrélation et différentes variantes du modèle autorégressif. Les analyses temporelles de leur côté vont plutôt se baser sur la détection d’événements dans le signal temporel comme par exemple les pics ou les croisements de l’origine.

Finalement, les rythmes estimés peuvent éventuellement être affinés en utilisant une fusion des rythmes ou un lissage temporel. La fusion des rythmes récupère un ensemble de rythmes estimés à un même instant  $t$  mais dépendant chacun d’une modulation différente (modulation de la ligne de base, modulation d’amplitude, modulation de fréquence). On trouve parmi ces fusions le *Pole Magnitude Criterion* et le *Pole Ranking Criterion*, qui se basent sur des rythmes respiratoires estimés via des modèles autorégressifs, la *Smart Fusion* qui effectue une fusion en fonction de la moyenne et de l’écart type des rythmes obtenus, et le *Spectral Peak Averaging* qui se base sur les spectres des signaux. Les méthodes de lissage temporel consistent à prendre en compte les rythmes précédents de façon à affiner la qualité de la mesure. Deux méthodes sont implémentées pour réaliser cette tâche, un filtre

médian sur les dix rythmes respiratoires précédents et un filtre passe-bas implémenté via  $0,8 \times rr(t-1) + 0,2 \times rr(t)$ .

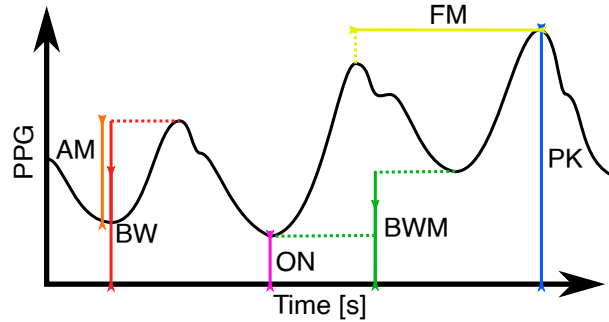


FIGURE 6.2 – Méthodes d’extraction du signal respiratoire depuis un signal cardiaque à l’aide de méthodes d’analyse temporelle présentées dans le tableau 6.1.

## 2 Résultats obtenus

Afin de tester cette chaîne de traitements, les signaux RGB sont tout d’abord obtenus par le traitement vidéo présenté dans la section 1 du chapitre 5. Ces signaux sont combinés à l’aide des algorithmes CHROM, PVM, PBV, EVM-filtre, WVM et DWVM de façon à obtenir soit un signal cardiaque soit un signal respiratoire (selon le filtrage ou la bande de fréquences à mettre en valeur). Ces signaux sont ensuite utilisés dans la chaîne de traitements selon le schéma 6.1. Une fois les rythmes respiratoires obtenus, ceux-ci sont comparés aux rythmes de la vérité terrain de façon à calculer les métriques de la section 2 (chapitre 5).

Pour limiter le nombre d’algorithmes à étudier par la suite, les MAE sont calculés sur les résultats des bases *Respiration*, *HandGrip*, *Movement* ainsi que sur l’ensemble des vidéos (*Global*). Pour chacun de ces groupes, les dix algorithmes avec les meilleurs MAE sont ensuite sélectionnés pour être affichés dans les tableaux de résultats. Les métriques de la section 2 du chapitre 5 sont présentés pour chacun de ces algorithmes dans les tableaux 6.2, 6.3, 6.4 et 6.5. Les algorithmes ont été triés en fonction du MAE, puisqu’il s’agit de la métrique retenue pour sélectionner les meilleures méthodes. Dans ce tableau, les colonnes Extr, Est, Fus et Liss correspondent aux différentes étapes de la chaîne de traitements présentés en orange dans la figure 6.1, à savoir respectivement l’extraction du signal respiratoire, l’estimation des rythmes respiratoires, la fusion des rythmes issus des différentes modulations et le lissage temporel. La colonne RC/RR indique si les rythmes respiratoires sont obtenus en paramétrant les algorithmes de combinaison pour extraire un signal cardiaque (Card., utilisant la modulation d’amplitude et de fréquence), un signal respiratoire (Respi., utilisant la modulation de la ligne de base) ou les deux (avec une fusion des rythmes).

Dans le tableau 6.2, on remarque que les meilleurs algorithmes de combinaison sont EVM-filtre et WVM ce qui montre une amélioration des résultats par rapport à CHROM, PBV et PVM qui restent néanmoins présents dans les meilleurs algorithmes. Du point de vue des algorithmes d’estimation du rythme respiratoire, les méthodes d’estimation basées sur des analyses fréquentielles semblent très présentes. On voit également que les méthodes de fusion des rythmes et de lissage temporel, améliorent la qualité de la mesure et se retrouvent notamment avec WVM et EVM-filtre. Enfin, les algorithmes sont tous basés sur une extraction du signal respiratoire directement via la combinaison des signaux RGB, à l’exception de la configuration utilisant CHROM pour la combinaison, ARM pour l’estimation du rythme respiratoire et SPA pour fusionner les rythmes (cette méthode effectuant une fusion des rythmes respiratoires et se basant donc sur les trois modulations).



TABLE 6.2 – Métriques de performances sur la base *Respiration* avec l'ensemble des modulations.

Combinaison	Est.	Fus.	Liss.	Card/Respi	MAE	r	Prec.1	RMSE
EVM-filtre	ACF		TFu	Respi.	<b>2,54</b>	0,64	0,34	3,72
WVM	ARM		TFu	Respi.	2,59	0,61	<b>0,36</b>	3,80
WVM	WCH		TFu	Respi.	2,66	0,60	<b>0,36</b>	3,96
EVM-filtre	ARS		TFu	Respi.	2,82	<b>0,72</b>	0,26	<b>3,59</b>
Chrom	ARS		TFu	Respi.	3,01	0,62	0,24	4,03
PBV	ARM		TFu	Respi.	3,14	0,54	0,28	4,58
PBV	WCH		TFu	Respi.	3,22	0,49	0,33	4,91
Chrom	ACF		TFu	Respi.	3,37	0,46	0,35	5,10
PVM	ARM		TFm	Respi.	3,37	0,47	0,23	4,82
Chrom	ARM	SPA	TFm	Les deux	3,69	0,37	0,20	5,36

TABLE 6.3 – Métriques de performances sur la base *Movement* avec l'ensemble des modulations.

Combinaison	Est.	Fus.	Liss.	Card/Respi	MAE	r	Prec.1	RMSE
PBV	CtA			Card.	<b>4,64</b>	0,37	0,10	5,92
PBV	ZeX	SFu		Les deux	4,93	0,16	0,05	<b>5,80</b>
WVM	ZeX	SFu	TFu	Les deux	4,93	0,36	0,11	6,14
EVM-filtre	PZX	SFu	TFu	Les deux	5,38	-0,11	<b>0,16</b>	6,72
PBV	PKS	SFu	TFu	Les deux	5,63	0,08	0,09	6,93
PVM	PKS	SFu	TFm	Les deux	6,07	-0,62	0,01	7,62
PVM	PKS	SFu	TFm	Les deux	6,19	0,27	0,03	7,21
WVM	PKS	SFu	TFu	Les deux	6,35	<b>0,55</b>	0,07	7,61
DWVM	PZX	SFu	TFm	Les deux	6,67	-0,07	0,03	7,65
DWVM	CtO	SFu	TFm	Les deux	6,94	-0,11	0,04	7,97

TABLE 6.4 – Métriques de performances sur la base *HandGrip* avec l'ensemble des modulations.

Combinaison	Est.	Fus.	Liss.	Card/Respi	MAE	r	Prec.1	RMSE
Chrom	ARS			Respi.	<b>4,16</b>	<b>0,52</b>	0,19	<b>5,55</b>
PBV	WCH			Respi.	4,25	0,47	<b>0,20</b>	5,90
DWVM	ARS		TFm	Respi.	4,99	0,25	0,13	6,25
WVM	WCH		TFu	Respi.	4,99	0,03	0,18	6,78
Chrom	ARM	SPA	TFu	Les deux	5,04	0,18	0,15	7,00
WVM	ARM	SPA	TFu	Les deux	5,05	0,04	0,17	6,81
DWVM	WCH	SPA	TFm	Les deux	5,23	0,18	0,13	7,15
EVM-filtre	ARS	SPA	TFu	Les deux	5,56	0,09	0,12	6,83
EVM-filtre	ACF	SPA	TFu	Les deux	5,76	0,02	0,14	7,38
PVM	ARM	SFu	TFm	Les deux	5,76	0,05	0,10	7,81

TABLE 6.5 – Métriques de performances *Global* avec l’ensemble des modulations.

Combinaison	Est.	Fus.	Liss.	Card/Respi	MAE	r	Prec.1	RMSE
Chrom	ARS			Respi.	<b>4,19</b>	<b>0,39</b>	0,20	5,83
DWVM	ARS		TFm	Respi.	4,26	0,28	0,19	<b>5,75</b>
WVM	ARM		TFu	Respi.	4,31	0,27	0,23	6,07
WVM	WCH		TFu	Respi.	4,31	0,26	<b>0,24</b>	6,09
EVM-filtre	ARS		TFu	Respi.	4,63	0,22	0,16	5,93
EVM-filtre	ACF		TFu	Respi.	4,71	0,23	0,21	6,40
Chrom	ARPz	SFu	TFu	Les deux	4,73	0,31	0,08	6,37
PBV	ZeX	SFu	TFu	Les deux	4,76	0,34	0,08	5,88
DWVM	WCH	SPA	TFm	Les deux	4,98	0,18	0,17	7,21
DWVM	CtO	SFu	TFm	Les deux	11,15	-0,11	0,02	12,94

Les vidéos contenant du mouvement (dans la base *Movement* et *HandGrip*) donnent de meilleurs résultats en extrayant d’abord un signal cardiaque et en calculant le signal respiratoire à partir de celui-ci. En considérant que le signal cardiaque est généralement plus présent dans le signal rPPG que le signal respiratoire d’une part, et que le signal respiratoire présent dans les signaux RGB va être porté en partie par la composante mouvement il semble logique que le signal cardiaque soit à privilégier dans ce genre de situations. Le scénario *HandGrip* présente des algorithmes basés sur le signal cardiaque et des algorithmes basés sur le signal respiratoire ce qui laisse à penser que le choix du type de signal dépend de la quantité de mouvement présent dans la vidéo.

Dans le tableau 6.3, PBV, EVM-filtre et WVM donnent les meilleurs résultats. Les algorithmes sont paramétrés pour obtenir un signal cardiaque. Le signal cardiaque est utilisé pour extraire un signal respiratoire à l’aide d’algorithmes d’analyse basés sur la détection de points d’intérêt puis le rythme respiratoire est obtenu à l’aide d’algorithmes de détection de *zero-crossing*, de *count-adv* et de transformée en ondelettes continue. La méthode de *zero-crossing* détecte les moments où le signal passe du positif au négatif ou inversement tandis que la méthode de *count-adv* détecte les respirations en se basant sur l’amplitude entre les pics et les creux adjacents.

La plupart des algorithmes retenus utilisent une étape de *Smart Fusion* qui consiste à garder la moyenne des trois rythmes détectés dès lors qu’ils ne sont pas trop distants les uns des autres (en fonction d’un seuil sur l’écart-type). Cette méthode permet de tirer parti des différentes modulations pour affiner le résultat. Une des configurations avec de bons résultats utilise une méthode de filtrage pour l’extraction du signal respiratoire, à savoir WVM avec une extraction du signal respiratoire basée sur Wam, une estimation du rythme respiratoire basée sur PKS et une étape de lissage temporelle TFu. Cette méthode présente notamment un  $r$  particulièrement élevé en comparaison des autres méthodes.

Concernant la base *HandGrip* (tableau 6.4), deux configurations donnent de bons résultats : CHROM avec une estimation basée sur ARS et PBV avec une estimation basée sur WCH. Les algorithmes de combinaison proviennent de la littérature et sont utilisés pour l’extraction directe du signal respiratoire. Les algorithmes utilisés pour l’estimation des rythmes respiratoires se basent tous les deux sur une analyse fréquentielle, soit à l’aide d’un modèle auto-régressif soit à l’aide d’un périodogramme de Welch.

En comparant les résultats obtenus dans ces tableaux avec ceux obtenus avec le pipeline simple (tableau 5.2), on remarque que l’utilisation des trois modulations n’améliore pas spécialement la qualité de la mesure sur le scénario *Respiration* (à l’exception de l’erreur systématique, ou Biais dans les tableaux de résultats). En regardant les signaux extraits de la base de vidéos *Respiration* on remarque que celle-ci contient peu de signal cardiaque et

que le signal respiratoire se retrouve donc principalement dans la modulation de la ligne de base. Une explication possible à ce phénomène serait liée au scénario qui impose une respiration contrôlée et pourrait donc augmenter les variations de pression intra-thoracique et la modulation de ligne de base. Une autre hypothèse serait que les conditions d'illumination ne sont pas les mêmes entre le scénario *Respiration* et les scénarios *Movement* et *HandGrip*. En effet, dans [91], une étude portant sur l'angle de la lumière a montré qu'en faisant varier celui-ci on obtenait plus de signal rPPG ou plus de signal ballistocardiographique (donc lié au mouvement). En supposant que la respiration fait varier les composantes rPPG et mouvement dans des proportions différentes des battements cardiaques, alors on peut considérer que le signal respiratoire sera plus facile à déterminer directement (modulation de la ligne de base) lorsque la composante mouvement est très présente dans les signaux RGB, cette dernière dépendant des conditions d'illumination.

Contrairement au scénario *Respiration*, on remarque dans le scénario *Movement* que plusieurs algorithmes, comme PBV ou WVM ont de meilleurs résultats que précédemment. Les métriques  $r$ , MAE, RMSE, Biais et  $\sigma_{\text{erreur}}$  sont améliorées avec l'utilisation des autres modulations (soit en prenant le signal cardiaque uniquement, soit en fusionnant les trois modulations). Cela nous montre principalement que le scénario *Movement* contient beaucoup de bruit associé au mouvement ce qui rend moins robuste l'extraction du signal respiratoire directement depuis la modulation de la ligne de base. Par ailleurs, PBV se retrouve dans les deux meilleures configurations, sachant que cette méthode est conçue pour extraire le signal cardiaque en supprimant la contribution associée au mouvement de la personne. Dans le scénario *HandGrip* il semble que les méthodes fonctionnent moins bien qu'avec le pipeline simple et qu'une bonne partie des configurations extrait le signal respiratoire uniquement via la modulation de la ligne de base.

D'après ces différents résultats, on remarque que les algorithmes donnant les meilleurs résultats vont dépendre du type de scénario. De manière générale, les algorithmes utilisant une estimation fréquentielle des rythmes respiratoires donnent de meilleurs résultats. Le choix d'extraire le signal respiratoire ou cardiaque via la combinaison va également dépendre du type de vidéo. Les signaux utilisés dans le scénario *Respiration* contiennent peu de signal cardiaque par rapport au signal respiratoire ce qui pourrait expliquer que l'utilisation d'une étape intermédiaire d'extraction du signal respiratoire depuis un signal cardiaque peut donner de moins bons résultats. Dans le cas de *HandGrip* et *Movement*, il y a plus de signal cardiaque et de bruits liés au mouvement. De ce fait, utiliser des méthodes qui extraient le signal cardiaque en supprimant la contribution du mouvement (comme CHROM et PBV) donnent de meilleurs résultats. Le scénario *Movement* bénéficie largement de la méthode de fusion des rythmes (*Smart Fusion*) ce qui laisse à penser que l'utilisation d'une modulation seule n'est pas suffisamment robuste dans un contexte où la personne bouge beaucoup.

De manière plus globale, il semble qu'il n'existe pas une chaîne de traitements idéale mais qu'un choix judicieux des algorithmes en fonction de la situation et des caractéristiques des signaux permet d'obtenir les meilleurs résultats. Bien que cette analyse doive être répliquée sur une base de vidéos plus grande afin de valider ces résultats, on peut voir que certains algorithmes donnent de très bons résultats sur certains types de vidéos et de très mauvais sur d'autres. Une manière d'améliorer la robustesse de la mesure rPPG serait alors de faire une analyse initiale des signaux pour quantifier la quantité de signal cardiaque et respiratoire, ainsi que la quantité de bruit lié aux mouvements et aux changements des conditions d'illumination. Ces différentes mesures pouvant varier en fonction des variations des conditions d'illumination, du mouvement de la personne, du sujet (par exemple les sujets âgés ont une modulation en fréquence plus faible) etc. Cette analyse serait ensuite réutilisée de façon à guider le choix des algorithmes pour choisir celui qui est le plus pertinent dans ce contexte.

# Chapitre 7

## Conclusion et projets

Ce chapitre résume les différents travaux réalisés au cours de la thèse et les apports qui en découlent. La qualité des algorithmes est rediscutée afin de mettre en évidence les éventuelles limitations associées aux méthodes ou à leur évaluation. Enfin, plusieurs possibilités de travaux futurs ont été détaillés afin de consolider ce qui a déjà été fait durant la thèse.

# 1 Conclusion

La photopléthysmographie sans contact (rPPG) est une technique assez récente qui présente un certain nombre de points forts pour la mesure de paramètres physiologiques. Elle se fait à distance du patient ce qui peut être nécessaire dans certains cas où l'hygiène est fondamentale ou lorsque la peau est fragile. Dans d'autres situations la mesure sans contact reste avantageuse en rendant moins contraignante la mesure en continu de la personne, là où les capteurs contact vont être moins confortables et plus gênants pour réaliser les gestes de tous les jours. Par ailleurs, la rPPG se base sur du matériel relativement simple et peu onéreux, à savoir une caméra et éventuellement une source de lumière. Son implémentation peut donc se faire sur des appareils de tous les jours comme un smartphone ou un ordinateur avec webcam, ou bien être réalisée sur des dispositifs dédiés pour pouvoir avoir une mesure plus précise ou des paramètres physiologiques moins simples à obtenir.

Ces caractéristiques font de la photopléthysmographie sans contact une méthode de mesure prometteuse à l'avenir dans beaucoup d'applications médicales ou de la vie de tous les jours. Néanmoins, il s'agit d'une méthode encore jeune et très peu utilisée dans des situations réelles à l'heure actuelle. Certains bruits présents dans la mesure vont en réduire la qualité et donc l'intérêt à utiliser cette méthode. Notamment, les mouvements de la personne et les changements dans les conditions d'illumination peuvent être difficiles à filtrer et fausser la mesure finale, notamment lorsque le bruit est périodique (mouvement de fitness, éclairage extérieur lors d'une conduite automobile de nuit...). Afin de résoudre ces problèmes, de nombreuses études se sont attachées au développement d'algorithmes pour améliorer la qualité de la mesure ou sa robustesse dans des scénarios difficiles.

La thèse avait pour but de contribuer à la photopléthysmographie sans contact en apportant des algorithmes supplémentaires et en validant leurs résultats. La mesure du rythme cardiaque faisant déjà l'objet de nombreuses publications, le choix a été fait de réaliser la mesure d'autres paramètres physiologiques comme le rythme respiratoire, la saturation en oxygène ou la tension artérielle. Ces paramètres sont importants pour le suivi des patients mais sont à l'heure actuelle moins étudiés que le rythme cardiaque dans le cadre de la photopléthysmographie sans contact.

Au fur et à mesure du temps, l'objectif s'est recentré sur l'extraction du rythme respiratoire, de façon à pouvoir contribuer au maximum à cette mesure. Celle-ci est très utile dans le suivi des patients mais reste souvent négligée par le personnel médical faute de temps. Les changements du rythme respiratoire peuvent pourtant être des prédicteurs de nombreuses complications vitales, comme les septicémies [129], les arrêts cardiaques [24], les accidents ischémiques [124] ou les états de choc [114]. Les contributions à cette mesure répondent donc à un besoin concret : celui de pouvoir réagir plus rapidement à un changement du rythme respiratoire tout en économisant du temps et de l'énergie au personnel médical.

Plusieurs méthodes ont été développées au fil du temps pour pouvoir pallier à ce manque : des capteurs au contact comme les ECG ou les oxymètres aux capteurs sans contact comme les ondes radio, l'analyse du mouvement ou la photopléthysmographie sans contact. Néanmoins, la mesure de la respiration à l'aide de la rPPG semble assez peu développée à l'heure actuelle. De ce fait, certains algorithmes comme ceux dédiés à la combinaison des signaux couleur se contentent de reprendre les travaux effectués pour l'estimation du rythme cardiaque, sans chercher à apporter de nouveautés particulières. Dans ce contexte, les travaux de la thèse ont eu pour but de voir s'il était possible de trouver des combinaisons couleur qui donnent un signal de bonne qualité pour l'estimation de la respiration.

Une première étude a permis de vérifier cette hypothèse en testant un grand nombre de combinaisons possibles. Ce travail permettait de représenter la qualité des signaux combinés sur une sphère de façon à déterminer des motifs dans les combinaisons à utiliser et de mettre en place des algorithmes qui exploiteront ces différents motifs. La représentation en sphère

TABLE 7.1 – Algorithmes proposés dans ce travail de thèse et basés sur la méthode GEVD.

Algorithme	Propriété d'intérêt	Bénéfice
EVM-FFT	SNR via une FFT	Comparaison avec Bruteforce
EVM-Filtre	SNR via un filtre passe-bande	Accélération, parallélisation et traitements en-ligne
WVM	Ratio des coefficients de la transformée en ondelettes continue	Analyse des signaux non-stationnaires
DWVM	Ratio des coefficients de la transformée en ondelettes discrète	Parallélisation et traitements en-ligne

consistait à indiquer sur une sphère la qualité de chaque combinaison RGB (les couleurs étant représentées sur chaque axe). Cette méthode a permis de vérifier que les combinaisons à utiliser pour maximiser le signal cardiaque et pour maximiser le signal respiratoire sont deux combinaisons distinctes, et que le choix fait par la littérature de réutiliser la combinaison cardiaque pour extraire un signal respiratoire fonctionne bien mais reste sous-optimal.

Par la suite, plusieurs évolutions de ce travail ont amélioré l'idée initiale de façon à pouvoir continuer à obtenir la meilleure combinaison possible tout en limitant le nombre de calculs à réaliser. Pour ce faire, l'algorithme GEVD est employé de façon à définir une combinaison qui met en évidence une propriété donnée. Plusieurs propriétés ont été testées comme le SNR ou les coefficients associés à une transformée en ondelettes. Cela a donné un ensemble d'algorithmes présentés dans le tableau 7.1. Dans ce tableau, on peut voir que les différents algorithmes sont des évolutions du précédent dont le but est de corriger un des points faibles de ce dernier. Ainsi, la méthode idéale était Bruteforce et a été améliorée via EVM-FFT pour avoir une méthode un peu plus rapide. Dans un deuxième temps, la méthode EVM-filtre a servi à accélérer le calcul de la combinaison, ce qui est un point crucial au vu de la vitesse des algorithmes comme CHROM ou PBV. Cette méthode permet également d'introduire la possibilité de paralléliser les traitements et de ne plus réaliser les analyses par segments mais en-ligne. Ce dernier point permet de ne plus se préoccuper de la question de la recombinaison des différents segments par la suite. Dans un troisième temps, WVM a été mise en place de façon à pouvoir réaliser une analyse basée sur une transformée en ondelettes. L'intérêt de ce choix est de pouvoir mieux analyser les signaux respiratoires qui sont généralement non-stationnaires. Enfin, DWVM avait pour but de consolider la méthode WVM en l'implémentant non plus à l'aide d'une transformée en ondelettes continue (ce qui nécessite de faire une analyse par segments) mais à l'aide d'une transformée en ondelettes discrète implémentée par une série de filtres passe bas et de sous-échantillonnages. De cette façon, il devient possible de faire une implémentation en-ligne et parallélisable tout en tenant compte de la forme du signal respiratoire (et non plus sa seule fréquence).

Au fur et à mesure des modifications, nous avons réussi à accélérer les traitements et à les rendre parallélisables sans pour autant dégrader la qualité de la mesure. La possibilité de paralléliser les traitements est particulièrement intéressante car elle permet de déporter les traitements sur des unités de calcul conçues pour le calcul parallèle, comme les GPU ou les FPGA. Cela permet de moins solliciter le CPU et donc de garder celui-ci pour les autres traitements. Le but final de cette parallélisation serait de pouvoir employer la rPPG dans des applications qui nécessitent une analyse en temps réel. Cet objectif de temps réel est important puisque de très nombreuses applications en tirent un bénéfice. Notamment le suivi médical des patients peut nécessiter de détecter le plus rapidement possible les détresses respiratoires afin de pouvoir y réagir au mieux. Enfin, un dernier travail a consisté à améliorer l'estimation des rythmes respiratoires en tirant parti des autres modulations et en testant d'autres types d'algorithmes pour l'estimation du rythme respiratoire et la fusion des rythmes obtenus. En étudiant les résultats obtenus à l'aide de cette méthode, il ressort que le meilleur choix d'algorithme dépend des propriétés des vidéos à analyser. Par exemple

les vidéos contenant peu de signal cardiaque doivent idéalement être traitées sans passer par le signal cardiaque qui pourrait introduire des erreurs.

## 2 Discussion et ouvertures

Les méthodes présentées dans le manuscrit avaient pour but d'améliorer la qualité de mesure du rythme respiratoire. En ne tirant pas parti du modèle d'interaction lumière-tissu, les méthodes obtenues ne nécessitent pas de calibration mais sont moins robustes aux bruits liés aux mouvements notamment. De nombreuses évolutions restent possibles à la suite de ces travaux. Tout d'abord la parallélisation, bien que possible, n'a pas été mise en place. Connaître le gain de temps pouvant être obtenu avec ce type d'implémentation est pourtant très important pour pouvoir garantir que la parallélisation apporte un réel intérêt. L'implémentation de certains algorithmes afin de paralléliser les calculs ne devrait pas donner d'accélération significative de par leur structure. En revanche, d'autres algorithmes comme DWVM et EVM-filtre devraient être en grande partie parallélisable et donc devenir plus rapide en faisant une implémentation sur GPU ou FPGA. Ensuite les méthodes présentées ne sont pas guidées dans leur combinaison, c'est-à-dire que contrairement à des méthodes comme CHROM ou PBV, elles ne tirent pas parti du modèle d'interaction lumière-tissu mais choisissent la combinaison uniquement en fonction des données d'entrée. Bien que cela soit très utile pour garantir de toujours avoir la meilleure combinaison possible, cela rend les algorithmes moins robustes en présence de mouvements périodiques notamment. Ces mouvements sont présents dans un grand nombre d'applications potentielles de la rPPG, ce qui réduit l'utilité de la méthode.

Afin d'améliorer la qualité de la combinaison, il serait possible d'utiliser une propriété qui serait capable de différencier le signal respiratoire du bruit. Une autre possibilité serait de guider la combinaison en utilisant des signaux intermédiaires comme dans la méthode CHROM ou PBV, de façon à supprimer la contribution du bruit. L'utilisation de la GEVD permet de prendre en compte plus de trois signaux couleur. Cela permettrait d'utiliser les algorithmes présentés dans la thèse sur des vidéos multispectrales ou sur des superpixels. En effectuant les combinaisons sur les signaux RGB de chaque superpixel cela permettrait d'améliorer la qualité de la mesure, en supposant que les signaux présents dans les superpixels ont des phases différentes [64] (contrairement à la plupart des publications qui utilisent une moyenne spatiale et supposent donc que toutes les signaux sont en phase). Enfin, le développement des algorithmes de combinaison s'est fait au détriment d'autres parties de la chaîne de traitement qui pourraient pourtant être améliorées avec des algorithmes plus spécifiques. Notamment, les analyses faites n'ont pas pris en compte les situations d'apnées qui sont pourtant des événements importants dans le suivi de patients. La détection d'apnées est pourtant un point important dans la mesure de la respiration puisqu'elles peuvent correspondre à des détresses vitales à traiter sans délai.

Au cours de cette thèse, plusieurs algorithmes pour la combinaison des signaux couleur ont été développés. Ces méthodes permettent d'améliorer la qualité de la mesure, notamment dans le cas de l'extraction du signal respiratoire. Elles ne sont néanmoins pas parfaites et plusieurs améliorations restent à entreprendre pour améliorer la qualité de la mesure.

---

## GLOSSAIRE MÉDICAL

- Accident ischémique** : Diminution brutale de la perfusion du cerveau, des rétines ou de la moelle épinière
- Aorte** : Artère en sortie du cœur qui transporte le sang dans le thorax et l'abdomen
- Artères** : Tuyaux élastiques qui conduisent le sang oxygéné et sous pression dans les différents organes et muscles du corps
- Artères carotidiennes** : Artères irriguant le cerveau
- Axe électrique** : Direction principale dans laquelle l'onde de dépolarisation du muscle cardiaque se dirige (généralement vers le bas et la gauche)
- Ballistocardiographie** : Méthode de mesure du signal cardiaque basée sur les mouvements provoqués par la contraction du cœur et le déplacement du sang dans le corps
- Battements ectopiques** : Battements cardiaques provoqués par des foyers ectopiques qui sont irritables ou sujets à un bloc d'entrée
- Bruits de Korotkoff** : Bruits provoqués par le déplacement du sang dans les artères lorsque celles-ci sont soumises à une certaine pression
- Capillaires** : Très petites artères / veines présentes dans l'ensemble du corps
- Capnographe** : Appareil mesurant la concentration en CO<sub>2</sub> dans l'air expiré
- Carboxyhémoglobine** : Hémoglobine transportant du monoxyde de carbone
- Choc anaphylactique** : Réaction allergique sévère pouvant être mortelle
- Choc septique** : Défaillance circulatoire aiguë provoquée par une infection
- COVID19** : Maladie mortelle provoquée par le SARS-CoV-2
- Délire paranoïaque** : Syndrome délirant présentant un thème paranoïaque unique, hautement systématisé et interprétatif
- Dérivations (ECG)** : Mesure de l'ECG faite en deux points et mesurant les dépolarisations qui se déplacent entre ces deux points
- Désoxyhémoglobine** : Hémoglobine ne transportant pas d'oxygène
- Détresse vitale** : Défaillance neurologique, respiratoire ou circulatoire pouvant provoquer d'autres détresses vitales ou le décès de la victime à court terme
- Diastole** : Étape de relâchement du muscle cardiaque permettant son remplissage
- Électrocardiographie** : Méthode de mesure basée sur les ondes de dépolarisation électrique provoquées par l'activité cardiaque
- Entaille dichrotique** : Légère diminution du signal photopléthysmographique qui correspond au passage de l'onde systolique à l'onde diastolique
- Épisode psychotique** : Syndrome délirant hallucinatoire non systématisé, organisé autour de thèmes de persécution, de mégalomanie ou d'idées mystiques (possession) avec adhésion totale de la victime
- État de choc** : Détresse circulatoire brutale caractérisée par une diminution importante du débit cardiaque
- Fibrillation** : Décharges électriques anarchiques et inefficaces par les foyers ectopiques soumis à des blocs d'entrée
- Foyer ectopique / d'automatisme** : Système présent dans le cœur en grande quantité et prenant le relais du nœud sinusal (ou d'un autre foyer) en cas de défaillance



**Foyers atriaux :** Foyers ectopiques situés dans les oreillettes et servants de relais en cas de défaillance du nœud sinusal

**Foyers jonctionnels :** Foyers ectopiques situés dans le nœud atrioventriculaire et servants de relais en cas de défaillance du nœud sinusal et des foyers atriaux

**Glycémie :** Concentration en glucose (sucres) dans le sang

**Hémoglobine :** Pigment rouge présent dans le sang et qui permet de transporter l'oxygène

**Hémorragie :** Perte de sang rapide, importante et qui ne peut pas s'arrêter spontanément

**Holter ECG :** Appareil de mesure ECG portable et simplifié utilisé pour réaliser des mesures sur du long terme

**Hypertension :** Maladie correspondant à une tension artérielle trop élevée

**Hypertrophie :** Organe (par exemple le cœur) d'une taille supérieure à la normale

**Intervalle inter-battements :** Délai entre deux battements cardiaques, utilisé pour le calcul du HRV

**Lésion :** Altération d'un tissu biologique ou d'un organe due à un traumatisme ou à une maladie

**Manomètre :** Dispositif utilisé pour mesurer une pression

**Mélanine :** Pigments foncés qui influent sur la couleur de la peau, des cheveux et des yeux

**Modulation d'amplitude :** Modulation du signal cardiaque provoquée par la respiration et se présentant sous la forme d'une variation de l'amplitude des pics cardiaques

**Modulation de fréquence :** Modulation du signal cardiaque provoquée par la respiration et se présentant sous la forme d'une variation du délai entre deux pics cardiaques

**Modulation de la ligne de base :** Modulation du signal cardiaque provoquée par la respiration et se présentant sous la forme de l'ajout d'un signal respiratoire

**Nœud atrioventriculaire (ou atrial) :** Point de jonction électrique entre les oreillettes et les ventricules du cœur. Il ralentit l'impulsion électrique afin que celle-ci contracte les ventricules au moment où ils seront sur remplis par la contraction des oreillettes

**Nœud sinusal :** Chef d'orchestre du cœur. Génère des impulsions électriques à une fréquence donnée

**Occlusion vasculaire :** Blocage brusque d'un vaisseau sanguin

**Ondes de Mayer :** Ondes visibles dans les signaux photopléthysmographiques qui correspondent à des oscillations de la pression artérielle due à la régulation de celle-ci

**Oreillette (cœur) :** Parties du cœur correspondant à deux petites cavités qui ont pour rôle de sur-remplir les ventricules

**Oxyhémoglobine :** Hémoglobine transportant de l'oxygène

**Paramètre physiologique :** Mesure utilisée pour diagnostiquer un problème de santé chez une personne

**Pathologie :** Maladie

**Pathologie vasculaire périphérique :** Maladie due à une obstruction des artères périphériques, généralement dans les jambes

**Perfusion (organe) :** Alimentation d'un organe en nutriments et en oxygène

**Plaie :** Effraction plus ou moins profonde de la peau

**Plan frontal :** Plan dont le vecteur normal est situé vers l'avant

**Plan longitudinal :** Plan dont le vecteur normal est situé vers le haut

**Plan sagittal :** Plan dont le vecteur normal est situé vers le côté

**Pneumotachographe :** Appareil mesurant le débit instantané d'air

**Point isobestique** : Longueur d'onde à partir de laquelle les spectres d'absorption de l'hémoglobine et de la désoxyhémoglobine se croisent

**Saturation en oxygène** : Paramètre physiologique représentant la concentration en oxyhémoglobine dans le sang

**Septicémie** : Infection généralisée

**Sphygmomanomètre** : Dispositif de mesure de la tension artérielle qui se base sur l'écoute des bruits de Korotkoff

**Sténose artérielle de la jambe** : Voir Pathologie vasculaire périphérique

**Stéthoscope** : Dispositif utilisé pour écouter des bruits (cardiaques ou respiratoires) facilement

**Système nerveux autonome** : Partie du système nerveux qui régule l'activité des organes en la stimulant ou en l'inhibant

**Systole** : Étape de contraction du muscle cardiaque permettant d'éjecter le sang dans le corps

**Télé médecine** : Domaine de la médecine consistant à réaliser des consultations à distance

**Tension artérielle** : Paramètre physiologique représentant la pression à l'intérieur des veines et des artères sous la forme de deux valeurs

**Trauma** : Blessure physique ou psychique infligée à l'organisme

**Valves cardiaques** : Parties du cœur dont le but est d'éviter le reflux sanguin et d'isoler électriquement les différentes parties du cœur (oreillettes et ventricules)

**Variabilité du rythme cardiaque** : Paramètre physiologique représentant les faibles variations du rythme cardiaque au fil du temps (dues à la respiration, à l'état émotionnel etc.)

**Veine porte** : Veine conduisant le sang désoxygéné du tube digestif (pancréas, rate...) vers le foie

**Veines** : Tuyaux élastiques qui conduisent le sang désoxygéné et sous basse pression dans le ventricule droit du cœur

**Ventricules (cœur)** : Parties du cœur correspondant à deux grandes cavités qui ont pour rôle d'éjecter le sang dans les artères



---

## TABLE DES ABRÉVIATIONS

**ASIC** : Application Specific Integrated Circuit  
**BIC** : Bayesian Information Criterion  
**BSS** : Blind Source Separation  
**CHROM** : CHROMinance based algorithm  
**CPU** : Central Processing Unit  
**CSD** : Correntropy Spectral Density  
**CSEM** : Centre Suisse d'Électronique et de Microtechnique  
**DEL** : Diode Électro-Luminescente  
**DWVM** : Discrete Wavelet Variance Maximization  
**ECG** : Électrocardiogramme  
**EMD** : Empirical Mode Decomposition  
**EVM** : Energy Variance Maximization  
**FFT** : Fast Fourier Transform  
**FPGA** : Field Programmable Gate Array  
**GAFAM** : Google Apple Facebook Amazon Microsoft  
**GEVD** : Generalized Eigen Value Decomposition  
**GPU** : Graphical Processing Unit  
**HRV** : Heart Rate Variability  
**IBI** : Inter Beat Interval  
**IBIS** : Iterative Boundaries implicit Identification for Segmentation  
**ICA** : Independent Component Analysis  
**KNN** : K-Nearest Neighbors  
**LF** : Low Frequencies  
**MAE** : Mean Absolute Error  
**NMD** : Non linear Mode Decomposition  
**PBV** : Pulse Blood Vector  
**PCA** : Principal Component Analysis  
**POS** : Plane Orthogonal to Skin  
**PPG** : Photoplethysmography  
**PVM** : Periodic Variance Maximization  
**RGB** : Red Green Blue  
**RMSE** : Root Mean Square Error  
**rPPG** : Remote Photoplethysmography  
**RR** : Rythme Respiratoire  
**S<sub>p</sub>O<sub>2</sub>** : Saturation Pulsée en Oxygène  
**SIDA** : Syndrome d'Immunodéficience Acquise  
**SLIC** : Simple Linear Iterative Clustering  
**SNR** : Signal Noise Ratio

**SOI** : Scale Of Interest  
**SSA** : Singular Spectrum Analysis  
**SSF** : Slope Sum Function  
**SSR** : Spatial Subspace Rotation  
**SST** : SynchroSqueezing Transform  
**STFT** : Short Term Fourier Transform  
**SVM** : Support Vector Machine  
**SVR** : Support Vector Regression  
**VFCDM** : Variable Frequency Complex Demodulation  
**VLf** : Very Low Frequencies  
**WVM** : Wavelet Variance Maximization

# Table des figures

1.1	Principes de fonctionnement de la PPG . . . . .	15
2.1	Exemple de spiromètre portatif . . . . .	27
2.2	Exemple de mesure ECG . . . . .	29
2.3	Étapes du battement cardiaque (ECG) . . . . .	29
2.4	Modulations respiratoires du signal ECG . . . . .	30
2.5	Modulation de fréquence . . . . .	31
2.6	Interaction lumière-tissu . . . . .	32
2.7	Le battement PPG . . . . .	32
2.8	Modulations respiratoires du signal PPG . . . . .	32
2.9	Modulations de ligne de base et d'amplitude . . . . .	33
2.10	Exemple d'oxymètre . . . . .	33
2.11	Spectres d'absorption du sang . . . . .	34
2.12	Image infrarouge . . . . .	36
2.13	Dispositif de mesure de la respiration à l'aide des ondes radio . . . . .	37
2.14	Les composantes de la lumière . . . . .	39
2.15	Traitements rPPG . . . . .	41
2.16	Filtres polariseurs et rPPG . . . . .	42
2.17	Lumière structurée et rPPG . . . . .	42
2.18	Segmentation en superpixels . . . . .	45
2.19	Sphère ICA . . . . .	49
2.20	Sphère PCA . . . . .	50
2.21	Sphère PVM . . . . .	51
2.22	Sphère <i>Green</i> . . . . .	52
2.23	Sphère <i>Green-Red</i> . . . . .	52
2.24	Sphère CHROM . . . . .	53

2.25	Sphère PBV . . . . .	54
2.26	Sphère POS . . . . .	55
3.1	Variation forte du SNR via une petite variation du vecteur de combinaison . . . . .	70
3.2	Exemple de sphère SNR irrégulière . . . . .	71
3.3	Utilisation d'une fenêtre gaussienne . . . . .	71
3.4	Utilisation de la convolution . . . . .	72
3.5	Sphères SNR . . . . .	73
3.6	Signaux Bruteforce . . . . .	74
3.7	Combinaison par segments . . . . .	75
3.8	La procédure d' <i>overlap add</i> découpe le signal en segments qui se chevauchent, applique une fenêtre et les recombine avec une somme. . . . .	75
3.9	Sphère SNR . . . . .	76
3.10	Retournement des segments via le SNR . . . . .	76
4.1	Filtrage avancé via la GEVD . . . . .	81
4.2	Traitements de la méthode EVM-FFT . . . . .	83
4.3	Traitements dans EVM-filtre . . . . .	84
4.4	EVM-filtre via une cascade de GEVDs . . . . .	86
4.5	Lien fréquence principale et ondelette . . . . .	87
4.6	Traitements et données de WVM . . . . .	89
4.7	Méthode de mise à jour WVM . . . . .	90
4.8	Ondelettes analytiques . . . . .	90
4.9	Transformée en ondelettes via des filtres . . . . .	92
4.10	Traitements dans DWVM . . . . .	92
5.1	Découpage en segments . . . . .	97
5.2	Estimation des rythmes respiratoires . . . . .	97
5.3	Base de vidéo . . . . .	98
5.4	Métriques de qualité . . . . .	99
5.5	Exemples de graphiques . . . . .	100
5.6	Comparaison des algorithmes de combinaison . . . . .	101
5.7	Résultats par classe de méthode . . . . .	104
5.8	Résultats entre l'état de l'art et les nouvelles méthodes . . . . .	105
5.9	Bland–Altman sur la base <i>Respiration</i> . . . . .	108
5.10	Graphe de corrélation sur la base <i>Respiration</i> . . . . .	109
6.1	Estimation de rythmes respiratoires via la rPPG . . . . .	113
6.2	Méthodes d'analyse temporelle. . . . .	115

---

## REFERENCES

### Publications personnelles

#### Articles de journaux internationaux avec comité de lecture

- [1] D. LUGUERN et al. « Wavelet Variance Maximization : A Contactless Respiration Rate Estimation Method Based On Remote Photoplethysmography ». In : *Biomedical Signal Processing and Control*. Elsevier (2021).

#### Articles de conférences et workshops internationaux avec comité de lecture

- [2] S. BOBBIA et al. « Real-Time Temporal Superpixels for Unsupervised Remote Photoplethysmography ». In : *Computer Vision and Pattern Recognition Workshops*. IEEE. 2018.
- [3] D. LUGUERN et al. « An Assessment of Algorithms to Estimate Respiratory Rate From the Remote Photoplethysmogram ». In : *Computer Vision and Pattern Recognition Workshops*. IEEE. 2020.
- [4] D. LUGUERN et al. « Remote Photoplethysmography Combining Color Channels with SNR Maximization for Respiratory Rate Assessment ». In : *ISMICT*. IEEE. 2020.



# Bibliographie générale

- [5] R. ACHANTA et al. « SLIC Superpixels Compared to State-of-the-Art Superpixel Methods ». In : *Pattern Analysis and Machine Intelligence*. IEEE (2012).
- [6] P. S. ADDISON et J. N. WATSON. « Secondary Transform Decoupling of Shifted Nonstationary Signal Modulation Components : Application to Photoplethysmography ». In : *International Journal of Wavelets, Multiresolution and Information Processing*. World Scientific (2004).
- [7] T. J. AKL et al. « In Vitro Performance of a Perfusion and Oxygenation Optical Sensor Using a Unique Liver Phantom ». In : *BiOS*. SPIE. 2012.
- [8] T. J. AKL et al. « Intestinal Perfusion Monitoring Using Photoplethysmography ». In : *Journal of Biomedical Optics*. SPIE (2013).
- [9] T. J. AKL et al. « Photoplethysmography beyond Perfusion and Oxygenation Monitoring : Pulse Wave Analysis for Hepatic Graft Monitoring ». In : *BiOS*. SPIE. 2014.
- [10] J. ALLEN. « Photoplethysmography and Its Application in Clinical Physiological Measurement ». In : *Physiological Measurement*. IOP Publishing (2007).
- [11] J. ANTONI. « The Spectral Kurtosis : A Useful Tool for Characterising Non-Stationary Signals ». In : *Mechanical Systems and Signal Processing*. Elsevier (2006).
- [12] G. M. AZMAL, A. AL-JUMAILY et M. AL-JAAFREH. « Continuous Measurement of Oxygen Saturation Level Using Photoplethysmography Signal ». In : *Biomedical and Pharmaceutical Engineering*. IEEE. 2006.
- [13] C. F. BABBS. « The Origin of Korotkoff Sounds and the Accuracy of Auscultatory Blood Pressure Measurements ». In : *Journal of the American Society of Hypertension*. Elsevier (2015).
- [14] A. BATES et al. « Respiratory Rate and Flow Waveform Estimation from Tri-Axial Accelerometer Data ». In : *International Conference on Bioelectric Sensors and Nanosensors*. IEEE. 2010.
- [15] Y. BENEZETH et al. « Probabilistic Signal Quality Metric for Reduced Complexity Unsupervised Remote Photoplethysmography ». In : *ISMICT*. IEEE. 2019.
- [16] N. BLANIK et al. « Hybrid Optical Imaging Technology for Long-Term Remote Monitoring of Skin Perfusion and Temperature Behavior ». In : *Journal of Biomedical Optics*. SPIE (2014).
- [17] N. BLANIK et al. « Remote Vital Parameter Monitoring in Neonatology – Robust, Unobtrusive Heart Rate Detection in a Realistic Clinical Scenario ». In : *Biomedical Engineering*. De Gruyter (2016).

- [18] S. BOBBIA, Y. BENEZETH et J. DUBOIS. « Remote Photoplethysmography Based on Implicit Living Skin Tissue Segmentation ». In : *ICPR*. IEEE. 2016.
- [19] S. BOBBIA et al. « Unsupervised Skin Tissue Segmentation for Remote Photoplethysmography ». In : *Pattern Recognition Letters*. Elsevier (2019).
- [20] F. BOUSEFSAF, C. MAAOUI et A. PRUSKI. « Continuous Wavelet Filtering on Webcam Photoplethysmographic Signals to Remotely Assess the Instantaneous Heart Rate ». In : *Biomedical Signal Processing and Control*. Elsevier (2013).
- [21] F. BOUSEFSAF, C. MAAOUI et A. PRUSKI. « Quantification de l'effort mental à partir d'une webcam en utilisant les amplitudes du flux sanguin et la variabilité de la fréquence cardiaque ». In : *Congrès Handicap*. 2014.
- [22] P. H. CHARLTON et al. « An Assessment of Algorithms to Estimate Respiratory Rate from the Electrocardiogram and Photoplethysmogram ». In : *Physiological Measurement*. IOP Publishing (2016).
- [23] P. H. CHARLTON et al. « Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram : A Review ». In : *Biomedical Engineering*. IEEE (2018).
- [24] L. CHELLURI. « Respiratory Deterioration and Cardiac Arrest ». In : *Critical Care Medicine*. Lippincott Williams & Wilkins (2019).
- [25] W. CHEN et D. MCDUFF. « DeepPhys : Video-Based Physiological Measurement Using Convolutional Attention Networks ». In : *European Conference on Computer Vision*. Springer (2018).
- [26] K. H. CHON, S. DASH et K. JU. « Estimation of Respiratory Rate From Photoplethysmogram Data Using Time-Frequency Spectral Estimation ». In : *Biomedical Engineering*. IEEE (2009).
- [27] U. A. CIFTCI, I. DEMIR et L. YIN. « FakeCatcher : Detection of Synthetic Portrait Videos Using Biological Signals ». In : *Pattern Analysis and Machine Intelligence*. IEEE (2020).
- [28] A. COHEN, I. DAUBECHIES et J.-C. FEAUVEAU. « Biorthogonal Bases of Compactly Supported Wavelets ». In : *Communications on Pure and Applied Mathematics* (1992).
- [29] I. DAUBECHIES. *Ten Lectures on Wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics. 1992.
- [30] H. E. F. DAVIES. « Respiratory Change in Heart Rate, Sinus Arrhythmia in the Elderly ». In : *Gerontologia Clinica*. Karger (1975).
- [31] G. de HAAN et V. JEANNE. « Robust Pulse Rate From Chrominance-Based rPPG ». In : *Biomedical Engineering*. IEEE (2013).
- [32] G. de HAAN et A. van LEEST. « Improved Motion Robustness of Remote-PPG by Using the Blood Volume Pulse Signature ». In : *Physiological Measurement*. IOP Publishing (2014).
- [33] P. DEHKORDI et al. « Estimating Instantaneous Respiratory Rate from the Photoplethysmogram ». In : *Engineering in Medicine and Biology Society*. IEEE. 2015.
- [34] K. de MIGUEL et al. « Home Camera-Based Fall Detection System for the Elderly ». In : *Sensors*. MDPI (2017).
- [35] H. DEMIREZEN et C. ERDEM. « Remote Photoplethysmography Using Nonlinear Mode Decomposition ». In : *International Conference on Acoustics, Speech, and Signal Processing*. IEEE. 2018.
- [36] P. E. DILAVERIS et al. « Simple Electrocardiographic Markers for the Prediction of Paroxysmal Idiopathic Atrial Fibrillation ». In : *American Heart Journal*. Elsevier (1998).

- [37] L. DRIJKONINGEN et al. « Validation of a Smartphone Based Photoplethysmographic Beat Detection Algorithm for Normal and Ectopic Complexes ». In : *Computing in Cardiology*. IEEE. 2014.
- [38] H. DUBEY, N. CONSTANT et K. MANKODIYA. « RESPIRE : A Spectral Kurtosis-Based Method to Extract Respiration Rate from Wearable PPG Signals ». In : *Connected Health : Applications, Systems and Engineering Technologies* (Philadelphia, Pennsylvania). IEEE. 2017.
- [39] M. ELLIOTT. « Why Is Respiratory Rate the Neglected Vital Sign? A Narrative Review ». In : *International Archives of Nursing and Health Care*. ClinMed (2016).
- [40] H. ELPHICK et al. « Thermal Imaging Method for Measurement of Respiratory Rate ». In : *European Respiratory Journal*. European Respiratory Society (2015).
- [41] S. FALLET et al. « Imaging Photoplethysmography : What Are the Best Locations on the Face to Estimate Heart Rate ? ». In : *Computing in Cardiology*. IEEE. 2016.
- [42] M. FATEMI et R. SAMENI. « An Online Subspace Denoising Algorithm for Maternal ECG Removal from Fetal ECG Signals ». In : *Iranian Journal of Science and Technology*. Springer (2017).
- [43] D. FELBER DIETRICH et al. « Heart Rate Variability in an Ageing Population and Its Association with Lifestyle and Cardiovascular Risk Factors : Results of the SAPALDIA Study ». In : *EP Europace*. Oxford Academic (2006).
- [44] L. FENG et al. « Dynamic ROI Based on K-Means for Remote Photoplethysmography ». In : *International Conference on Acoustics, Speech, and Signal Processing*. IEEE. 2015.
- [45] L. FENG et al. « Motion-Resistant Remote Imaging Photoplethysmography Based on the Optical Properties of Skin ». In : *Circuits and Systems for Video Technology*. IEEE (2015).
- [46] M. M. FLECK, D. A. FORSYTH et C. BREGLER. « Finding Naked People ». In : *European Conference on Computer Vision*. Sous la dir. de B. BUXTON et R. CIPOLLA. Springer. 1996.
- [47] S. FLEMING. « Measurement and Fusion of Non-Invasive Vital Signs for Routine Triage of Acute Paediatric Illness ». Thesis. Oxford University, UK, 2010.
- [48] S. FLEMING et al. « Non-Invasive Measurement of Respiratory Rate in Children Using the Photoplethysmogram ». In : *Medicine and Biology Society*. IEEE. 2008.
- [49] S. G. FLEMING et L. TARASSENKO. « A Comparison of Signal Processing Techniques for the Extraction of Breathing Rate from the Photoplethysmogram ». In : *International Journal of Biological and Medical Sciences*. OMICS International (2007).
- [50] M. GALINIER et al. « Intérêt du holter ECG dans la surveillance de la cardiopathie hypertensive ». In : *Annales de Cardiologie et d'Angéiologie*. Elsevier (2002).
- [51] A. GARDE et al. « Estimating Respiratory and Heart Rates from the Correntropy Spectral Density of the Photoplethysmogram ». In : *PLOS ONE* (2014).
- [52] R. GONZÁLEZ-LANDAETA, O. CASAS et R. PALLÀS-ARENY. « Heart Rate Detection from an Electronic Weighing Scale ». In : *Physiological Measurement*. IOP Publishing (2008).
- [53] A. GROSSMANN, R. KRONLAND-MARTINET et J. MORLET. « Reading and Understanding Continuous Wavelet Transforms ». In : *Wavelets*. Inverse Problems and Theoretical Imaging. 1989.
- [54] A. GROSSMANN et J. MORLET. « Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape ». In : *SIAM Journal on Mathematical Analysis*. Society for Industrial and Applied Mathematics (1984).

- [55] J. GUNTHER, N. RUBEN et T. MOON. « Model-Based (Passive) Heart Rate Estimation Using Remote Video Recording of Moving Human Subjects Illuminated by Ambient Light ». In : *International Conference on Image Processing*. IEEE. 2015.
- [56] A. HAAR. « Zur Theorie Der Orthogonalen Funktionensysteme ». In : *Mathematische Annalen*. Springer (1910).
- [57] Y. HSU, Y.-L. LIN et W. HSU. « Learning-Based Heart Rate Detection from Remote Photoplethysmography Features ». In : *International Conference on Acoustics, Speech, and Signal Processing*. IEEE. 2014.
- [58] R.-Y. HUANG et L.-R. DUNG. « A Motion-Robust Contactless Photoplethysmography Using Chrominance and Adaptive Filtering ». In : *Biomedical Circuits and Systems Conference*. IEEE. 2015.
- [59] M. JACQUET-LAGRÈZE et al. « A New Device for Continuous Assessment of Gut Perfusion : Proof of Concept on a Porcine Model of Septic Shock ». In : *Critical Care*. Springer (2014).
- [60] A. JOHANSSON. « Neural Network for Photoplethysmographic Respiratory Rate Monitoring ». In : *Medical and Biological Engineering and Computing*. Springer (2003).
- [61] R. JOSHI et al. « Pattern Discovery in Critical Alarms Originating from Neonates under Intensive Care ». In : *Physiological Measurement*. IOP Publishing (2016).
- [62] C. JULIEN. « The Enigma of Mayer Waves : Facts and Models ». In : *Cardiovascular Research*. Elsevier (2006).
- [63] A. A. KAMSHILIN et N. B. MARGARYANTS. « Origin of Photoplethysmographic Waveform at Green Light ». In : *Physics Procedia*. Elsevier (2017).
- [64] A. A. KAMSHILIN et al. « A New Look at the Essence of the Imaging Photoplethysmography ». In : *Scientific Reports*. Nature (2015).
- [65] W. KARLEN et al. « Multiparameter Respiratory Rate Estimation From the Photoplethysmogram ». In : *Biomedical Engineering*. IEEE (2013).
- [66] V. KESSLER et al. « Multimodal Fusion Including Camera Photoplethysmography for Pain Recognition ». In : *International Conference on Companion Technology*. IEEE. 2017.
- [67] F. AL-KHALIDI et al. « Respiratory Rate Measurement in Children Using a Thermal Imaging Camera ». In : *International Journal of Scientific & Engineering Research* (2015).
- [68] R. KHAN et al. « Color Based Skin Classification ». In : *Pattern Recognition Letters*. Elsevier (2012).
- [69] M. KOPELIOVICH, Y. MIRONENKO et M. PETRUSHAN. « Architectural Tricks for Deep Learning in Remote Photoplethysmography ». In : *International Conference on Computer Vision Workshop*. IEEE. 2019.
- [70] J. KUUTTI et al. « Evaluation of a Doppler Radar Sensor System for Vital Signs Detection and Activity Monitoring in a Radio-Frequency Shielded Room ». In : *Measurement*. Elsevier (2015).
- [71] S. KWON et al. « ROI Analysis for Remote Photoplethysmography on Facial Video ». In : *Engineering in Medicine and Biology Society*. IEEE. 2015.
- [72] J. LÁZARO et al. « Deriving Respiration from Photoplethysmographic Pulse Width ». In : *Medical and Biological Engineering & Computing*. Springer (2013).
- [73] J. LÁZARO PLAZA. « Non-invasive techniques for respiratory information extraction based on pulse photoplethysmogram and electrocardiogram ». Universidad Zaragoza, 2015.

- [74] R. LE PAGE et al. « Comparison Of Two Methods For Demodulation Of Pulse Signals - Application In Case Of Central Sleep Apnea ». In : *Journal of Science and Technology*. Vietnam Academy of Science and Technology (2020).
- [75] D. LEE et al. « Heart Rate Estimation from Facial Photoplethysmography during Dynamic Illuminance Changes ». In : *Engineering in Medicine and Biology Society*. IEEE. 2015.
- [76] H. LEE et al. « A New Algorithm for Detecting Central Apnea in Neonates ». In : *Physiological Measurement*. IOP Publishing (2011).
- [77] M. LEWANDOWSKA et al. « Measuring Pulse Rate with a Webcam - a Non-Contact Method for Evaluating Cardiac Activity ». In : *Federated Conference on Computer Science and Information Systems*. IEEE. 2011.
- [78] M. H. LI, A. YADOLLAHI et B. TAATI. « A Non-Contact Vision-Based System for Respiratory Rate Estimation ». In : *Engineering in Medicine and Biology Society*. IEEE. 2014.
- [79] J. M. LILLY et S. C. OLHEDE. « On the Analytic Wavelet Transform ». In : *IEEE Transactions on Information Theory*. IEEE (2010).
- [80] L.-G. LINDBERG, H. UGNELL et P. Å. ÖBERG. « Monitoring of Respiratory and Heart Rates Using a Fibre-Optic Sensor ». In : *Medical and Biological Engineering and Computing*. Springer (1992).
- [81] S. LIU et al. « 3D Mask Face Anti-Spoofing with Remote Photoplethysmography ». In : *European Conference on Computer Vision*. Sous la dir. de B. LEIBE et al. Springer. 2016.
- [82] I. LORATO et al. « Camera-Based on-Line Short Cessation of Breathing Detection ». In : *International Conference on Computer Vision Workshop*. IEEE. 2019.
- [83] T. LUGUEV, D. SEUSS et J.-U. GARBAS. « Deep Learning Based Affective Sensing with Remote Photoplethysmography ». In : *Conference on Information Sciences and Systems*. IEEE. 2020.
- [84] R. MACWAN, Y. BENEZETH et A. MANSOURI. « Remote Photoplethysmography with Constrained ICA Using Periodicity and Chrominance Constraints ». In : *Biomedical Engineering*. Springer (2018).
- [85] R. MACWAN et al. « Periodic Variance Maximization Using Generalized Eigenvalue Decomposition Applied to Remote Photoplethysmography Estimation ». In : *Computer Vision and Pattern Recognition Workshops*. IEEE. 2018.
- [86] R. MARROQUIN, J. DUBOIS et C. NICOLLE. « Ontology for a Panoptes Building : Exploiting Contextual Information and a Smart Camera Network ». In : *Semantic Web*. IOS Press (2018).
- [87] MATLAB. *Version 9.5.0 (R2018b)*. 2018.
- [88] D. MCDUFF, S. GONTAREK et R. W. PICARD. « Improvements in Remote Cardiopulmonary Measurement Using a Five Band Digital Camera ». In : *Biomedical Engineering*. IEEE (2014).
- [89] Y. MEYER. « Orthonormal Wavelets ». In : *Wavelets*. Sous la dir. de J.-M. COMBES, A. GROSSMANN et P. TCHAMITCHIAN. Inverse Problems and Theoretical Imaging. 1989.
- [90] L. MIRMOHAMADSADEGHI et al. « Real-Time Respiratory Rate Estimation Using Imaging Photoplethysmography Inter-Beat Intervals ». In : *Computing in Cardiology*. IEEE. 2016.
- [91] A. V. MOÇO, S. STUIJK et G. de HAAN. « Ballistocardiographic Artifacts in PPG Imaging ». In : *Biomedical Engineering*. IEEE (2016).

- [92] H. MONKARESI, R. A. CALVO et H. YAN. « A Machine Learning Approach to Improve Contactless Heart Rate Monitoring Using a Webcam ». In : *Journal of Biomedical and Health Informatics*. IEEE (2014).
- [93] J. P. MONTANI et al. « Advantages of Continuous Measurement of Cardiac Output 24 h a Day ». In : *Heart and Circulatory Physiology*. American Physiology Society (1995).
- [94] J. MORLET et al. « Wave Propagation and Sampling Theory—Part I : Complex Signal and Scattering in Multilayered Media ». In : *Geophysics*. Society of Exploration Geophysicists (1982).
- [95] A. A. AL-NAJI et al. « Monitoring of Cardiorespiratory Signal : Principles of Remote Measurements and Review of Methods ». In : *IEEE Access* (2017).
- [96] Y. NAM, J. LEE et K. H. CHON. « Respiratory Rate Estimation from the Built-in Cameras of Smartphones and Tablets ». In : *Annals of Biomedical Engineering*. BMES (2014).
- [97] R. NANDAKUMAR, S. GOLLAKOTA et N. WATSON. « Contactless Sleep Apnea Detection on Smartphones ». In : *International Conference on Mobile Systems, Applications, and Services*. Association for Computing Machinery. 2015.
- [98] Y.-H. NHO et al. « User-Adaptive Fall Detection for Patients Using Wristband ». In : *International Conference on Intelligent Robots and Systems*. IEEE. 2016.
- [99] M. NOGAWA et al. « Development of a Tissue Oxygen Consumption Measurement Method Based on Near-Infrared Photoplethysmography ». In : *Engineering in Medicine and Biology Society*. IEEE. 2004.
- [100] C. Ó CONAIRE, N. O’CONNOR et A. SMEATON. « Detector Adaptation by Maximising Agreement between Independent Data Sources ». In : *Object Tracking and Classification in and Beyond the Visible Spectrum*. IEEE (2019).
- [101] S. C. OLHEDE et A. T. WALDEN. « Generalized Morse Wavelets ». In : *IEEE Transactions on Signal Processing*. IEEE (2002).
- [102] C. ORPHANIDOU et al. « Data Fusion for Estimating Respiratory Rate from a Single-Lead ECG ». In : *Biomedical Signal Processing and Control*. Elsevier (2013).
- [103] C. ORPHANIDOU et al. « Spectral Fusion for Estimating Respiratory Rate from the ECG ». In : *Information Technology and Applications in Biomedicine*. IEEE. 2009.
- [104] T. L. PALMIERI. « Infection Prevention : Unique Aspects of Burn Units ». In : *Surgical Infections*. Mary Ann Liebert Publishers (2019).
- [105] M. PARACCHINI et al. « Remote PhotoPlethysmoGraphy Using SPAD Camera for Automotive Health Monitoring Application ». In : *AEIT Automotive*. IEEE. 2019.
- [106] N. PARADKAR et S. R. CHOWDHURY. « Cardiac Arrhythmia Detection Using Photoplethysmography ». In : *Engineering in Medicine and Biology Society*. IEEE. 2017.
- [107] G. PARATI et al. « Advantages and Disadvantages of Non-Invasive Ambulatory Blood Pressure Monitoring. - Abstract - Europe PMC ». In : *Journal of Hypertension*. International Society of Hypertension (1990).
- [108] J. PARK, A. SABHARWAL et A. VEERARAGHAVAN. « Direct-Global Separation for Improved Imaging Photoplethysmography ». In : *Computer Vision and Pattern Recognition Workshops*. IEEE. 2018.
- [109] S. PARK et al. « Remote Pulse Rate Measurement From Near-Infrared Videos ». In : *Signal Processing Letters*. IEEE (2018).
- [110] D. B. PERCIVAL et A. T. WALDEN. *Wavelet Methods for Time Series Analysis*. 2000. 628 p.

- [111] M. A. F. PIMENTEL et al. « Toward a Robust Estimation of Respiratory Rate From Pulse Oximeters ». In : *Biomedical Engineering*. IEEE (2017).
- [112] M.-Z. POH, D. J. MCDUFF et R. W. PICARD. « Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam ». In : *Biomedical Engineering*. IEEE (2011).
- [113] P. PRAVDIN et J. SMORODNIKOVA. *COVID-19 and Wearables Open Data Research*. Welltory, 2020.
- [114] C. PRESCOTT et S. RUFF. « The Shocked Patient ». In : *Medicine*. Elsevier (2017).
- [115] A. QAYYUM et al. « Estimation of Non-Contact Smartphone Video-Based Vital Sign Monitoring Using Filtering and Standard Color Conversion Techniques ». In : *Life Sciences Conference*. IEEE. 2017.
- [116] R. RIMDEIKA et R. BAGDONAS. « Major Full Thickness Skin Burn Injuries in Premature Neonate Twins ». In : *Burns*. Elsevier (2005).
- [117] K. M. ROSEN, S. H. RAHIMTOOLA et R. M. GUNNAR. « Pseudo A-V Block Secondary to Premature Nonpropagated His Bundle Depolarizations ». In : *Circulation*. AHA Journals (1970).
- [118] R. RUANGSUWANA, G. VELIKIC et M. F. BOCKO. « Methods to Extract Respiration Information from ECG Signals ». In : *Acoustics, Speech and Signal Processing*. IEEE. 2010.
- [119] J. C. RUIZ-RODRÍGUEZ et al. « Innovative Continuous Non-Invasive Cuffless Blood Pressure Monitoring Based on Photoplethysmography Technology ». In : *Intensive Care Medicine*. Springer (2013).
- [120] R. SAMENI, C. JUTTEN et M. B. SHAMSOLLAHI. « A Deflation Procedure for Subspace Decomposition ». In : *Signal Processing*. IEEE (2010).
- [121] A. SCHÄFER et K. W. KRATKY. « Estimation of Breathing Rate from Respiratory Sinus Arrhythmia : Comparison of Various Methods ». In : *Annals of Biomedical Engineering*. Springer (2008).
- [122] S. A. SCHUCKERS. « Spoofing and Anti-Spoofing Measures ». In : *Information Security Technical Report*. Elsevier (2002).
- [123] B. SENOUCI et al. « Fast Prototyping of a SoC-Based Smart Camera :A Real-Time Fall Detection Case Study ». In : *Journal of Real-Time Image Processing*. Springer (2014).
- [124] B. SHAH et al. « Predictors of In-Hospital Mortality of Acute Ischemic Stroke in Adult Population ». In : *Journal of Neurosciences in Rural Practice*. NCBI (2017).
- [125] S. A. SHAH et al. « Respiratory Rate Estimation during Triage of Children in Hospitals ». In : *Journal of Medical Engineering & Technology*. Taylor & Francis Online (2015).
- [126] K. H. SHELLEY. « Photoplethysmography : Beyond the Calculation of Arterial Oxygen Saturation and Heart Rate ». In : *Anesthesia & Analgesia*. IARS (2007).
- [127] C.-L. SHEN et al. « Respiratory Rate Estimation by Using ECG, Impedance, and Motion Sensing in Smart Clothing ». In : *Journal of Medical and Biological Engineering*. Springer (2017).
- [128] I. S. SIDOROV, M. A. VOLYNSKY et A. A. KAMSHILIN. « Influence of Polarization Filtration on the Information Readout from Pulsating Blood Vessels ». In : *Biomedical Optics Express*. OSA (2016).
- [129] A. J. SINGER et al. « Quick SOFA Scores Predict Mortality in Adult Emergency Department Patients With and Without Suspected Infection ». In : *Annals of Emergency Medicine*. PubMed (2017).

- [130] K. SOBOTKA et I. PITAS. « Segmentation and Tracking of Faces in Color Images ». In : *International Conference on Automatic Face and Gesture Recognition*. IEEE. 1996.
- [131] R. SONG et al. « Heart Rate Estimation from Facial Videos Using a Spatiotemporal Representation with Convolutional Neural Networks ». In : *Instrumentation and Measurement*. IEEE (2020).
- [132] J. SPIGULIS. « Multispectral, Fluorescent and Photoplethysmographic Imaging for Remote Skin Assessment ». In : *Sensors*. MDPI (2017).
- [133] R. STOJANOVIC et D. KARADAGLIC. « A LED-LED-Based Photoplethysmography Sensor ». In : *Physiological Measurement*. IOP Publishing (2007).
- [134] E. STRICKLAND. « Facebook Takes on Deepfakes ». In : *Spectrum*. IEEE (2020).
- [135] S. SUN et al. « Systolic Blood Pressure Estimation Using PPG and ECG during Physical Exercise ». In : *Physiological Measurement*. IOP Publishing (2016).
- [136] X. SUN et al. « SleepMonitor : Monitoring Respiratory Rate and Body Position During Sleep Using Smartwatch ». In : *Interactive, Mobile, Wearable and Ubiquitous Technologies*. ACM Journals (2017).
- [137] X. SUN et al. « Robust Heart Beat Detection From Photoplethysmography Interlaced with Motion Artifacts Based on Empirical Mode Decomposition ». In : *International Conference on Biomedical and Health Informatics*. IEEE. 2012.
- [138] Y. SUN. « Motion-Compensated Noncontact Imaging Photoplethysmography to Monitor Cardiorespiratory Status during Exercise ». In : *Journal of Biomedical Optics*. SPIE (2011).
- [139] Y. SUN et al. « Use of Ambient Light in Remote Photoplethysmographic Systems : Comparison between a High-Performance Camera and a Low-Cost Webcam ». In : *Journal of Biomedical Optics*. SPIE (2012).
- [140] D. F. SWINEHART. « The Beer-Lambert Law ». In : *Journal of chemical education*. ACS Publications (1962).
- [141] C. TAKANO et Y. OHTA. « Heart Rate Measurement Based on a Time-Lapse Image ». In : *Medical Engineering & Physics*. Elsevier (2007).
- [142] T. TAMURA et al. « Wearable Photoplethysmographic Sensors—Past and Present ». In : *Electronics*. MDPI (2014).
- [143] L. TARASSENKO et al. « Non-Contact Video-Based Vital Sign Monitoring Using Ambient Light and Auto-Regressive Models ». In : *Physiological Measurement*. IOP Publishing (2014).
- [144] R. S. THAKUR. *Vital Sign Monitoring Based on Remote PPG and WIFI*. California State University, Sacramento, 2019.
- [145] J. E. THATCHER et al. « Multispectral and Photoplethysmography Optical Imaging Techniques Identify Important Tissue Characteristics in an Animal Model of Tangential Burn Excision ». In : *Journal of Burn Care & Research*. Oxford Academic (2016).
- [146] R. THAVOT et al. « Dataflow design of a co-processor architecture for image processing ». In : *Design and Architectures for Signal and Image Processing*. EPFL. 2008.
- [147] J. F. THAYER et al. « Estimating Respiratory Frequency from Autoregressive Spectral Analysis of Heart Period ». In : *Medicine and Biology Magazine*. IEEE (2002).
- [148] A. TRUMPP et al. « The Value of Polarization in Camera-Based Photoplethysmography ». In : *Biomedical Optics Express*. OSA (2017).



- [149] S. TULYAKOV et al. « Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions ». In : *Computer Vision and Pattern Recognition Workshops*. IEEE. 2016.
- [150] M. van GASTEL, S. STUIJK et G. de HAAN. « Motion Robust Remote-PPG in Infrared ». In : *Biomedical Engineering*. IEEE (2015).
- [151] M. van GASTEL, S. STUIJK et G. de HAAN. « Robust Respiration Detection from Remote Photoplethysmography ». In : *Biomedical Optics Express*. OSA (2016).
- [152] W. VERKRUYSSSE, L. O. SVAASAND et J. S. NELSON. « Remote Plethysmographic Imaging Using Ambient Light ». In : *Optics Express*. OSA (2008).
- [153] L. WANG et A. WONG. « COVID-Net : A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images ». In : *Electrical Engineering and Systems Science* (2020).
- [154] W. WANG, B. BALMAEKERS et G. de HAAN. « Quality Metric for Camera-Based Pulse Rate Monitoring in Fitness Exercise ». In : *International Conference on Image Processing*. IEEE. 2016.
- [155] W. WANG, S. STUIJK et G. de HAAN. « A Novel Algorithm for Remote Photoplethysmography : Spatial Subspace Rotation ». In : *Biomedical Engineering*. IEEE (2016).
- [156] W. WANG, S. STUIJK et G. de HAAN. « Exploiting Spatial Redundancy of Image Sensor for Motion Robust rPPG ». In : *Biomedical Engineering*. IEEE (2015).
- [157] W. WANG et al. « Algorithmic Principles of Remote PPG ». In : *Biomedical Engineering*. IEEE (2017).
- [158] W. WANG et al. « Robust Heart Rate from Fitness Videos ». In : *Physiological Measurement*. IOP Publishing (2017).
- [159] D. WIDJAJA et al. « Application of Kernel Principal Component Analysis for Single-Lead-ECG-Derived Respiration ». In : *Biomedical Engineering*. IEEE (2012).
- [160] A. D. WIENS et al. « Toward Continuous, Noninvasive Assessment of Ventricular Function and Hemodynamics : Wearable Ballistocardiography ». In : *Biomedical and Health Informatics*. IEEE (2015).
- [161] B.-F. WU et al. « Neural Network Based Luminance Variation Resistant Remote-Photoplethysmography for Driver’s Heart Rate Monitoring ». In : *IEEE Access* (2019).
- [162] J.-X. WU et al. « Bilateral Photoplethysmography Analysis for Peripheral Arterial Stenosis Screening With a Fractional-Order Integrator and Info-Gap Decision-Making ». In : *Sensors*. IEEE (2016).
- [163] J.-M. WU et al. « A Unified Framework for Automatic Detection of Wound Infection with Artificial Intelligence ». In : *Applied Sciences*. MDPI (2020).
- [164] J. XIE et al. « Association Between Hypoxemia and Mortality in Patients With COVID-19 ». In : *Mayo Clinic Proceedings*. Elsevier (2020).
- [165] K. XIE et al. « Non-Contact Heart Rate Monitoring for Intensive Exercise Based on Singular Spectrum Analysis ». In : *Multimedia Information Processing and Retrieval*. IEEE. 2019.
- [166] W.-Y. XU, J. MENG et X.-M. ZHAO. « Non-contact monitoring of ambulatory blood pressure based on high speed camera ». In : *Engineering Science*. ZheJiang University (2017).
- [167] K. YAMAMOTO et al. « Health Observation App for COVID-19 Symptom Tracking Integrated With Personal Health Records : Proof of Concept and Practical Use Study ». In : *Journal of Medical Internet Research*. JMIR Publications (2020).

- [168] M. K. YAPICI et T. E. ALKHIDIR. « Intelligent Medical Garments with Graphene-Functionalized Smart-Cloth ECG Sensors ». In : *Sensors*. MDPI (2017).
- [169] C. ZHAO et al. « A Novel Framework for Remote Photoplethysmography Pulse Extraction on Compressed Videos ». In : *Computer Vision and Pattern Recognition Workshops*. IEEE. 2018.
- [170] J. ZHENG et al. « Feasibility of Imaging Photoplethysmography ». In : *BioMedical Engineering and Informatics*. IEEE. 2008.
- [171] J. ZHENG et al. « Remote Simultaneous Dual Wavelength Imaging Photoplethysmography : A Further Step towards 3-D Mapping of Skin Blood Microcirculation ». In : *Multimodal Biomedical Imaging III*. SPIE. 2008.