Blind source separation of single-sensor recordings : Application to ground reaction force signals

Ramzi El Halabi

N°d'ordre NNT : 2018LYSES 031

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
L'Université Jean Monnet de Saint-Etienne

Ecole Doctorale N° 488
Sciences Ingénierie Santé

Spécialité :

Image, Vision et Santé

Soutenue publiquement le 19/10/2018, par :
Ramzi EL HALABI

BLIND SOURCE SEPARATION OF SINGLE-SENSOR RECORDINGS: APPLICATION TO GROUND REACTION FORCE SIGNALS

Devant le jury composé de :

GUILLET, François, Professeur, IUT de Roanne	*Président*
CARRAULT, Guy, Professeur, Université Rennes 1	*Rapporteur*
NAIT-ALI, Amine, Professeur, Université Paris-Est Créteil	*Rapporteur*
SERHAL, Dina, Maîtresse de conférences, Université Rafik Hariri	*Examinatrice*
MOSLEM, Bassam, Maître de conférences, Université Rafik Hariri	*Examinateur*
O DIAB, Mohamed, Professeur, Université Rafik Hariri	*Examinateur*
EL BADAOUI, Mohamed, Professeur, IUT de Roanne	*Directeur de thèse*

BLIND SOURCE SEPARATION OF SINGLE-SENSOR RECORDINGS: APPLICATION TO GROUND REACTION FORCE SIGNALS

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Engineering

At

UNIVERSITY OF LYON, JEAN MONNET SAINT-ÉTIENNE UNIVERSITY, LASPI, FRANCE

Under a Codirecting Convention with

RAFIK HARIRI UNIVERSITY- COLLEGE OF ENGINEERING –

ELECTRICAL AND COMPUTER ENGINEERING DEPARTMENT

BY
HALABI Ramzi

ROANNE, 2018

ACKNOWLEDGEMENT

First and foremost, it is my duty as a human being and from the deepest point of my heart, I would like to thank Allah (SWT) for all that I have been blessed with before and during the past few years, from a supportive family to a lovely wife to a wonderful baby girl throughout an unforgettable PhD journey that took a huge effort to accomplish.

Second of all, I would like to state my deep gratitude to all of my advisors Dr. Mohamad Diab, Dr. Mohamed El Badaoui, Dr. Bassam Moslem and Dr. Francois Guillet, for being a true backbone for this work to get accomplished the way it did; not only I have had the chance to share technical knowledge with them, but I have been blessed with a team of patient and caring advisors who sacrificed their time to provide advice and lead the way.

Thirdly, the administrative team around me throughout this work has been outstandingly supportive, starting with the Rafik Hariri University administration, to each and every member of the LASPI at UJM from administrative assistants to IT technicians to doctoral school administrators.

Last but not least, I would like to express deep appreciation to my dearest mother, brother, aunt and wife for their immense support that words may not describe, being there for me during my hard times and keeping my hopes high to proceed. Not to forget my mentor and one of my role models, Mr. Samir Berjaoui, my ex-boss at RHU, who always provided support and encouragement, both technically and spiritually. May Allah (SWT) bless them all and fill their hearts with joy and guide them to the best of this life and the afterlife.

UN RÉSUMÉ DE LA THÈSE DE

Ramzi Halabi Pour obtenir le grade de Docteur en Science de l'Ingénieur

Titre de la thèse : Séparation Aveugle de Sources des Signaux Monocanaux: application aux signaux de Force de Réaction de Terre.

Les signaux multicanaux sont réellement des signaux captes à travers plusieurs canal ou capteur, portant chacun un mélange de sources, une partie desquelles est connue alors que le reste des sources reste inconnu, et parfois complètement inconnues. Les méthodes à l'aide desquelles l'isolement ou la séparation des sources est accomplie sont connues par les méthodes de séparation de sources en général, et au cas où le degré d'inconnu est large, ces méthodes sont connues par la séparation aveugle des sources (SAS).

Cependant, la SAS appliquée aux signaux multicanaux est en fait plus facile de point de vue mathématique que l'application de la SAS sur des signaux monocanaux, ou un seul capteur existe et tous les signaux arrivent au même point pour enfin produire un mélange de sources inconnues. Ce domaine est effectivement celui de la thèse que nous présentons. En effet, nous avons développé une nouvelle technique de SAS qui est une combinaison de plusieurs méthodes de séparation et d'optimisation, basée sur la factorisation non-négative des matrices (NMF). Cette méthode peut être utilisée dans de nombreux domaines comme l'analyse des sons et de la parole, les variations de la bourse, et même les séismographes. Néanmoins, dans cette thèse, les signaux de force de réaction de terre verticaux (VGRF) monocanaux d'un groupe d'athlètes coureurs d'ultra-marathon sont analysés et séparés pour l'extraction du *peak* passif du *peak* actif d'une nouvelle manière adaptée à la nature de ces signaux.

Les signaux VGRF sont des signaux cyclo-stationnaires caractérisés par des *double-peaks* l'un desquels étant très rapide et parcimonieux, indiquant les phases de course de l'athlète, et l'analyse de ces *peaks* est extrêmement importante pour déterminer et prédire la condition du coureur : problème physiologique, problème anatomique, fatigue…etc. De plus, un grand nombre de chercheurs ont prouvé que l'impact du pied postérieur avec la terre d'une manière brutale, l'analyse de ce phénomène peut nous ramener à une prédiction de blessure interne, et meme ils essayent d'adopter une technique de course dont ils ont appelé *Non-Heel-strike Running (NHS),* par laquelle ils obligent les coureurs de courir sur le pied-antérieur seulement. Donc pour mieux étudier ce phénomène, la séparation du *peak* d'impact du VGRF est une méthode à travers laquelle on peut isoler la source portant les informations patho-physiologiques et le degré de la fatigue.

D'un premier coup, nous avons introduit de nouvelles méthodes de prétraitement et de traitement des signaux VGRF pour remplacer le filtrage de bruit traditionnel utilisé partout, et qui peut parfois détruire les *peaks* d'impact qui sont nos sources à séparer, base sur le concept de soustraction spectrale pour le filtrage, utilisée avec les signaux de parole, après l'application d'un algorithme d'échantillonnage intelligence et adaptatif qui décompose les signaux en pas isolés.

En second lieu, une analyse des signaux VGRF en fonction du temps a été faite pour la détection et la quantification de la fatigue des coureurs durant les 24 heures de courses. Cette analyse a été accomplie au domaine fréquentiel/spectral ou nous avons détecté un décalage clair du contenu fréquentiel avec la progression de la course indiquant la progression de la fatigue.

Troisièmement, nous avons défini les signaux *cyclosparse* au domaine temporel, puis on a traduit cette définition à son équivalent au domaine temps-fréquence utilisant la transformée

Fourier a court-temps (STFT). Cette représentation a été décomposée a travers une nouvelle méthode qu'on a appelé *Cyclosparse Non-negative Matrix Factorisation* (Cyclosparse-NMF), basée sur l'optimisation de la minimisation de la divergence Kullback-Leibler (KL) avec pénalisation liée à la périodicité et la parcimonie des sources, ayant comme but final d'extraire les sources *cyclosparse* du mélange monocanal appliquée aux signaux VGRF monocanaux.

La méthode a été testée de plus sur des signaux analytiques pour pouvoir prouver l'efficacité de l'algorithme et les résultats ont été satisfaisants, et le *peak* impact a été séparé du mélange VGRF monocanal.

AN ABSTRACT OF THE THESIS OF

Ramzi Halabi for the degree of Doctor of Philisophy

Title of the Thesis : Blind Source Separation of Single-sensor Recordings : Application to Ground Reaction Force Signals.

The purpose of the presented work is to develop a customized Single-channel Blind Source Separation technique that aims to separate cyclostationary and transient pulse-like patterns/sources from a linear instantaneous mixture of unknown sources.

For that endeavor, synthetic signals of the mentioned characteristic were created to confirm the separation success, in addition to real life signals acquired throughout an experiment in which experienced athletes were asked to participate in a 24-hour ultra-marathon in a lab environment on an instrumented treadmill through which their VGRF, which carries a *cyclosparse* Impact Peak, is continuously recorded with very short discontinuities during which blood is drawn for in-run testing, short enough not to provide rest to the athletes. The synthetic and VGRF signals were then pre-processed, processed for Impact Pattern extraction via a customized Single-channel Blind Source Separation technique that we termed Cyclo-sparse Non-negative Matrix Factorization and analyzed for fatigue assessment.

As a result, the Impact Patterns for all of the participating athletes were extracted at 10 different time intervals indicating the progression of the ultra-marathon for 24 hours, and further analysis and comparison of the resulting signals proved major significance in the field of fatigue assessment; the Impact Pattern power monotonically increased for 90% of the subjects by an average of 24.4 ± 15% with the progression of the ultra-marathon during the 24-hour period.

Upon computation of the Impact Pattern separation algorithm, fatigue progression showed to be manifested by an increase in reliance on heel-strike impact to push to the bodyweight as a compensation for the decrease in muscle power during propulsion at toe-off.

This study among other presented work in the field of VGRF processing forms methods that could be implemented in wearable devices to assess and track runners' gait as a part of sports performance analysis, rehabilitation phase tracking and classification of healthy vs. unhealthy gait.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AOS	Alternating Optimization Strategy
BSS	Blind Source Separation
SCBSS	Single-channel Blind Source Separation
NMF	Non-negative Matrix Factorization
VGRF	Vertical Ground Reaction Force
NHS	Non-heel-strike
STFT	Short-time Fourier Transform
KL	Kullback-Leibler
COG	Center of Gravity
GRF	Ground Reaction Force
EMG	Electromyography
VCR	Video Cassette Recorder
APGRF	Antero-posterior Ground Reaction Force
MLGRF	Medio-lateral Ground Reaction Force
COP	Center of Pressure
HS	Heel-strike
PP	Passive Peak
SS:	Source Separation
AI	Artifical Intelligence
ASA	Auditory Scene Analysis
f-MRI	Functional Magnetic Resonance Imaging
LTI	Linear Time-invariant
PNL	Post Non-linear

CMM	Convolutive Mixture Model
FFT	Fast Fourier Transform
DFT	Discrete Fourier Transform
FIR	Finite Impulse Response
IIR	Infinite Impulse Response
SOS	Second Order Statistics
HOS	Higher-order Statistics
JADE	Joint Approximate Diagonalization of Eigenmatrices
PDF	Probability Density Function
HMM	Hidden Markov Model
MAP	Maximum a posteriori Probability
CS	Cyclostationarity
FFT	Fast Fourier Transform
DFT	Discrete Fourier Transform
PSD	Power Spectral Density
IDFT	Inverse Discrete Fourier Transform
MHPV	Mean High-frequency Power Variation
SMIPS	Spectral Mean Inter-peak Shift
TF	Time-frequency
CASA	Computational Auditory Scene Analysis
EMD	Empirical Mode Decomposition
IMF	Intrinsic Mode Function
ISA	Independent Subspace Analysis
ICA	Independent Component Analysis

GENERAL INTRODUCTION

Blind source separation (BSS) is a challenging yet highly required field of research in signal and image processing, in Artificial Intelligence and in Big Data Analytics. It studies the unmixing and separation of signal or pattern mixtures within a complex multivariate system, having minimal to no a priori knowledge of the target source. The target source is usually a buried pattern of valuable data into a mixture of different sorts of patterns, generally considered independent, that shall be separated and recovered for further analysis. Despite the tremendous effort being put into the development of BSS algorithms, the area is still being developed for customized techniques that are precisely designed for specific classes of data.

The work being presented throughout this thesis is a contribution into the field of BSS, and more specifically single-channel BSS (SCBSS), where a single observation/data vector is available as a muti-component mixture, one of which shall be recovered. Upon reviewing the available techniques, we realized the near-absence of methods that are tailored to tackle cyclostationary single-channel signals. That fact stimulated the design of the Cyclosparse Non-negative Matrix Factorization (NMF), which is an adapted version of the well-known NMF set of algorithms. The thesis not only presents the theoretical aspect of the method, but also presents the results of applying the methods of processing, analysis and SCBSS to the Vertical Ground Reaction Force (VGRF) signals.

With the fact that the frequency of falls increases with age and frailty level, statistical studies prepared by the World Health Organization stated the following rates: 6-31% in China, 21.6% in Barbados, 34% in Chile, while the rate of hospital admissions due to falls at the age of 60 and older reached 1.6 to 3 per 10,000 in Australia, Canada, and the United Kingdom. This led

numerous research scientists to dedicate their efforts in predicting falling through Gait analysis. Being one of our two main goals, the presented work in this paper aims to model the running VGRF in a highly dynamic activity where the loss of balance is more probable and of greater significance, leading to a better understanding of the normal running patterns in contrast with loss-of-balance patterns, if any. Furthermore, in the field of athletic performance assessment, running is a common activity for a wide range of elite sports such as marathons, sprints, soccer, football, rugby and many others, whose athletes suffer from a relatively high foot and knee injury rate. Considering the cases where the cause of eventual injury is gradual deterioration of the anatomical system through which running is performed e.g. foot and knee, an accurate model describing the different patterns in a VGRF signal provides a clearer view of the progression of the injury, thereby the underlying biomechanical manifestations serving as tools to be analyzed by medical experts for injury evolution prevention.

Chapter 1 in this thesis consists of a well-rounded summary on the two main domains that are addressed throughout the thesis which are Gait analysis and Blind Source Separation (BSS), since the eventual goal is the BSS of the Vertical Ground Reaction Force (VGRF) signal's patterns. Furthermore, light is shed on a specific class of BSS techniques, which is the single-channel BSS (SCBSS) that deals with observations/mixture signals recorded on a single channel; although it is still a BSS class, the implemented algorithms require special attention and familiarization.

Chapter 2 provides the first and most important stage in the process that leads the way to subsequent analysis, and that is the signal processing phase applied on the signals of interest in this thesis being the VGRF signals. This stage is actually a multi-stage process that goes as follows: Pre-processing (sampling and segmentation) and processing (filtering). Furthermore,

new methods for fatigue detection from those signals are presented. Novel techniques in VGRF processing and analysis are presented in this chapter, which were published in 3 international conferences.

Chapter 3 introduces and goes in-depth into the novel SCBSS technique Cyclosparse Non-negative Matrix Factorization (NMF) in a block-based manner: Problem formulation, followed by the cost function's design along with the formulation of the penalty functions within, and finally the optimization algorithm.

Chapter 4 in this thesis is where all results in terms of pre-processing, processing and analysis are presented, analyzed and discussed in a way to justify the answers this thesis tackle. Not only were the methods applied on real acquired data (VGRF) but also on synthetic signals to prove the effectiveness of the method and unbiased aspect of it.

Chapter 5 is the concluding chapter where general conclusions are stated and future perspectives and improvement proposals are presented, along with the obstacles that were encountered throughout the process and possible solutions to avoid getting into them in future applications.

CHAPTER 1

LITERATURE REVIEW

1.1 Gait Analysis

1. 1. 1 Definition

Walking and Gait are commonly interchangeably used while there is a major difference between these two terms; normal walking and running are defined as methods of locomotion involving the use of two legs, alternately, to provide support and propulsion [1]. On the other hand, Gait is defined as the manner or style of walking rather than the process itself, which makes Gait analysis a field in which comparison between individuals is conducted.

That being cleared, the Gait Cycle is defined as the time interval between two successive occurrences of one of the repetitive events of walking or running, most commonly considering the initial contact as the reference point for the cycle [2]. The conventional way is to start with the initial contact of the right foot with the ground until it contacts the ground again, which is depicted in Fig. 1.1.

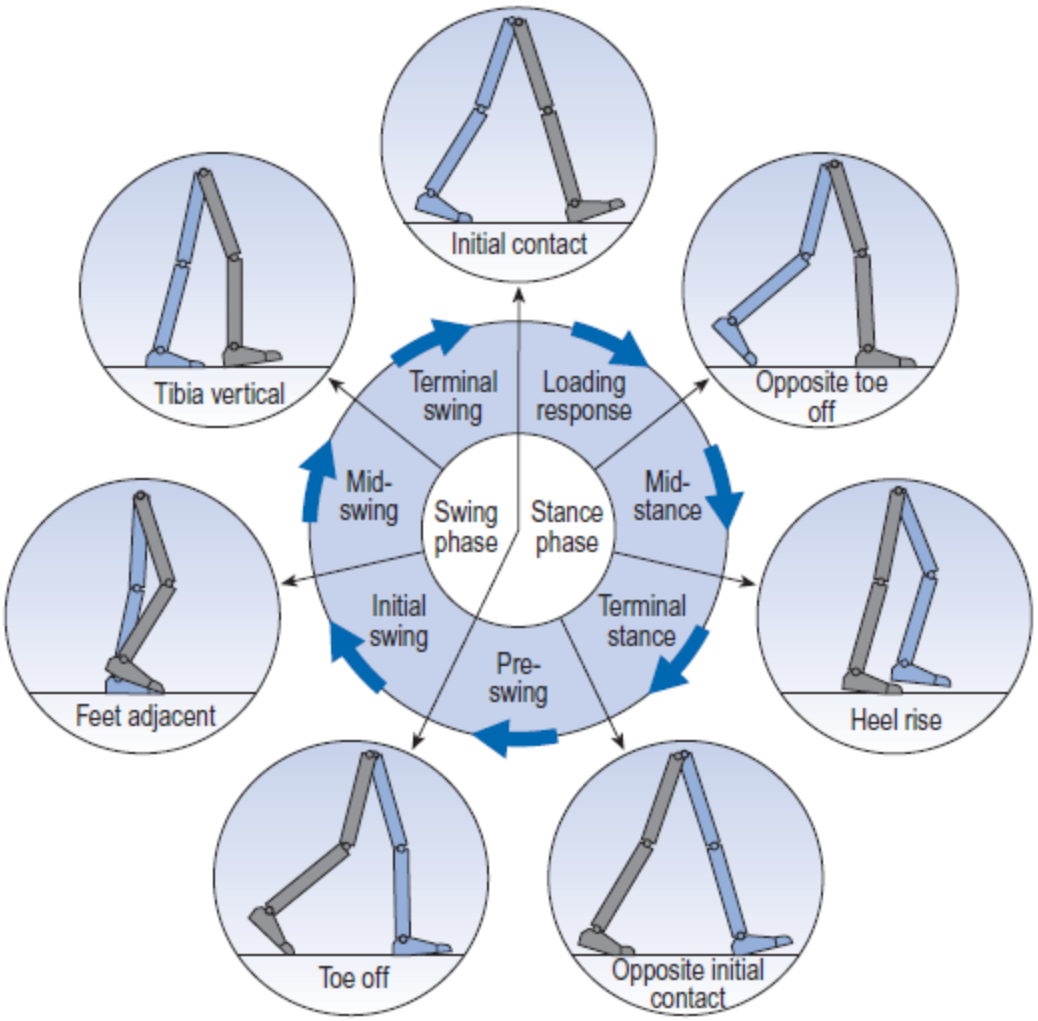

Figure 1.1: Gait Cycle Illustration (image courtesy Michael W. Whittle's "An Introduction to Gait Analysis")

The gait cycle phases are stated as follows:

- Initial Contact
- Opposite toe-off
- Heel-rise
- Opposite initial contact
- Toe-off
- Feet adjacent

- Vertical Tibia
- Repeat.

The stated events sum up to seven total phases constituting the gait cycle, the first four of which occur in the stance phase with the foot touching the ground, while the other three belong to the swing phase in which the foot is not in contact with the ground.

Moreover, each of the phases (stance and swing) is likewise subdivided into phases; the contact phase or support phase that covers the phases from initial contact till toe-off is subdivided into four stages [3]:

- Loading response
- Mid-stance
- Terminal stance
- Pre-swing

As for the swing phase, three phases exist:

- Initial swing
- Mid-swing
- Terminal swing

In fact, the time to complete a whole gait cycle is termed the Cycle Time in the majority of the reputable literature, disregarding the minor differences that exist in the field while referring to the same concepts and mechanisms in different words and terminology.

Now concerning the timing aspect of gait and the synchronization of shifts between different phases and sub-phases that we discussed earlier, see Fig.1.2.

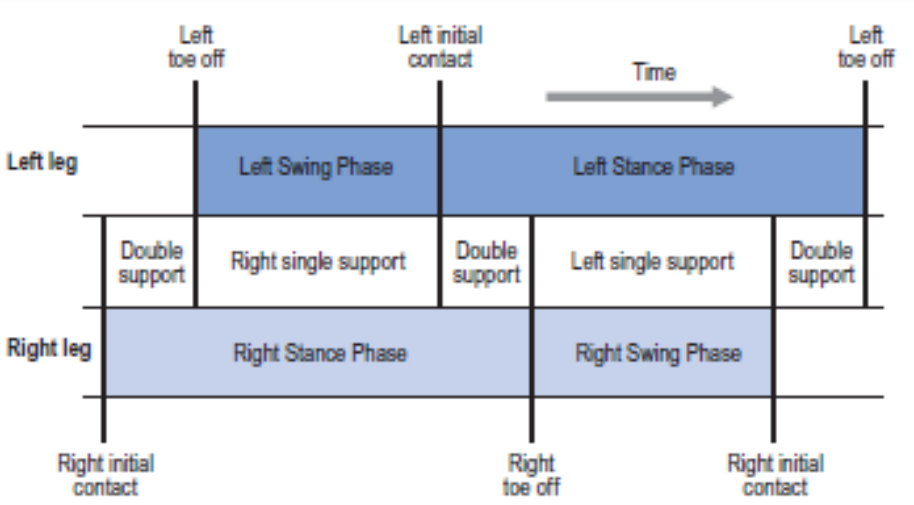

Figure 1.2: Timing Aspect of the Gait Cycle (image courtesy Michael W. Whittle's "An Introduction to Gait Analysis")

As shown in Fig.1.2, each gait cycle contains two periods of double support and two phases of single-support phase. The stance phase occupies around 60% of the gait cycle while the swing phase occupies the remaining 40%, out of which each double support phase lasts for about 10%. Noteworthy, the stated percentages are variable in function of walking speed (walking, speed-walking and running), such that the swing phase varies proportionally with speed while the stance and double support phases vary inversely proportionally, which was reported in [4] until complete absence of double-stance phases when transitioning into running. The periods in between successive steps during which the runner's feet are not in contact with the ground are termed flight phases/float/double-float/non-support.

In order to describe foot placement on the ground, two main terms are used: step and stride [5]. The step length is defined as the distance at which one foot moves forward in front of the other foot, while the stride length is the distance between two successive placements of the same foot i.e. cycle duration [6]. The measurement of such lengths aids in the understanding and quantification of the differences between normal and pathological gait, which may show manifestations of hopping, or step length variability in the same stride.

Moreover, the number of performed steps during a given duration is termed cadence, having the unit of steps per minute, which is actually a metric related to the half-cycles rather than full cycles that are defined as strides. However, the cycle-time is often used instead of the cadence, being the inverse of cadence, also known as the cycle time (1.1).

$$cycle\ time\ (s) = \frac{120}{cadence\ \left(\frac{steps}{\min}\right)} \tag{1.1}$$

In addition to the cycle time, gait is quantified and characterized by speed, which is defined as the distance covered by the whole body in a given time, most commonly in meters per second. This parameter may be calculated as the product of cadence and stride length with proper unit usage (1.2).

$$speed\ \left(\frac{m}{s}\right) = stride\ length\ (m)\ \times \frac{cadence\ (\frac{steps}{\min})}{120} \tag{1.2}$$

Alternatively, if cycle time is used instead of cadence, the equation becomes (1.3):

$$speed\ \left(\frac{m}{s}\right) = \frac{stride\ length\ (m)}{cycle\ time\ (s)} \tag{1.3}$$

1.1.2 History

Actually, history shows that the study of Gait appeared from as early as the Renaissance when it got accurately studied by scientists e.g. Leonardo Da Vinci, Galileo and Newton described walking. However, the earliest considerable scientific approach was performed by Galileo's student Borelli in 1682 when he published his famous *De Motu Animalum*. His method consisted of tracking the center of gravity (COG) of the body to interpret how balance is maintained during

walking via constant forward movement of the feet being the support area. Borelli's effort was followed by the Weber brothers who first described the gait cycle in 1836 [7] by taking precise timing measurements considering a pendulum-like model of the leg of cadaver. Later on during the 1870s, Marey and Muybridge were the pioneers in the field; Marey published a study of human limb movement via photographic analysis of subjects wearing black and limb-attached white markers while standing on a platform at which the exerted pressure as well as the location of the COG. As for Muybridge, analysis was performed on horse trotting in 1878 based on a 24-camera photographic system with fast trigger for successive frames, and his studies progressed onto human beings walking and running and performing various activities [8].

Figure 1.3: Photographic Gait Markers on Horse

At a later stage in 1895, a major breakthrough in gait analysis was published by Braune and Fischer who implemented a photographic technique based on limb-attached fluorescent markers, which led to the first three-dimensional trajectory, velocity and acceleration analysis of body

segments, combined with masses and accelerations of the body segments to eventually estimate the walking cycle stages' forces.

Figure 1.4: Limb-attached Fluorescent Markers

Valuable and significant work was further developed by Bernstein in the 1930s where he performed a multitude of photographic methods to further assess kinematic parameters of a large 150-subject database mainly focusing on the limb-specific COG and overall body [9].

One of the greatest breakthroughs in the field of gait analysis was the development of the force platform or force plate that was conceptualized in 1924 by Amar and enhanced and implemented in 1938 by Elftman [10], both systems being purely mechanical; the first design consisted of pointer movement in response to force application to the platform, while the latter consisted of high-speed photography of the basic pointer.

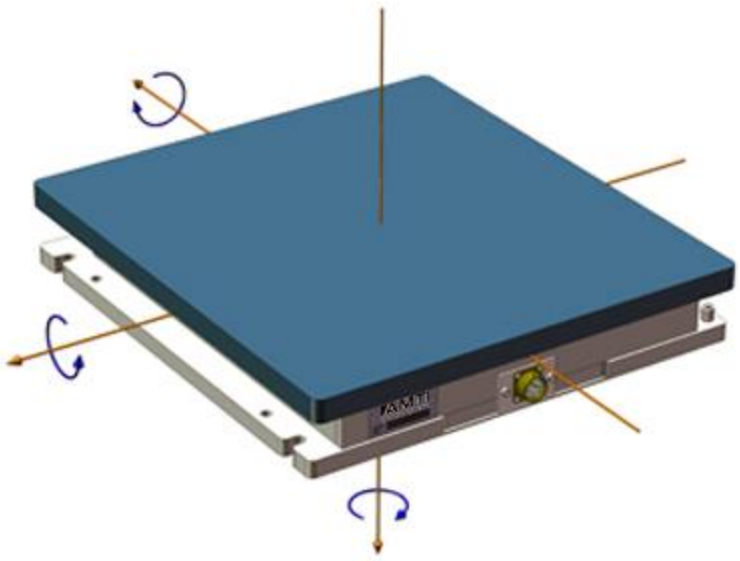

Figure 010.5: 3-dimensional Force Plate

Later on in the 1950s, Bresler and Frankel computed free-body calculations of the main concerned joints being the hip, knee and ankle which led to the calculation of ground reaction forces (GRF) that are the effects of gravitational forces on the limb segments and inertial forces [11].

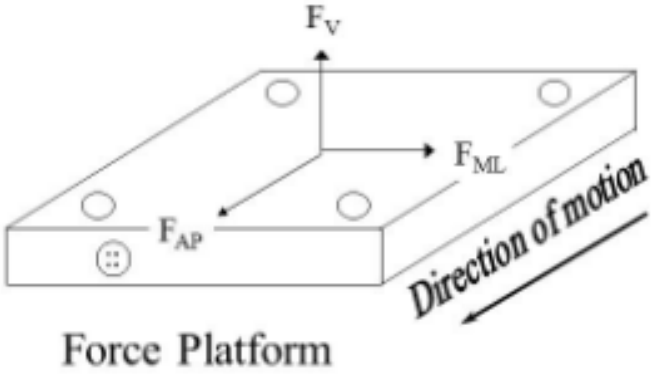

Figure 01.6: 3-dimensional Force Platform

The body-segments' actions and the role of the different muscles became clearly modeled and documented, which transferred the researchers' interest into the quantification and modeling of

the underlying forces being generated at the joints. One of the pioneers in that field was Paul who first calculated hip joint forces as well as knee joint forces in the 1965-66 period, which opened the door for extensive mathematical research in the modeling of joint forces.

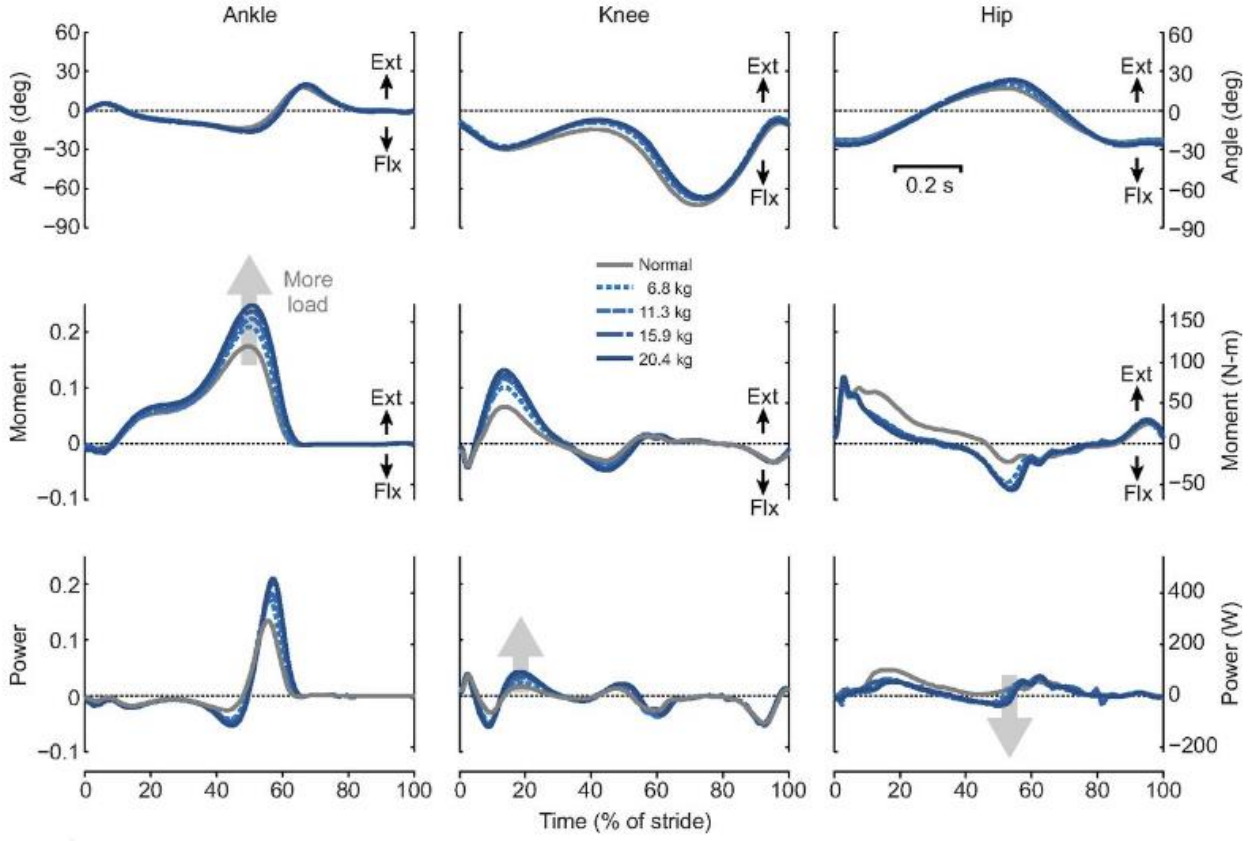

Figure 01.7: Jont Forces' Graphical Representation

For that endeavor, more precise hardware and computational tools were required, which led the way to precision electronics instead of basic photography e.g. reliable force platforms [12], instrumented treadmills [13], Electromyography (EMG) [14], 3D kinematics modeling software [15]…etc. The main methods used for gait measurement and analysis will be presented in section 1.1.2.

1.1.3 Methods

Gait analysis has been studied for decades, as described in the history description above, and this implies the use of multiple measurement and analysis hardware and software schemes to cope with the technological era in which the study is being performed. However, it is incorrect to say that some methods are better or more reliable than others, each of which having advantages and disadvantages. Moreover, the availability of high-end technologies in this field among others does not imply clinicians and experts are obliged to perform their clinical studies using the top-notch equipment; efficiency being crucial in all domains, one may not have to deal with the huge cost of gait analysis equipment if the applications or experiments do not require such advanced tools and low-end accurate tools would suffice.

In this section, the different techniques implemented in gait analysis will be covered including the state-of-the-art technology.

1.1. 3. a Photographic

Being the least technologically-complicated gait analysis system of all, photographic methods are the most challenging for the clinician or the researcher requiring vast knowledge and observation accuracy and definitely experience in the field to be able to analyze the recorded gait sequence and diagnose. However, research has shown that the performance of human observers versus gait analysis combined kinetic/kinematic systems is weaker in terms of abnormality detection. Ergo, the visual observation or video observation technique with no advanced software to accompany the system or advanced sensing tools is superficial and suffers from many disadvantages, of which we state: Short duration of recordings limited by the observers

observation span, incapability of detecting high-frequency events, limitation of analysis to movements rather than forces, and dependence on the observer's skills.

In fact, the manner in which this type of gait observation is performed is the distribution of different observers/video cameras around a subject performing gait activities e.g. walking, one to the right, the other to the left, a third one in front and a fourth on the backside, and observations are noted down for further analysis. The subject actually walks on a walkway that is 8 meters long on average, and 4 meters wide, as shown in Fig.1.8.

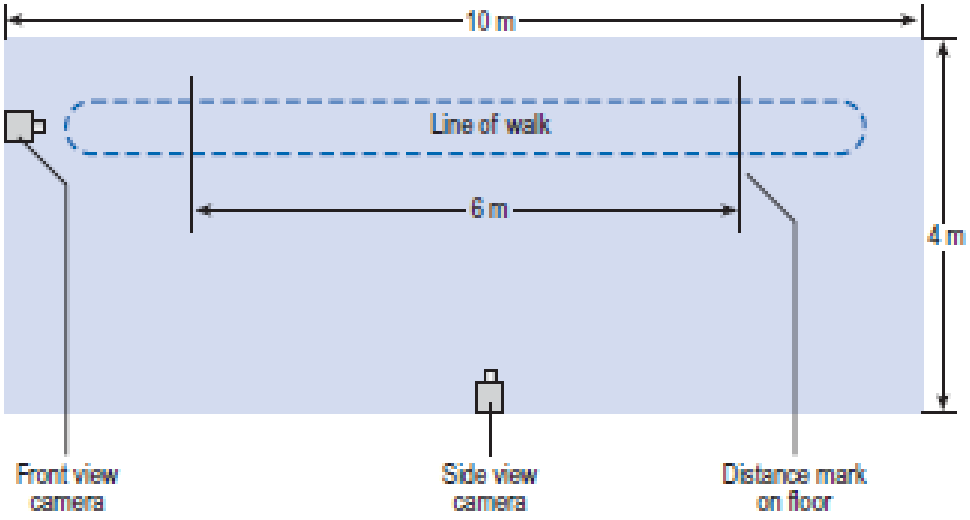

Figure 01.8: Gait Laboratory Layout for Visual Gait Analysis via Video Recording

Following the recording, assessment is definitely required for the analysis part to be fulfilled, which requires the data to be stored and remain permanently available. In the previously explained video observation, before the invention of the Video Cassette Recorder (VCR) and other video storage devices, many disadvantages were stated, whereas with data saving capabitlities developed in later stages, the following added advantages were realized: Reduced number of walk repetititons requested from the subject, ability of the subject to observe his/her

actual walk for more accurate feedback, and ease of transfer of gait analysis techniques to other practitioners in terms of teaching.

1.1.3. b Footswitches

In order to solely focus on the gait cycle timing, a multitude of systems has been developed, of which we state the footswitch, which is actually simple yet accurate in the endeavor that it was designed to fulfill. The actual architecture of the hardware is fairly simple: one switch is fixed underneath the heel of the subject for heel-strike detection, and a second footswitch is fixed underneath the forefoot. Using such technology, one may capture the timing of the following parameters: Initial contact, foot flat, heel-rise, toe-off and the duration of the stance phase, as shown in Fig. 1.9.

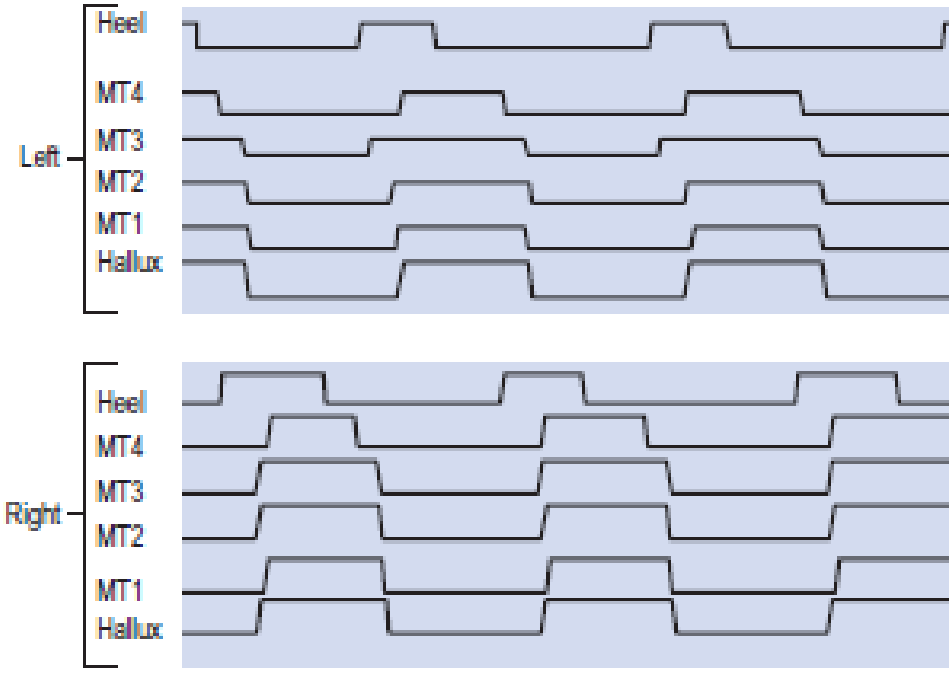

Figure 010.9:Footswitch System Sample Output

The high state indicates the occurrence of foot contact while the low state indicates the absence of contact, which makes it fairly simple to measure the pulse durations in order to compute the duration of steps, strides, phases…etc. Moreover, if the double support parameters are required, simultaneous sensor recordings from both feet is implemented. In order to be able to visualize the data in real-time, process it and interpret it, wireless communication interfaces are usually attached to the measurement hardware so that data gets wirelessly accessed in real-time.

As any other device, drawbacks exist, which are manifested as weak mechanical integrity. In fact, sensors like the footswitch are prone to getting broken at any instant due to the repetitive and continuous pressure load applied to it, being the subject walking or running. Despite the advancement of such designs by duplicating or adding thickness to the switch metals or so, this system is still weak and not completely reliable. However, the main advantage behind the use of such hardware is the minimal design complexity and the relatively compact size, making it embeddable in shoes and other footwear in a way not to create discomfort to the subject wearing it. See Fig.1.10 for a typical device configuration.

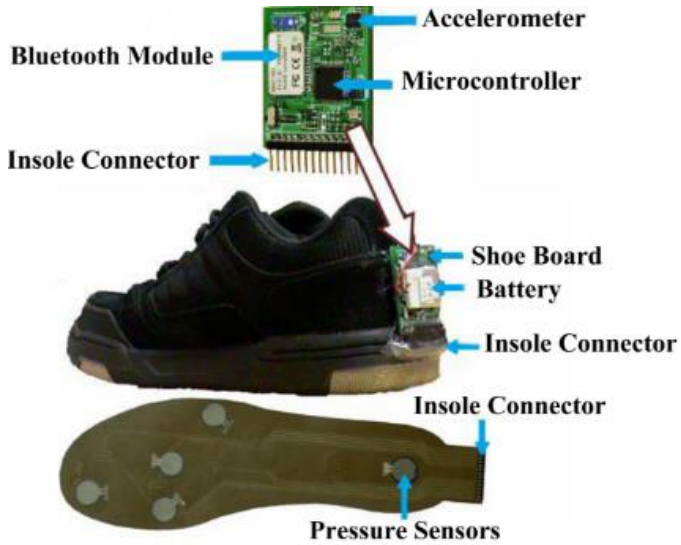

Figure 01.10: In-shoe Embedded Pressure Sensors with Acquisition and Transmission Circuitry

1.1.3. c Instrumented Walkway

As an additional feature to the presented footswitches that solely provide timing information, instrumented walkways are implemented to measure the timing and the positioning of the feet on the ground during gait. Fig. 1.11 shows a typical instrumented walkway along with the output's illustration.

Figure 01.11:Typical Instrumented Walkway

The design concept behind this technology is commonly based on an electrically-conductive e.g. copper, aluminum, a metalic mesh or even custom conductive rubber covering a walkway, which acts similarly to the footswitch with added positional value to output data. Whenever the subject steps on the walkway, contact is performed leading to a closed electrical loop that joins the foot to the conductive layer, detecting thereby the exact timing of the steps. Morevover, in order to determine the step and stride length, speed measurement via opto-interrruptor system is

measured, combined with time-related data, provides distance-related data through simple calculation that relates velocity to distance and time as in (1.4)

$$V\ (m/s)\ =\ \frac{d(m)}{t(s)} \tag{1.4}$$

Alternatively, a substitute design is commonly implemented, which functions according to the same concepts as the previously explained system, but having the footswitches embedded in the walkway itself as an array of swiches, which adds the localization feature based on the detected signals' respective sensors. One of the other advantages of this system is the fact that data needs not to be transferred from the shoe-embedded footswitches to the processing and analysis unit, in fact, the walkway itself is the center of data acquisition that transmits the data to the processing and analysis unts. Of the most popular designs of that functional concept, we state the GAITRite system developed by Bilney *et al.* in 2003 [16], which is being vastly used in research centers and advanced clinics.

1.1.3 d. Electrogoniometry

Electrogoniometry is the continuous measurement of the angle at a specific joint, providing thereby an output of angle per time, which aids in the illustration of the joint's kinematics with the help of computer software or even the use of the output data without further manipulation to track joint kinematics and assess them. The sensor configuration itself is obviously termed the Electrogoniometer. Furthermore, if measurement is performed on more than one joint, the scatter plot of unit angle versus angle may sometimes be computed to assess relative joint kinematics resulting in an angle-angle diagram or cyclogram.

Two main categories of Electrogoniometers are most commonly used: Potentiometer-based and Flexible Strain-gauge-based.

The rotary potentiometer-type electrogoniometer as the term indicates is based on a macro-scale variable resistor attached to the joint under-study where any movement of the adjacent body segments surrounding the joint are connected in a way to have the central spindle or wiper of the potentiometer attached to one of the segments and the body of the potentiometer attached to the other segments. The sensors are attached to the joints via cuff fixation wrapped around the limb segment above and below the joint under test, and each potentiometer may cover a single degree of motion. This configuration creates a straightforward measurement system that outputs resistance value proportional to the joint angle.

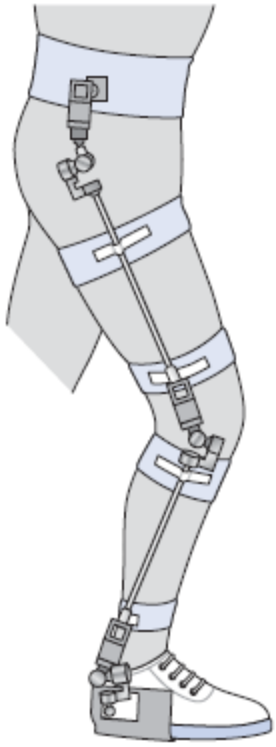

Figure 1.12: Worn Triaxial Goiniometer over the Hip, Knee and Angle (Chattecx Corporation)

On the other hand, the flexible strain-gauge electrogoniometer is based on the use of a flat thin strip of metal fixed to the adjacent limb segments across the joint under test; any strain applied to the metal manifested through its bending during limb movement, which is depicted in Fig. 1.13.

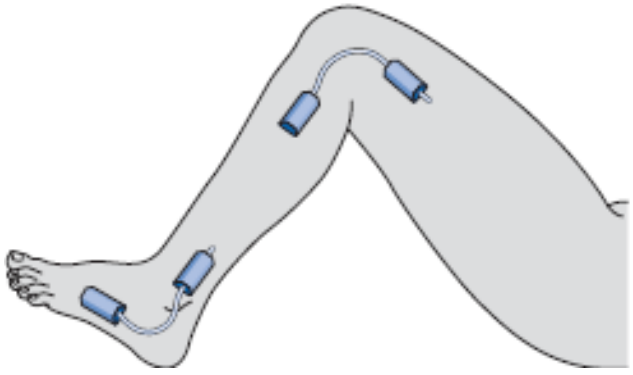

Figure 1.13: Ankle and Knee Flexible Goniometer Illustration (Biometrics Ltd)

1.1.3. e. Electromyography

Electromyograhy (EMG) is the measurement of the electrical activity of a muscle group at the macro-scale or specific muscle fibers at the micro-scale in form of voltage variation in function of time, which we call the Electromyogram (see Fig. 1.14.). The Electromyogram is in fact a non-stationary signal that is relatively challenging to process and analyze due to its nature in terms of randomness and low amplitude [17].

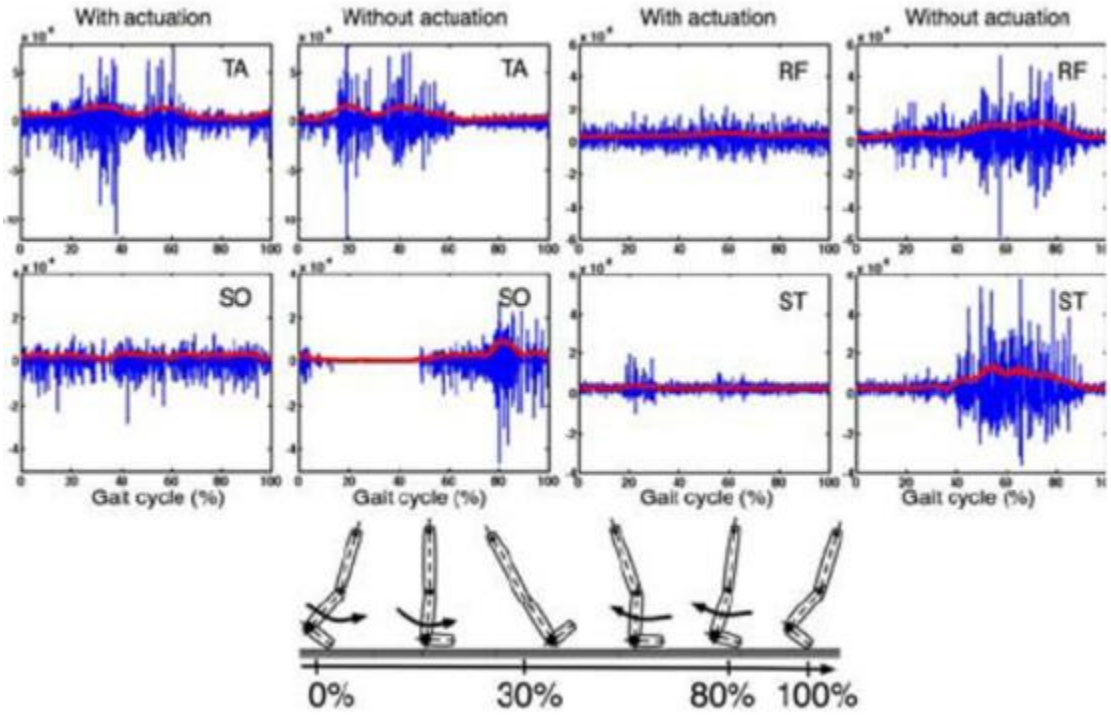

Figure 01.14: EMG Activations for the Tibialis Anterior (TA), Soleus (SO), Rectus Femoris (RF) and Semitendinosus (ST) for the actuated and non-actuated leg during one full gait cycle. Raw (blue thin line) and processed (red thick)

As for the contribution of EMG analysis to gait analysis, it is used to assess the contribution of individual muscle groups as well as synchronization in the performance of gait tasks e.g. walking and running [18], which is indicative of normal/pathological gait.

1.1.3. f. Kinematic and Kinetic/Kinematic Systems

Kinematics is defined as the study of the geometrical description of motion in terms of displacements, velocities and accelerations, and it is commonly used in gait analysis to record the displacements and other mentioned parameters of body segments and joints e.g. linear and angular velocities and accelerations at specific contributing joints.

This class of gait analysis systems provides data in form of 2D or 3D measurements depending on the number of used cameras that will take photographs of the subject in motion, wearing markers, from different angles and then save the data for further analysis. The images are captures at standard frame rates of which we state 50Hz, 60 Hz and 200 Hz and even greater in new system, and the markers may be reflective markers, active markers such as sequentially-flashing LEDs, electromagnetic sensors e.g. coils, ultrasound transmitters and receivers. Concerning the assessment of such systems, they are convenient for measuring position but weak in determining accelerations and velocities, being computationally costly, which may be compensated by added accelerometers to the system to acquire accurate acceleration data e.g. accelerometers. In terms of resources to dig deeply into the available technologies for that class of systems, refer to [19].

In order to enhance this class of gait analysis systems, a hybrid or combined class of analysis was devised, which is known as the combined kinetic/kinematic analysis system. Actually, the kinematic aspect of it is the introduction of a force platform, being a kinetic system that measures forces, which are added to the kinematic data captured by the cameras or any of the previously enumerated technologies. This sort of systems is actually self-corrective in a way that the lack of details in the kinetic part is enriched by the information provided by the kinematic system, creating thereby a full gait analysis system.

1.1.3. g. Under-foot Pressure/Force Measurement

Another method of gait analysis is presented in this section, however mentioned in the previously stated hybrid kinetic/kinematic systems, being the measurement of actual pressure applied underneath the foot. Just like all other methods, different approaches have been devised to provide under-foot pressure, of which we state: Floor-mounted pressure platform and shoe-insole embedded pressure sensor array. Typical under-foot pressure values differ from activity to activity; 80-100 kPa while standing still, 200-500 kPa while walking, and up to 1500 kPa while running. The most popular of this class of systems is the force sensor system based on the walking of the subject across a pressure platform underneath which an array of resistive or capacitive strain gauge sensors is fixated.

Another commonly used system is the Force Platform or Force Plate which measures the Ground Reaction Force (GRF). This class of systems is commercially available in the market (see Fig.1.15) but nevertheless, a multitude of laboratories and research centers have these platforms custom-designed to meet specific needs such as accuracy and force range.

In fact, the database that is processed and analyzed in this thesis is based on an instrumented treadmill as will be discussed in later sections, which belongs to this class of force sensor systems.

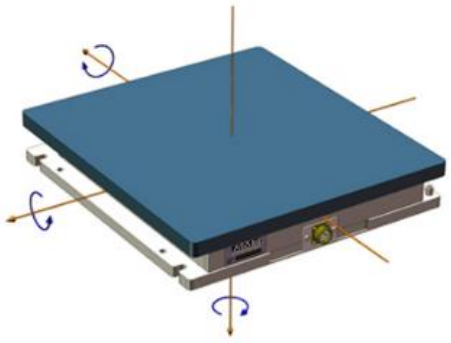

Figure 1.15: 3-axis Force Platform Illustration

However, the limited surface area of a force platform defies the purpose of gait analysis which is the analysis of the walking/running activity for a multitude of strides or gait cycles, triggering the innovation in this field by the creation of the Instrumented Treadmill [20].

In Fig. 1.16, a typical Instrumented Treadmill's sensor architecture is illustrated, which happens to be the actual system used by our research group to acquire the VGRF signals that will be processed and analyzed thoroughly in this thesis.

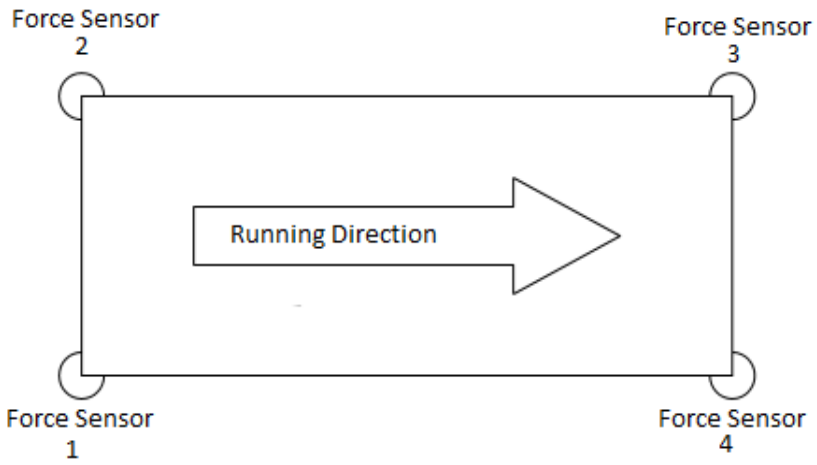

Figure 010.16: Instrumented Treadmill VGRF Sensor Localization

This class of measurement schemes actually relies on corner-based force sensors to create a bi- or tri-directional force transducer or platform. In fact, the most common force acting on the body is the Ground Reaction Force (GRF) which is active on the foot while the subject is standing, walking, or running. The resulting force vector is actually three-dimensional: Vertical (VGRF), Antero-posterior (APGRF) and Medio-lateral (MLGRF) [21]. A sample recording of a 3-dimensional instrumented treadmill is shown in Fig. 1.17.

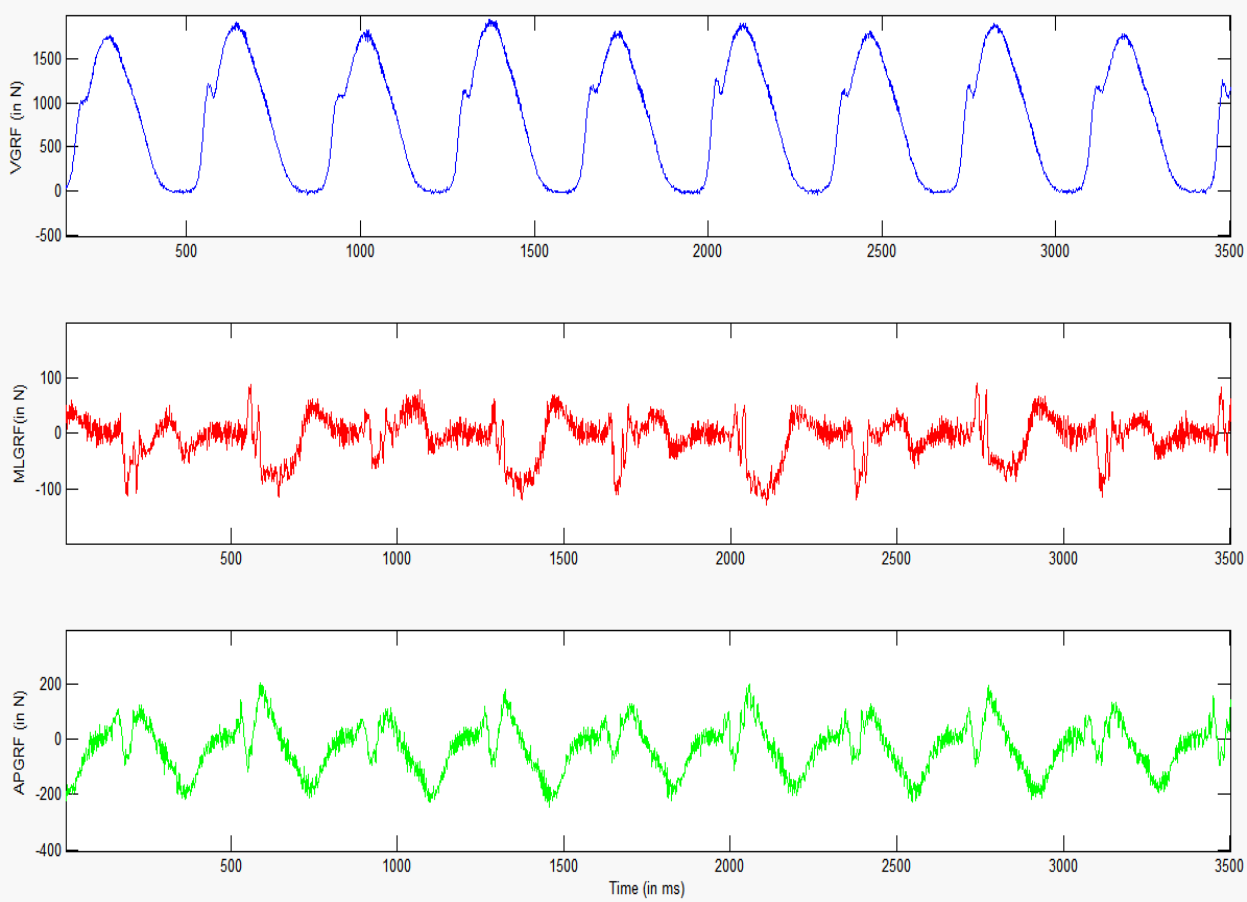

Figure 01.17: 3D GRF (Vertical GRF in Blue, Medio-lateral GRF in Red, Antero-posterior GRF in Green)

The actual data output of such systems is distributed among 6 to 8 channels; the 6-channel system outputs 3 force vector magnitudes (VGRF, APGRF, and MLGRF) and 3 force moments referenced against the center of the platform, while the 8-channel system outputs 4 corner-based VGRF signals, 2 side-based MLGRF signals, and 2 front-back based APGRF signal. This output data gets transferred to an acquisition system that digitizes the data and gives the possibility to manipulate the data as follows: Individual GRF component plot, Butterfly Diagram plot, and Center of Pressure (COP).

Typically, the VGRF only is taken for further processing and analysis since it is the most expressed in terms of power level, and this is the reason why we decided to perform our computation over VGRF signals.

More focus will be given to the VGRF signals throughout the thesis, being the main signal upon which the algorithms are implemented, for many reasons that will get clearer with the progression of this work.

1.1.4 Vertical Ground Reaction Force (VGRF):

1.1.4. a. Pattern Overview

The VGRF, a sample of which is shown in Fig. 1.18, is very characteristic in a sense that it shows clear manifestations of every event within the gait cycle as depicted in Fig. 1.19.

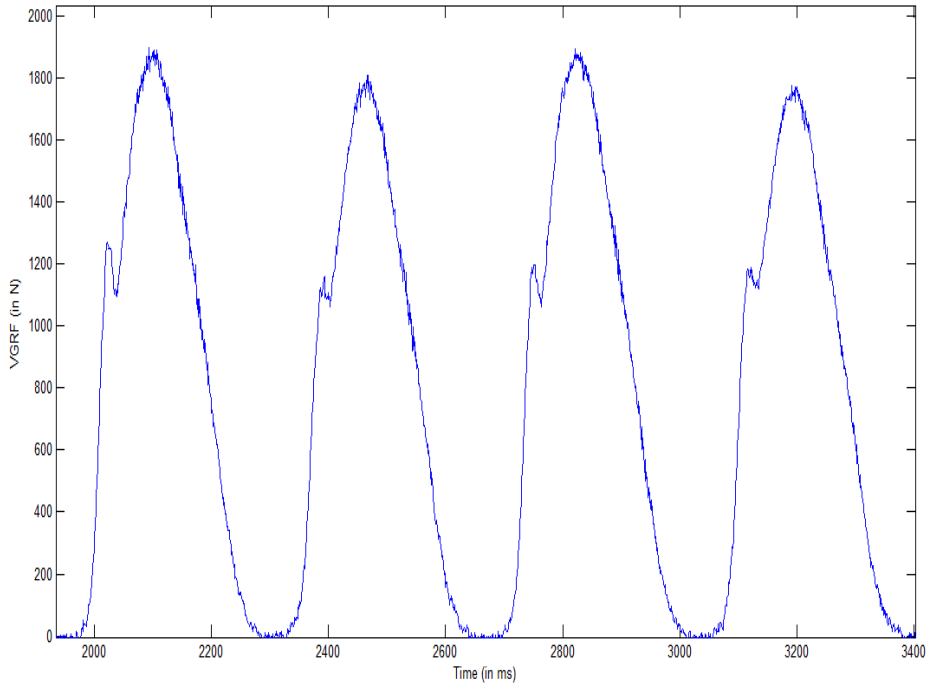

Figure 1.18: Sample 4-step VGRF Signal

Put into words, the VGRF signal is defined a sequence of double-peaked step-signals, each couple of which is termed stride, being the right-foot step and the left-foot step. The double-peak pattern is in fact the manifestation of the heel-strike, during which the foot is in contact with the ground, followed by the toe-off, during which the toes are pressing on the ground to push the body forward. That being stated, it is expected to have a force signal, during running, that has serially-defined couples of peaks, the first peak being lower in amplitude and sharper than the second body-pushing peak; these two peaks are termed Impact Peak and Propulsive Peak, respectively [22].

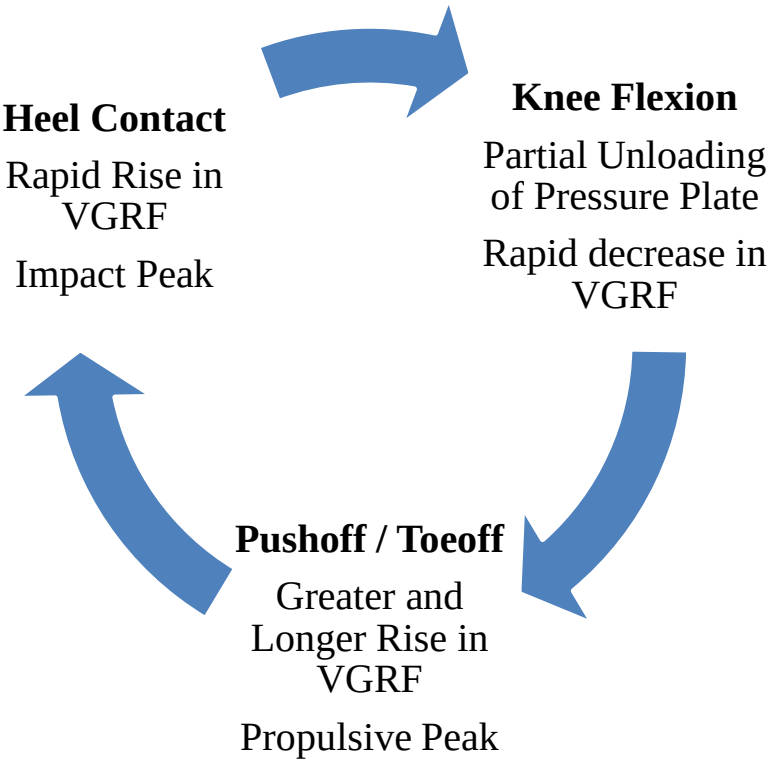

Figure 1.19: Sequential VGRF Events

The transient peak at heel-strike, which is termed the heel-strike peak (HS) or the Impact Peak (IP) or even the Passive Peak (PP) is manifested as an sharp increase to a value that exceeds the

bodyweight, followed by partial unloading on the knee that drops the VGRF by a considerable amount, then a second and higher rise takes place at toe-off exceeding bodyweight of the subject. The VGRF signal has been extensively referred to as a reliable tool for the purpose of analyzing the gait cycle, mainly due to the ease of acquisition and the simplicity of the waveform. Actually, VGRF has been studied for decades in various domains e.g. walking and running biomechanics [23], athletics [24] and rehabilitation [25].

1.1.4. b. Database Description

As for this thesis, the VGRF signals were collected during a special case of running termed Ultra-marathon. In our research, we acquired the VGRF of fourteen male volunteers (mean _ SD: age 41.1 ± 8.9 years; weight 73.6 ± 8.2 kg, height 176.9 ± 5.8 cm, body mass index (BMI) 23.5 ± 1.9 kg/m2 and body fat: 17.7 ± 4.3%). They were recruited among experienced ultra-endurance runners and all of them had run at least a race longer than 24 h or 4100 km. On average, they had 15.3 ± 7.1 years of training history in running and 7.1 ± 4.4 years of ultra-endurance experience. They reported to run an average of 80.5 ± 11.7 km/week. Written informed consent was obtained from the subjects. The study was conducted according to the Declaration of Helsinki. The protocol has been approved by the local ethics committee (Comité de Protection des Personnes Sud-Est 1, France) and registered in http:// clinicaltrial.gov (# NCT 00428779). Among the 14 subjects, 12 were able to complete the 24TR. One subject was excluded by the physician because of a hematoma due to the initial muscle biopsy procedure and the other one was excluded because of low blood pressure. Each subject was asked to run for the duration of 20 seconds every 2 hours for 24 consecutive hours, and the VGRF of the 20-second runs (12 recordings for each of the 12 subjects) were recorded and saved. The acquisition of the desired VGRF signal was performed using 4 force sensors located in the corners of an instrumented treadmill

described as previously shown in Fig. 1.16. The actual and detailed architecture of the hardware system, which is an instrumented treadmill, is thoroughly presented in [13].

1.2 Blind Source Separation

1.2.1 Background

In the past decade or so, source separation (SS) has been the center of attention in the fields of Signal Processing and Artificial Intelligence (AI). The field of SS deals with what is known as signal mixtures that are to be separated into their independent components or sources.

This process of SS can be categorized into two man categories: Supervised and Unsupervised. Supervised SS is a class of separation techniques where training information is required for the separation to be feasible, while unsupervised SS is the other class of separation techniques where no training information is required. Training information is actually known as labeled data, i.e. isolated sources are available a priori to train the separation system to better recognize the sources in the separation phase. The commonly used term denoting the unsupervised class is Blind Source Separation (BSS), which is a whole field of study joining a variety of mathematical techniques implemented to separate a mixture of sources with no a priori knowledge about the source-mixing system nor the sources themselves.

The field of BSS has been going through an increasingly powerful progress in terms of implemented algorithms and even in the area of application. BSS is widely is used in advanced statistics, signal processing, neural computation and modeling, econometrics, remote sensing and many more areas that require discovery of underlying or hidden patterns [26]. However, the area in which BSS is most commonly used is speech and music processing [27], being implemented as an algorithm or set of algorithms aiming to separate audio mixtures or even monaural sounds

into their independent components e.g. beat detection, musical instrument separation, speech enhancement [28]…etc. The technical terminology describing the separation of audio mixtures is Auditory Scene Analysis (ASA), which is the extraction of individual audio sources from a mixture of sources e.g. separation of speech from interfering background sounds and separation of individual musical instruments from a polyphonic ensemble [29]. And since signal and image processing tools and algorithms are much similar in terms of the way data is handled, BSS has found its way into image processing, medical imaging and machine vision [30].

To summarize BSS applications, the main applications in which BSS algorithms are being implemented mainly revolve around the following fields of research:

- Reception for single- and multi-user communications [31, 32, 33, 34]
- Biomedical Signal Analysis e.g. Electroencephalogram (see Fig. 1.20) [35], Electrocardiogram [36]…etc;
- Functional Magnetic Resonance (f-MRI) Image Processing and Analysis (see Fig 1.24) [37];
- Image Restoration [38];
- Ultrasonography [39];
- Astronomical Imaging [40];
- Feature Extraction [41, 42];
- Sensor Array Data Processing [43, 44];
- Geophysical Exploration [45];
- Denoising [46];
- Voice-controlled Machines (see Fig. 1.23) [47];

- Semiconductor Manufacturing and Circuit Testing [47];

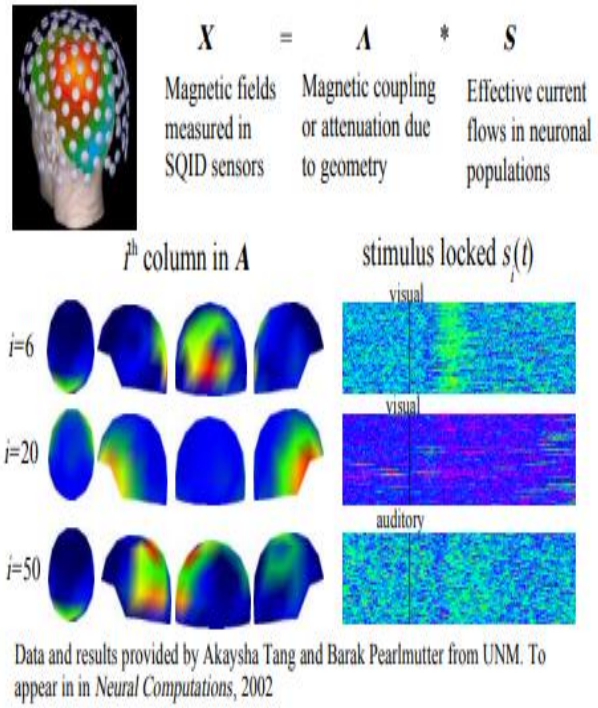

Figure 01.20: BSS in MEG Decomposition and Brain Activity Analysis

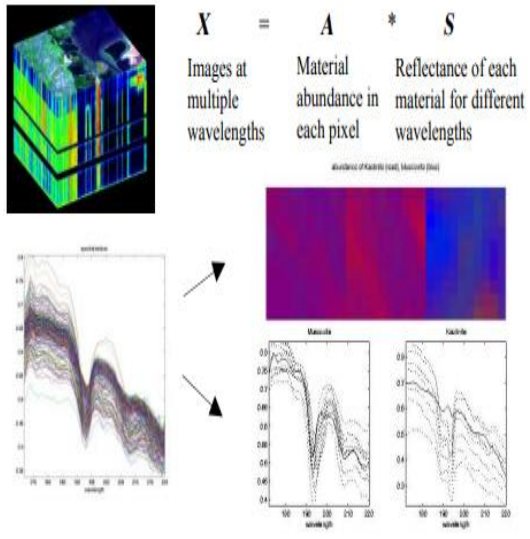

Figure 01.21: BSS in Hyperspectral Imaging

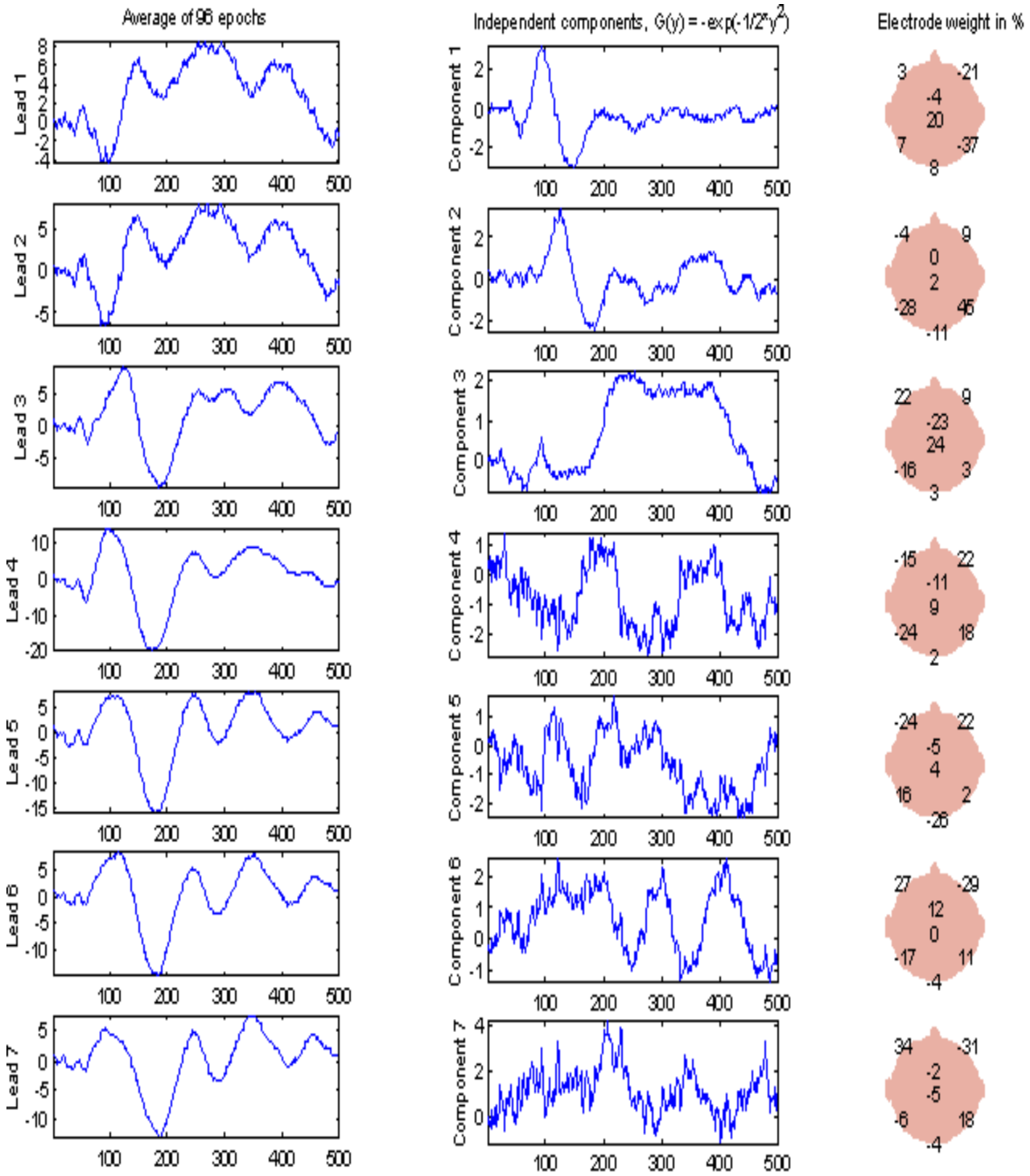

Figure 01.22:BSS in Electroencephalography (EEG)

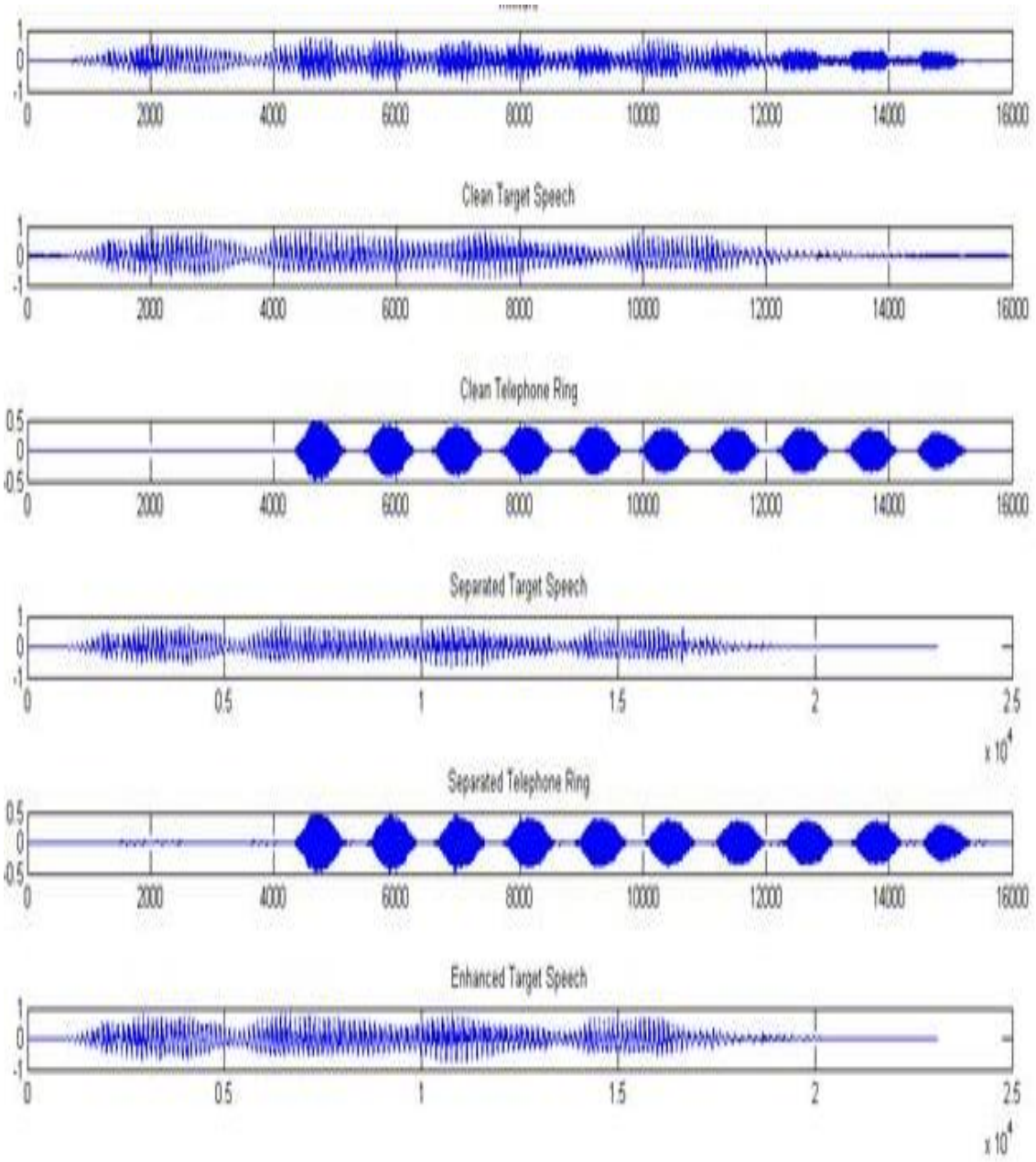

Figure 01.23: BSS in Speech/Sound Processing

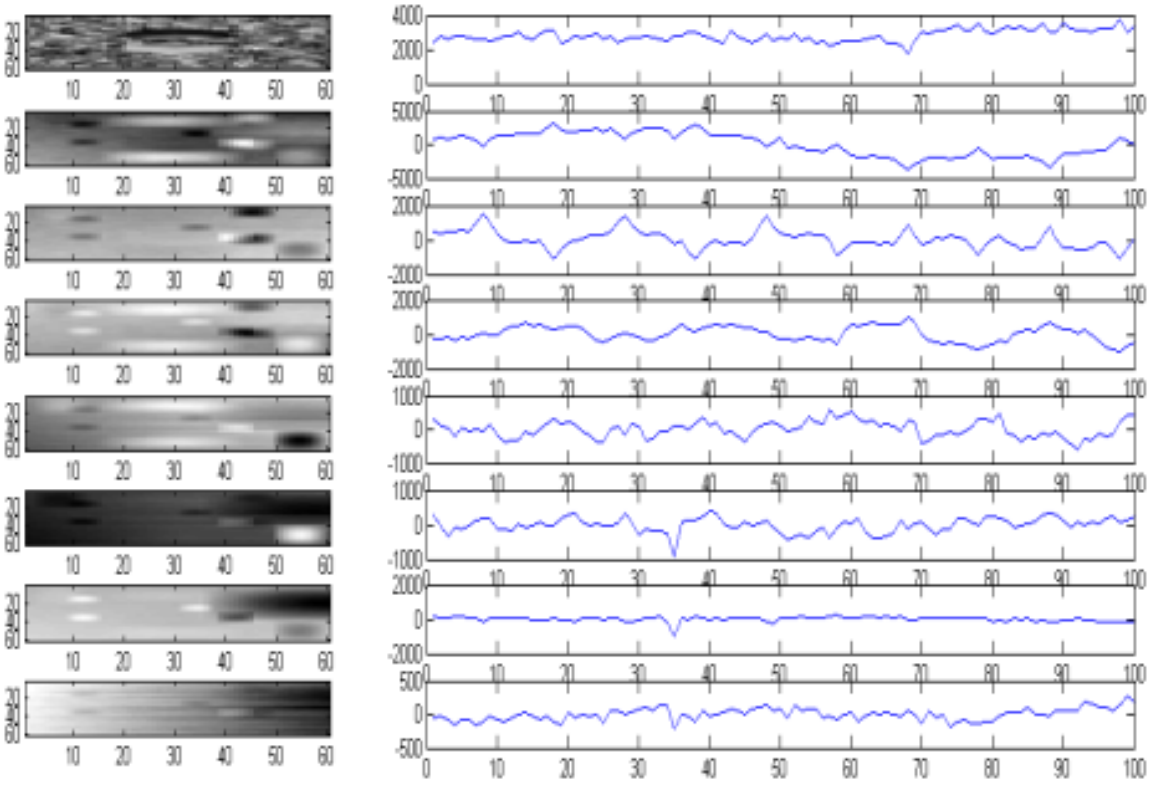

Figure 01.24: BSS in f-MRI Image Processing

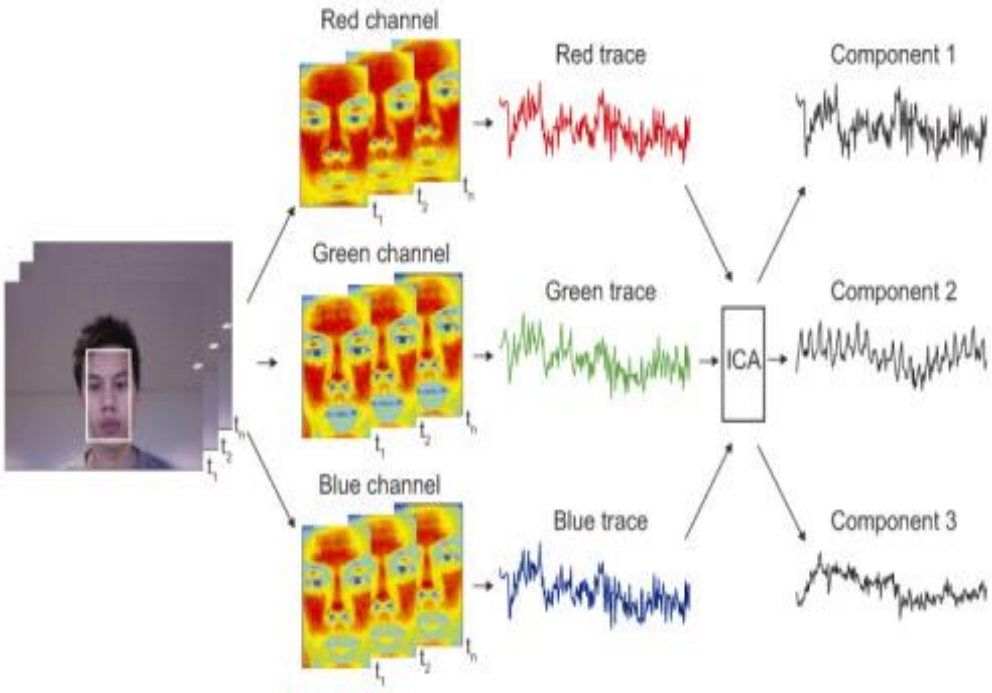

Figure 1.25: BSS in Facial Imaging for Non-contact Heart Rate Measurement

1.2.2 Problem Formulation

BSS problems are governed by a general mathematical definition described in (1.5) below:

$$x(t) = [x_1(t)\, x_2(t) \ldots x_{N_0}(t)] \tag{1.5}$$

which is a random mixture of underlying source signals (1.6):

$$s(t) = [s_1(t)\, s_2(t) \ldots s_{N_s}(t)]^{\mathrm{T}} \tag{1.6}$$

according to the following linear representation (1.7):

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{N_0}(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N_s} \\ a_{21} & a_{22} & \cdots & a_{2N_s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_0 1} & a_{N_0 2} & \cdots & a_{N_0 N_s} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_{N_s}(t) \end{bmatrix} \Leftrightarrow x(t) = \mathbf{A}\mathbf{s}(t) + \mathrm{u(t)} \tag{1.7}$$

where **A** is the unknown mixing matrix of dimension N_o *x* N_s, *t* is the time or sample index and *u* is additive noise.

BSS is actually implemented to estimate the original sources as well as the mixing matrix having no prior information, i.e. relying solely on the observations. Noteworthy, the presented mathematical model for BSS is nothing but a simplification and idealization of the linear approximation of the model, which in real life applications is too ideal; other major factors play a role in adding complexity to the mixing process e.g. noise, propagation delay of signals, which is why BSS is still an ongoing field of research being in continuous progress.

1.2.3 BSS Method Categorization

Upon reviewing available literature in the field of BSS, one may categorize the problems into three main classes according to three main criteria: Linearity, Time-delay and Determinism. This classification leads to the following categories: Linear versus Non-linear BSS, Instantaneous versus Convolutive BSS, and Overcomplete / Overdetermined versus Underdetermined BSS.

1.2.3. a. Linearity in BSS Problems

The most commonly and widely researched field of BSS is the linear time-invariant (LTI) model, and that is due to the analysis simplicity and explicit separability. The Linear BSS system is based on the assumption that the source mixture is represented by a linear combination as shown in (1.8).

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{N_0}(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N_s} \\ a_{21} & a_{22} & \cdots & a_{2N_s} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N_0 1} & a_{N_0 2} & \cdots & a_{N_0 N_s} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_{N_s}(t) \end{bmatrix} \Leftrightarrow x(t) = \mathbf{As}(t) + \mathrm{u(t)} \qquad (1.8)$$

The assumptions surrounding this method may be summarized as follows:

a. The signals $x_i(t)$ are stationary and zero-mean
b. The sources $s_i(t)$ are statistically independent;
c. The noise components $u_i(t)$ are statistically independent of the sources;

Each of the assumptions will be explained and covered in later sections. Noteworthy, unless additional information is supplied as a priori characteristics of the sources, the scaling and order of the separated sources will remain imprecise and confusing, which is normal in linear systems

where multiplication is no longer commutative and the mixing coefficients according to which the mixing process took place are non-evident.

As for the Non-linear model, a more realistic environment is considered, where the observed signals are non-linear distorted signals. Despite the fact that literature holds a large number of studies and proposed algorithms on Linear BSS, nonlinear BSS has not been well developed, a review of which is presented in [48]. Linear BSS techniques have shown their inability to separate the sources which are non-linearly mixed, which led to the emerging of the Non-linear techniques, of which we state Post Non-linear (PNL) [49] and Bi-Linear (or Linear Quadratic) mixtures [50], Convolutive Post Non-linear Mixtures [51] and Conformal mappings [52]. The mathematical representation of such Non-linear models is stated in (1.9) below:

$$x(t) = f(s(t)) \tag{1.9}$$

where f: $\mathfrak{R}^N \rightarrow \mathfrak{R}^M$is the unknown non-linear mixing function, *N* and *M* are the number of source and observation signals, respectively. The BSS problem may also be written in the following form (1.10):

$$\text{Find g such that: } g(x(t)) = g(f(s(t))) = \hat{s}(t) \tag{1.10}$$

where g: $\mathfrak{R}^M \rightarrow \mathfrak{R}^N$ is the separating function to be estimated and *x(t)* the vector of reconstructed sources, which is the output of the BSS algorithm.

1.2.3. b. Mixing Delay in BSS Problems

Regarding the time-wise mixing aspect of the sources, BSS may be oriented to deal with sources that are simultaneously mixed i.e. with the absence of time delays, which defines the

Instantaneous Mixture Model, or to deal with mixtures that were formed out of different source signals in form of combinations of several time-delayed versions of the sources themselves and/or mixed signals themselves which defines the Convolutive Mixture Model (CMM).

In the simpler mixing models, the mixture is regarded as the sum of differently weighted source signals, however in most real-world applications e.g. acoustics, the individual sources are weighted and delayed contributing to the sum or mixture with multiple delays indicating for instance the manifestation of the multiple paths by which acoustic signals propagate to a microphone. Such mixtures are termed convolutive mixtures that vary in the number of delay elements, which are application-specific, reaching thousands of delay elements in acoustics. The mentioned acoustic signals might be speech or music or underwater sonar signals, radio signals captured by antenna arrays as mixtures, astronomical data, and functional brain imaging data and bio-potentials.

The mathematical mixing model of the Instantaneous Mixture Model is shown in (1.11)

$$x(t) = \mathbf{A}s(t) + u(t) \tag{1.11}$$

This model is also known as the delayless (linear) mixture model, s.t **M**=$\mathbf{M_0}$ is an **M**x**N** matrix containing the mixing coefficients.

As for the delayed sources' case, assuming a reverberation-free environment with propagation delays, the mixing model may be represented as follows in (1.12):

$$x_m(t) = \sum_{n=1}^{N} a_{mn} s_n(t - k_{mn}) + u_m(t) \tag{1.12}$$

Where $k_{mn}$is the propagation delay between source *n* and sensor *m*.

$$x(t) = \sum_{k=0}^{K-1} A_k s(t-k) + u(t) \tag{1.13}$$

However, in the derivation of many algorithms, for simplification purposes, the environment is considered to be noise-free, which reduces to (1.14):

$$y(t) = \sum_{k=0}^{K-1} A_k x(t-k) \tag{1.14}$$

1.2.3. c. Determinism in BSS Problems

As for determinism, it is actually a classification criterion for BSS problems related to the comparison of the sources' number *N* against the observations' number *M*. That being said, three distinct situations are possible:

- N > M: The number of sources is greater than the number of observations.

Such systems are termed over-determined systems and are easily solved using linear BSS methods, the mixing matrix being invertible.

- N = M: The number of sources is equal to the number of observations.

Such systems are termed determined systems and are also easily solved using linear BSS methods, the mixing matrix being an invertible square matrix.

- N < M: The number of sources is greater than the number of observations.

Such systems are termed over-determined systems and may not be solved using linear methods even under perfect knowledge of the mixing system, i.e. the sources may never be recovered via linear methods.

The mentioned issue concerning under-determined systems leads us to the field of single-channel BSS (SCBSS), being the hardest problem to solve in BSS, having a single observation of the mixture data and requiring the separation of the multitude of sources, whether instantaneously or in a convolutive manner. This class of BSS problems will be discussed in the next section.

1.2.4 BSS Process

1.2.4.a. The Mixing Model

As previously stated in terms of classification of BSS problems, convolutive mixtures and instantaneous mixtures are the two classes to which all mixtures belong according to the delay between source mixing, which implies two different mixing models.

Firstly, the convolutive mixture is modeled as shown in (1.15):

$$x_m(t) = \sum_{n=1}^{N} \sum_{k=0}^{K-1} a_{mnk} s_n(t-k) + v_m(t) \tag{1.15}$$

where $s(t) = (s_1(t),\ldots,s_N(t))$ is a mixture of N source signals at time index t assumedly received at a sensor array of M sensors, $x(t) = (x1(t),\ldots,xM(t))$ are the received mixture signal, and a_{mnk} are the mixing filter coefficients. While being a time-variant system with variable coefficients, simplicity of calculations imposes the assumption of stationarity.

In linear terms i.e. matrix notation, the current model may be represented as in (1.16):

$$x(t) = \sum_{k=0}^{K-1} A_K s(t-k) + v(t) \tag{1.16}$$

where $\mathbf{A_k}$ is an *M* x *N* matrix containing the k^{th} filter coefficients and *v(t)* is the *M* x 1 noise vector.

The instantaneous mixing model is considered to some extent a special case of the convolutive mixture model, assuming all source signals are captured by the sensors simultaneously, simplifying the model to (1.17):

$$x(t) = \mathbf{A}s(t) + v(t) \tag{1.17}$$

where $\mathbf{A=A_0}$ is an *M* x *N* matrix containing the instantaneous mixing coefficients.

Another simplification of the mixing model is the delayed sources model that assumes a reverberation-free environment with propagation delay, represented as follows in (1.18):

$$x_m(t) = \sum_{n-1}^{N} a_{mn} s_n(t - k_{mn}) + v_m(t) \tag{1.18}$$

where k_{mn} is the propagation delay between the source and the signal *n* and observation *m*.

Assuming an ideal environment where no noise coexists with the signals of interest, the model becomes (1.19):

$$x(t) = \sum_{k=0}^{K-1} \mathbf{A_k} s(t-k) \tag{1.19}$$

However, it is often assumed that the number of observations is greater or equal to the number of mixed sources, which may be solved using the techniques described above with minimal complications, and these systems are called overdetermined mixing systems. On the other hand, if this assumption is not fulfilled, i.e. the number of observations is less than the number of sources, the system is termed underdetermined and the linear methods are no longer sufficient to solve the problems and separate the sources.

As previously stated, a common problem simplification scheme is implemented in the pre-separation phase, which is the domain transformation e.g. from time domain to frequency domain, leading to a system modeled as follows in (1.20):

$$\mathbf{X}(\omega) = \mathbf{A}(\omega)\mathbf{S}(\omega) + \mathbf{V}(\omega) \tag{1.20}$$

where ω is the angular frequency s.t. ω =2πF, **A(*ω*)** is a complex *M* x *N* matrix, **X(*ω*)** and **V(*ω*)** are complex *M* x 1 vectors, and **S(*ω*)** is a complex *N* x 1 vector.

The domain transformation from the time domain into the frequency domain is implemented using the Fast Fourier Transform (FFT), which is a faster implementation of the Discrete Fourier Transform (DFT) within a T-point user-set time segment with a sliding window over the entire signal of interest (1.21).

$$\begin{aligned} X(\omega, t) &= \mathbf{DFT}([x(t), \dots, x(t+T-1)]) \\ X(\omega, t) &= \sum_{\tau=0}^{T-1} \omega(\tau) x(t+\tau) e^{-j\omega\tau/T} \end{aligned} \tag{1.21}$$

where the window function $\omega(\tau)$ is chosen to minimize band-overlap caused by narrow time gaps in between windows.

For simpler modeling and subsequent computation, the system may be seen in form of blocks instead of point-by-point form. The block would consist of T samples as shown in (1.22):

$$\begin{aligned} x(t) &= \mathbf{A_0}s(t) + \cdots + \mathbf{A_{K-1}}s(t-K+1) \\ x(t-1) &= \mathbf{A_0}s(t-1) + \cdots + \mathbf{A_{K-1}}s(t-K) \\ x(t-2) &= \mathbf{A_0}s(t-2) + \cdots + \mathbf{A_{K-1}}s(t-K-1) \end{aligned} \tag{1.22}$$

This leads to an *M*-dimensional output sequence being written as an *MT* x 1 vector as in (1.23):

$$\hat{x}(t) = [x^T(t), x^T(t-1), \dots, x^T(t-T+1)]^T \tag{1.23}$$

where $x^T(t) = [x_1(t), \dots, x_N(t)]$.

In a similar fashion, the N-dimensional input sequence may be represented as an *N(T+K-1)* x 1 vector as in (1.24):

$$\hat{s}(t) = [s^T(t), s^T(t-1), \dots, s^T(t-T-K+2)]^T \tag{1.24}$$

This implies the convolutive model representation shown in (1.25):

$$\hat{x}(t) = \widehat{\mathbf{A}}\hat{s}(t) + \hat{v}(t) \tag{1.25}$$

where A has the following form (1.26):

$$\widehat{\mathbf{A}} = \begin{bmatrix} \mathbf{A_0} & \cdots & A_{K-1} & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \mathbf{A_0} & \cdots & A_{k-1} \end{bmatrix} \tag{1.26}$$

From a mathematical perspective, matrix A is termed block-Toeplitz matrix of dimensions *MT* x *N(T+K-1)*. This representation seems from a first superficial look like the equation describing the instantaneous mixture model, which is not true knowing the exact block structure of the block matrices that carry the convolutive aspect i.e. time lags within.

1.2.4. b. The Separation Model

The separation of sources is the ultimate purpose behind BSS methods, which is in more detail the process of finding an estimate *y(t)* of the original source signals *s(t)*. One way to go for this endeavor is to find the mixing filters A_k in an explicit manner, which is not always necessary. Instead, one might sufficiently estimate what is termed separation filters W_l aiming to mixing-induced cross-talk, i.e. isolating the sources by eliminating the mixing effect. The stated filter may take one of two forms as is known in Digital Signal Processing (DSP) basics: Finite Impulse Response (FIR) and Infinite Impulse Response (IIR), with or without feedback, respectively.

One of two structures is the feed-forward structure, where the FIR separation system is represented as in (1.27):

$$y_n(t) = \sum_{m=1}^{M} \sum_{l=0}^{L-1} w_{nml} x_m(t-l) \tag{1.27}$$

Or in a matrix form as in (1.28):

$$y(t) = \sum_{l=0}^{L-1} \mathbf{W_l} x(t-l) \tag{1.28}$$

And in the block Toeplitz form, the separation system is expressed as in (1.29) and explained in more detail in [53]:

$$\hat{y}(t) = \widehat{\mathbf{W}}\hat{x}(t) \tag{1.29}$$

The z-domain representation of the described separation system becomes (1.30):

$$\mathbf{Y}(\boldsymbol{z}) = \mathbf{W}(z)\mathbf{X}(z) \tag{1.30}$$

Noteworthy, the ultimate purpose behind BSS is the recovery of sources that have no interference from other sources in the mixtures, and not necessarily an identical version of the original source signal before the mixing process. This is where the concept of filters takes role, the recovered source signal being a filtered version of the original source as follows in (1.31):

$$\mathbf{Y}(z) = \mathbf{W}(z)\mathbf{A}(z)\mathbf{S}(z) = \mathbf{G}(z)\mathbf{S}(z) \tag{1.31}$$

where **A(z)** is the z-domain representation of the filter matrix, and **G(z)** is the combined mixing and filtering matrix.

In fact, the separation is considered successful when the interference-free signals are recovered even if they turn out to be permuted and scaled versions of the original source signals, leading to the following representation (1.32):

$$\mathbf{G}(z) = \mathbf{P}\boldsymbol{\Lambda}(z) \tag{1.32}$$

where **P** is the permutation matrix, and $\boldsymbol{\Lambda}$ is a diagonal matrix holding the scaling filters as its diagonal.

Hence, if **A**(z) is accurately identified, and **W**(z) is chosen to be the inverse of **A**(z), then **Λ**(z) would be an identity matrix and the source signals would be exactly recovered. And this is where the difference between blind source identification and BSS is manifested; in source separation, it is sufficient to recover convolved versions of the original sources, i.e. no restrictions on **Λ**(z).

A summarizing table of the described methods is presented, which is shown in Fig 1.26.

	Mixing Process	Separation Model
Time	$x_m(t) = \sum_{n=1}^{N} \sum_{k=0}^{K-1} a_{mnk} s_n(t-k) + v_m(t)$	$y_n(t) = \sum_{m=1}^{M} \sum_{l=0}^{L-1} w_{nml} x_m(t-l)$
	$\boldsymbol{x}(t) = \sum_{k=0}^{K-1} \boldsymbol{A}_k \boldsymbol{s}(t-k) + \boldsymbol{v}(t)$	$\boldsymbol{y}(t) = \sum_{l=0}^{L-1} \boldsymbol{W}_l \boldsymbol{x}(t-l)$
z-domain	$\boldsymbol{X}(z) = \boldsymbol{A}(z)\boldsymbol{S}(z) + \boldsymbol{V}(z),$	$\boldsymbol{Y}(z) = \boldsymbol{W}(z)\boldsymbol{X}(z)$
Frequency domain	$\boldsymbol{X}(\omega) = \boldsymbol{A}(\omega)\boldsymbol{S}(\omega) + \boldsymbol{V}(\omega)$	$\boldsymbol{Y}(\omega) = \boldsymbol{W}(\omega)\boldsymbol{X}(\omega)$
Block Toeplitz Form	$\widehat{\boldsymbol{x}}(t) = \widehat{\boldsymbol{A}}\widehat{\boldsymbol{s}}(t)$	$\widehat{\boldsymbol{y}}(t) = \widehat{\boldsymbol{W}}\widehat{\boldsymbol{x}}(t)$

Figure 01.26: Summarizing Table of Source Separation Processes and Separation Models

While the system described above is a feed-forward system, the other class of systems is the feedback system implemented via Infinite Impulse Response (IIR) filters that are precisely used to invert the effect of Finite Impulse Response (FIR) filters. The IIR filter-based structure of unmixing is presented in (1.33):

$$y_n(t) = x_n(t) + \sum_{l=0}^{L-1} \sum_{m=1}^{M} u_{nml} y_m(t-l) \tag{1.33}$$

As presented in (1.33), the number of sources is considered to the number of observations, which we already explained as being a determined system, s.t. u_{nml} are the IIR filter coefficients.

In matrix form, (1.33) is represented as follows in (1.34):

$$y(t) = x(t) + \sum_{l=0}^{L-1} \mathbf{U}(l) y(t-l) \tag{1.34}$$

And in z-domain representation, (1.34) is represented as follows in (1.35), provided $(I + U(z))^{-1}$:

$$\mathbf{Y}(z) = \big(\mathbf{I} + \mathbf{U}(z)\big)^{-1} \boldsymbol{X}(z) \tag{1.35}$$

In some cases, the feed-forward and the feedback structures are combined in a hybrid network, in which a feedback network follows a feed-forward network.

1.2.4.c. The Separation Criteria

As previously discussed, BSS algorithms are based on certain assumptions that are considered the building block that leads the way in terms of sources and mixing system. Considering the sources themselves, independence and decorrelation are the main criteria according to which the separation takes action. The two main classes of methods implemented in BSS are Higher-order statistics (HOS) and Second Order Statistics (SOS).

In Table 1-1, a summary of the classified separation criteria is presented:

Table 1-1: Separation Criteria

Class	Criterion	Reference
HOS	4^{th} Order Statistics	[54]
	Non-linear Cross-moments	[55]
	Information Theoretic	[56]
SOS	Minimum-phase Mixing	[57]
	Non-stationarity	[58]
	Cyclostationarity	[59]
	Non-whiteness	[60]

- **Higher Order Statistics:**

The core element and factor upon which this class of BSS is based is statistical independence of the sources. Statistical dependence may be computed at the 2nd order and the 4th order between model signals, and it is generally expressed as follows in (1.36):

$$E\left[y_n(t)^{\alpha}, y_{n'}(t-\tau)^{\beta} = 0\right] \quad n \neq n', \alpha, \beta = \{1,2,\dots\}, \forall_T \tag{1.36}$$

where E[.] is the expectation, assuming non-Gaussian sources which have null higher cumulants.

- **4th Order Statistics**:

While the minimization of cross-moments is required for effective separation, most algorithms do not aim to minimize all cross-moments, rather, they aim to minimize 2nd and 4th order dependence between the source signals by minimizing cross-moments and cross-cumulants. This is actually manifested through the off-diagonal elements of cross-cumulants in case of statistical independence of signals. This method has been implemented by a wide number of researchers in case of convolutive mixtures, the most widely used method being the Joint Approximate Diagonalization of Eigenmatrices (JADE) algorithm for complex valued signals in the frequency domain. On the other hand, 2nd and 3rd order cumulants were studied less commonly to deal with asymmetrical signals. Moreover, a common parameter in the 4th order, which is the kurtosis, is also commonly referred to for the separation of convolutive mixtures, as in [61].

- **Non-linear Cross Moments**:

Other algorithms make use of HOS indirectly for the separation of convolutive mixtures via non-linear functions. The condition assessed by this type of algorithms is modeled as follows in (1.37):

$$E\left[f\left(y_n(t)\right), g(y_{n'}(t-\tau))\right] = 0\Big] \tag{1.37}$$

where *f(.)* and *g(.)* are odd non-linear functions.

The separation of convolutive mixtures using this type of algorithms is - accomplished upon computation of the Taylor expansion of the non-linear functions, which lead to the capturing of the higher order moments through which the statistical independence is assessed. Originally, this method was implemented on convolutive mixtures, as in [62], which got enhanced in further stages in [63] to cover instantaneous mixtures as well.

- **Information Theoretic**:

Another way to interpret statistical independence is the assessment of the probability density function (PDF). In fact, independence of model sources *y* manifests into the following joint PDF (1.38):

$$p(y) = \prod_n p(y_n) \tag{1.38}$$

The above-stated equality implies the absense of mutual information between model sources y_n. From another perspective, the information theoretic methods are based on entropy maximization in each of the variables, i.e. having the sum of each variable's entropy equal to the total joint-entropy in *y*. This also means that

no mutual information is carried by the variables, which implies their mutual independence. If we were to state the most popular algorithm that is based on the current principle of entropy maximization, Infomax is robust algorithm designed by Bell and Sejnowski [64] characterized by effective separation and fast convergence. This algorithm may also be derived and implemented using (1.38) based on the principle of Maximum Likelihood as in [65] or even the Kullback-Leibler divergence between the empirical distribution and the independence model as in [66].The concept of divergence measures and optimization shall be discussed in more detail in the upcoming chapters. In such methods, the PDF $p_s(s_n)$ of the sources sn should necesarrily be assumed or modeled, which leads to the capturing of the HOS. The PDF is thereby classified into parametric or non-parametric, being based on an assumption or on estimation from real data, respectively. Most methods are actually based on parametric representations of the PDF, however, some major work has been implemented based on non-parametric representations. The resulting analytic non-linear functions are derived with (1.39):

$$f(Y) = -\frac{\partial \ln p(|Y|)}{\partial |Y|} e^{j.\arg(Y)} \tag{1.39}$$

where *p(Y)* is the probability density of the model source $Y \in C$.

In fact, many algorithms were derived based on (1.39), using the ML approach directly as in [67], or using closely related probabilistic approaches e.g. Maximum a Posteriori (MAP) methods where prior information about the model

parameters are considered as in [68]. Other closely related methods e.g. Hidden Markov Models (HMM) and Bayesian approaches were left unexplained in this chapter because of the fact that BSS is our main target and those methods are mainly supervised SS methods that required significant training on prior isolated source signals.

- **Second Order Statistics**:

Having explained the concept behind HOS-based BSS, it should be stated that not all cases require that amount of complexity; in some situations, SOS are sufficient to do the job. Actually, the criterion of separability that is considered in SOS is simply the uncorrelatedness of the source signals, which is less computationally demanding than mutual independence and entropy assessment. The assumptions on which these methods are stationarity-related: non-stationarity or cyclostationarity (CS) and minimum-phase mixing. The assumptions are necessary for any SOS technique that is not sufficient by itself, but that does not deny the fact that SOS possess some considerable advantages such as reduced computational complexity, reduced noise and outlier sensitivity, and reduced requirement for data to compute estimations, which is detailed in [69].

- **Minimum Phase**

Minimum phase mixing states that two source signals are separable by decorrelation if the mixing system is minimum phase, meaning the FIR coupling filters are to be strictly causal and have stable inverses. This concept is detailed in [70].

- **Non-stationarity**

Non-stationarity of the signals is used as well as a separability criterion in many applications e.g. speech, which is a non-stationary signal on time scales greater than 10 ms as stated in [71].

In fact, non-stationarity is a statistical state of a signal having varying statistics over time, which turned out to be a useful criterion in the separation process, and it was first proposed that the minimization of estimated cross-powers within locally stationary segments would lead to sufficient separation conditions. Of these methods we state Joint-diagonalization algorithms that extend to instantaneous mixtures as well as convultive mixtures, in both the time domain and the frequency domain. Noteworthy, the terms uncorrelatedness and whiteness are used interchangebly. Whiteness is achieved, in a linear algebra point of view, by minimization of the off-diagonal cross-powers, which translates into the following cost function in the frequency domain (1.40):

$$J = \sum_{t,\omega} \left\| \mathbf{R}_{yy}(\omega, t) - \mathbf{\Lambda}_y(\omega, t) \right\|^2 \tag{1.40}$$

where $\Lambda_y(w, t)$ is an estimate of the cross-power spectrum of the model sources, assumebly diagonal. This cost function holds two variables, time and frequency, and is to be minimized w.r.t. $\mathbf{W}(\omega)$ and $\mathbf{\Lambda_y}(\omega, t)$ subject to one or more normalization constraint.

Multiple variations on this algorithm were implemented,whether by changing the domain of analysis into the time domain [72], or by adding more constraints e.g. combined non-stationarity, non-gaussianity and non-whiteness as in [73].

- **Cyclostationarity**

As for cyclostationarity (CS), it is a considerable criterion in BSS, where the signals' cumulative distribution is invariant w.r.t time shifts of period *T* and its multiples. In simpler words, cyclostationarity is a statistical state in which the statistical parameters of a signal are cyclically or periodically variable w.r.t. time. To be more specific, CS is assessed on two levels: wide-sense CS and strict-sense CS.

When the first order statistical parameters are periodic with period *T*, the signal is said to be cyclostationary of 1st order (1.41):

$$m_x(t) = E[x(t)] = E[x(t+T)] = m_x(t+T) \tag{1.41}$$

where $m_x(t)$is the 1st order moment, i.e. expected value, which is periodic with period *T*.

When 2nd order statistical parameters are periodic with period *T*, the signal is said to be cyclostationary of 2nd order (1.42):

$$R_x(t,\tau) = E[x(t).x^*(t-\tau)] = R_x(t+T,\tau+T) \tag{1.42}$$

where $R_x(t,\tau)$ is the 2nd order moment known as the autocorrelation function which is periodic with period *T*.

One example of CS signals is a sinusoidal signal with random amplitude e.g. communications signals and voiced speech as in [74], and more importantly gait signals like the walking VGRF and running VGRF as detailed in [75] and as will follow in this work.

The CS property has been extensively studied in the field of BSS to explicitely recover mixed sources. Some researchers even used combined features like cyclostationarity and sparseness which forms a cyclosparse system as termed by *Sabri et.al* in [76] and defined as follows: *"*The signal object of this study is assumed to be cyclostationary with random impulses i.e. consists of *d* periodic random impulses where *d* represents the number of impulses by cycle. On the one hand, only few elements are nonzero by cycle, so the signal is considered to be sparse. Furthermore, the positions of these nonzero elements (impulses) are cyclic/periodic i.e. they keep the same positions whatever the cycle.

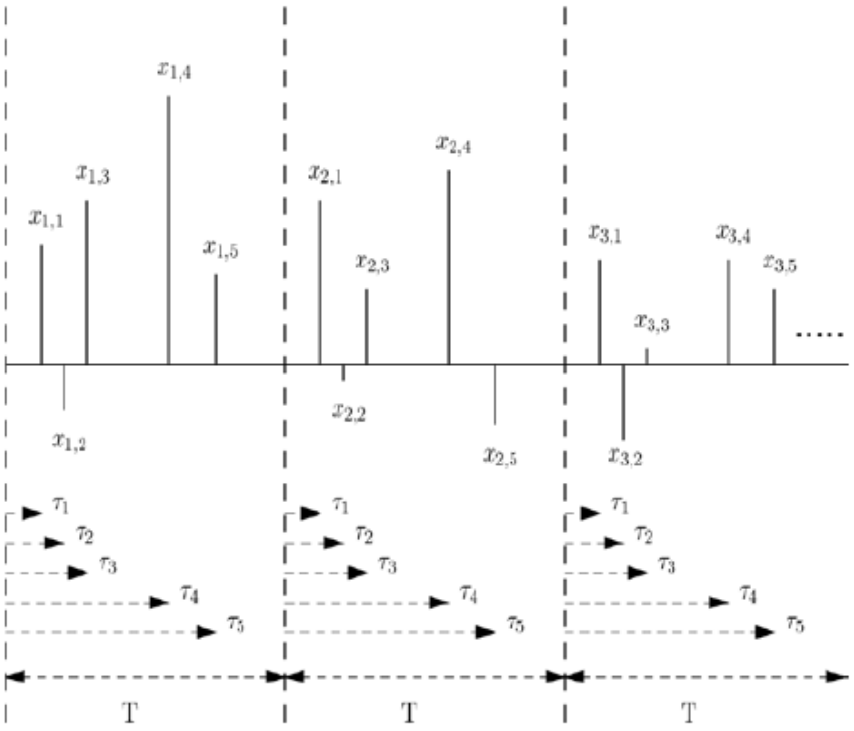

Figure 1.27: Sample Cyclosparse Signal (courtesy *Sabri et al.*)

In Fig. 1.27, d is the number of effective impulses in the period T with $x_{i,}$ and τ_k being their amplitude and delay factors respectively. K denotes the number of periods per signal and the sub-index i stands for the period index, so $x_{i,}$ represents the impulse with τ_k as delay factor in the i^{th}-period. (t) represents the random noise of the system.

- **Non-whiteness**

  Moreover, non-whiteness is also a criterion used in BSS algorithms, implying the presence of temporal correlation between source signals e.g. acoustic signals, which is a considerable property that has many advantages. Of these advantages, we state the possible reduction of a convolutive mixture problem to an instantaneous mixture problem, with added constraints of course, as in [77].

CHAPTER 2

VGRF PROCESSING AND ANALYSIS

2.1 Pre-processing

Before any processing is performed on any sort of signal, and upon acquisition, a pre-processing phase is mandatory for reasons that shall become clearer throughout this chapter.

To be more specific, VGRF signals that form the core of our database require delicate and customized pre-processing before any processing or analysis may be reliably performed. That being said, the purpose behind the pre-processing phase is the preparation of the datasets or studied signals by precise manipulation to have them custom-ready for particular planned analysis techniques to be performed in upcoming phases.

Going deeper into the building blocks of the pre-processing phase, three main steps are most commonly referred to in order to have any dataset ready for processing and analysis, summarized as follows: sampling, segmentation and labeling, as shown in Fig. 2.1.

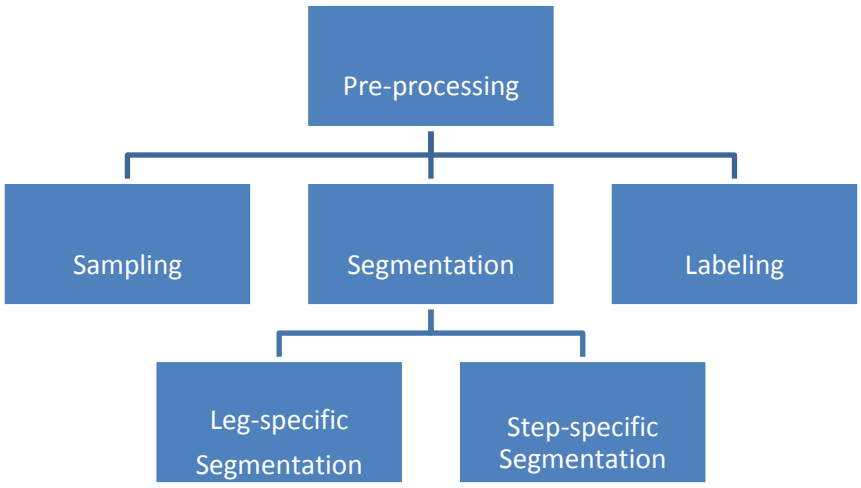

Figure 2.1: Pre-processing Phase Chart

2.1.1 Sampling

In Signal Processing, sampling is the process of reduction of a continuous-time signal to a discrete-time signal, i.e. the conversion of a continous dataset to a sequence of samples forming thereby a dataset of samples separated from each other by a precise period of time termed the sampling time T_s, hence a sampling frequency $f_s = 1 / T_s$. Shanon states:

"If a function *x(t)* contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced 1/(2B) seconds apart.".

When the time interval between two consecutive samples is too high, imperfections known as aliasing are exhibited during reconstruction. Modern statements of the theorem are sometimes careful to explicitly state that *x(t)* must contain no sinusoidal component at exactly frequency B, or that B must be strictly less than ½ the sample rate. The two thresholds, 2B and $f_s/2$ are respectively called the Nyquist rate and Nyquist frequency. And respectively, they are attributes of *x(t)* and of the sampling equipment. The condition described by these inequalities is called the Nyquist criterion. The symbol $T = 1/f_s$ is customarily used to represent the interval between samples and is called the sample period or sampling interval. And the samples of function *x(t)* are commonly denoted by $x[n] = x(nT)$ (alternatively "x_n" in older signal processing literature), for all integer values of *n*.

In particular, VGRF signals possess a frequency content that varies from activity to activity; in an increasing order of maximum frequency, VGRF signals are classified as follows: walking, speed-walking and running. The relationship between the frequency content of the VGRF signal and the activity is clearly the variation in speed, i.e. walking is performed in slower step rate than speed-walking which in its turn is slower in step rate than running. Normally, the range of

relevant data including stride rate and step rate is bounded by a maximum frequency of 5 Hz meaning 280 steps per minute,which is the equivalent of 4.7 steps per second. However, the pattern of the VGRF signal especially during running activities shows a transient high-frequency spike as described earlier by the Impact Peak, which is the manifestation of the heel-strike regardless of the running speed. That being said, in order to determine the optimal sampling frequency for running VGRF signals, the periodogram of that signal shall be computed, describng thereby its detailed frequency content, a sample of which is shown in Fig. 2.2.

The algorithm used to compute the Fourier spectrum was the Welch Periodogram that makes use of the Fast Fourier Transform (FFT) algorithm to generate the Power Spectral Density (PSD) of the analyzed signal. A fast Fourier transform (FFT) is an algorithm that samples a signal over a period of time (or space) and divides it into its frequency components. These components are single sinusoidal oscillations at distinct frequencies each with their own amplitude and phase.

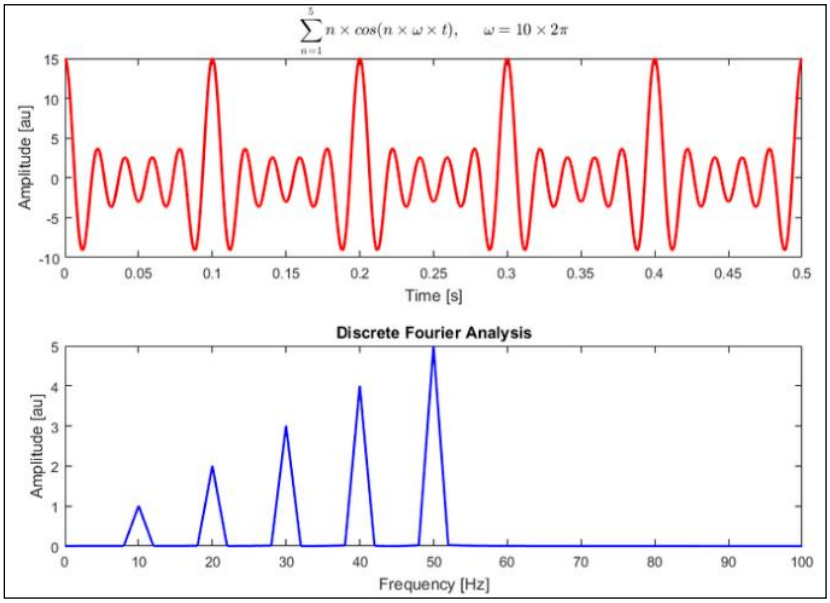

Figure 2.02: Sample DFT Computation

An FFT algorithm computes the discrete Fourier transform (DFT) of a sequence, or its inverse (IFFT). The Fourier analysis converts a signal from its original domain to a representation in the frequency domain and vice versa. Hence, an FFT rapidly computes such transformations by factorizing the DFT matrix into a product of sparse (mostly zero) factors.

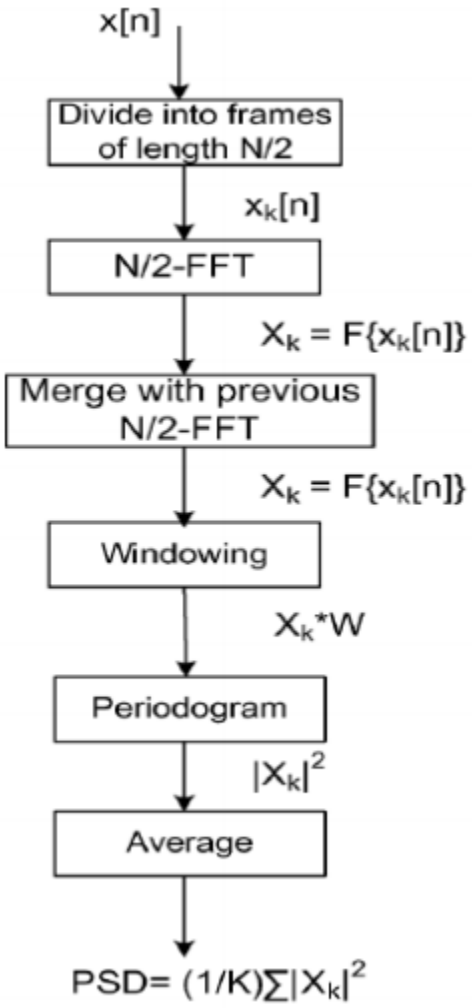

Figure 2.3: FFT Flow Chart

In the resulting periodogram (see Fig. 2.4.), a series of neighboring peaks in the low-frequency domain, at frequencies less than 30 Hz, are present in a cyclic fashion, which was analyzed as a marker for cyclostationarity [75]. The stride rate (1.4 Hz) and the step rate (2.8 Hz) both lie along with their harmonics in the form of spectral lines. On the other hand, noticeable spectral activity is present in the high frequency domain between 100 and 180 Hz suspected to be system noise, and at 250, 300, 350 and 450 Hz which are the manifestations of the mechanical resonance of the instrumented treadmill, which is usually disregarded since it does not carry subject-related biomechanical information. However, the detailed filtering technique adopted in our work will be detailed in the processing section.

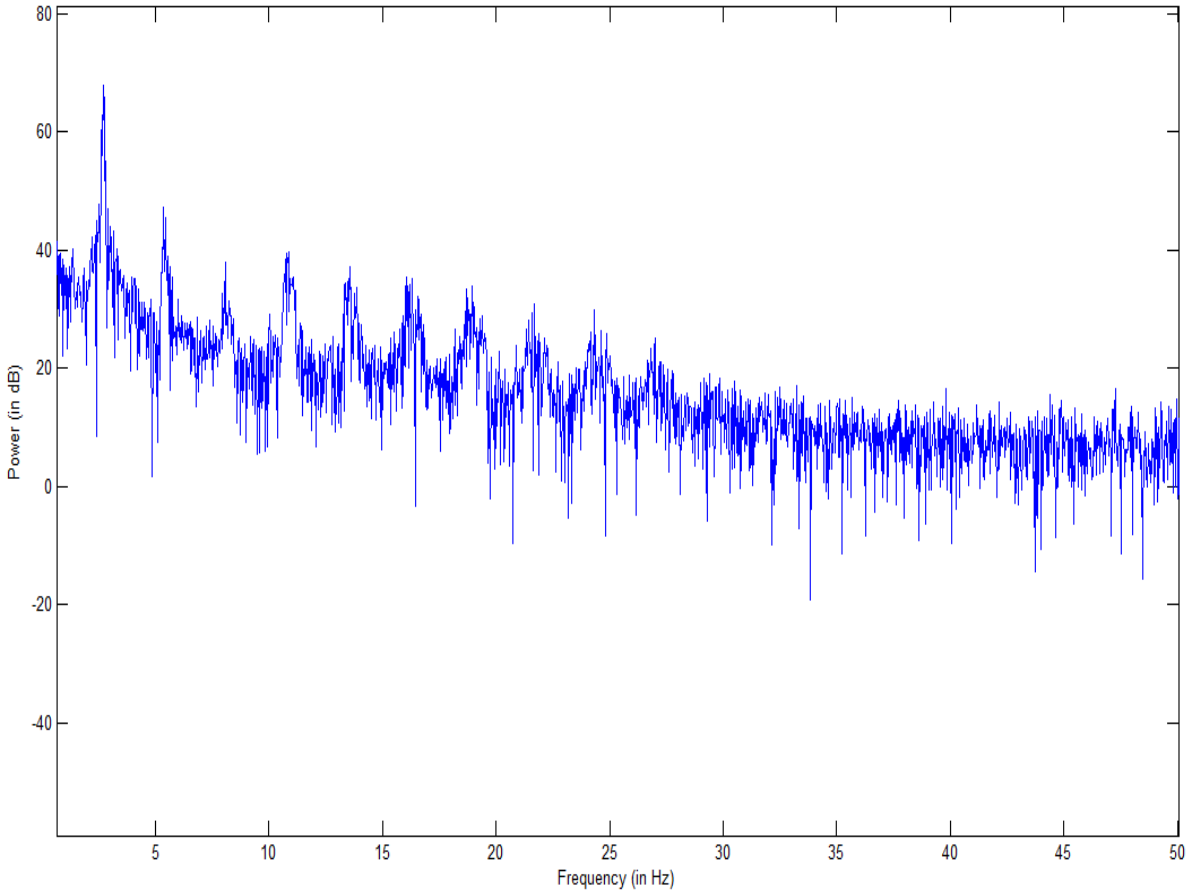

Figure 2.04: VGRF Welch Periodogram in the frequency range [0:50]

That being described, it is noteworthy to state the reason behind the computation of the periodogram of the signal before the segmentation of the signal, which was theoritized by Nyquist. The Nyquist theorem in sampling is actually the core of the sampling, aiming to prevent any undersampling of the signals, leading thereby to missed data and erroneous patterns, and it is governed by (2.1):

$$f_s \geq 2.f_{max} \tag{2.43}$$

2.1.2 Segmentation

Segmentation is defined as the division of a signal into a sequence of discrete segments of finite length that lie in between important events. In case of VGRF, segmentation may take two levels based on the eventual analysis techniques and purpose of the study. On a first level, in case the VGRF is measured via force platform or an instrumented treadmill, as described in previous chapters, the acquired data takes the form of a force vector in function of time such that the contribution of both legs is simultaneously recorded. In case a general overview of the subjects' gait is desired, a combined-leg dataset is good enough, however, if leg-specific analysis is to be performed e.g. for the assessment of inter-leg symmetry as in [78], leg-specific separation is required. On the other hand, if the assessment of gait consistency, balance or fall predicition is desired, step-specific segmentation is required generating a set of isolated steps to be individually assessed and compared. The main branching of the stated concepts is shown in Fig. 2.5.

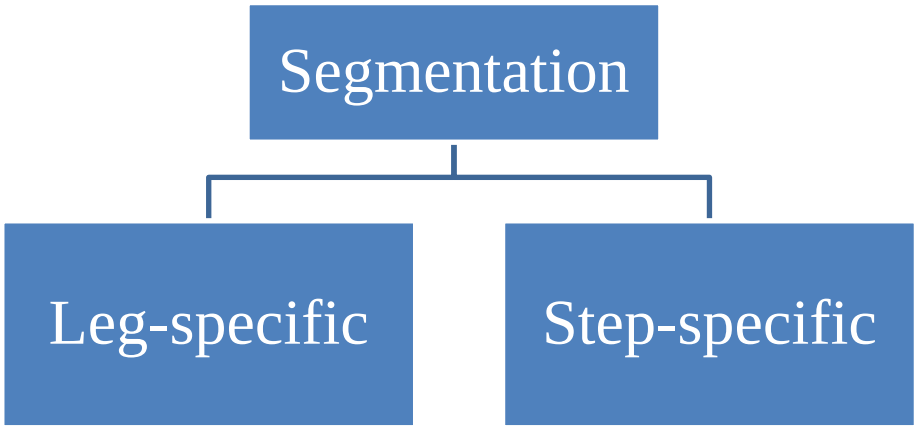

Figure 2.5: VGRF Segmentation Diagram

2.1.2.a Leg-specific Segmentation:

As previously stated, the segmentation of any signal is to be customized based on the desired and intended analysis. In fact, signals of the sort of the instrumented treadmill output carry a combined force pattern showing the manifestations of the entire gait cycle, left and right leg activity, which is of great advantage in most cases except when an assessment of the inter-leg symmetry is aimed for. In cases where the analysis requires a comparative aspect between leg-specific VGRF signals, the separation of the signals becomes mandatory for reliable analysis. For that reason, a simple yet effective and efficient algorithm was designed to adaptively separate the leg-specfic VGRF patterns based on the concept of peak detection and accurate time-localization of rising and falling edges with maximum care to local maxima avoidance. The details of the algorithm are stated below as well as a graphical representation of the process (see Fig. 2.6.).

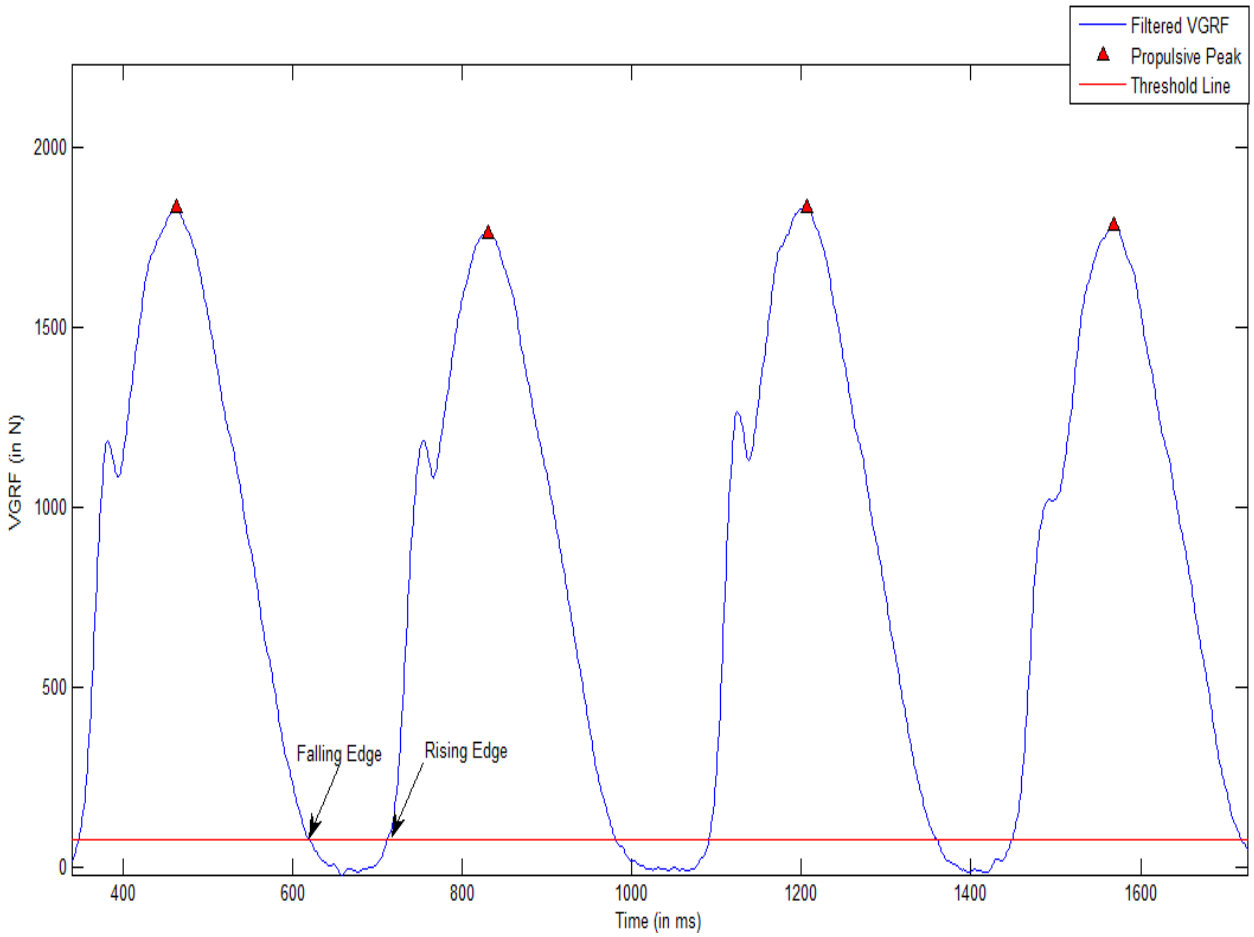

Figure 2.06: Graphical Representation of the Leg-specific Segmentation Process

Algorithm 1: Left-Right Leg VGRF Isolation

Input: **(**1 x *N*)-vector **VGRF**

Output: Two (1 x *N*)-vectors: **leg1_VGRF** and **leg2_VGRF**

Procedure:

1. Filter **VGRF** using a smoothing 4th order Butterworth low pass filter;
2. Find the propulsive peaks after setting minimum peak height and minimum peak distance;
3. Set a 5% **threshold** of the smallest peak;
4. Compute the difference between **filtered_VGRF** and **threshold**;
5. Set **leading_edge_found** flag to 0;
6. Set **j** to 0 and **k** to 2;
7. Initialize l**eg1_VGRF=VGRF**;
8. Initialize l**eg2_VGRF** = **VGRF**;
9. **for** i ← 1 to (*N*-1) **do**

repeat

if (leading_edge_found = 0 *and* d**ifference(i)<0** *and* **difference(i+1)>0)**

leading_edge_found=1;

j=**j**+1;

step_edges(j)=i;

end

if (l**eading_edge_found** = 1 *and* **difference(i)>0** *and* d**ifference(i+1)<0)**

leading_edge_found=0;

j=**j**+1;

step_edges(j)=i;

end

10. Locate the midpoints between respective falling and rising **step_edges;**
11. Adjust **leg1_VGRF** by zero-padding it between **step_edges** in a step-alternating fashion;
12. Set **leg2_VGRF = VGRF – leg1_VGRF;**

All VGRF signals of our Ultra-marathon database underwent the leg-specific separation successfully, which doubled the database and gave it more specificity for more advanced analysis. Samples of the original VGRF signals and leg-specfic VGRF signals are shown in Fig. 2.7.

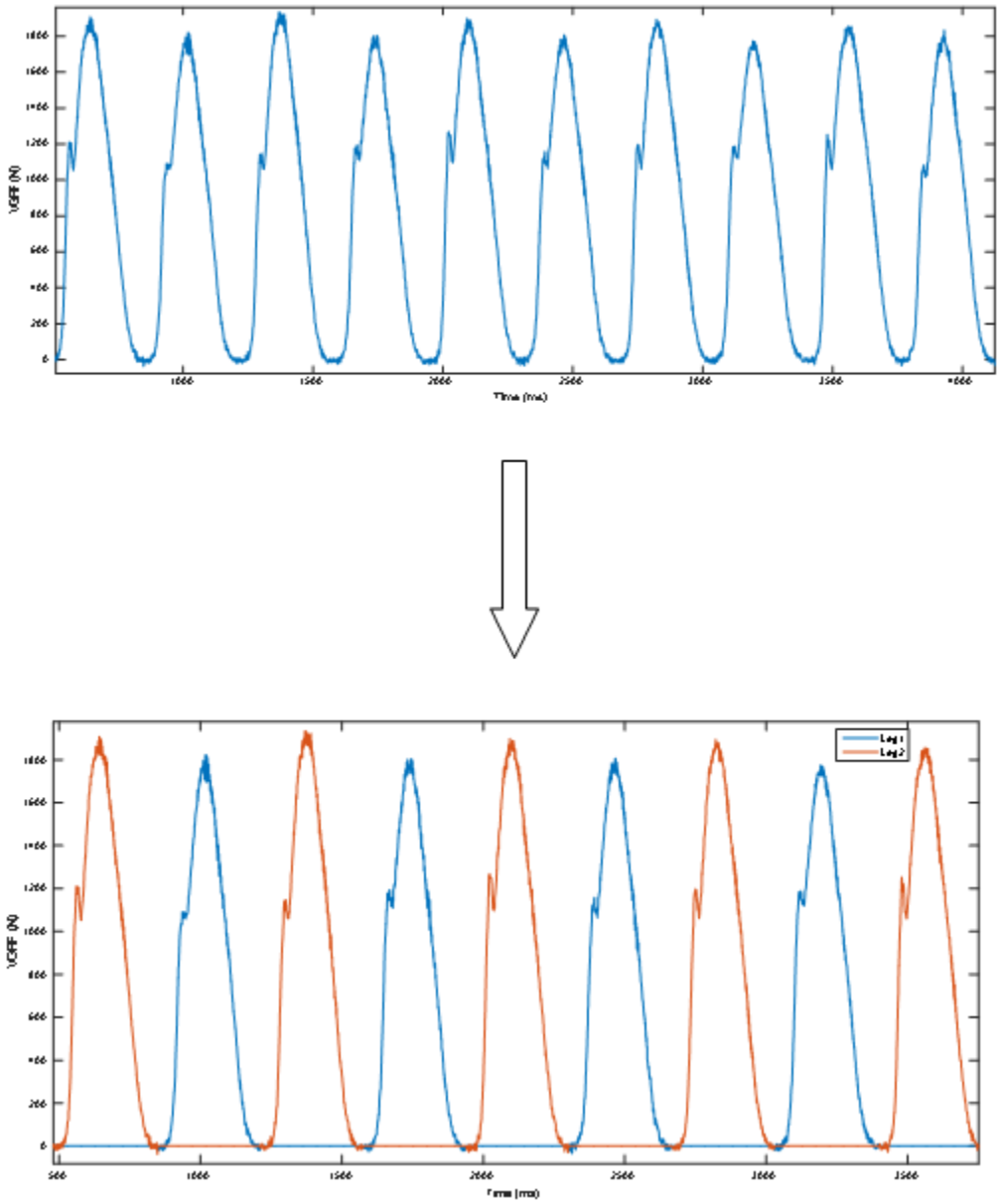

Figure 2.7: Graphical Representation of a Sample Leg-separated VGRF Signal

2.1.2.b Step-specific Segmentation

Moreover, in order to have the chance to exploit other major aspects of VGRF analysis, regardless of the leg-specficity, the evolution of VGRF parameters from step to step is indicative of the progression of biomechanical and physiological manifestations if accurately analyzed, of which we state: fatigue, injury, stability…etc. For that reason, we designed an algorithm that takes in a raw VGRF signal as a time vector and return a series of concatenated vectors representing successive steps performed by the subject/athlete, allowing us to analyze the run on a step-basis and track the progression of the above-mentioned parameters. The detailed description of the designed algorithm is explicitely presented in Algorithm 2 below, as well as a graphical representation of the separated steps in Fig. 2.8.

Algorithm 2: Step-specific VGRF Segmentation

Input: **(**1 x *N*)-vector **VGRF**

Output: K (1 x $\frac{N}{K}$)-vectors: **step_1_VGRF… step_K_VGRF**

Procedure:

1. Determine initial rising edge and final falling edge;
2. Locate rising edge above slope threshold indicating index of the beginning of the step;
3. Set a minimum time delay in between edges;
4. Locate falling edge followed by near-zero slope indicating index of ending of step;
5. Segment and save the step in between indeces as Step_1;
6. Repeat until reaching the last point of the VGRF vector;

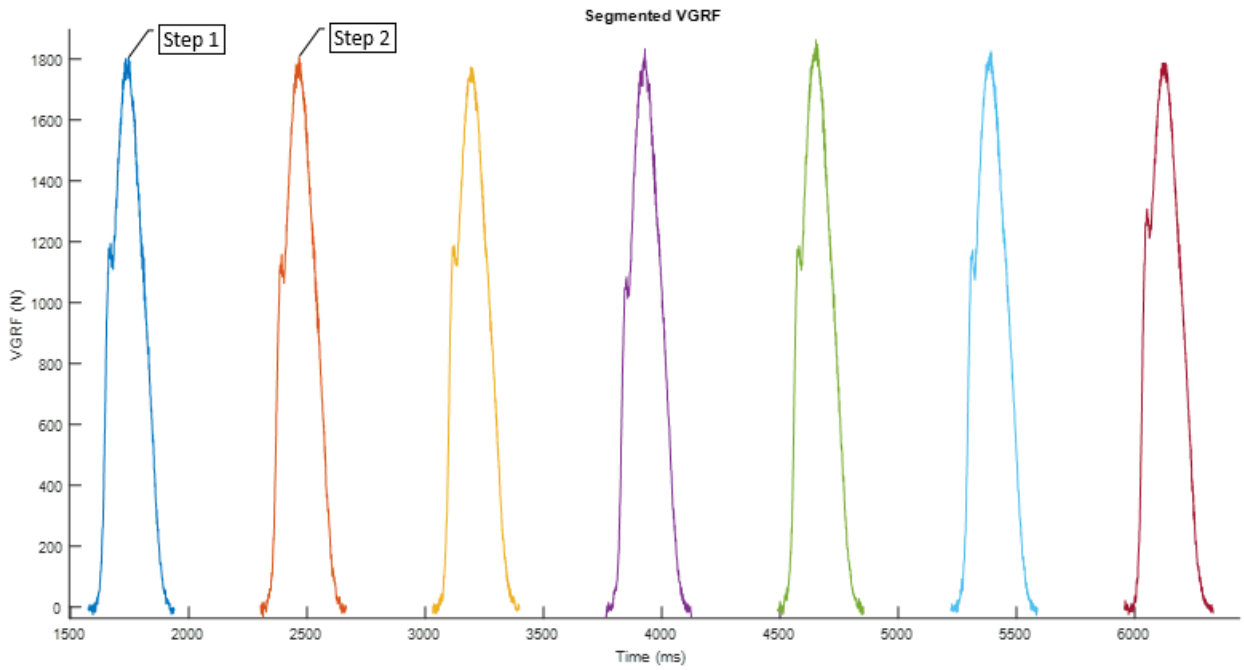

Figure 2.8: Graphical Representation of a Sample Step-separated VGRF Signal

Once we have separated the steps and stored each step in a time vector, the average step pattern may be computed by superposition, zero padding of the left and right tail and averaging. This process is optional in case an overall perspective of the VGRF signal is required to track inter-subject parameter variability or the progression of features in function of time. A sample of the mean VGRF pattern is shown in Fig. 2.9. below,

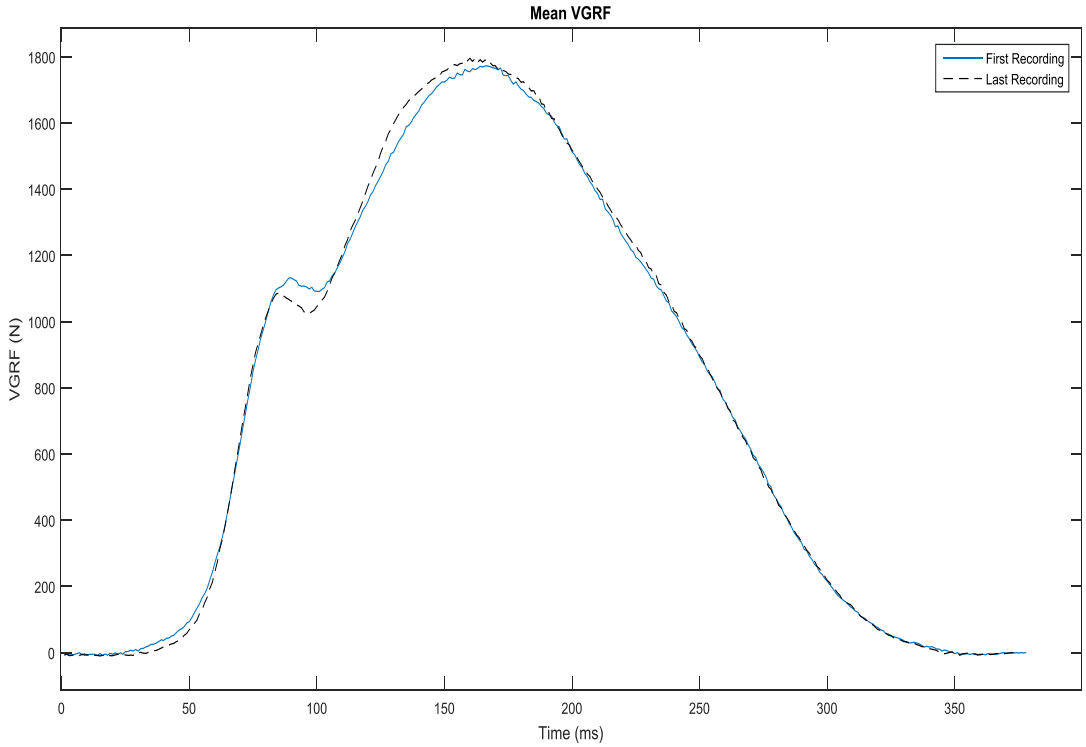

Figure 2.9: Mean step-VGRF at the beginning of the Ultra-marathon (blue) and at the end of the Ultra-marathon (black dashed line)

2.2 Processing

While signal processing encompasses a multitude of techniques implemented for different purposes, we will be focusing on the filtering of VGRF signals. In fact, having discussed the nature of the frequency content of the running VGRF signals, we realized the presence of two distinguishable peaks or patterns of different frequency levels; the heel-strike transient (HST) or impact pattern is of higher frequency that the propulsive pattern and the propulsive pattern being a smooth and relatively prolonged pattern. Moreover, the mechanical aspect as well as the electronics aspect of the measurement hardware i.e. instrumented treadmill is a noise-susceptible environment that should be accurately dealt with. Just like any other processing technique, under-filtering the signals would keep high noise levels that corrupt the VGRF-related pattern, while on the other hand, over-filtering the signals would remove the transient peak and all sharp edges that could possibly carry useful data.

Upon review of the existing literature in that matter, the vast majority of researchers filtered the VGRF signals using simple low-pass filters aiming to remove high-frequency noise e.g. electronic noise from circuitry, surrounding electromagnetic waves' interference, treadmill's motor noise…etc. However, imprecise low-pass filtering affects the impact pattern which is crucial for subsequent analysis that shall be revealed in later sections. Other researchers implemented more advanced techniques e.g. wavelet de-noising and total variation filtering [79], which goes back to the same concept of cancelling out a complete range of frequencies assuming it is nothing but noise or undesired data.

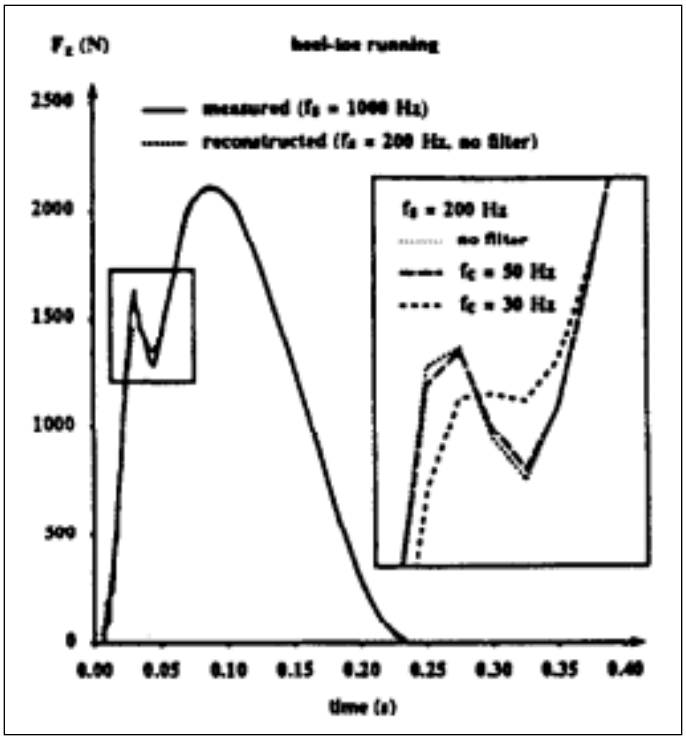

Figure 2.10: Illustrated Over-filtering of VGRF

That being cleared out, the need for a customized filtering technique in such sensitive cases is obvious, and for that endeavor we decided to implement the spectral subtraction technique. Spectral subtraction is a denoising method based on the restoration of the power spectrum or the magnitude spectrum of a signal buried in noise via subtraction of an estimate of the average noise spectrum from the noisy signal spectrum. Commonly, the magnitude or power spectrum of the noise estimated from the segments or periods of time where the signal of interest is definitely absent and only the noise is present. It is assumed that noise is stationary or slowly varying, and that the noise spectrum is more or less invariant from inactivity period the other. Being the ultimate goal, the restoration of the time-domain signals is performed through estimation of the instantaneous magnitude spectrum along with the phase of the noisy signal, and then transformed via an inverse discrete Fourier transform (IDFT) to the time domain.

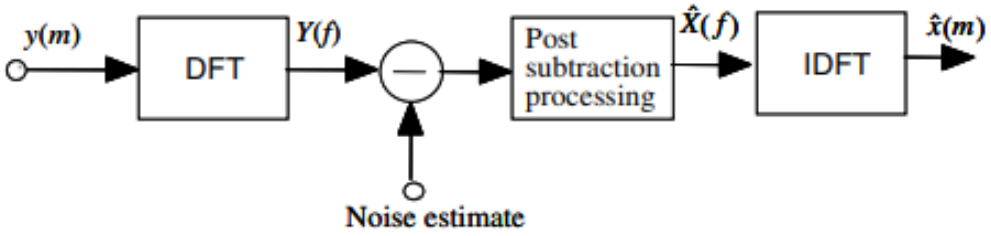

Figure 2.11: Spectral Subtraction Process Illustration

In terms of computational complexity, spectral subtraction is relatively inexpensive. However, in terms of accuracy, the random variation of noise w.r.t. time causes the spectral subtraction to result in negative estimates of the short-time magnitude or power spectrum. The magnitude and power spectrum are in fact non-negative variables, implying the mapping of their estimates into non-negative values. The nonlinear rectification process leads to the distortion of restored signal's distribution, and this becomes more visible and significant with the decrease of the signal-to-noise ratio.

Just like supervised and unsupervised systems, this class of filtering may be based on a priori knowledge of the noise spectrum when noise is separately available, or may be totally unsupervised in the sense that the "pause" or inactivity periods' spectra are computed and subtracted from the overall signals' spectra.

In fact, in many applications, including ours, the only signal that is available is the noisy signal. In these situations, it is not possible to cancel out the random noise, but it may be possible to reduce the average effects of the noise on the signal spectrum. The increase in the mean of the signal spectrum can be removed by subtraction of an estimate of the mean of the noise spectrum from the noisy signal spectrum. The noisy signal model in the time domain is given by (2.2):

$$y(m) = x(m) + n(m) \tag{2.44}$$

where *y(m)*, *x(m)* and *n(m)* are the noisy signal, the additive noise and the original signal respectively, and *m* is the discrete time index.

In the frequency domain, the noisy signal model of (2.2) is expressed as (2.3):

$$Y(f) = X(f) + N(f) \tag{2.45}$$

where *Y(f)*, *X(f)* and *N(f)* are the Fourier transforms of the noisy signal *y(m)*, the original signal *x(m)* and the noise *n(m)* respectively, and *f* is the frequency. In spectral subtraction, the incoming signal *x(m)* is buffered and divided into segments of *N* samples length. Each segment is windowed, using a Hanning or a Hamming window, and then transformed via discrete Fourier transform (DFT) to *N* spectral samples. The windows alleviate the effects of the discontinuities at the endpoints of each segment. The windowed signal is given by (2.4):

$$y_w(m) = w(m)y(m) = w(m)[x(m) + n(m)] = x_w(m) + n_w(m) \tag{2.46}$$

The windowing operation can be expressed in the frequency domain as (2.5):

$$Y_w(f) = W(f) * Y(f) = X_w(f) + N_w(f) \tag{2.47}$$

where the operator * denotes convolution.

Figure X illustrates a block diagram configuration of the spectral subtraction method, and the equation describing spectral subtraction may be generally expressed as (2.6):

$$\left|\hat{X}(f)\right|^b = |Y(f)|^b - \alpha\overline{|N(f)|^b} \tag{2.48}$$

where $\left|\hat{X}(f)\right|^b$ is an estimate of the original signal spectrum $|X(f)|^b$and $|N(f)|^b$is the time-averaged noise spectra. It is assumed that the noise is a wide-sense stationary random process. For magnitude spectral subtraction, the exponent b=1, and for power spectral subtraction, b=2. The parameter α in (2.6) controls the level of subtracted from the contaminated signal. For full

noise subtraction, α=1 and for over-subtraction α>1. The time-averaged noise spectrum is obtained from the periods when the signal is absent and only the noise is present as (2.7):

$$\overline{|N(f)|^b} = \frac{1}{K}\sum_{i=0}^{K-1}|N_i(f)|^b \tag{2.49}$$

where $|N_i(f)|^b$ is the spectrum of the i_{th} noise frame, assuming there are *K* frames in a noise-only period, where *K* is a variable. Alternatively, the averaged noise spectrum can be obtained as the output of a first order digital low-pass filter as (2.8):

$$|N_i(f)|^b = \rho\overline{|N_{i-1}(f)|^b} + (1-\rho)|N_i(f)|^b \tag{2.50}$$

where the low-pass filter coefficient ρ is typically set between 0.85 and 0.99. For restoration of a time-domain signal, the magnitude spectrum estimate $\left|\hat{X}(f)\right|$ is combined with the phase of the noisy signal, and then transformed into the time domain via the IDFT as (2.9):

$$\hat{x}(m) = \sum_{k=0}^{N-1}\left|\hat{X}(k)\right|e^{j\theta_Y(k)}e^{-j\frac{2\pi}{N}km} \tag{2.51}$$

To avoid negative magnitude estimates caused by random variations of noise, the output of the spectral subtraction algorithm is post-processed via mapping function T[·] of the form (2.10):

$$T\left[\left|\hat{X}(f)\right|\right] = \begin{cases} |X(f)| & if\ \left|\hat{X}(f)\right| > \beta|Y(f)| \\ \text{fn}[|Y(f)|] & otherwise \end{cases} \tag{2.52}$$

In its simplest form, fn[Y(f)] = noise floor, where the noise floor is a positive constant. An alternative is to take $fn[|Y(f)|] = \beta\,|Y(f)|$. In that case (2.11):

$$T\left[\left|\hat{X}(f)\right|\right] = \begin{cases} |X(f)| & if\ \left|\hat{X}(f)\right| > \beta|Y(f)| \\ \beta[|Y(f)|] & otherwise \end{cases} \tag{2.53}$$

As previously stated, spectral subtraction may be implemented in the power or the magnitude spectral domains. The two methods are theoretically analogous however they differ performance-wise. The detailed sequence of the described algorithm is illustrated in Fig. 2.12 and explicitly stated in Algorithm 3 below.

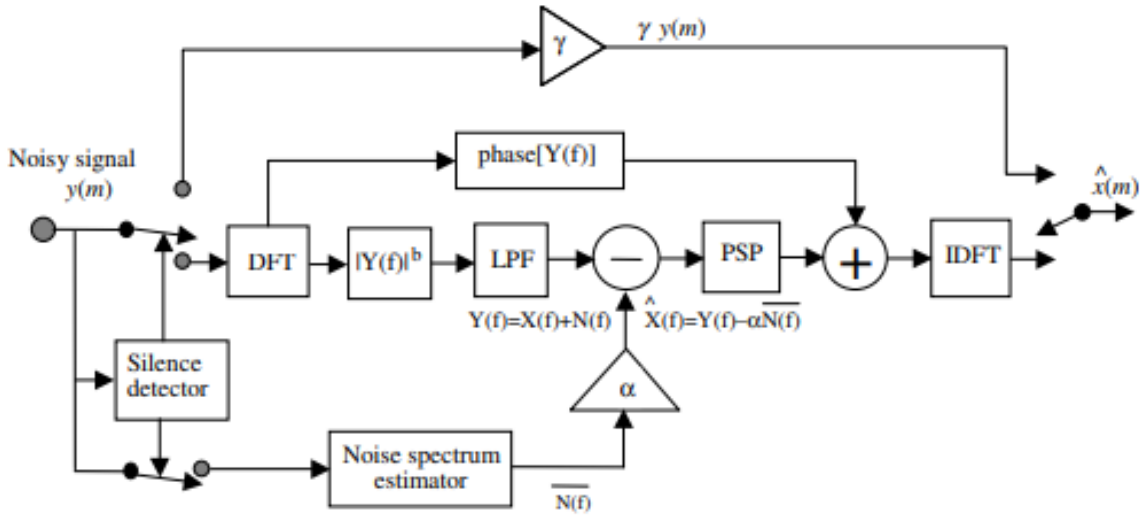

Figure 2.12: Block Diagram Illustration of the Full Spectral Subtraction Process

In case of VGRF, the silence/inactivity periods of no activity are actually the inter-step leap or flight time during which no foot contact with the floor is manifested, termed Leap Periods, and this case may be not be found in walking VGRF signals, unlike our database where the subjects are running (see Fig. 2.13). In fact, the leaping periods were detected as the inter-step periods, having already detected the steps, then concatenated into a single vector, then the spectrum of the resulting concatenation is computed. Finally, spectral subtraction is performed between the noisy VGRF signal and the noise spectrum.

Algorithm 3: Spectral Subtraction Denoising

Input: **(**1 x *N*)-vector **VGRF**

Output: (1 x *N*)-vector: **Filtered_VGRF**

Procedure:

1. Extract the inter-step segments (Leap Periods);
2. Create a signal as the concatenation of the extracted inter-step segments;
3. Compute the Discrete Fourier Transform on **VGRF** followed by a magnitude operator;
4. Perform Low-pass filtering to reduce noise variance;
5. Perform post-processing to remove processing distortions;
6. Perform spectral subtraction **PSD_Filtered_VGRF** = **PSD_VGRF**–**PSD_noise;**
7. Compute the Inverse Discrete Fourier Transform to the processed signal;
8. Apply attenuator γ for noise attenuation during silent periods.

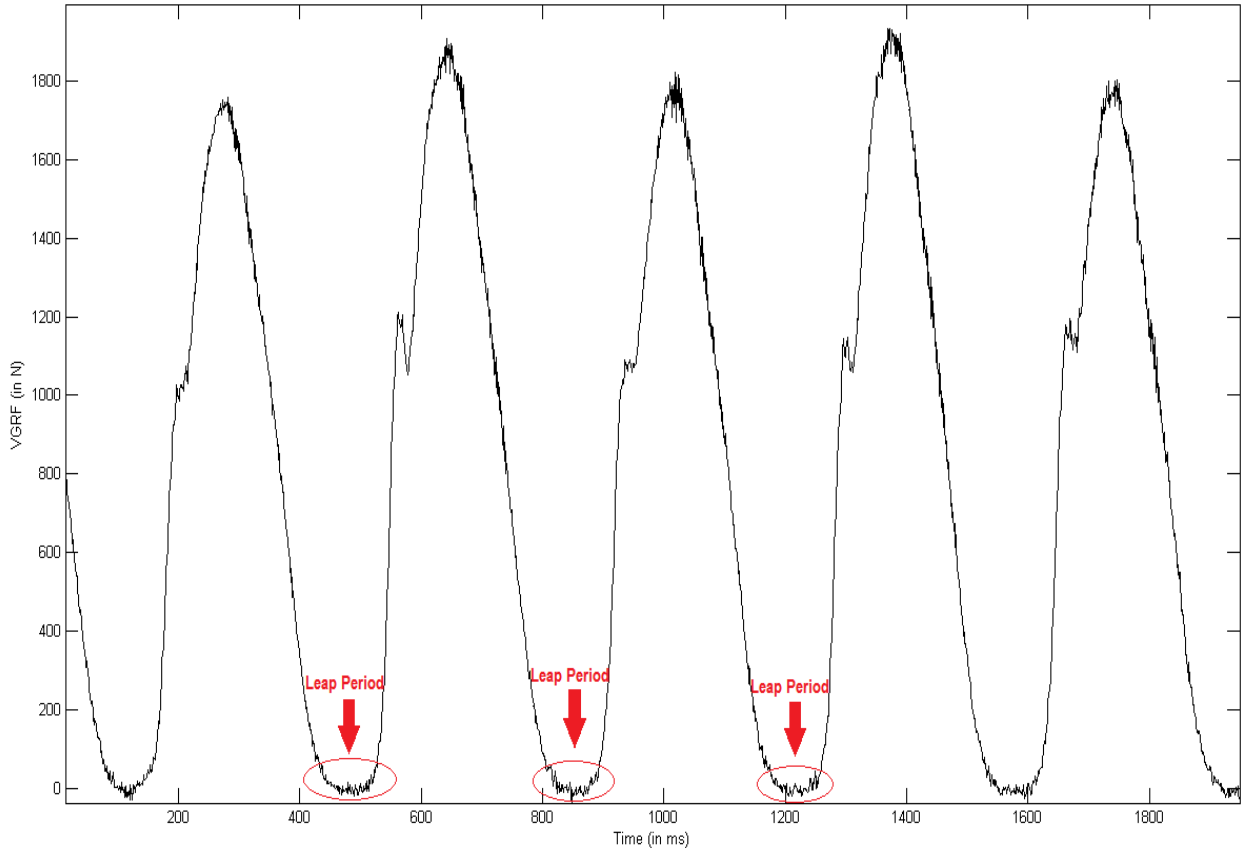

Figure 2.13: Graphical Representation of the Leap Periods in a Sample VGRF Signal

2.3 FATIGUE ANALYSIS

In order to transition from the time domain into the frequency domain, the most widely used technique is the Fourier transform, described in (2.12):

$$F\{x_k\} = X_k \stackrel{\text{def}}{=} \sum_{n=0}^{N-1} x_n \,.\, e^{-2\pi i k n/N}, k \in \mathbb{Z} \tag{2.54}$$

where F is the Fourier operator, X_k is an N-periodic sequence of complex numbers $x_0, x_1, \ldots, x_{N-1}$.

The algorithm used to compute the Fourier spectrum was the Welch Periodogram illustrated in Fig. 2.3 that makes use of the Fast Fourier Transform (FFT) algorithm to generate the Power Spectral Density (PSD) of the analyzed signal. In the resulting periodogram shown in Fig. 2.14, a series of neighboring peaks in the low-frequency domain, at frequencies less than 30 Hz, are present in a cyclic fashion, which

was analyzed as a marker for cyclostationarity. The stride rate (1.4 Hz) and the step rate (2.8 Hz) both lie along with their harmonics in the form of spectral lines. On the other hand, noticeable spectral activity is present in the high frequency domain between 100 and 180 Hz suspected to be system noise, and at 250, 300, 350 and 450 Hz which are the manifestations of the mechanical resonance of the instrumented treadmill, which will be discussed along with the results of the spectral subtraction.

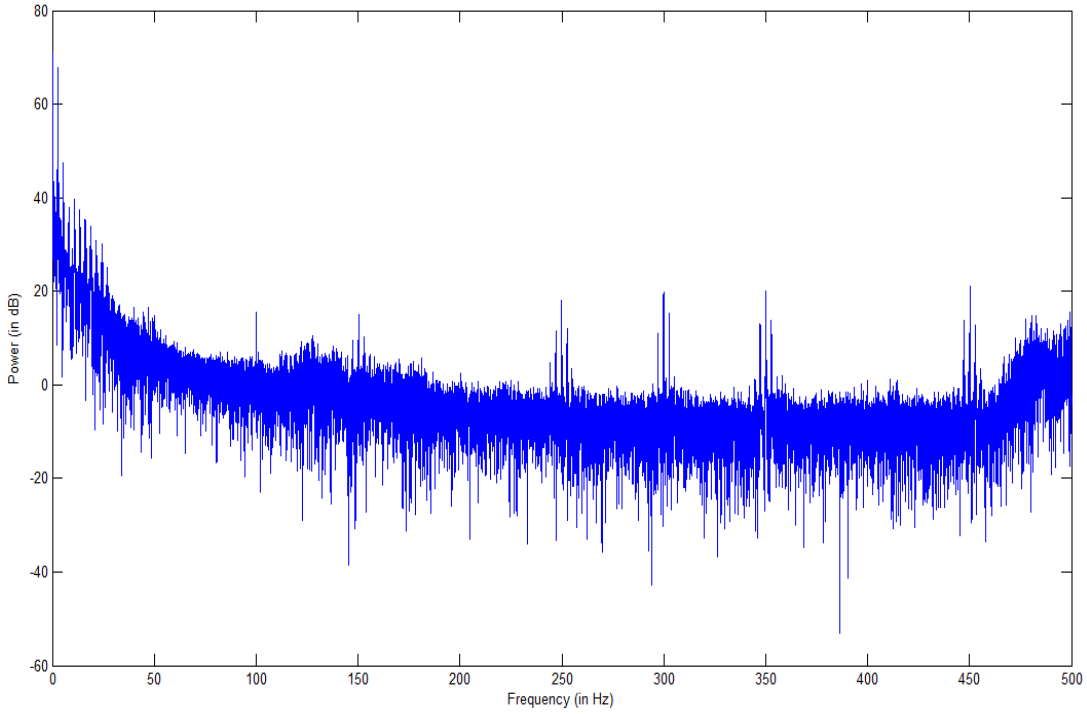

Figure 2.14: Full Range VGRF Welch Periodogram

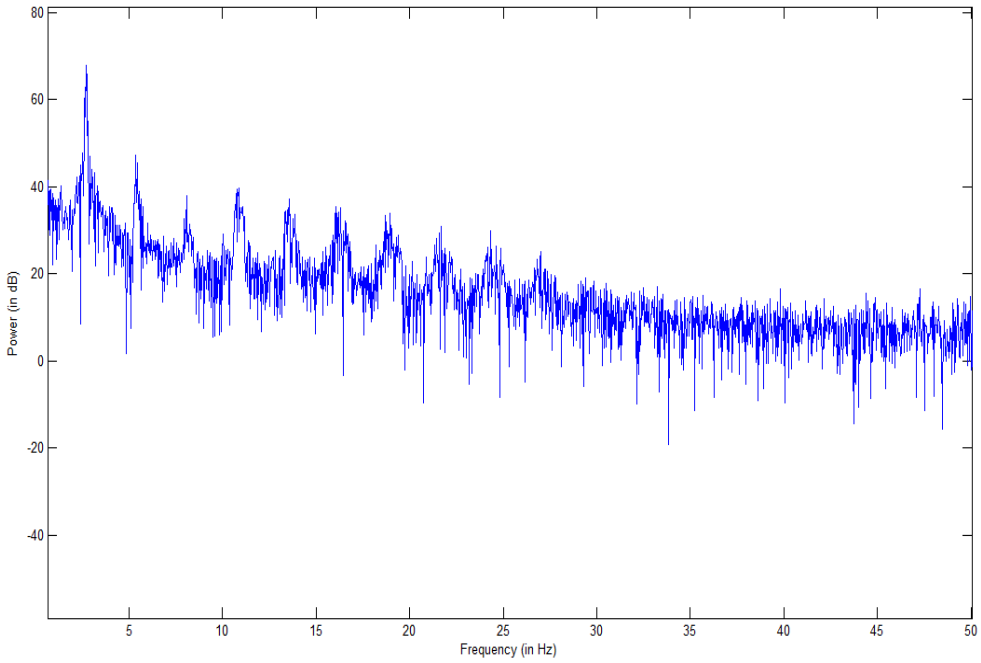

Figure 2.15 Low Frequency Periodogram of a Sample VGRF Signal

That being presented, the idea behind this work was to compare the spectral content of VGRF signals progressively in time and study the effect of fatigue on the spectral characteristics for each subject, hoping to eventually quantify the progression of fatigue and assess the endurance capability of the athletes. For that endeavor, two Welch Periodograms were computed for each subject, at the beginning of the ultra-marathon and after 24 hours of running i.e. the final recording. The purpose behind this analysis was to capture a fatigue indicator with as much accuracy as possible, and the proposed metrics are the Spectral Mean Inter-peak Shift (SMIPS) and the Mean High-frequency Power Variation (MHFPV) defined in (2.13) and (2.14), respectively. The technique using which the frequency-specific power is quantified is the Short-time Fourier Transform (STFT), which generates a numerical descriptor of the variation of frequency-specific power in function of time. Our method compares the high frequency spectral

power at the beginning of the 24-hour ultra-marathon and at the end, describing thereby the progression of fatigue.

$$SMIPS(\%) = \frac{\sum_{n=1}^{N} f(Pk_{n+1}) - f(Pk_n)}{N-1} \tag{2.55}$$

Where $f(Pk_n)$is the frequency at which peak $Pk_n$is manifested, *n* is the peak index, and *N* is the number of peaks in the periodogram.

$$MHFPV(\%) = \frac{P_{end}(f > f_{th}) - P_{start}(f > f_{th})}{P_{start}(f > f_{th})} \tag{2.56}$$

Where $P_{end}(f > f_{th})$ is the end-of-run power at a user-selected high-frequency range f_{th} being the threshold frequency, and $P_{start}(f > f_{th})$ is the beginning-of-run power for the same frequency range.

CHAPTER 3

SINGLE-CHANNEL BLIND SOURCE SEPARATION (SCBSS)

3.1 Problem Formulation

In this thesis, a specific class of BSS is studied, which is the Single-channel BSS (SCBSS). Having stated the class of BSS in previous sections, SCBSS is the extreme case of underdetermined source separation problems, requiring the unsupervised or blind separation of sources from a single observation signal, which combines two obstacles: the absence of prior knowledge of the sources and the least amount of available data to interpret. Despite being a rare scenario if theoretically interpreted, this sort of problems is very common in a variety of fields, and mostly pronounced in the field of audio processing in the analysis of monaural sound waves and isolation of individual instruments, in speech processing [80] in presence of random noise, and in Neuroscience and more specifically in spike sorting [81].

The problem may be formulated in the time domain as in (3.1), and in the time-frequency domain as in (3.2):

$$y(t) = \sum_{i=1}^{N_s} x_i(t) \tag{3.1}$$

where $i = 1,\ldots,N$ represents the number of sources and the aim is to estimate the sources $x_i(t)$ when only the observation signal $y(t)$ is available. This representation is that of an undetermined system.

$$Y(f,\ t_s) = \sum_{i=1}^{N_s} X_i(f, t_s) \tag{3.2}$$

where $Y(f,\ t_s)$ and $X_i(f, t_s)$ denote Time-frequency (TF) components obtained any TF analysis method.

Upon reviewing the major signal processing techniques being used for the characterization of signals, it may be stated that the three main categories under which they all fall are: Time Domain, Frequency Domain and Joint Time-Frequency Domain. Time Domain analysis provides insight on the signal's characteristics in function of time, regardless of the spectral content, while Frequency Domain analysis provides insight on the spectral characteristics of the signal regardless of the temporal aspect. The two stated domains lack the ability of analyzing the variation of the spectral content in function of time, which led to the introduction of joint Time-Frequency analysis in which both spectral and temporal features are simultaneously computed. The methods for joint Time-Frequency Analysis are numerous, of which we state: Cohen Class Bilinear TFD [82], Cohen-Posch TFD [83], Spectrogram, Wavelet Scalogram [84], Matching Pursuit TFD [85] and Adaptive TFD [86].

The purpose behind the characterization of the signals in time-domain and then in the joint time-frequency domain lies in the fact that the implemented technique is a SCBSS technique based on matrix factorization. In other words, in order to separate the sources from a single-channel data vector, a matrix representation of this signal is to be computed for the factorization to take place. One of the most common techniques used in speech analysis and communication, as stated in the above-cited references, is the transformation of the time-domain signal into the joint time-frequency domain, which is a 3-dimensional representation that characterizes the signal in function of time, frequency and power.

Firstly, the transform from the time domain into the frequency domain is performed via Fourier transform is defined as follows:

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t).\, e^{-j\omega t} dt \tag{3.3}$$

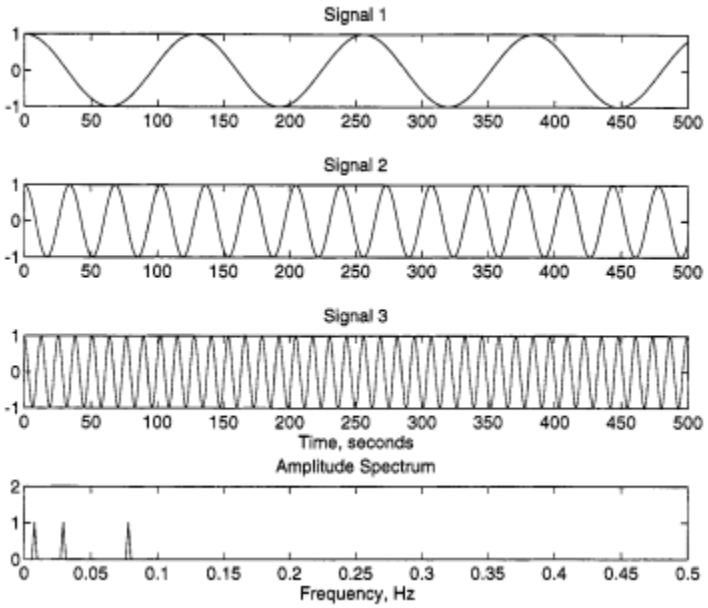

Figure 0.1: Spectral Decomposition of an Arbitrary Signal

However, despite the identification of the spectral content of the signal *x(t)*, the onset time of each frequency component remains unknown, outputting magnitude versus frequency. This calls for a more advanced technique to be able to describe the variation of the spectral content in function of time, of which we define the STFT as follows:

$$STFT_x(t, \omega) = \int x(\tau).\, h(\tau - t).\, e^{-j\omega\tau} d\tau \tag{3.4}$$

where $STFT_x(t, \omega)$ is the magnitude spectrogram of signal *x* and *h(t)* is the window function.

.

Now that the magnitude spectrogram of the studied mixture signal is computed, we have a choice to either keep it in the magnitude form or transform it into the power spectrogram (3.5) by simply squaring up (3.4):

$$SP_x(t,\omega) = |STFT_x(t,\omega)|^2 \tag{3.57}$$

The spectrogram however, as all methods, suffers from a drawback that is window-related; the shorter the window, the higher the time-resolution and the worse the frequency resolution, and vice versa, which implies the existence of a serious trade-off in terms of resolution. That led to the development of more accurate methods from that perspective, which is outside the scope of this work, since the STFT is the implemented TF representation throughout the thesis.

That being cleared away, subsequent information that will include manipulation of the TF domain version of the signals is in fact nothing but the STFT of that signal.

3.2 Problem Classification

SCSS methods are numerous and fall under different categories based on different criteria. The signal mixture being the single-channel observation is transformed from the time domain into another domain, mainly the TF domain, as a basic transform into multi-dimensional data that can be separated into its sources more easily.

An illustration of the general framework for supervised SCSS is presented in Fig. 3.2., which is outside the scope of our work, while an illustration of that of SCBSS is presented in Fig. 3.3.

In supervised methods, the input to the separation system is the single-channel data mixture in addition to the source models' training data. The mixture is then transformed into another

domain and another representation which, when combined with the source and mixing models, facilitates the computation of the separated source estimates.

On the other hand, unsupervised or SCBSS methods deal with the separation of entirely unknown sources having no training data at all, which calls for assumptions that are made to reduce the complexity of the system. Out of these assumptions, we state decorrelation, statistical independence, or the minimum description length principle.

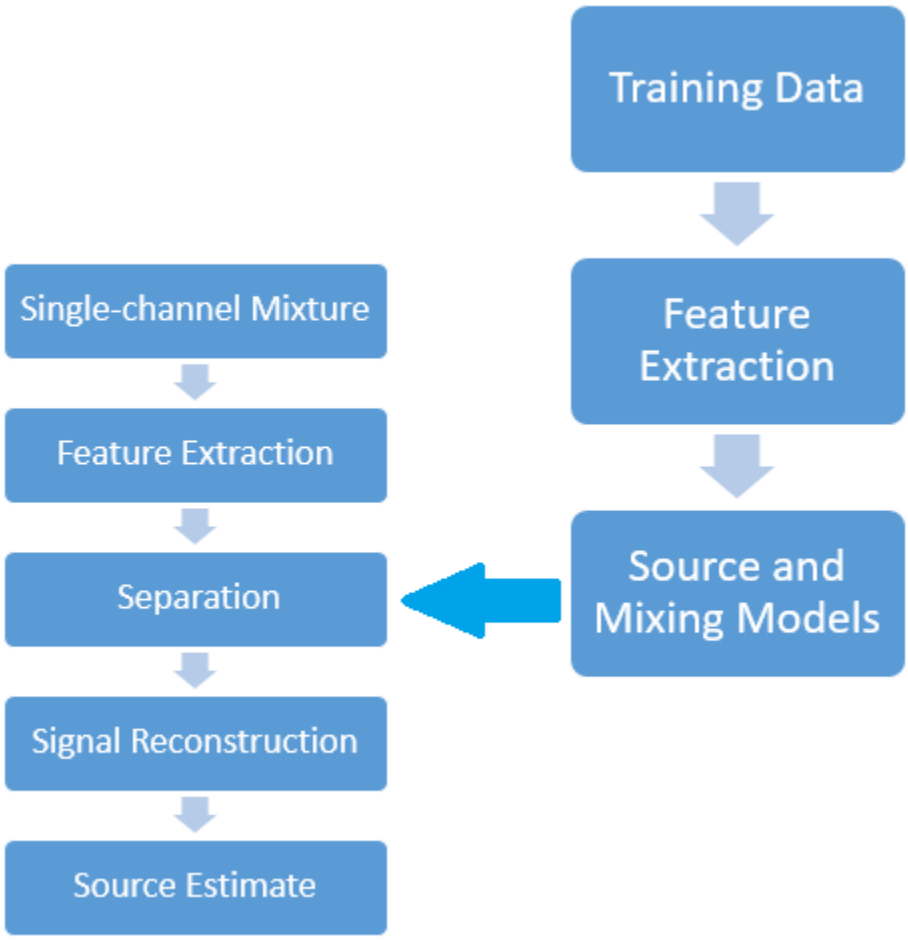

Figure 3.2: Unsupervised SCSS Framework

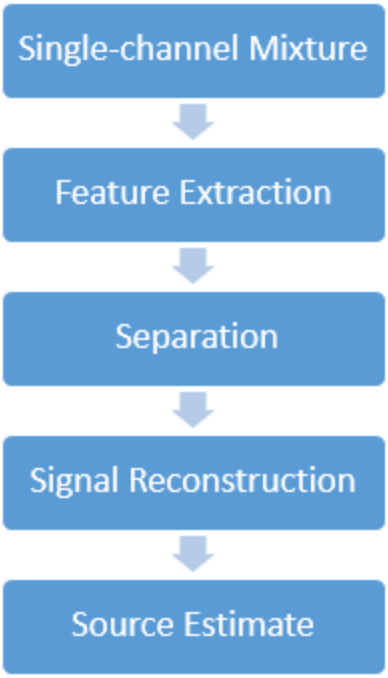

Figure 3.3: SCBSS Framework

The notable difference between unsupervised and supervised SCBSS is the fact that in unsupervised SCBSS the only input to the system is the single-channel observation or signal mixture. And knowing that a 1-D vector is inseparable as it is in the time domain, a domain transformation is necessarily required to create a multi-dimensional data set in form of a matrix so that feature extraction, decomposition and separation are applied upon it, which enters a signal reconstruction process to eventually have a source estimate of separated sources.

In order to accomplish the stated process, a multitude of methods were designed and implemented by researchers in all fields, but the methods are classified according to the separation technique, as shown in Fig. 3.4 leaving us with the following classes: Computational Auditory Scene Analysis (CASA)-based SCBSS [87], EMD-based SCBSS [88], Independent Subspace-based SCBSS [89], and Non-negative Matrix Factorization (NMF)-based SCBSS [90].

Noteworthy, all the stated methods, supervised and unsupervised SCSS, follow the linear instantaneous model which was explained in previous sections.

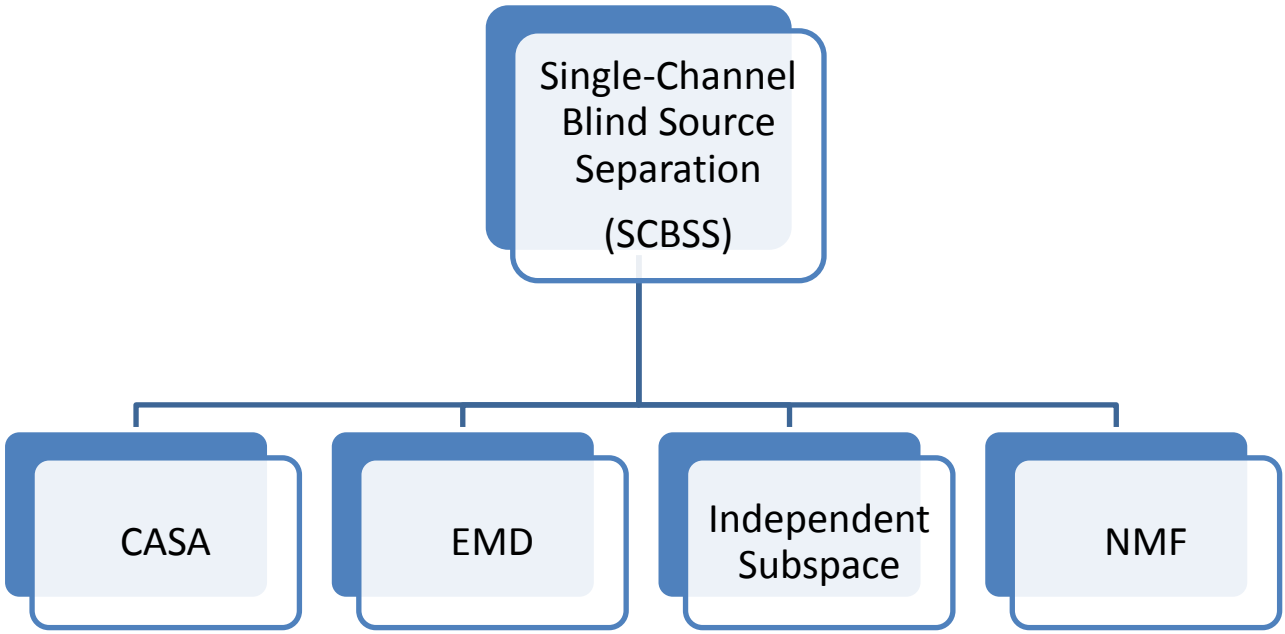

Figure 3.4: SCBSS Class-tree

3.2.1 Computational Auditory Scene Analysis (CASA)-based

The purpose behind the CASA-based SCBSS methods is the replication of the human auditory system's physiology via selection of appropriate domain transformation mainly to the TF domain via STFT or Cochleogram, followed by the segmentation of the observation mixture into TF cells which will then be grouped and associated in terms of specific criteria e.g. harmonicity, common onset, correlated modulation, and duration of sinusoidal parts. Streams of notes (in musical signals) based on pitch proximity are then built, and finally the estimated sources are reconstructed via highly-correlated cluster grouping in the TF domain. It is clear that this class of methods is mainly implemented in applications related to musical signal and audio processing. One of the drawbacks of this technique is the inability to separate musical signals of different streams and sharing the same pitch range, in addition to the fact that the replication of the role of the nervous system in the interpretation of sounds was never replicated, which makes it a partially effective and accurate method.

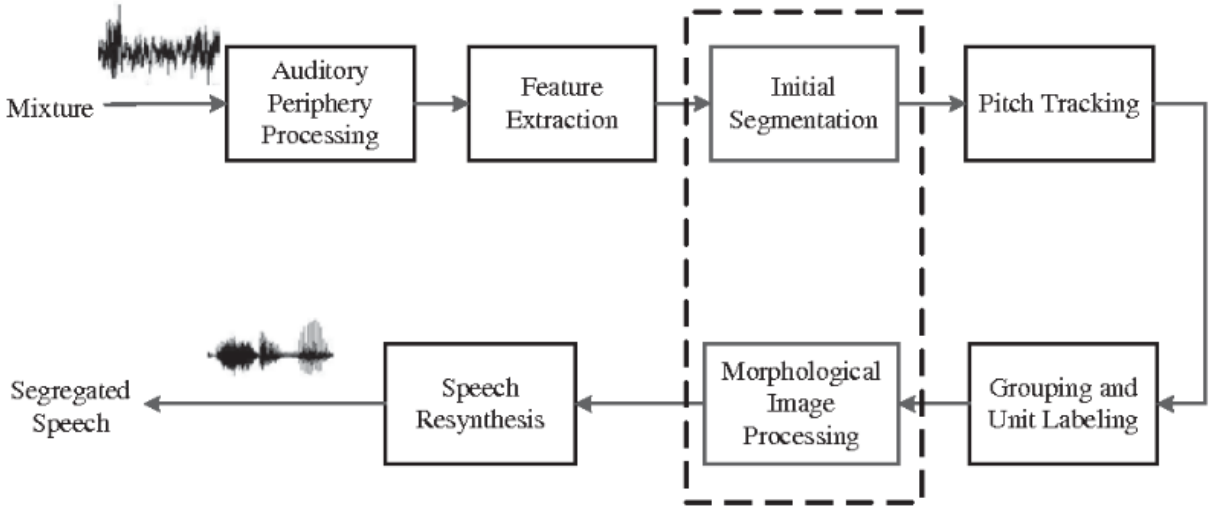

Figure 3.5: CASA Process Block Diagram Illustration

3.2.2 Non-negative Matrix Factorization (NMF)-based

3.2.2. a. Introduction

Factorization-based techniques have been an area of interest for researchers in various fields and especially in BSS. In case of SCBSS, this class of algorithms begins its role upon domain transformation i.e. transition from the time domain into the TF domain, resulting in a TF representation decomposed as follows in (3.6) and (3.7):

$$Y(f,\ t_s) = \sum_{i=1}^{N_s} X_i(f, t_s) \tag{3.6}$$

$$|Y|^{\cdot 2} \approx \sum_{i=1}^{N_i} |X|_i^{\cdot 2} \tag{3.7}$$

where $|\mathrm{Y}|^{\cdot 2} = |\mathrm{Y}(\mathrm{f}, \mathrm{t_s})|^2$ and $|\mathrm{X}|_\mathrm{i}^{\cdot 2} = |\mathrm{X_i}(\mathrm{f}, \mathrm{t_s})|^2$ are 2-dimensional matrices with time instants as row vectors and frequency bins as column vectors. Note that the superscript "$\cdot$" indicates element-wise operations.

In order to get multidimensional data representation out of a single-channel time series, the Power Spectrogram of the signal of interest is computed, and then the NMF algorithm is applied upon the Spectrogram to separate the dominant underlying time-frequency patterns in an unsupervised manner, having minimal a priori knowledge of the sources to be extracted.

NMF is actually an unsupervised data decomposition technique that belongs to latent variable analysis techniques, and it is most commonly implemented in the following fields: topics recovery, feature learning, clustering, temporal segmentation, filtering and source separation, and

coding. It has been successfully applied in text mining [91], in image processing and analysis: unsupervised object discovery [92], object and face recognition [93], tagging [94], denoising and inpainting [95], texture classification [96], spectral data [97], hashing [98], and in EEG processing and analysis: feature extraction [99], and artifact rejection [100]…etc.

But most importantly, NMF is state-of-the-art in BSS, where it has been used in speech separation [101], in music separation [102], in signal enhancement/denoising [103], in compression [104], in music transcription [105].

3.2.2. b. NMF Models

Basically, the Power Spectrogram Y of the signal is factorized into a Basis Vector Matrix W and a Weight Matrix H, as follows in (3.8):

$$V = |Y|^{.2} \approx W.H \tag{58}$$

where $V \in R_{+}^{F \times T_s}$ is the power time-frequency representation of the mixture *y(t)* which is also factorized into two non-negative matrices, $W \in R_{+}^{F \times K}$ and $H \in R_{+}^{K \times T_s}$, W being the data matrix containing a set of spectral basis vectors and H being an encoding matrix that describes the amplitude or activation of each basis vector at each time instant.

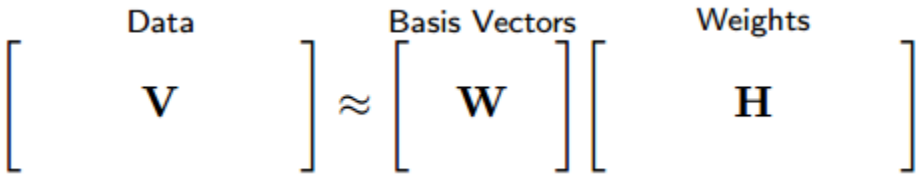

Figure 3.6: NMF Matrix Decomposition Illustration

If K is chosen to be $K=T_s$, factorization will take place, leading to pattern or source separation, while dimensionality remains intact. In fact, K is normally chosen to be a realistic estimation of the number of sources or hidden patterns embedded in the mixture signal.

The product of the two matrices **W** and **H** is randomly initialized and then iteratively optimized via linear divergence-based cost function that expresses the distance between the desired ideal spectrogram **V** and the approximated spectrogram **WH** as follows in (3.9):

$$\min_{W,H\geq 0} D\{\mathbf{V}|\mathbf{WH}\} \tag{59}$$

where D{**V**|**WH**} is a separable matrix divergence (3.10):

$$D\left(V\middle|\hat{V}\right) = \sum_{f=1}^{F}\sum_{n=1}^{N} d\left(v_{fn}\middle|\hat{v}_{fn}\right) \tag{60}$$

And d(x|y) defined for all x, y ≥ 0 is a scalar divergence such that:

- d(x|y) is continuous over x and y;
- d(x|y) ≥ 0 for all x, y ≥ 0;
- d(x|y) = 0 iff x = y;

where *x* and *y* are the terms being compared in the divergence metric.

Various error functions have been developed to assess the reconstruction error between **V** and **WH**, of which we state: the Euclidean Distance (squared error), the Kullback-Leibler (KL) Divergence and the Itakura Saito (IS) [106] as in (3.11), (3.12), and (3.13), respectively:

$$E_{EUC}(\mathbf{W},\mathbf{H}) = \|\mathbf{V}-\mathbf{WH}\|^2 = \frac{1}{2}.\sum_{i,j}(V_{ij}-(\mathbf{WH})_{ij})^2 \tag{61}$$

$$E_{KL}(\mathbf{W},\mathbf{H}) = \mathbf{V}.\log_{10}\frac{\mathbf{V}}{\mathbf{WH}} - \mathbf{V} + \mathbf{WH} \tag{62}$$

$$E_{IS}(\mathbf{W},\mathbf{H}) = \frac{\mathbf{V}}{\mathbf{WH}} - \log_{10}\frac{\mathbf{V}}{\mathbf{WH}} - \mathbf{1} \tag{63}$$

Most commonly, the STFT is the TF representation referred to whenever NMF-based SCBSS algorithms are designed. The NMF-based SCBSS became popular in the music industry, where it is used extensively for filtering of audio signals and separation of instrumental sounds from a monaural signal.

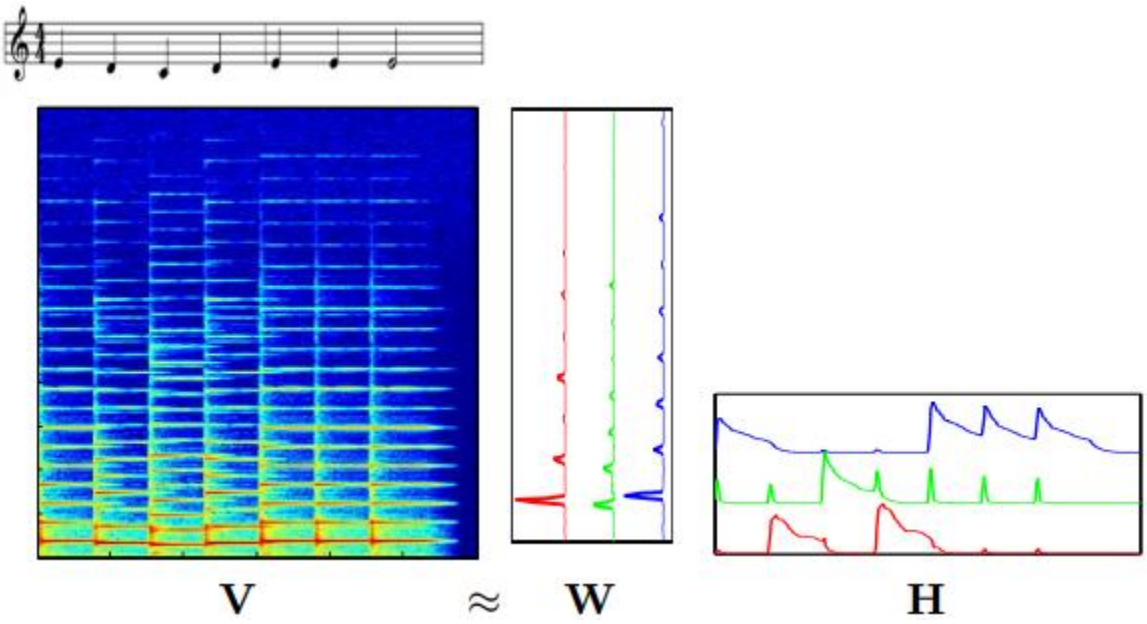

Figure 3.7: Spectrogram Decomposition via NMF Illustration

This divergence measure *d(x|y)* is assessed in terms of convexity properties w.r.t to the variables being x and y. Convexity is a property describing the shape of the curve of divergence versus x and y, independently, as illustrated in Fig. 3.8.

Divergence $d(x\|y)$	EUC	KL	IS
Convex on x	yes	yes	yes
Convex on y	yes	yes	no

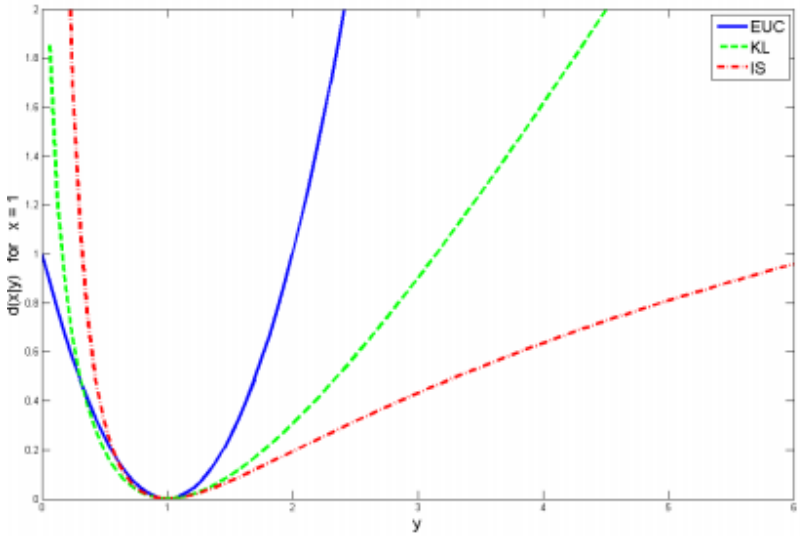

Figure 03.8: Divergence Convexity Graphical Representation

Convexity is a major property in divergence analysis that indicates the shape of the divergence curve in terms of the presence of local and global minima, which aids in the prediction of the performance of the subsequent optimization algorithm through which the minimum of the cost function is determined. In fact, convexity is studied in terms of x and in terms of y, the best case of which is the convexity on both levels as is the case with the Euclidean Distance.

Another property to be assessed in cost function minimization is scale invariance, which describes the effect of scaling one or all variables in the divergence functions on the whole divergence measure.

$$d_{EUC}(\lambda x|\lambda y) = \lambda^2 . d_{EUC}(x|y) \tag{3.14}$$

$$d_{KL}(\lambda x|\lambda y) = \lambda . d_{KL}(x|y) \tag{3.15}$$

$$d_{IS}(\lambda x|\lambda y) = d_{IS}(x|y) \tag{3.16}$$

The IS divergence is scale invariant, which means it provides higher accuracy in the representation of data with large dynamic range e.g. audio spectra.

Actually, the two main divergence families are the β-divergence and the α-divergence, having the following representations:

$$d_\beta(x|y) = \begin{cases} \frac{1}{\beta(\beta-1)}\left(x^\beta + (\beta-1)y^\beta - \beta x y^{\beta-1}\right) & \beta \epsilon \Re\{0,1\} \\ x\log\frac{x}{y} - x + y & \beta = 1 \\ \frac{x}{y} - \log\frac{x}{y} - 1 & \beta = 0 \end{cases} \tag{3.17}$$

$$d_\alpha(x|y) = \frac{1}{\alpha(\alpha-1)}(\alpha x + (1-\alpha)y - x^\alpha y^{1-\alpha}) \tag{3.18}$$

Therefore, when β=0, the IS divergence is computed, when β=1 the KL divergence is computed, and when β=2 the EUC divergence is computed.

The NMF divergence is chosen based on the data under analysis as well as the application. It can be chosen by intuition, or by prior knowledge of the studied system, or by invariance need.

Noteworthy, any divergence out of the mentioned divergences has a probabilistic model equivalent, where the divergence minimization is equivalent to a maximum likelihood criterion [107]:

$$D\left(V\middle|\hat{V}\right) = -\log p\left(V\middle|\hat{V}\right) + const \tag{3.19}$$

In Table 3.1, examples of equivalency are presented.

Table 3-1: Divergence Measures and Related Probabilistic Models

Divergence $D(\mathbf{V}\|\widehat{\mathbf{V}})$		Probability distribution	p.d.f. $p(\mathbf{V}\|\widehat{\mathbf{V}})$
EUC	$\sum_{f,n}(v_{fn}-\hat{v}_{fn})^2$	$v_{fn} \sim \text{Gaussian}\left(\hat{v}_{fn}, \sigma^2\right)$	$\prod_{f,n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(v_{fn}-\hat{v}_{fn})^2}{2\sigma^2}\right)$
KL	$\sum_{f,n}\left(v_{fn}\log\frac{v_{fn}}{\hat{v}_{fn}} - v_{fn} + \hat{v}_{fn}\right)$	$v_{fn} \sim \text{Poisson}\left(\hat{v}_{fn}\right)$	$\prod_{f,n} \frac{1}{\Gamma(v_{fn}+1)} \hat{v}_{fn}^{v_{fn}} \exp\left(-\hat{v}_{fn}\right)$
IS	$\sum_{f,n}\left(\frac{v_{fn}}{\hat{v}_{fn}} - \log\frac{v_{fn}}{\hat{v}_{fn}} - 1\right)$	$v_{fn} \sim \text{Exponential}\left(\frac{1}{\hat{v}_{fn}}\right)$	$\prod_{f,n} \frac{1}{\hat{v}_{fn}} \exp\left(-\frac{v_{fn}}{\hat{v}_{fn}}\right)$

3.2.2. c. Algorithms for solving the NMF problem

The solution to the NMF optimization problem is modeled as follows:

$$\min_{\mathbf{W},\mathbf{H}\geq 0} D(\mathbf{V}|\mathbf{WH}) \Leftrightarrow \min_{\theta} C(\theta) \stackrel{\text{def}}{=} D\left(\mathbf{V}|\mathbf{WH}\right) \tag{3.20}$$

where $\theta \stackrel{\text{def}}{=} \{\mathbf{W}, \mathbf{H}\}$ denotes the NMF parameters, bounded by the following conditions:

- $\mathbf{W}, \mathbf{H} \geq 0$;
- Uniqueness of the solution is not guaranteed;
- Multiple local and global minima might be encountered.

This implies that optimization in our case is a double optimization: optimization of **W** and optimization of **H.** Hence, two options are available: Optimize **W** and **H** in parallel, or optimize over one matrix e.g. **W** while keeping the other matrix **H** known and fixed, which is termed the Alternating Optimization Strategy (AOS). In fact, in most divergences, convexity is fulfilled separately w.r.t. **W** and w.r.t. **H** but not w.r.t {**W**, **H**}.

That led to the state-of-the-art NMF optimization algorithms based on iterative IOS whose algorithm translates into the following iteration-wise steps: Update **W** given **H** is fixed, then update **H** given **W** is fixed. In some references, this method is termed Block-coordinate Descent.

As for the iterative updates of the variables, which is a core element in optimization problems, a wide range of techniques was proposed and implemented, some more commonly than the others, most of which we present in Table 3.2.

On each iteration of the algorithm, a new value for **W** and **H** is computed by either additive update (AU) rules or multiplicative update (MU) rules. In both cases, the goal is to improve the quality of the approximation monotonically by applying updates to the variables, **W** and **H** in our case. If so, convergence is guaranteed, at least locally, to an optimal matrix factorization.

Gradient descent update is probably the simplest of all techniques in terms of implementation difficulty, being a first order iterative optimization algorithm; however it lacks the speed of convergence required by most applications. The concept behind this rule is finding the local minimum of a function by taking steps that are proportional to the negative of the gradient of the function at the current point as follows in (3.21):

$$a_{n+1} = a_n - \gamma \nabla F(a_n) \tag{3.21}$$

where *F(x)* is a multivariable function that is defined and differentiable in the vicinity of a point a and $\nabla F(a)$ is the gradient of F at a.

For a small enoughγ, $F(a_n) \geq F(a_{n+1})$, which means the function is in monotonic decrease until reaching the desired optimum.

Starting with a guess x_0 as a local minimum of F, a sequence is created as follows in (3.22):

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n - \gamma_n \nabla F(\boldsymbol{x}_n), \qquad n \geq 0 \tag{3.22}$$

As for the Newton-like algorithms, Newton's method is fairly simple to implement as well, which is based on the successive finding of better root approximations of the roots of a real-

valued function. The method actually begins with a function *f* defined over real numbers *x*, the function's derivative *f'* and an initial value x_0 of a function root, approximating thereby x_1 as follows in (3.23):

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{3.23}$$

And the process is repeated until an acceptable value is reached.

As for the MU rules proposed by Lee and Seung, which is nothing but a diagonally rescaled gradient descent rule, they chose it to be a good compromise in terms of convergence speed and implementation simplicity, and they proved its convergence to the optimal local minimum of the cost function. However, it slightly differs from divergence metric to the other. For instance, the update rule for LS is:

$$\mathbf{W} \leftarrow \mathbf{W}.\frac{|\mathbf{Y}|^{.2}\mathbf{H}^{\mathbf{T}}}{\mathbf{W}\mathbf{H}\mathbf{H}^{\mathbf{T}}} \; ; \mathbf{H} \leftarrow \mathbf{H}.\frac{\mathbf{W}^{\mathbf{T}}|\mathbf{Y}|^{.2}}{\mathbf{W}^{\mathbf{T}}\mathbf{W}\mathbf{H}} \tag{64}$$

while the update rule for KL is:

$$\mathbf{W} \leftarrow \mathbf{W}.\frac{(|\mathbf{Y}|^{.2}./\mathbf{W}\mathbf{H})\mathbf{H}^{\mathbf{T}}}{\mathbf{1}\mathbf{H}^{\mathbf{T}}} \text{ and } \mathbf{H} \leftarrow \mathbf{H}.\frac{(|\mathbf{Y}|^{.2}./\mathbf{W}\mathbf{H})\mathbf{W}^{\mathbf{T}}}{\mathbf{1}\mathbf{W}^{\mathbf{T}}} \tag{3.25}$$

where '.' and './' denote the element-wise multiplication and division, respectively, and '1' is an all-one *F* by T_s matrix.

Table 3-2: Update Rule Descriptive Table

Algorithm	Advantages	Drawbacks
Multiplicative Update (MU)	• Easy implementation; • Guaranteed non-negativity of **W** and **H**	• Unguaranteed decreasing monotonicity of cost; • Relatively slow convergence rate;
Gradient-like	Faster convergence than MU	Non-negativity constraints must be explicitly handled;
Newton-like	Faster convergence than gradient-like and MU algorithms.	Non-negativity constraints must be explicitly handled;
Expectation-Maximization	• Non-negativity constraints are implicitly handled; • Possibility of introducing other constraints via	• Slower convergence than MU rules; • Limited to NMF with probabilistic

	probabilistic priors;	formulation;
Stochastic Gradient-like	Ability to handle big data and continuous data streams.	N/A

3.2.2. d. Model Order Choice, Initialization and Stopping Criteria

Another major factor in the NMF algorithms is the model order *k*, which determines the number of columns of **W** and the number of rows of **H**, which is most commonly user-settable and less commonly automatically computed or predicted by the algorithm itself. In fact, *k* is usually fixed throughout the NMF decomposition based on a value chosen by intuition or on prior knowledge of the expected number of sources/clusters, and in a few cases, the model order is automatically set within the algorithm.

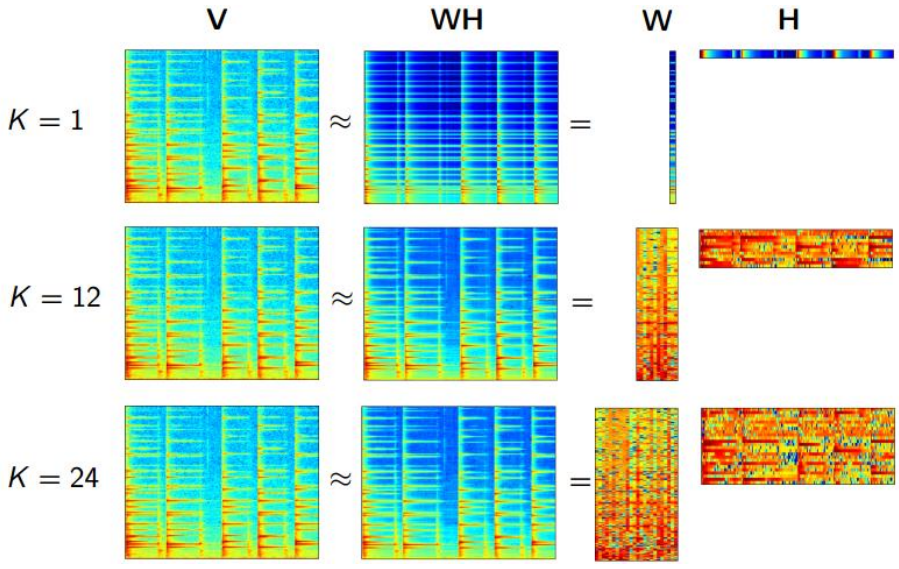

Figure 0.9: Effect of Model Order Choice on Matrix Decomposition

As for the initialization of **W** and **H,** it has high importance in all local optimization problems due to the fact that local minima exist. However, common practice dictates the random initialization of non-negative matrices several times and keeping the solution with the lowest

final cost. Nevertheless, some structural data-driven initializations were devised e.g. initialization of **W** by the clustering of data points **V**, and the initialization of **W** by Singular Value Decomposition (SVD) of data points **V**, and many other similar techniques.

In order for the algorithm to stop, stopping criteria are required for the loop to end at wherever the user programs it to end at, for the number of iterations is crucial in iterative optimization; the iteration number in fact raises the issue of a tradeoff between computational cost from one side, and the approximation error and model quality form the other side. That being said, the following are the main stopping criteria in iterative optimization:

- Fixed number of iterations (user-settable) e.g. 1000;
- Approximation error (cost) threshold e.g 2.5%;
- Approximation error relative decrease threshold e.g. 1%;

3.2.2. e. Regularization

The main problem with NMF is the non-uniqueness of the solution; in fact, NMF and other matrix factorization techniques rely on divergence minimization as discussed, which optimally reaches zero, however, the exactness of the linear product does not imply that each of the two matrices in the product are the target matrices. This leads to an unguaranteed extraction of latent components as desired within the target application.

The aforementioned drawback of NMF is possibly solved by imposing some knowledge-based constraints on **W**, on **H**, or on both **W** and **H**, which usually add uniqueness to the decomposition and the extraction of more suitable latent components.
The more the constraints are realistic and customized to the system under study, the higher the

probability of successful separation. This is done by adding terms to the objective function that mathematically translate the imposed constraint, according to the following model:

$$C(\mathbf{W},\mathbf{H}) = D(\mathbf{V}|\mathbf{WH}) + \lambda\varphi(\mathbf{H}) + \mu\theta(\mathbf{W}) \tag{3.26}$$

where φ and θ encode the desired structures termed penalty functions and λ and μ control the strength of the penalty.

This method of penalization actually acts oppositely to the direction of the minimization by adding cost, which forces the optimization algorithm to go through more iterations until satisfactory values of the penalty functions are reached, meaning the constraints are met up to an acceptable extent.

Out of the various classes of penalties, we state the sparsity-inducing penalties [108], group sparsity-inducing penalties [109], smoothness-inducing penalties [110], graph regularized NMF [111], orthogonal NMF [112], tri-NMF [113], and much more.

Upon matrix decomposition, the recovery of the sources being the ultimate goal is to be performed. This process is termed re-synthesis of the time-frequency representation and is composed of two main phases: Getting the mixture phase via overlap-add (OLA) method and reversing the STFT using the Inverse STFT (ISTFT) for each of the separated source signals.

In fact, the OLA method is explained as follows: Each successive window is taken back into the time domain using the IFFT, then each window is shifted by the step size, and added to the result of the previous shift. The OLA is mandatory because of the degree of window overlap performed during the forward STFT as previously described. The concept of OLA is depicted in Fig. 3.10.

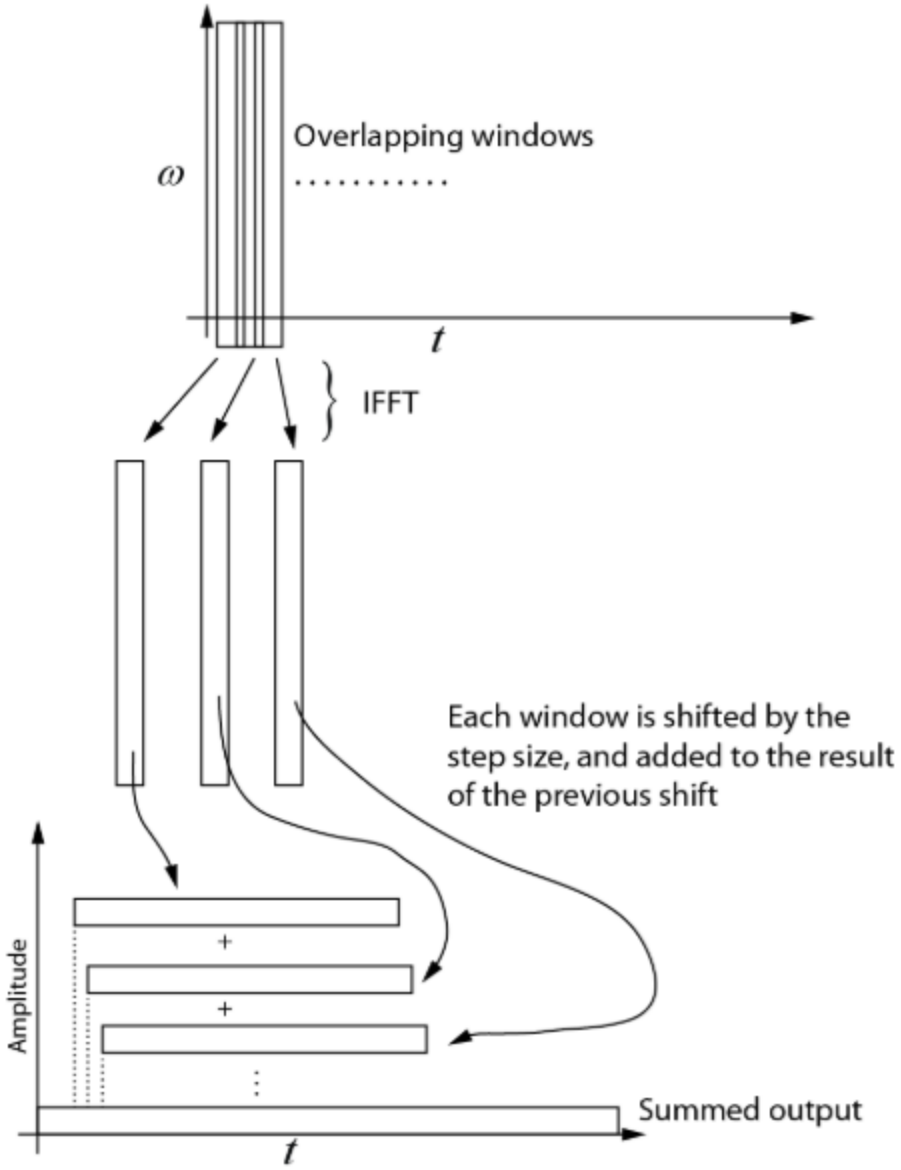

Figure 0.10: OLA Illustration

An illustration of the entire NMF SCBSS process is shown in Fig. 3.11.

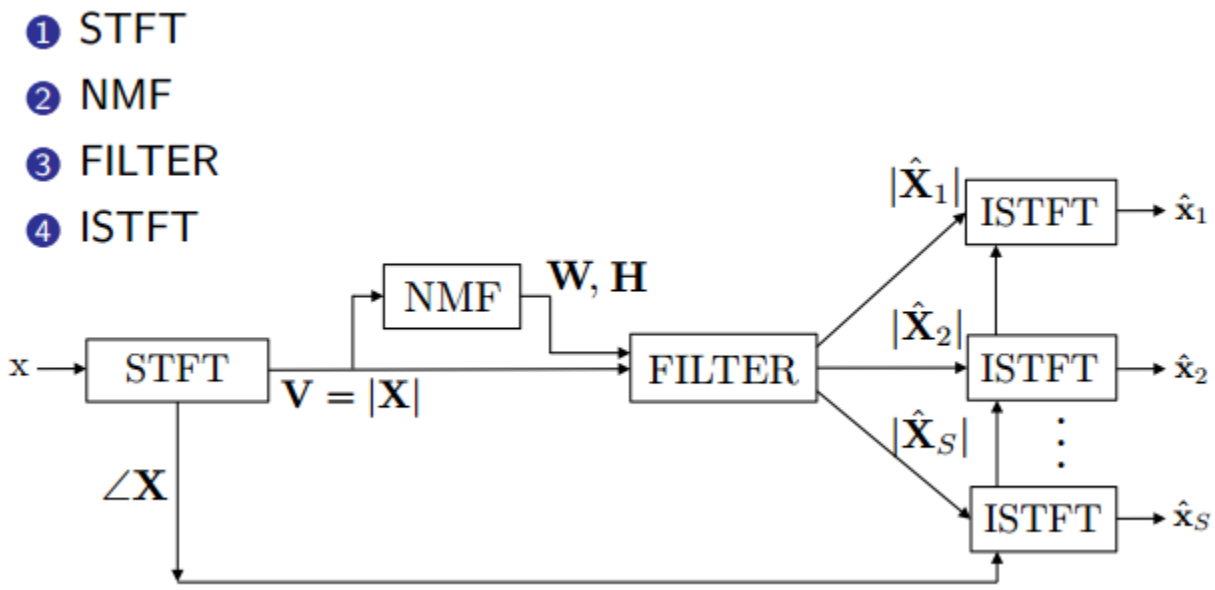

Figure 0.11: NMF SCBSS Block Diagram Illustration

3.2.3 Empirical Mode Decomposition (EMD)-based

The EMD or Hilbert-Huang data-driven technique is a non-linear technique pioneered by N. E. Huang *et al*. for the adaptive representation of non-stationary signals as sums of zero-mean AM-FM components to be effective [114].

EMD is based on the consideration of local level oscillations in non-stationary non-linear signals, decomposing each signal into simple oscillatory functions termed Intrinsic Mode Functions (IMFs). The IMFs are classified as local *details* and local *trends*, being the high-frequency and low-frequency components, respectively. In mathematical terms, the EMD decomposition is modeled as follows in (3.27):

$$y(t) = \sum_{n=1}^{N} c_n(t) + {r_N}^{EMD}(t) \tag{3.27}$$

where c_n(t) is the n^{th} IMF, N is the total number of IMFs, and ${r_N}^{EMD}(t)$ is the final residue.

The EMD basic algorithm is presented as follows:

Algorithm: EMD

Input: **(**1 x *N*)-vector *x(t)*

Output: IMFs

Procedure:

1. Identify all extrema of *x(t)*;
2. Interpolate between minima to get lower envelope $e_{min}(t)$
3. Interpolate between maxima to get upper envelope $e_{max}(t)$
4. Extract the detail d(t) = x(t) – m(t)
5. Iterate on the residual *m(t)*

A more refined version of this process is termed sifting which is the performance of steps 1 to 4 upon the detail signal *d(t)* until a zero-mean signal is reached according to a stopping criterion. Once this is reached, the detail is then considered an IMF and step 5 is applied and the residual is computed. Obviously, the number of extrema decreases gradually from residual to the other and completion of the decomposition is guaranteed, outputting a finite number of modes.

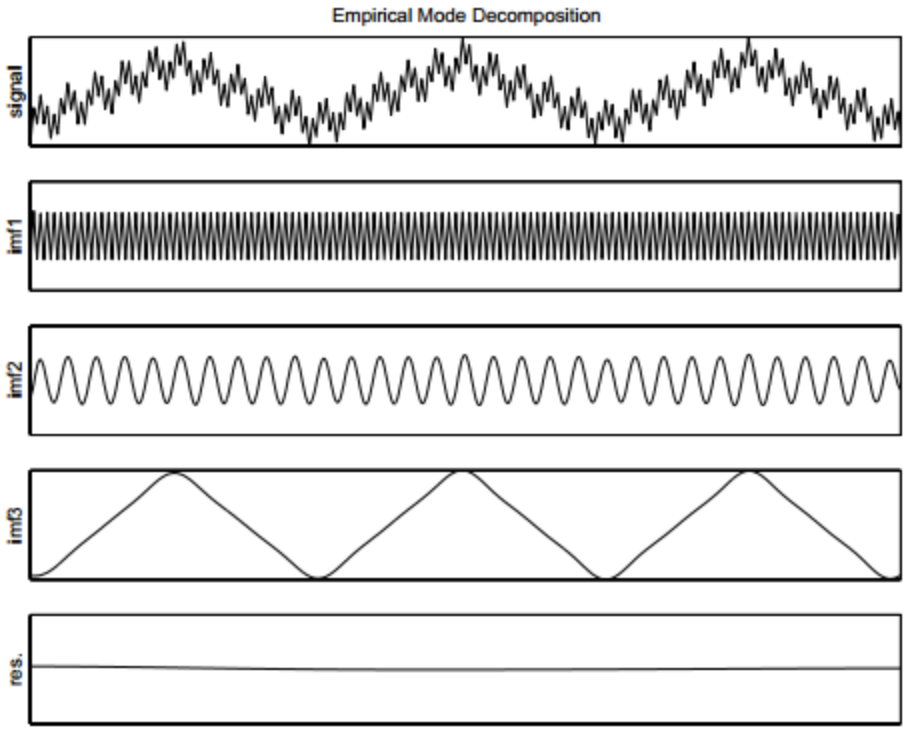

Figure 3.12: EMD of a 3-component Signal – Non-linear Oscillations

Upon computation of the EMD of a single-channel observation signal, a successive SCBSS step is performed upon the resulting IMFs, each of which consisting of a sub-band of frequencies with a reduced degree of mixing. Subsequently, IMF spectrograms are computed either via STFT or Hilbert Transform, and matrix decomposition of the spectrogram takes place via a multitude of techniques of which we state EMD-NMF [115], EMD-2D-Sparse NMF [89], EMD-Hilbert Spectrum ISA [116]. However, the EMD 2D-Sparse NMF developed by Gao [89] is of major significance to the field due to its high effectiveness in comparison with all other algorithms, and its algorithm is illustrated in Fig. 3.13 below. Noteworthy, 2D-sparse NMF is a regularized NMF approach having a sparsity-inducing penalty embedded in its cost function, as discussed earlier.

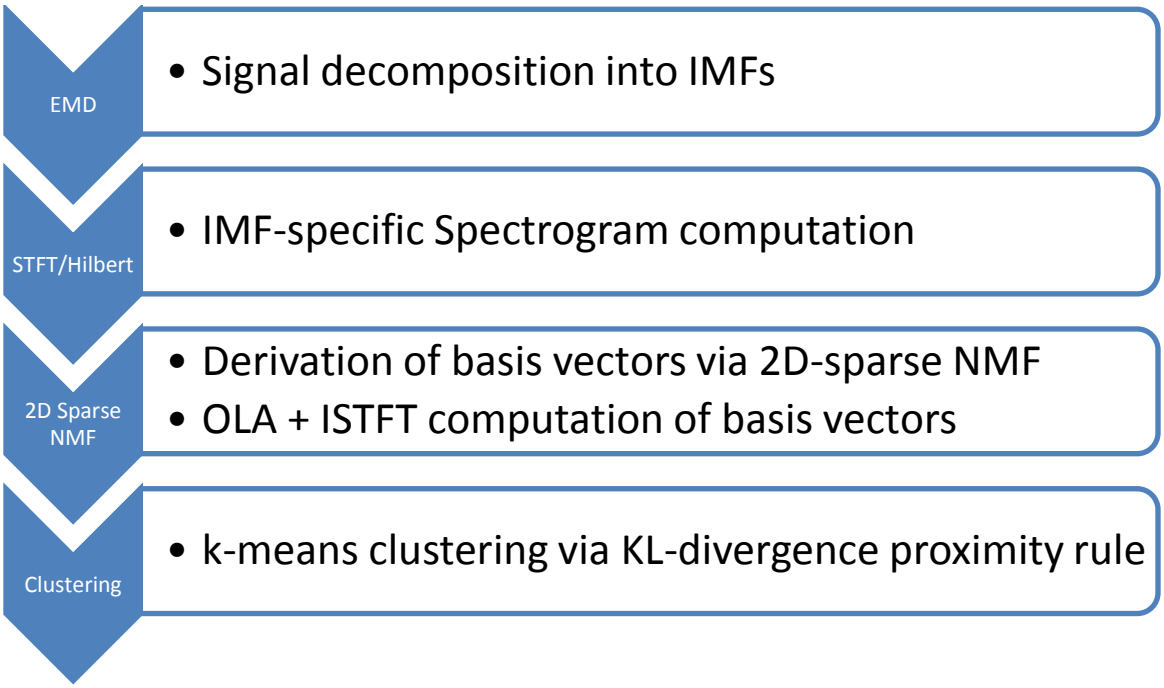

Figure 0.13: EMD-2D Sparse-NMF Diagram

3.2.4 Independent Subspace Analysis (ISA)-based

An additional approach in the field of SCBSS is based on ISA techniques. Briefly, the main idea behind the concept of ISA is the decomposition of the TF space of mixture signals as being the sum of independent source subspaces. In other words, having computed the mixture TF representation $|\mathbf{Y}|^{\cdot 2}$, each time frame of it is modeled as a weighted sum of N_{ISA} independent basis vectors $\mathbf{z}_{\boldsymbol{\rho}}^{\boldsymbol{ISA}}$ as follows:

$$\underline{\mathbf{y}}_{t_s} = \sum_{\rho=1}^{N_{ISA}} \mathbf{w}_{\rho,t_s}^{ISA} \mathbf{z}_{\rho,t_s}^{ISA} \tag{3.28}$$

where each basis vector is weighed by a time-varying scalar w_{ρ,t_s}^{ISA}. Hence, each source is spanned by basis vector subset defining thereby a subspace, being a matrix with basis vectors in columns$\mathbf{Z}_i^{ISA} = [\mathbf{z}_{1,i}^{ISA}, \dots, \mathbf{z}_{N_{ISA}^i,i}^{ISA}]$.

In ISA techniques, projection of the input $\underline{y}_{t_s}$ onto each basis component in the subspace results in the computation of the weight coefficients as follows, assuming orthonormal components:

$$\mathbf{w}_i^{ISA^T} = \mathbf{Z}_i^{ISA^T} \underline{\mathbf{y}}_{t_s} \tag{3.29}$$

This is now the projection of $\underline{\mathbf{y}}_{t_i}$on the subspace spanned by the basis vectors$\mathbf{Z}_i^{ISA}$, thereby, through successive and iterative projections onto the *I* sets of basis vectors, $\underline{\mathbf{y}}_{t_s}$ is decomposed into independent subspaces as follows:

$$\underline{\mathbf{y}}_{t_s} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{I}} \mathbf{Z}_i^{ISA} \mathbf{w}_i^{ISA^T} \tag{3.30}$$

Now customizing the model to the required spectrogram decomposition, all time frames are to be covered and it is estimated as follows:

$$|\widetilde{\mathbf{Y}}|_i^2 = \mathbf{Z}_i^{ISA} \mathbf{W}_i^{ISA^T} \tag{3.31}$$

where $\mathbf{W}_i^{ISA^T} = [\mathbf{w}_{i,t_s}^{ISA}, \dots, \mathbf{w}_{\rho,T_s}^{ISA}]$.

As a final step, OLA and ISTFT are applied to reconstruct the time domain source signal from $|\widetilde{\mathbf{Y}}|_i^2$, as previously described. However, the fact that segment overlapping is a part of the STFT algorithms i.e. a considerable amount of cross-spectral terms is encountered, it is arguable that STFT is not the best way to go in order to compute the time-frequency representation in ISA-based SCBSS techniques because that would lead to weak separation efficiency. Instead, some have used the Hilbert Spectrum where the cross-spectral terms are practically absent.

CHAPTER 4

CYCLOSPARSE NON-NEGATIVE MATRIX FACTORIZATION

4.1 Problem Formulation

In this thesis, a novel approach is introduced to handle SCBSS tasks adapted to applications where single-channel signals are of periodic and cyclically-repeating pattern and most importantly occurring as short bursts or transient spikes that occur once or few times per cycle. We summarized this concept by describing the target signals as cyclosparse signals, as introduced by Sabri *et al.* [76], combining thereby two features that are cyclic-behavior and sparseness. Despite the common name, the presented method in this work is based on other concepts and hence different mathematical model than in [76]. As a first thought, one would think of signals that have the following features in the time domain, which is incorrect thinking; in fact, since the target mixture signals are single-channel signals, as previously detailed, the BSS section of the algorithm requires a multidimensional signal representation e.g. TF spectrogram matrix as input, and that matrix should be characterized by the user beforehand so that prior knowledge of the expected sources creates some quantifiable constraints to be added to the cost function for optimal separation to take place.

First of all, an investigation on how a cyclosparse signal in the time domain is actually represented in the TF domain and it is over the TF parameters that the BSS code is required to act. That being said, let's define a cyclosparse signal: A cyclosparse signal is a signal that demonstrates cyclically-repeating transient peak-like oscillations of minimum occurrence in a single cycle, ideally once, surrounded by mostly null values.

An example of such a signal may be a periodic pulse train having a single sharp pulse in each cycle or period, as shown in Fig. 4.1

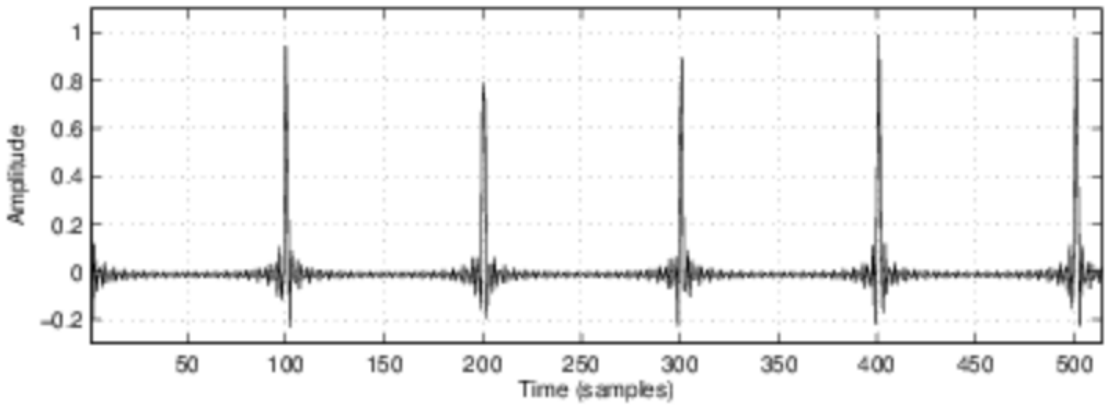

Figure 0.1: Cyclosparse Pulse Train

Since the TF representation actually preserves the time characteristics of the signal before transformation, any time-related characteristic for any frequency component shall be preserved, i.e. a cyclic 100 Hz peak appearing every 10 seconds shall lead to an equivalent 100 Hz component repeating every 10 seconds on the spectrogram along with any coexisting frequency component at that particular time window. An illustration of that concept is shown in Fig. 4.2.

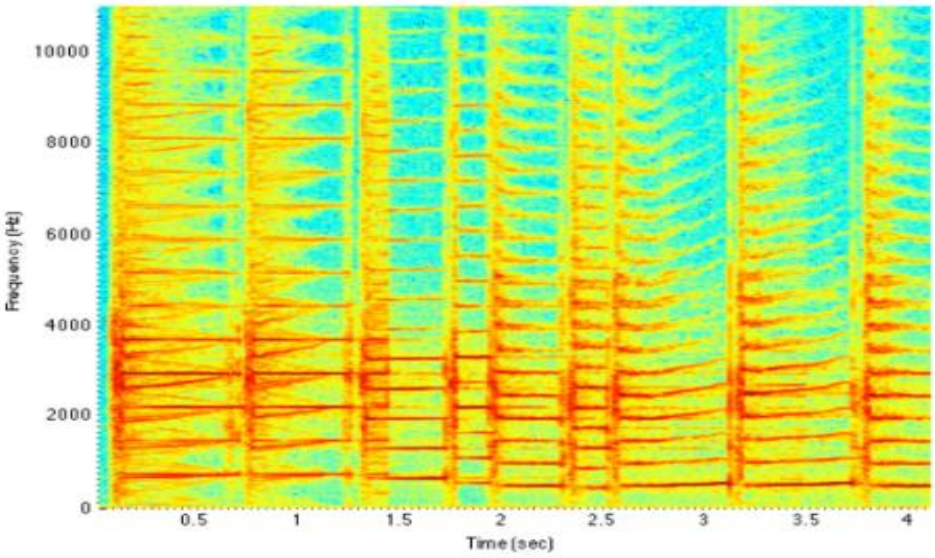

Figure 0.2: Cyclic Spectrum in the TF Domain

4.2 Cost Function Design

As explained in previous chapters, in order to reach the desired matrix factorization in NMF that guarantees a high level of solution uniqueness, unconstrained divergence minimization is not enough due to the fact that the resulting **W** and **H** despite being as close as possible to the power spectrogram $|\mathbf{Y}|^{\cdot 2}$ are not necessarily the desired matrices for the specific application. As a solution to the standard divergence measure, constrained optimization via penalty functions is implemented such that the added penalty functions target source-specific characteristics known about the source signals, which are in our case periodicity and sparseness.

Optimization is the process of minimizing or maximizing the value of an objective function by updating one or more variables in an accurate and customized manner until reaching the desired output as generally presented in (4.1).

$$\min_{x \in M} f(x) \tag{4.65}$$

Where the function $f(x)$: $\mathfrak{R}^n \rightarrow \mathfrak{R}$ is called the objective function and the set M is the feasible set of (O).

Based on the description of function *f* and the feasible set *M*, the problem (O) may be classified into one of the following classes: linear, quadratic, semi-infinite, semi-definite, multiple-objective, discrete optimization…etc.

In cases as ours, the objective function is the divergence function that measures the error between the original spectrogram and the approximated/estimated spectrogram, as presented in (4.2):

$$\min_{\mathbf{H} \geq \mathbf{0}} C(\mathbf{H}) \stackrel{\text{def}}{=} D(\mathbf{V}|\mathbf{WH}) \tag{4.66}$$

Where *D* is the divergence function as previously explained.

The algorithm, however, shall update both **W** and **H**; for that endeavor, sequential update is performed by updating **W** keeping **H** fixed, and then updating **H** keeping **W** fixed, sequentially via alternation as described.

The criterion function *C(***H***)* separates into $\sum_n D(\mathbf{v}\,|\mathbf{Wh})$ where $\mathbf{v_n}$ and $\mathbf{h_n}$ are the nth column of **V** and **H** respectively, which leaves us with (4.3):

$$\min_{\mathbf{h}\geq\mathbf{0}} C(\mathbf{h}) \stackrel{\text{def}}{=} D(\mathbf{v}|\mathbf{Wh}) \tag{4.67}$$

Where $\mathbf{v} \in \mathfrak{R}_+^F, \mathbf{W} \in \mathfrak{R}_+^{F\mathbf{x}K}$ and $\mathbf{h} \in \mathfrak{R}_+^K$.

Moreover, supplementary function of W and H (or both) are seldom added to the basic cost function C(.) in order to add regularization of the factor matrix estimates, reflecting thereby some sort of prior belief or assumption or constraint on the desired matrices. The added terms are termed penalty function, and the NMF algorithm becomes a penalized NMF algorithm, as shown in (4.4):

$$\min_{\mathbf{h}\geq\mathbf{0}} C(\mathbf{h}) \stackrel{\text{def}}{=} D(\mathbf{v}|\mathbf{Wh}) + L(\mathbf{h})\ subject\ to\ \mathbf{h} \geq 0 \tag{4.68}$$

Where *L*(**h**) is the penalty term.

4.2.1 Divergence Measure

As a divergence measure, the KL divergence (3.5) was adopted for its accuracy, scale properties, and convexity properties that were discussed in Chapter 1.

$$E_{KL}(\mathbf{W},\mathbf{H}) = \mathbf{V}.\mathbf{log_{10}}\frac{\mathbf{V}}{\mathbf{WH}} - \mathbf{V} + \mathbf{WH} \tag{4.69}$$

4.2.2 Periodicity Penalty Function

$P(H, \alpha, \beta)$ is the penalty function or regularization term that induces periodicity of the basis activation function $\mathbf{H}_{k,1}$, …, $\mathbf{H}_{k,T}$; T is the total number of time frames in the TF representation. The iterative optimization updates **H** leading to simultaneous variation in the periodicity regularization term which assesses its degree of periodicity: the more periodic the activation is, the lower $\mathrm{P}(\mathrm{H}, \alpha, \beta)$becomes, the less contribution this term has to the cost, and vice versa.

In NMF, matrices **W** and **H** are non-negative, making it possible to approximate the logarithm of $\mathbf{H}_{k,t}$ by the weighted sum of harmonically constrained sinusoids [117] as follows in (4.6), which actually is nothing but the Fourier series:

$$\log H_{k,t} \cong \sum_{m,n} (\alpha_{m,n}^{k} \cos nP_m t + \beta_{m,n}^{k} \sin nP_m t) \tag{4.70}$$

where $\alpha_{m,n}^{k}$, $\beta_{m,n}^{k} \in \mathbb{R}$ determine the repetitive pattern's shape on the $1/P_m$ interval, P_m being the fundamental frequency which is in our case the cyclic frequency of the cyclostationary signal computed as in [75], and n denotes the order of the harmonic partial. This guarantees that the extracted source is periodic with frequency P_m. Hence, $\log U_{k,t}$ is the sum of M arbitrary periodic functions.

In order to recover $H_{k,t}$ and using (3.3), it may be approximated as the exponential of the sum of the sinusoids as in (4.7):

$$H_{k,t} = \mathrm{e}^{\log H_{k,t}} \cong \sum_{m,n} (\alpha_{m,n}^{k} \cos nP_m t + \beta_{m,n}^{k} \sin nP_m t) \tag{4.71}$$

The implemented method automatically computes an appropriate cyclic frequency according to the best model fit, and it may be input by the user, as stated, if the cyclic frequency is known a priori.

Most importantly, the cost function $P(\mathbf{H}, \alpha, \beta)$is defined as the Euclidean distance-based divergence between the two terms of the equality in (4.3):

$$P(\mathbf{H}, \alpha, \beta) = \sum_{k,t} \left| \log U_{k,t} - \sum_{m,n} (\alpha_{m,n}^{k} \cos nP_m t + \beta_{m,n}^{k} \sin nP_m t) \right|^2 \tag{4.72}$$

4.2.3 Sparseness Constraint

As for the sparse behavior of the impact waveform, a data vector signal is said to be sparse if the majority of its values is zero; the more non-zero elements are expressed, the less sparse the data vector becomes. In mathematical terms, sparseness is mostly computed using the l_p-*norm*, which is defined as follows:

$$l_p = \|x_i\|_p = \sqrt[p]{\sum_i |x_i|^p} \tag{4.73}$$

where x_i are the elements of the data vector and p is the order of norm l.

The l_p norm is actually the cardinal of non-zero elements in the data vector x which directly indicates its degree of sparseness.

From a visual perspective, the sparse behavior of any signal is manifested in the form of a sharp impulse or a few sharp impulses embedded in a mostly *all-zero* vector.

Our main objective is to be able to impose degrees of sparseness on one or both of matrices **W** and **H**, and the concerned matrix is to be known beforehand through intuition and prior vision of

the expected system. That being said, we decided to impose sparseness on the activation matrix **H**, since we assume the system to have sparsely activated spectral patterns, but not necessarily sparsely-shaped spectral patterns.

Hence, our problem is formulated as follows:

Given a non-negative spectrogram matrix **V** of size *N*x*T*, find the non-negative matrices **W** and **H** of sizes *N*x*K* and *K*x*T*, respectively, such that the divergence function is minimized under user-settable sparseness levels as in (4.10) and (4.11):

$$Sparseness(x) = \|x\|_{1,2} = \frac{\sqrt{n} - (\sum_{i=1}^{N}|x_i|)/\sqrt{\sum_{i=1}^{N} x_i{}^2}}{\sqrt{n} - 1} \tag{4.74}$$

where $\|x\|_{1,2}$is a mixture of the l_1 and l_2 norms of signal *x*, *n* is the dimensionality of *x* This function reaches unity i.e. a value equal to 1 if and only if *x* contains a single non-zero component, and on the other extreme reaches a null value if and only if all components are different than zero. An illustration of this metric is shown in Fig. 4.3.

Figure 0.3: Illustration of Different Sparseness Leveled Vectors in Histogram Format

$$Sparseness(h_i) = S_h, \forall i \tag{4.75}$$

Where $\mathbf{h}_i$ is the i[th] row of **H**, and S_h is the desired level of sparseness to be imposed on **H**, and these two in addition to *K* are user-settable.

For simplicity purposes, scale is of no particular interest since $\mathbf{w}_i\mathbf{h}_i = (\mathbf{w}_i\lambda)(\frac{\mathbf{h_i}}{\lambda})$ and L_2 is fixed to unity.

Combining the two observed characteristics of the VGRF signal, cyclostationarity and sparseness, this type of signals is termed *cyclosparse* signal, as introduced by Sabri [76]. as a periodic random impulse signal *x(t)* of period *T* having the cardinal of nonzero elements of the entire signal equal to the cardinal of nonzero elements of the signal over any period multiplied by the number of cycles *k* as follows in (4.12).

$$\|x\|_{1,2} = \|x_{1,:}\|_{1,2} + \|x_{2,:}\|_{1,2} + \cdots + \|x_{k,:}\|_{1,2} \tag{4.76}$$

In the present study, each cycle *k* is known to have a single impulse, the cycle being a single step having one impact pattern, represented in Fig. 4.4 for *k*=3.

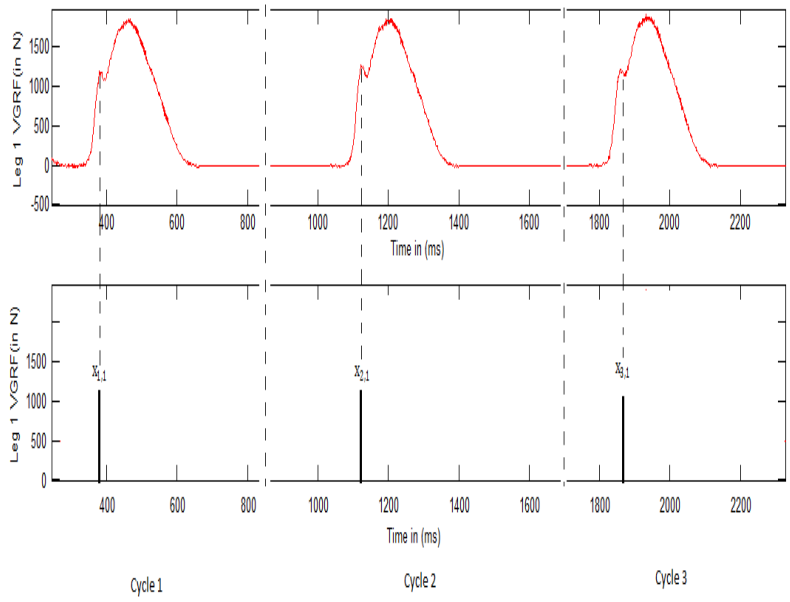

Figure 0.4: Single-peaked Cyclosparse Signal Representation of the Impact Pattern

That being explained, our final cost function should have the following form (4.13):

$$C(\mathbf{W},\mathbf{H}) = E_{KL}(\mathbf{W},\mathbf{H}) + \omega_P . P(\mathbf{H},\alpha,\beta)$$
$$\textit{Such that sparseness}(h_i)=S_{\mathbf{h}},\ \forall \boldsymbol{i} \tag{4.77}$$

where $\omega_P > 0$ is a regularization parameter, $D(\mathbf{V}|\mathbf{WH})$ is the divergence measure between the observed spectrogram and the current NMF model, $P(\mathbf{H}, \alpha, \beta)$ is the periodicity-inducing penalty function, and

4.3 Optimization Algorithm

4.3.1 Sparseness Constraint Optimization

The cost function being developed, an algorithm is required to update the values inside the matrix factor **H** until the stopping criterion is reached. The adopted algorithm was a modified gradient descent algorithm updated and devised by Hoyer [108], which he termed projected gradient descent algorithm with sparseness constraints, onto which we have added the periodicity penalty function whose optimization will be explained in the next section. The gradient descent as previously explained is an algorithm that updates the variables by taking steps in the direction of the negative gradient, and the addition of the sparseness constraint required it to be a projected method i.e. projects onto the constraint space, which is the sparseness criterion imposed by the user. The algorithm is described in Algorithm 3 as follows:

Algorithm 3: NMF with Activation Sparseness Constraint

1. Initialize **H** to a random positive matrix;
2. If the set sparseness constraint applies on **H,** project each row of **H** to be non-negative, have unit L_2 norm and L_1 set to achieve desired sparseness;
3. Iterate
   (a) If sparseness on **H** apply:
      i. Set $\mathbf{H} := \mathbf{H} - \mu_H \mathbf{W}^T(\mathbf{WH} - \mathbf{V})$

ii. Project each row of H to be non-negative, have unit L2 norm and L1 norm set to achieve the desired sparseness

else take standard multiplicative step $\mathbf{H := H \otimes (W^T V) \oslash (W^T W H)}$

where $\otimes$ and $\oslash$ denote element-wise multiplication and division, respectively, and $\mu_{\mathbf{H}}$ is a small positive constant step size which is automatically set by the software.

However, enforcing sparseness and enforcing non-negativity are performed through the following scheme in Algorithm 4.

Algorithm 4: Activity Sparseness and Non-negativity Projection

1. Set $s_i := x_i + \frac{L_1 - \sum x_i}{length(\mathbf{x})}, \forall \mathbf{i}$;
2. Set Z:={};
3. Iterate:
   (a) Set $m_i := \begin{cases} L_1/(\dim(\mathbf{x}) - size(Z) & if\ i \notin Z \\ 0 & if\ i \in Z \end{cases}$
   (b) Set $s := m + \propto (s - m)$ where $\propto \geq 0$ is selected s.t. s satisfies the L_2 norm constraint through solving a quadratic equation;
   (c) If all components of **s** are non-negative, return **s**, end;
   (d) Set $Z := Z \cup \{i; s_i < 0\}$;
   (e) Set $s_i := 0, \forall i \in Z$;
   (f) Calculate $c := \frac{\sum s_i - L_1}{\dim x - size(Z)}$
   (g) Set $s_i := s_i - c, \forall i \notin Z$
   (h) Go to (a)

However, the update process is not straightforward if **W** and **H** were to be simultaneously updated. In fact, an alternating method is most commonly referred to e.g. Alternating Least Squares (ALS), Alternating Least Squares with Optimal Brain Surgeon (ALS-OBS), Alternating Non-negative Least Squares using Projected Gradients (ALS-PG)…etc.

The vast majority of the mentioned algorithms take 3 inputs which are: Original Matrix to be factorized (spectrogram), Number of Components (anticipated number of sources to be separated), the Maximum Number of Iterations or the minimum error. The codes for the stated update methods are found in the Appendix.

Noteworthy, not all cost functions are convex in nature, i.e. convergence of the algorithm is not guaranteed. In fact, if the function is not convex, multiple extrema would exist, which raises the probability of getting stuck in a local extremum and ending the process at an early stage. When such a problem arises, common practice states the necessity to design an auxiliary function.

The $\mathfrak{N}_+^K$ x $\mathfrak{N}_+^K \rightarrow \mathfrak{N}_+$mapping G($\mathbf{h}|\tilde{\mathbf{h}}$) is said to be an auxiliary function to C(**h**) if and only if:

- $\forall \mathbf{h} \in \mathfrak{N}_+^K, C(\mathbf{h}) = G(\mathbf{h}|\tilde{\mathbf{h}})$
- $\forall (\mathbf{h}, \tilde{\mathbf{h}}) \in \mathfrak{N}_+^K$ x $\mathfrak{N}_+^K, C(\mathbf{h}) \leq G(\mathbf{h}|\tilde{\mathbf{h}})$

Clearly stated, the auxiliary function $G(\mathbf{h}|\tilde{\mathbf{h}})$ is an upper bound to *C*(**h**), which is tight for $\mathbf{h} = \tilde{\mathbf{h}}$ **(see Fig. X).** Hence, the iterative optimization of $G(\mathbf{h}|\tilde{\mathbf{h}})$ replaces the optimization of C(**h)** since any iterate $\mathbf{h}^{(i+1)}$ satisfying G($\mathbf{h}^{(i+1)}|\mathbf{h}^{(i)}) \leq G(\mathbf{h}^{(i)}|\mathbf{h}^{(i)})$ satisfies $C(\mathbf{h}^{(i+1)}) \leq C(\mathbf{h}^{(i)}$**.**

Therefore, the typical choice of the iterate is in (4.14):

$$\mathbf{h}^{(i+1)} = \arg\min_{\mathbf{h} \geq 0} G(\mathbf{h}|\mathbf{h}^{(i)}) \tag{4.78}$$

The mathematical proof of the auxiliary function concept is found in Appendix B.

4.3.2 Periodicity Term Optimization

In order to optimize the periodicity function, we aim to find optimal values for $\mathbf{H}, \alpha, and\ \beta$, and since the optimization process is not straight forward, the use of an auxiliary function that is optimized instead of the original cost function is performed as described in the previous section. The choice of the auxiliary function is clearly explained in [117] as in (4.15) using Jensen's equality:

$$I^{+}(\mathbf{W},\mathbf{H}) = \sum_{w,t} V_{w,t} \log V_{w,t} - V_{w,t} \sum_{k} \lambda_{w,k,t} \log \frac{W_{w,k} H_{k,t}}{\lambda_{w,k,t}} - V_{w,t} + X_{w,t} \qquad (4.1579)$$

where $\lambda_{w,k,t}$ is a positive weight satisfying $\sum_k \lambda_{w,k,t} = 1$.

An exact upper bound is achieved when (4.16) is satisfied:

$$\lambda_{w,k,t} = \frac{W_{w,k} H_{k,t}}{\sum_k W_{w,k} H_{k,t}} \qquad (80)$$

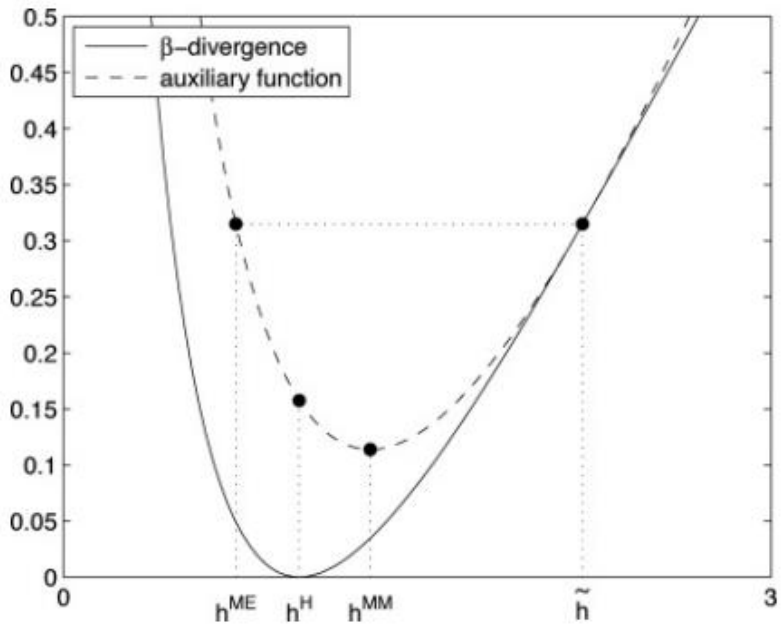

Figure 4.5 Auxiliary Function Upper Bound Graphical Illustration

Whenever the stopping criterion is discussed, convergence should be the first thing to come to mind. In fact, logically speaking, the algorithm should not stop until an acceptable error value has been reached implying the convergence, i.e. the desired factorization according to the imposed penalty-based constraints is met. However, in some applications where computational cost is to be minimized, the maximum number of iterations is imposed onto the algorithm so that regardless of the error value, the algorithm shall stop and output whatever separation results are computed. An illustration of the convergence tracking curve is shown in Fig. 4.6.

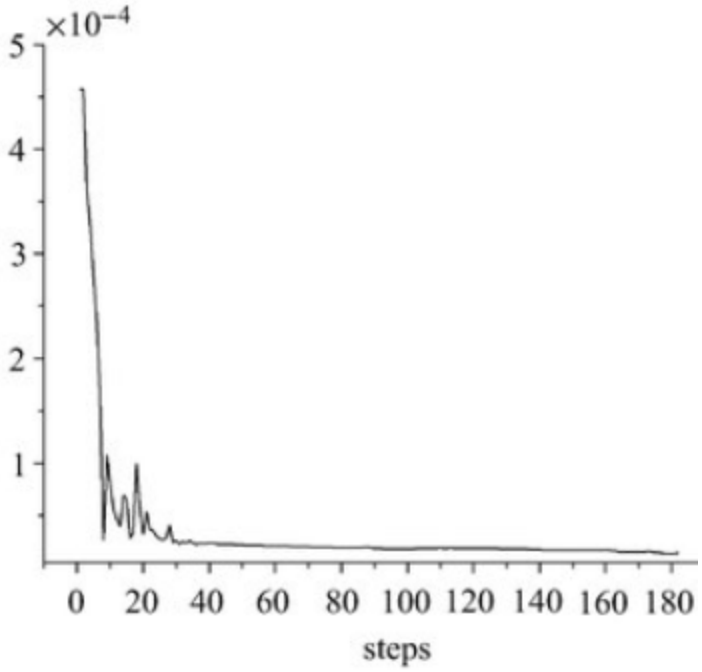

Figure 04.6: Convergence Curve Illustration

CHAPTER 5

RESULTS AND DISCUSSION

5.1 VGRF Processing and Analysis

5.1.1 VGRF Processing

To recall, in order to optimally process the VGRF signals with the highest SNR and lowest degradation level of the sharp edges of the Impact Pattern, the traditional filtering techniques that require a user-set cutoff frequency or band-pass/reject frequency range were avoided upon thorough assessment of the published results in the literature. Instead, we have implemented the Spectral Subtraction technique commonly implemented to reduce noise levels in speech or music data, and adopted it the VGRF signals. The inputs to this method are the noisy VGRF signal vector in addition to the leap periods as shown in Fig. 5.1.

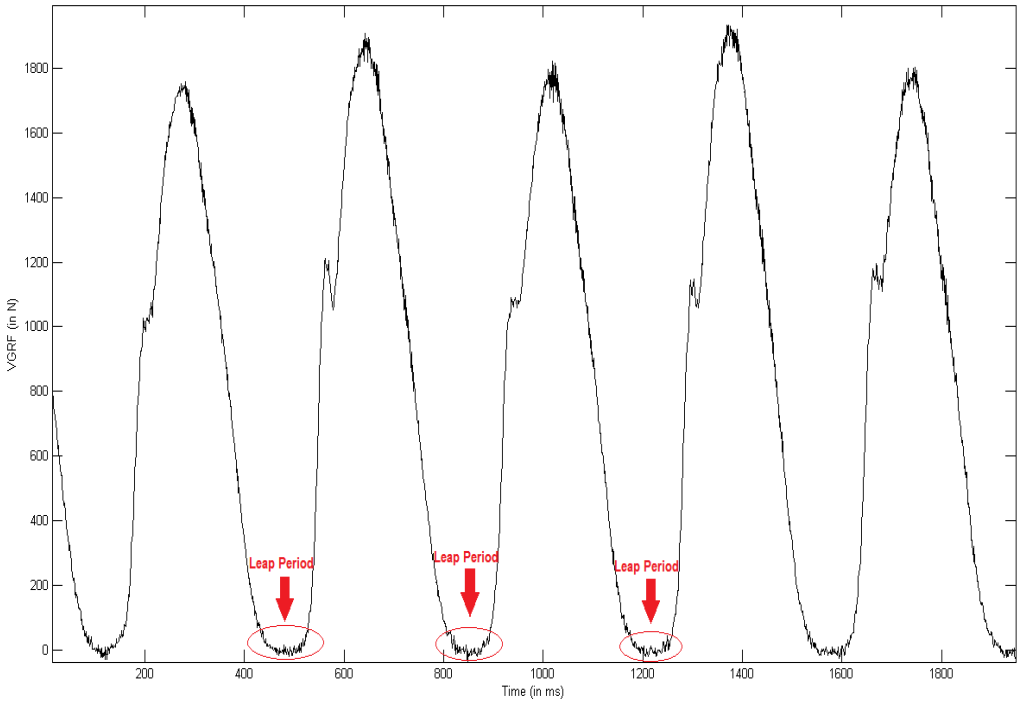

Figure 5.1: Leap Periods in VGRF Signals

Once extracted, the leap periods that are the no-activity periods during the run were concatenated into a single vector as shown in Fig. 5.2.

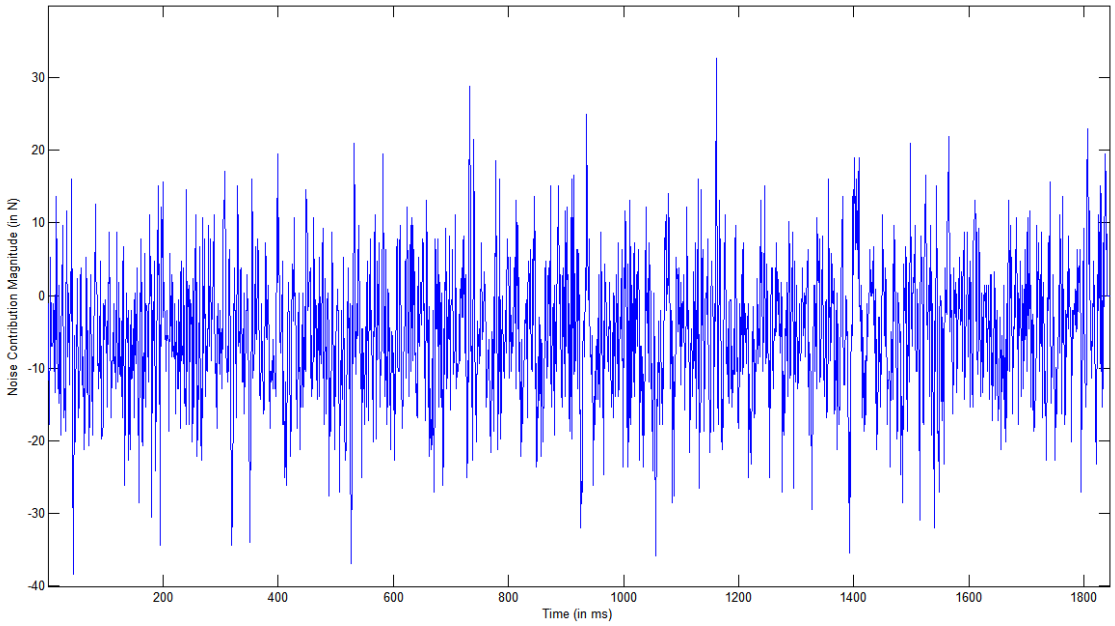

Figure 5.2: Sample Concatenated Leap Period Signals

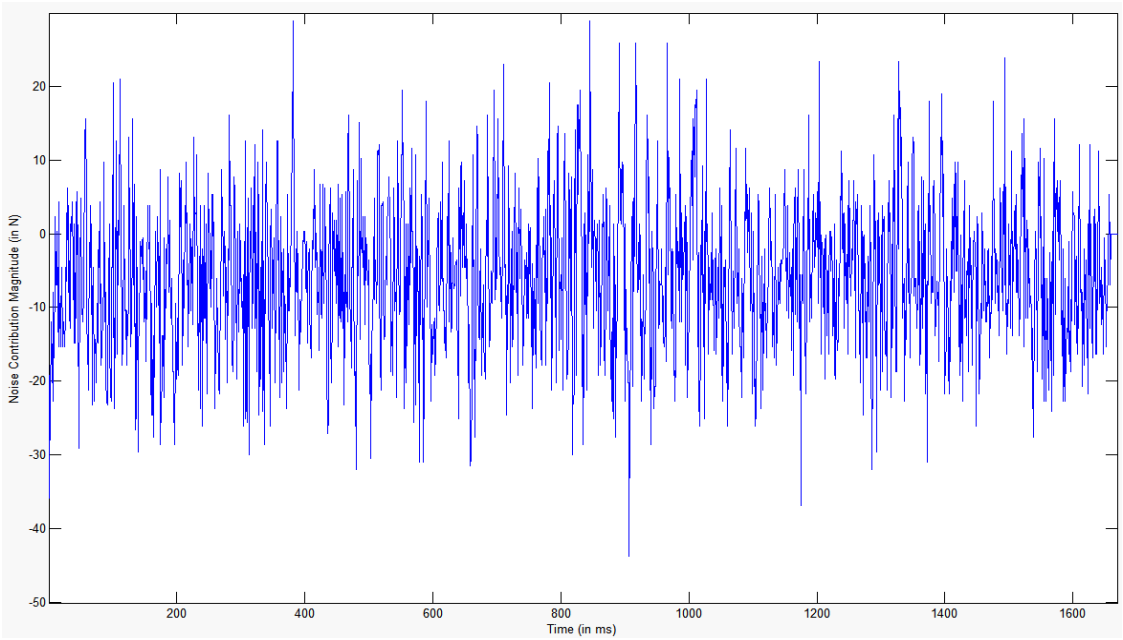

Figure 5.3: Sample Concatenated Leap Period Signals

As seen in Fig. 5.2 and Fig. 5.3, the leap period noise is random in nature, and its magnitude in comparison with the noisy VGRF signal is low, however makes a remarkable difference when

performing time-frequency analysis and adds high-frequency noise to the signal in a way that causes misinterpretation of the impact pattern. In fact, two alternatives were available for us to deal with the noise floor: leap-period concatenation or leap-period averaging; in case of concatenation, the different leap periods are extracted and placed adjacent to one another such that the spectrum of the concatenated vector is subsequently computed, while in case of averaging, the different leap periods are extracted and averaged such that the spectrum of the averaged signal is computed. The reason why we chose to go for the concatenation technique lies within the random nature of the noise component, which if averaged might lead to significant data cancelling out to zero, while concatenation preserves the patterns as well as all noise data points for accurate spectral representation.

As a next step, the Welch periodogram of the concatenated noise signal is performed, a sample of which is presented in Fig. 5.4 and Fig. 5.5.

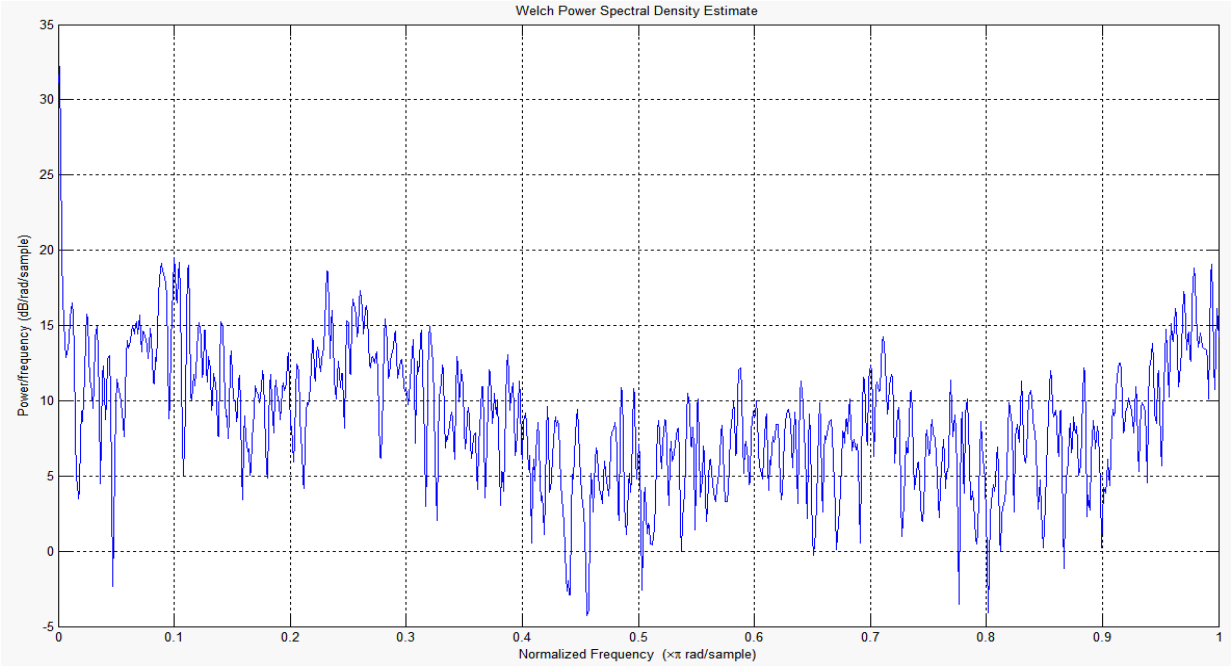

Figure 5.4: Sample Leap Period Welch Periodogram

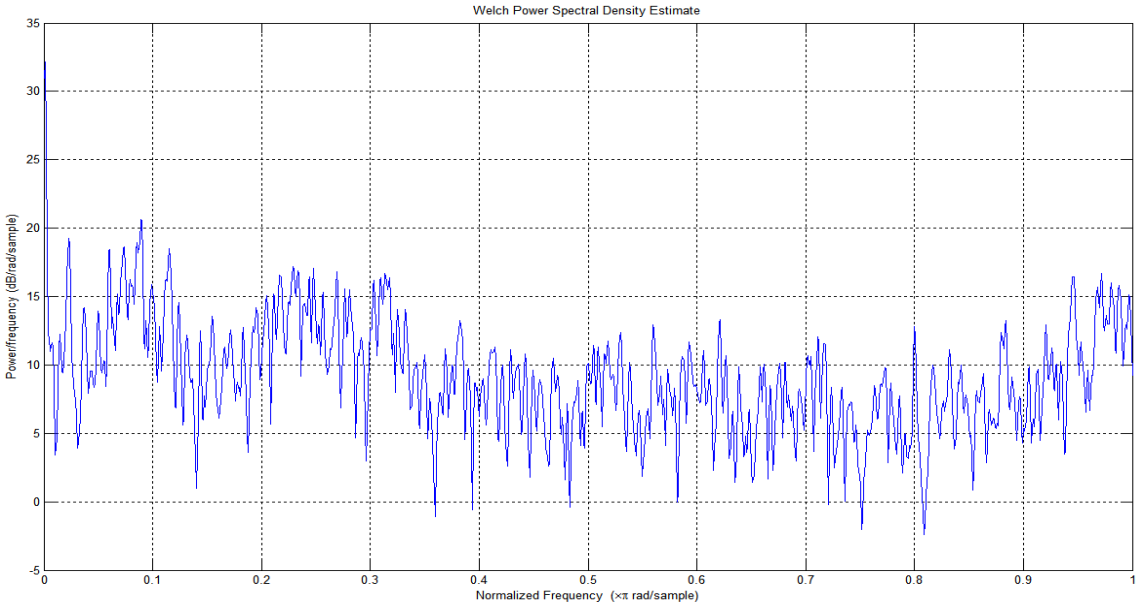

Figure 5.5: Sample Leap Period Welch Periodogram

At this stage, the noise spectrum is subtracted from the original VGRF spectrum shown in Fig. 5.6.

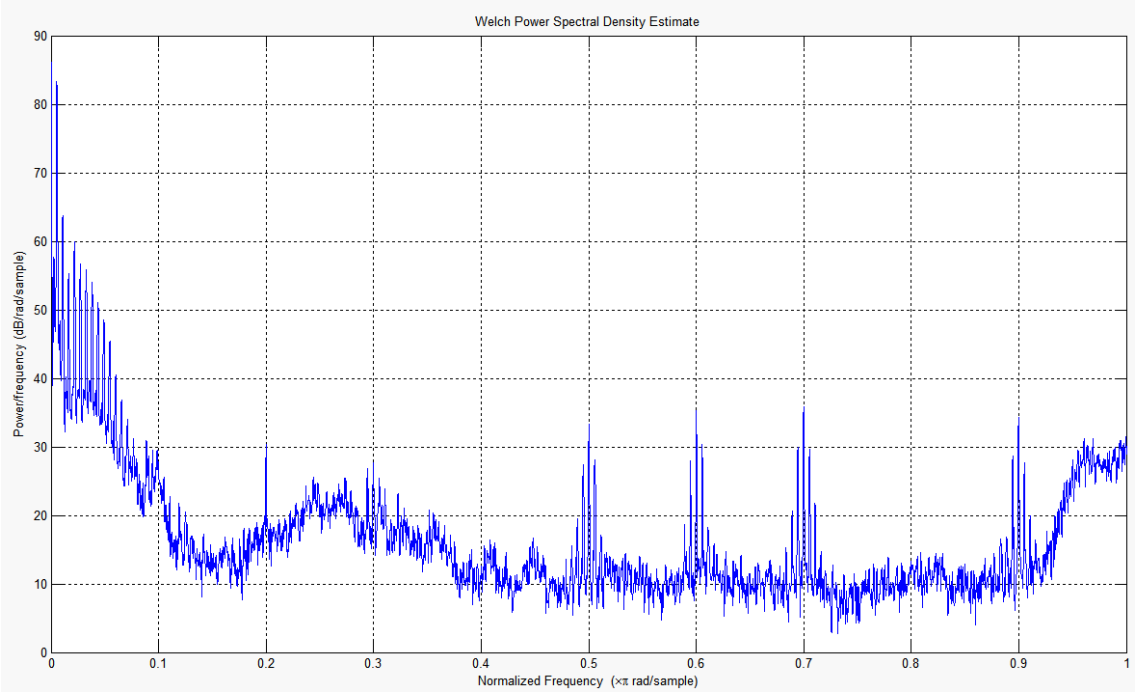

Figure 0.6: Sample Running VGRF Welch Periodogram

The observation of all noise periodograms showed a high noise level at the surrounding of 100 Hz, which we know for a fact to be the standard treadmill's motor vibration noise, along with its harmonics that clearly show in Fig 5.6, and this phenomenon is actually a drawback in the instrumented treadmill measurement environment. In fact, upon review of the quasi-totality of the instrumented treadmill systems used in research, avoidance of noise contamination of the recorded signals was anticipated and prevented at the very first stage in the treadmill building phase where the treadmill is activated with no individual standing on it or running, and the sensor outputs are assessed for noise level; if the noise is beyond tolerance limit, additional mechanical fixation of the treadmill is performed to reduce vibrations and hence unwanted high frequency noise.

Upon computation of the noise spectrum, spectral subtraction was performed as described in the respective section, leading to a filtered VGRF spectrum. However, the effectiveness of the filtering process may not be assessed before the re-transformation of this spectrum into the time domain and assessing the preservation of the main pattern traits of the VGRF signals e.g. impact pattern and propulsive pattern. This fact was stated while explaining the method, where we said that the main issue against traditional filtering is the deterioration of the pattern traits and especially the sharp-edged impact pattern.

That being said, and upon computation of the filtered VGRF spectrum, the IDFT was applied to the amplitude periodogram with phase taken into the consideration as explained in previous sections, and a comparison between the noisy VGRF and Filtered VGRF is shown in Fig. 5.7.

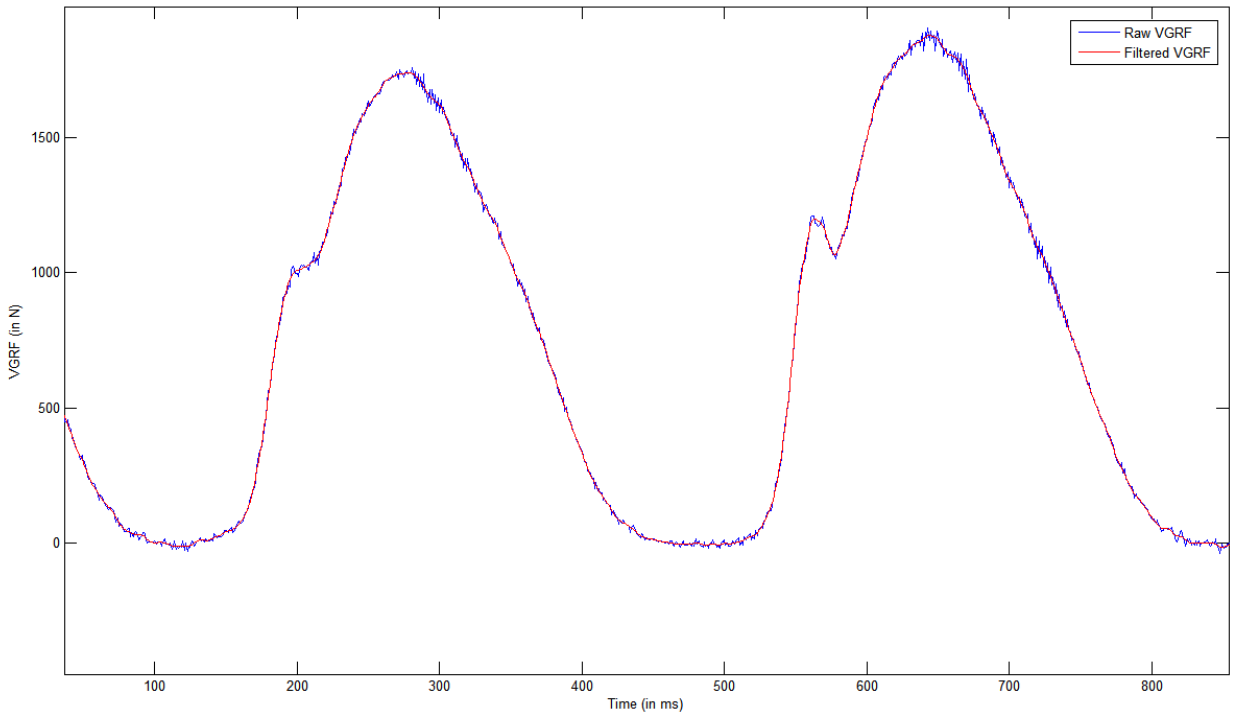

Figure 0.7: Fltered vs Raw VGRF Signal

As seen in Fig. 5.7, the technique we used to filter out the noise did the required task in an optimal way, being able to remove the leap period noise almost completely and keeping the impact patterns intact for further analysis. However, quantitatively speaking, we do not have metric to assess the level of deterioration of the impact pattern, having no isolated impact recordings in our database. In fact, this points out to a crucial point in our work that is to be investigated further by coming up with a measurement scheme to acquire isolated impact VGRF, which shall be used to prove our results, whether in terms of filtering or in terms of subsequent BSS.

5.1.2 VGRF Analysis

Concerning the analysis of VGRF signals, as explained in Chapter 2, the main purpose to reach upon performing the analysis was the quantification of fatigue progression in running activities.

Firstly, a descriptor was introduced for that endeavor, shown in (2.13), which we termed Spectral Mean Inter-peak Shift (SMIPS), the illustration of which is shown in Fig. 5.8.

$$SMIPS(\%) = \frac{\sum_{n=1}^{N} f(Pk_{n+1}) - f(Pk_n)}{N-1} \tag{2.81}$$

where $f(Pk_n)$is the frequency at which peak $Pk_n$is manifested, n is the peak index, and N is the number of peaks in the periodogram.

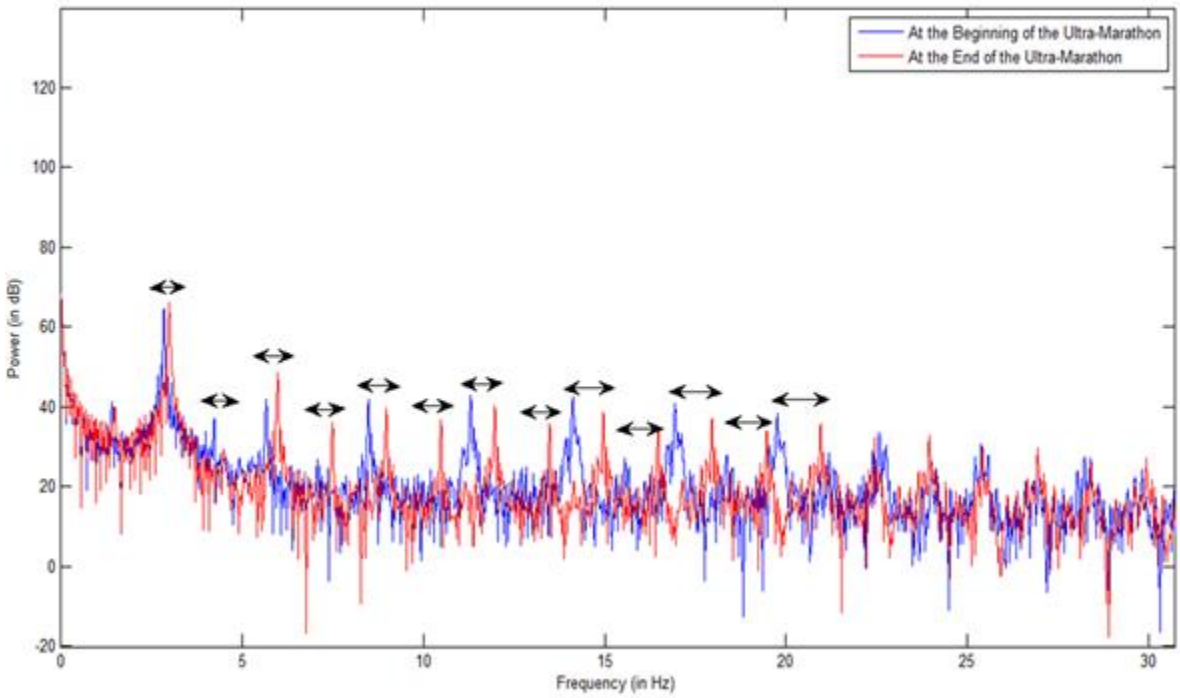

Figure 0.8: Comparative Graphical Representation of Welch Periodograms of Running VGRF Signals

Furthermore, another descriptor was introduced in this thesis, as presented in (2.14), which is the Mean High-frequency Power Variation (MHFPV):

$$MHFPV(\%) = \frac{P_{end}(f > f_{th}) - P_{start}(f > f_{th})}{P_{start}(f > f_{th})} \tag{2.82}$$

where $P_{end}(f > f_{th})$ is the end-of-run power at a user-selected high-frequency range f_{th} being the threshold frequency, and $P_{start}(f > f_{th})$ is the beginning-of-run power for the same frequency range.

As clearly shown in Fig. 5.8, and along the progression of the ultra-marathon, two main observations are made: firstly, the cyclic spectral peaks show an increasing inter-peak shift throughout the entire spectrum, pronounced more evidently at higher frequencies, and secondly the entire spectrum manifests a considerable shift towards higher frequencies. From an interpretational perspective, the shift of the spectrum towards higher frequencies knowing that the spectral peaks are manifestations of step and stride repetition rate indicates the fact that with the progression of fatigue, the runners shift naturally shift their running technique from adopting smooth wide hops into rapid narrow hops; this phenomenon is expected and even personally sensed whenever a subject is running on a treadmill or on the ground and fatigue starts manifesting, any individual would naturally loose the capability of performing wide hops that require high effort to lift the body off the ground for long distances, referring thereby to more moderate and quick stepping technique that would decrease the strain on the muscles. The SMIPS results are shown in Table 5-1.

Upon interpretation of the periodogram and realizing the above-stated remarks, it was mandatory for us to go further with the spectral content analysis but with the addition of the time factor, being thereby able to track the progression of the spectrum in terms of its spectral components; a shift towards higher frequency would logically mean a higher high-frequency power. The tracking of the frequency-specific power was computed via STFT spectrogram computation and comparison of power of the different spectral components, as shown in Fig. 5.9, Fig. 5.10 and

5.11 illustrating the spectrogram in color map format, in 3D format, and in contour format, respectively.

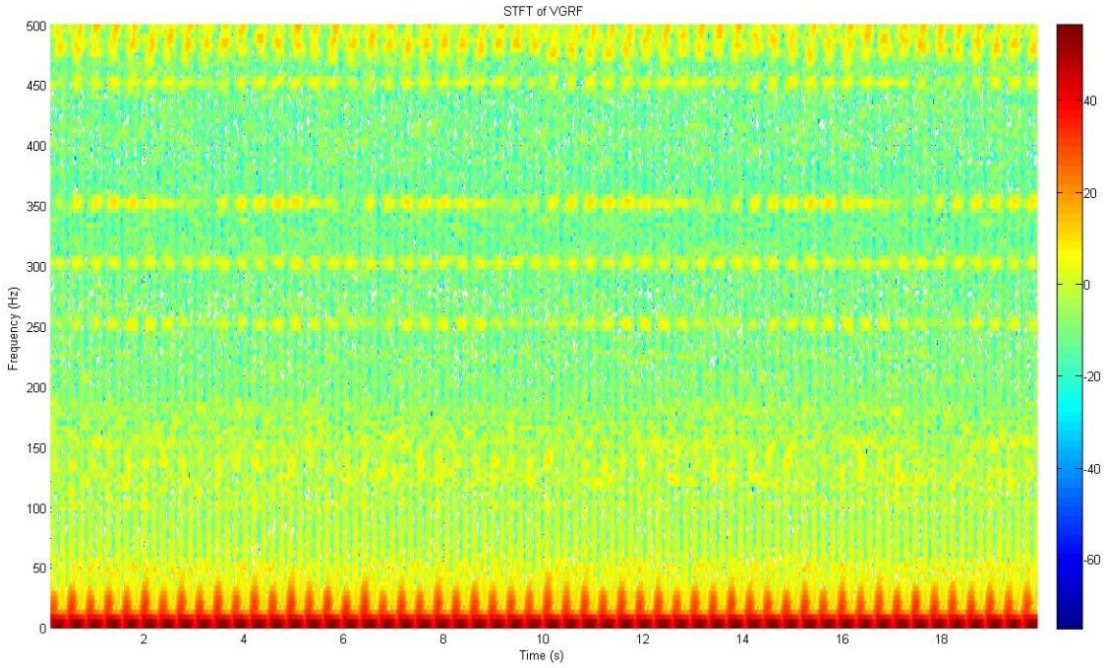

Figure 0.9: STFT Spectrogram of a Sample Running VGRF Signal (color map)

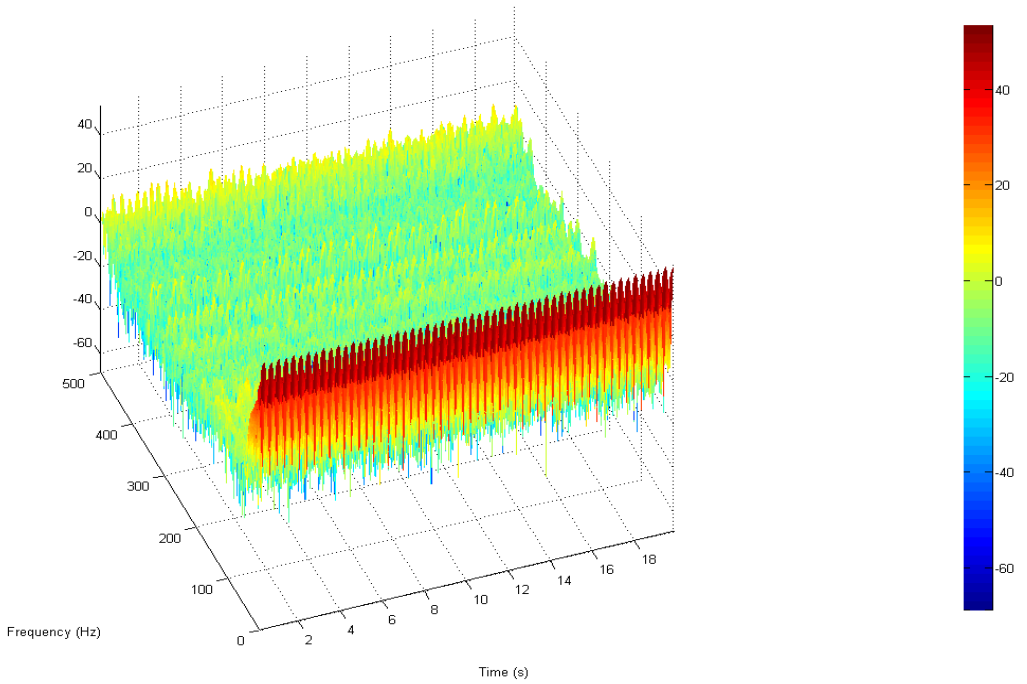

Figure 0.10: STFT Spectrogram of a Sample Running VGRF Signal (3D view)

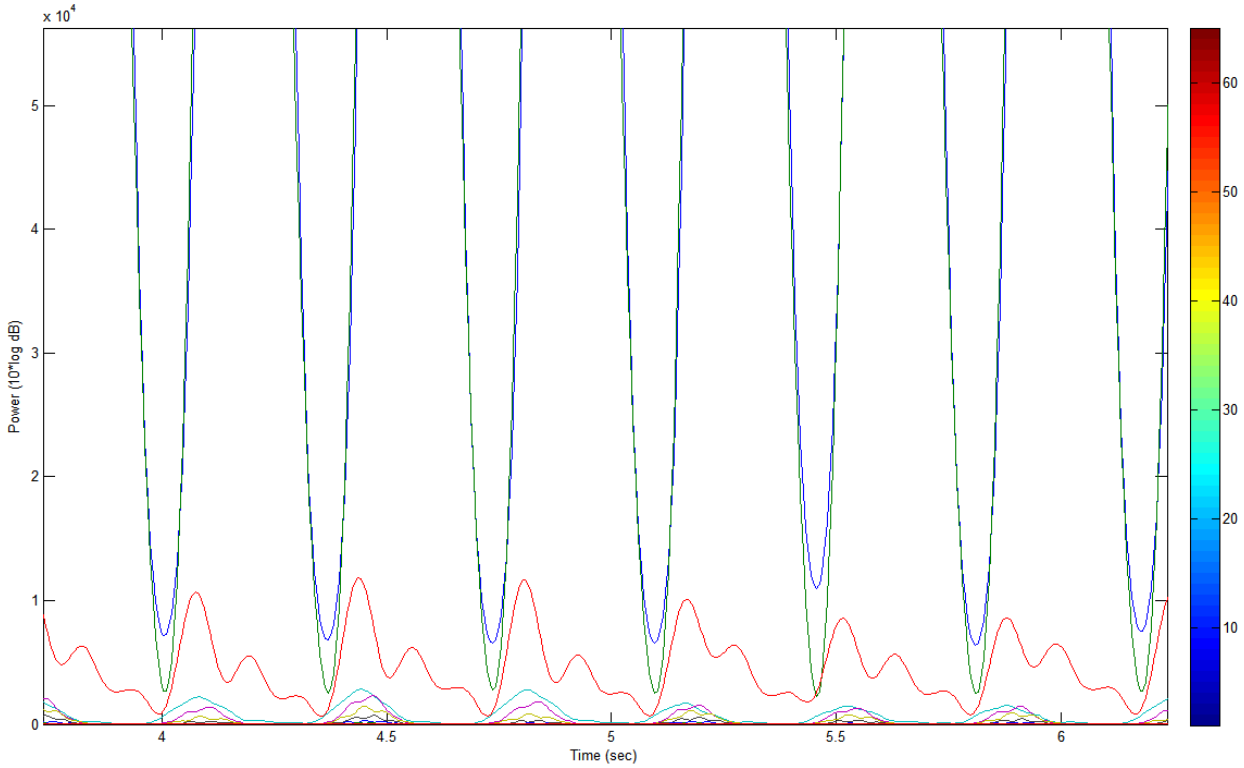

Figure 0.11: Frequency Components of an STFT Spectrogram of a Sample Running VGRF Signal (Highest Frequency Component in Red)

Upon computation of the high frequency VGRF component, manifested by the impact peak during heel-strike, and tracking its mean power throughout the ultra-marathon for the ten athletes, a monotonic increase of the mean high frequency power of 24.44±13.40 % was detected among the studied population, with only one athlete with a nearly constant impact power throughout the ultra-marathon. A sample of the high frequency component's power variation in progression with the ultra-marathon is shown in Fig. 5.12, showing a clear increase in the high frequency power during the 24 hours of running.

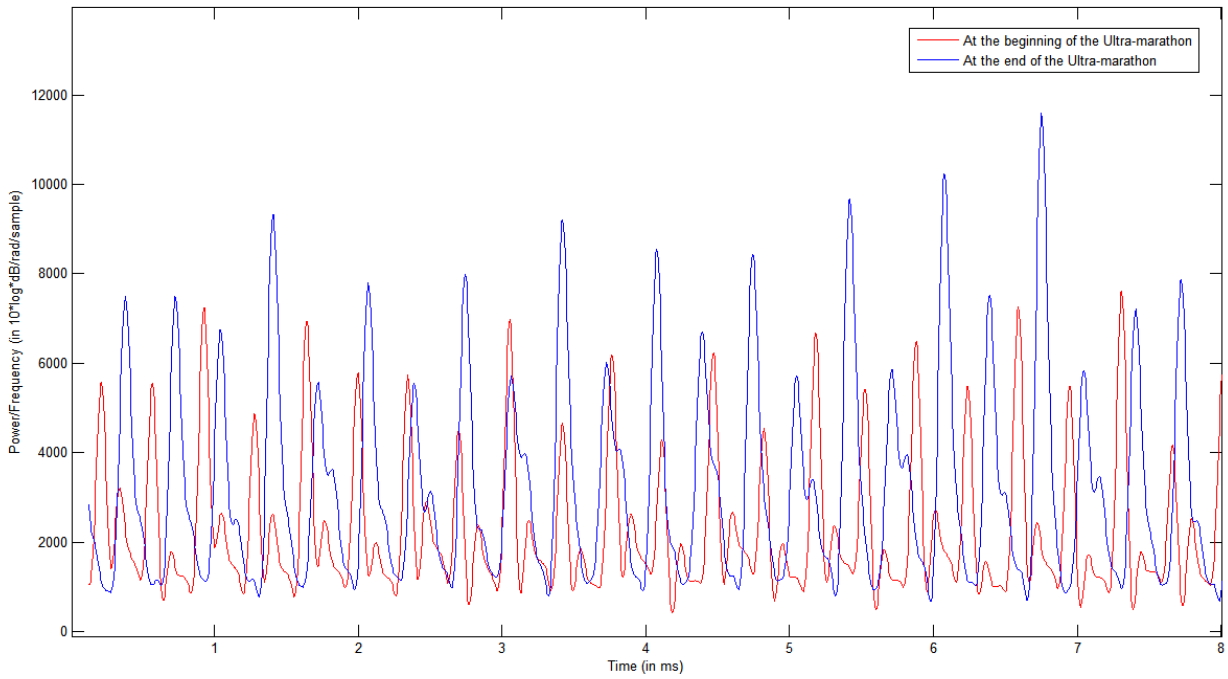

Figure 0.12: Comparative Graphical Representation of the High Frequency Component of the STFT Spectrograms of Sample Running VGRF Signals

Table 5.1 Spectral Analysis Results

Spectral Parameter	
SMIPS (mean±std)	MHFPV (mean±std)
0.66±0.38 Hz	24.44±13.40 %

The computed results were in conformity with our initial theory stating the fact that the body biomechanically compensates for lost muscle power with higher stress on the heel, tibia and knee through higher impact during heel-strikes, which was clearly indicated by a considerable increase in high frequency power.

5.2 VGRF SCBSS

As previously explained, the NMF algorithm in general requires two main user inputs: the magnitude/power spectrogram of the single-channel signal of interest, as well as the model order number denoted *k*, which corresponds to the estimated number of spectral bases. In our application, we chose *k*=3 based on our intuition since the VGRF signal carries three main spectral patterns: the impact pattern, the propulsion pattern, and the modulation pattern that carries the first two throughout the steps being the step modulation pattern.

Firstly, we created different synthetic sparse signals and assessed the metric we used in the sparseness penalty function in (3.6):

$$\|x\|_{1,2} = \frac{\sqrt{n} - (\sum_{i=1}^{N}|x_i|)/\sqrt{\sum_{i=1}^{N} x_i{}^2}}{\sqrt{n} - 1} \tag{3.83}$$

The resulting sparseness of the synthetic signal was computed to be 0.8, which is a relatively high degree of sparseness, as shown in Fig. 5.13.

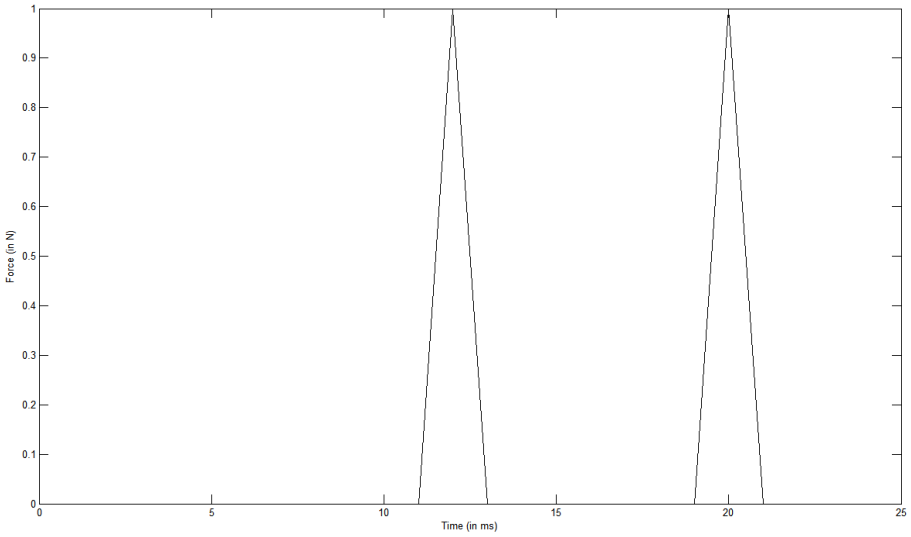

Figure 0.13: Sample Synthetic Sparse Signal with sparseness=0.8

As far as periodicity is concerned, we tested the fitting algorithm of periodicity onto the high frequency components of the VGRF STFT spectrogram, which we supposed to be similar to the resulting spectral pattern of the impact pattern, being a high frequency pattern. The results are graphically shown in Fig. 5.14 and numerically shown (fit results) below.

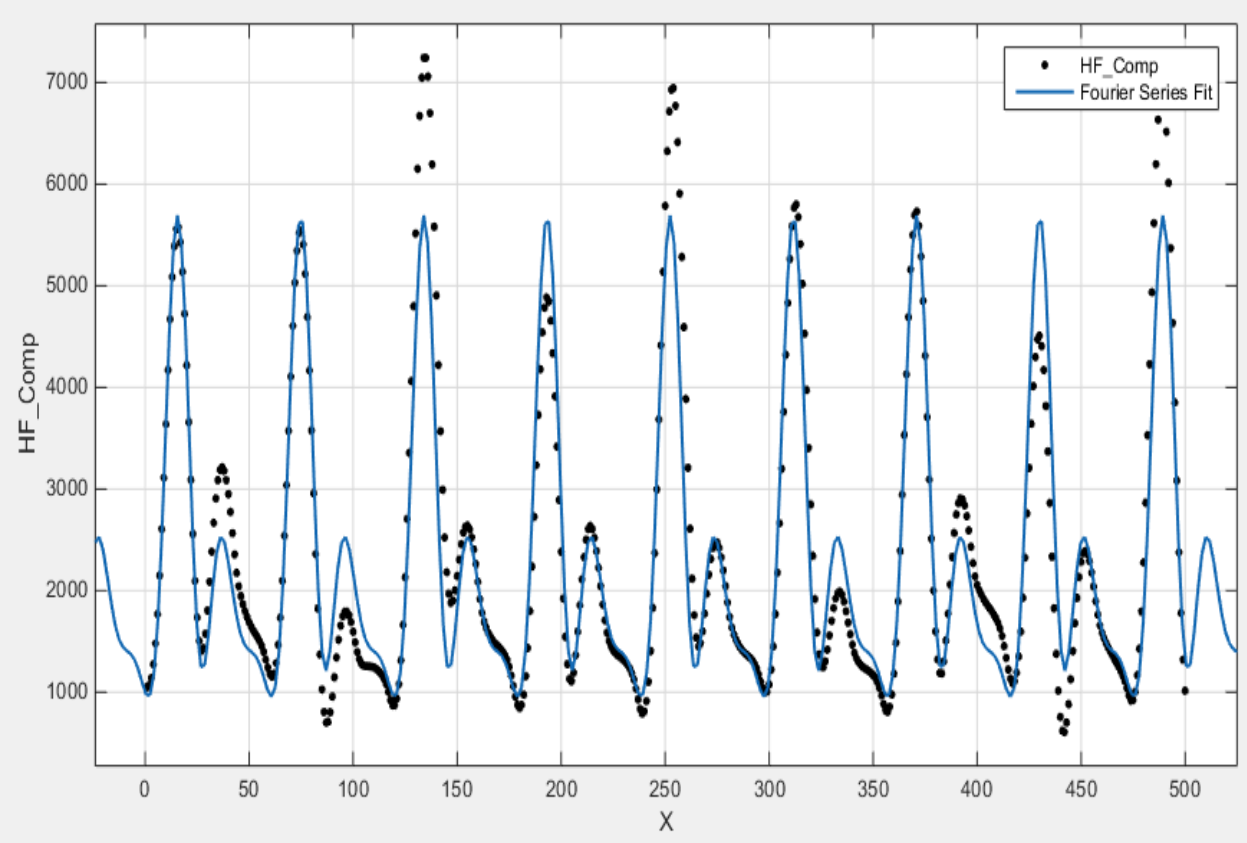

Figure 5.14: Graphical Representation of the Fourier Coefficient Fit of the HF Component of a sample Running VGRF Signal

The optimization algorithm stopped when the goodness of fit termed R^2 reached 0.982, which is a very good fit, showing the robustness of the periodicity tracking algorithm.

General model Fourier:

$$f(x) = a0 + a1*\cos(x*w) + b1*\sin(x*w) + a2*\cos(2*x*w) + b2*\sin(2*x*w) + a3*\cos(3*x*w) + b3*\sin(3*x*w) + a4*\cos(4*x*w) + b4*\sin(4*x*w)$$

where x is normalized by mean 250.5 and std 144.5

Coefficients (with 95% confidence bounds):

a0 =	2355	(2336, 2374)
a1 =	1144	(1118, 1171)
b1 =	575.4	(548.9, 601.8)
a2 =	1224	(1197, 1250)
b2 =	200.6	(174.4, 226.9)
a3 =	624.6	(598.3, 650.9)
b3 =	490.9	(464.4, 517.3)
a4 =	29.99	(3.572, 56.4)
b4 =	115	(88.76, 141.3)
w =	15.33	(15.32, 15.34)

Goodness of fit:

SSE: 2.2e+07

R-square: 0.982

Adjusted R-square: 0.9816

RMSE: 211.9

The above-presented results show the robustness of the penalty functions and constraints designed for the endeavor of separating cyclosparse impact patterns from the overall single-channel VGRF signals. In terms of sparseness assessment, the used metric proved its robustness in quantitatively describing the sparseness level of any given signal, and our use of it in order to force sparseness onto the separated sources is confirmed. As for the periodicity assessment penalty function, we were able to model a TF component using a Fourier-inspired modeling technique with linear optimization via EUC divergence minimization, resulting in a 98.16% goodness of fit. A question might come to mind is about the reason behind wanting to prove the effectiveness of the penalty functions and constraints instead of assessing the overall separation at the end of the road, and the answer is as follows: SCBSS deals with signals that are mixtures of completely unknown sources, and the only way to prove the effectiveness or separation success of such systems is having readily-available isolated sources e.g. wanting to separate the singer's voice from background music while having his voice readily-separate beforehand as a reference signal to compare with. Having explained that, the VGRF signals' nature dictate that no impact pattern may solely recorded from a natural run, unless we have the runner forced to run on his heels, which makes no sense and causes subject discomfort, forcing us thereby to rely on the robustness of the technique itself on synthetic signals rather than real isolated sources.

Finally, upon implementation of the overall code, the results turned out to be in harmony with what literature states about the heel-strike transient, which was never separated in other references for us to use as a goodness of separation criterion. The resulting impact and propulsion patterns in two of the randomly selected sample signals are shown in Figs. 5.15-5.16, showing the separated impact patterns from the overall VGRF patterns.

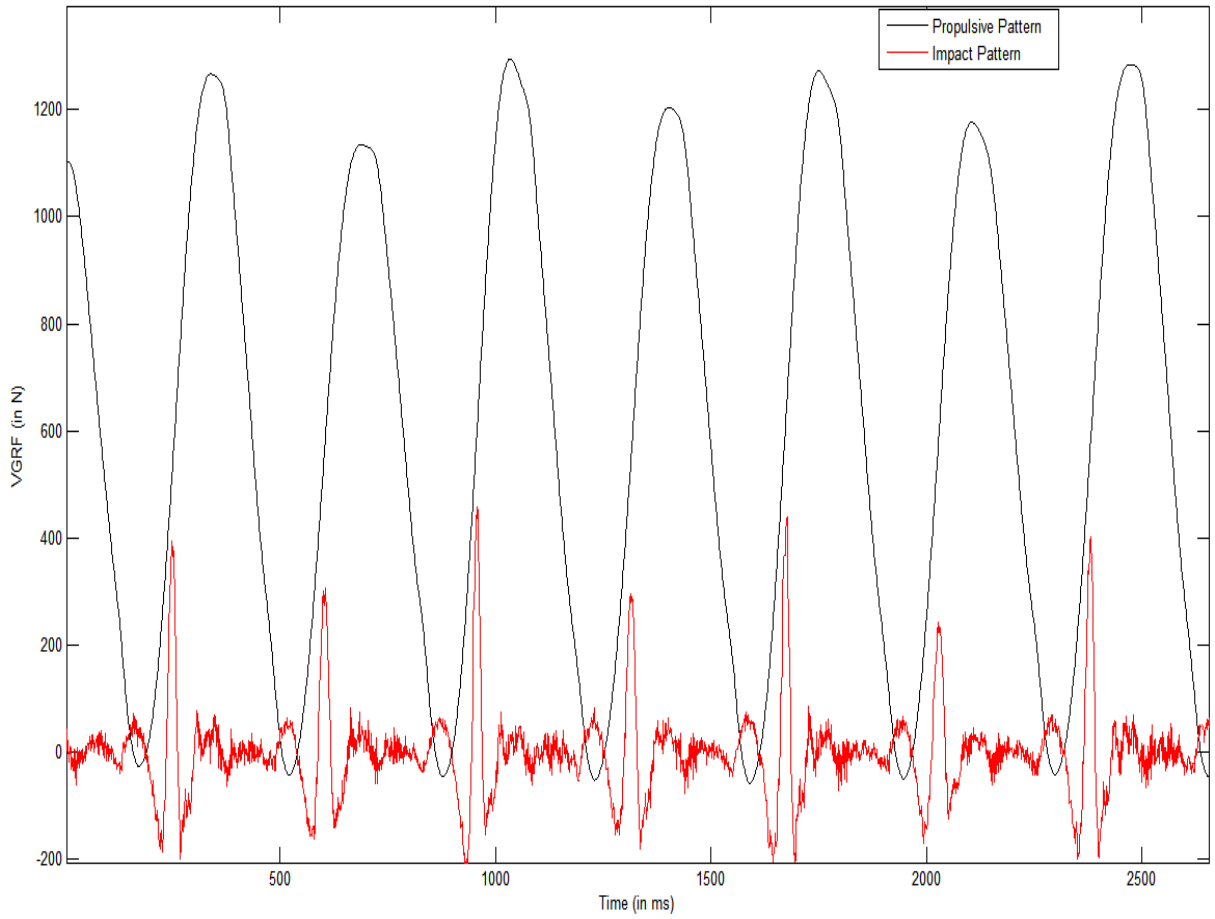

Figure 0.15: Sample of Separated Running VGRF Patterns

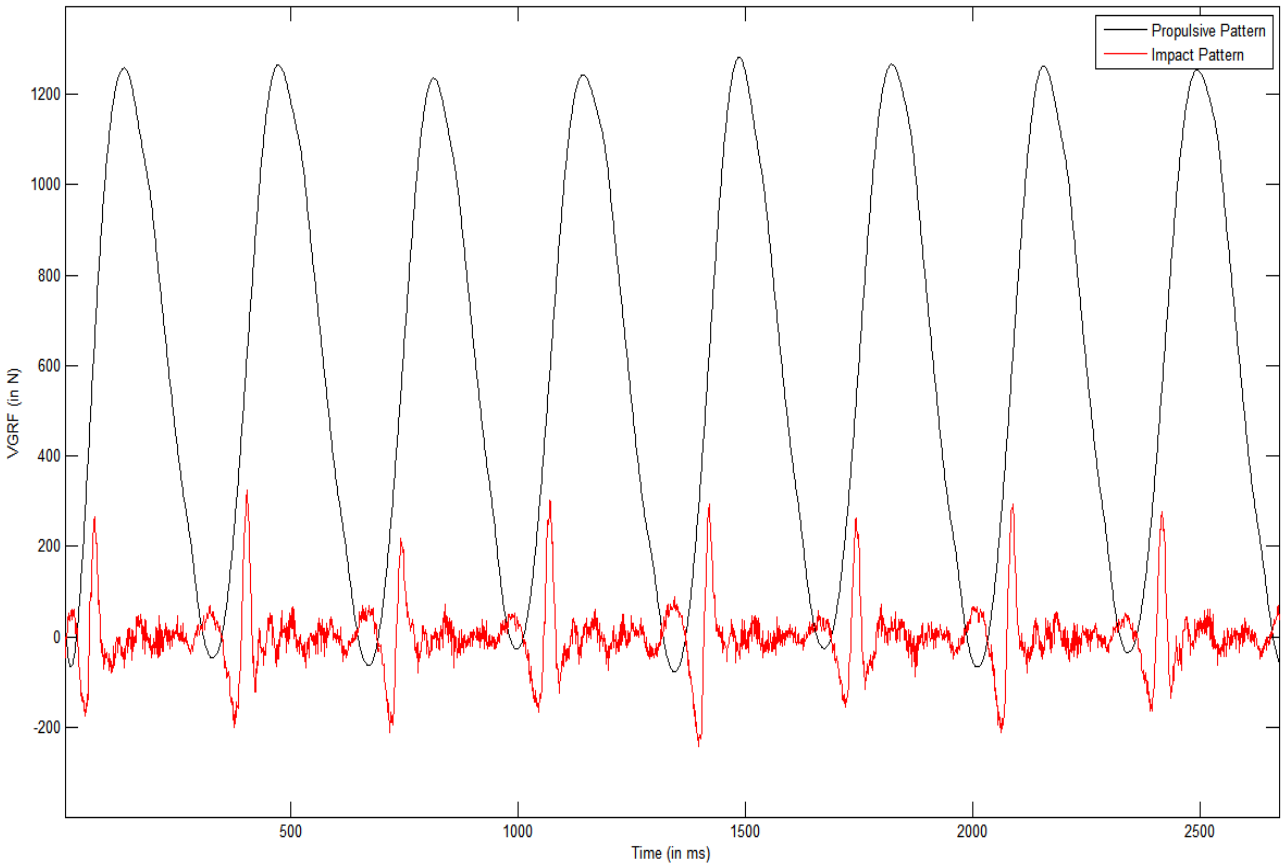

Figure 0.16: Sample of Separated Running VGRF Patterns

From the above shown graphical representations of the separated VGRF patterns, the impact pattern turned out to be an oscillatory pattern rather than a single peak within the VGRF signal, as some researchers modeled it. This impact pattern is initiated at heel-strike and propagated throughout the entire step rather than appearing shortly on heel-strike and vanishing away. In terms of signal description, we may describe the impact pattern as a sharp transient damped peak manifested through broadband high-frequency oscillations. And from a biomechanical perspective, a wide range of interpretations is possible upon the separation of the impact pattern; for instance, we realize that the sharper the impact is at heel-strike, the stronger and more harmful the hit is over the tibia and the associated musculoskeletal system, which led some

researchers as stated in the literature to consider training athletes over NHS running, preventing thereby the strong hit against the ground and its subsequent damage to the bones and ligaments. Furthermore, the separation showed clear inter-leg impact asymmetry, which if accurately analyzed and quantified may open the door for the diagnosis of the underlying musculoskeletal system of the runner, whether in terms of athletic performance or even pathological run.

Moreover, we realized that the spectral content of the impact is quite complex and ergo may not be considered as being a single-high-frequency component that can be easily extracted via standard filters as performed in the literature.

The real work in terms of biomechanical study actually begins upon separation, being now able to model

CHAPTER 6

CONCLUSION AND FUTURE PERSPECTIVE

In this dissertation, we explored the field of Single-channel Blind Source Separation (SCBSS) via customized and regularized Non-negative Matrix Factorization (NMF) that we adapted to separate sparse pulse-like periodic/cyclic source signals from a single-channel mixture; the devised algorithm was termed the Cyclosparse NMF.

Throughout the dissertation, we went through the state-of-the art techniques and algorithms implemented in BSS in general and SCBSS in specific, as well as the state-of-the-art measurement systems in the field of Gait analysis, leading the way to our main application in this thesis which is the application of SCBSS on Vertical Ground Reaction Force (VGRF) signals for the eventual separation of the Impact Pattern that offers valuable insight on the anatomical and physiological state of runners.

The first contribution made in this thesis was the optimization of the pre-processing and processing techniques of VGRF signals in order to have the best Signal-to-noise Ratio (SNR) and minimal loss of relevant data that is specific to the VGRF pattern e.g. losing sharp edges from the impact pattern via the Spectral Subtraction method. The stated contribution was published and presented in an international conference.

The second contribution made in this thesis was the quantification of runners' fatigue progression via comparative spectral and joint time-frequency analysis of the VGRF signals between the beginning of the running task and its end; the Welch periodogram was analyzed for spectral feature extraction and the Short-time Fourier Transform (STFT) spectrogram was analyzed for joint time-frequency feature extraction. The stated contribution was accepted for

publication in 2 international conferences, presented in the first and shall be presented in the near future in the second conference.

The third contribution made in this thesis was the design of an SCBSS algorithm to separate cyclosparse sources from a single recording and without any a priori information, i.e. totally unsupervised. The method was termed the Cyclosparse NMF, and it is based on optimized and regularized non-negative matrix factorization. The algorithm was implemented on single-channel VGRF signals and the Impact patterns were successfully extracted, showing rather major characteristics of the pattern that might carry valuable biomechanical interpretations. The stated contribution was submitted to an international peer-reviewed journal for publication.

As future perspectives, and upon performing our in-depth analysis of VGRF signals, we suggest taking our work into another level in terms of online / real-time fatigue analysis to improve the reliability and efficiency of rehabilitation programs as well as athletic performance metrics. Furthermore, with the impact pattern being separated, focus shall be applied onto the mathematical modeling of the impact, which would eventually lead to the true effect of the heel-strike on the anatomical and physiological state of the runners. This would be of great value in terms of injury prediction and prevention.

Bibliography

[1] R. Katiyar, "Clinical Gait Data Analysis based on Spacio-Temporal Features," *International Journal of Computer Science and Information Security,* vol. VII, no. II, pp. 174-183, 2010.

[2] A. Kharb, V. Saini, Y. K. Jain and S. Dhiman, "A Review of Gait Cycle and Its Parameters," *International Journal of Computational Engineering and Management,* vol. 13, pp. 78-83, 2011.

[3] J. Richards, The Comprehensive Textbook of Biomechanics, China: ELSEVIER, 2018.

[4] T. Marasovic, M. Cecic and V. Zanchi, "Analysis and Interpretation of Ground Reaction Forces in Normal Gait," *WSEAS Transactions on Systems,* vol. 8, no. 9, pp. 1105-1115, 2009.

[5] J. W. Kim, H. J. Jang, D. H. Hwang and C. Park, "A Step, Stride and Heading Determination for the Pedestrian Navigation System," *Journal of Global Positioning Systems,* vol. 3, pp. 273-279, 2004.

[6] M. W. Whittle, Gait Analysis: An Introduction, Butterworth-Heinemann, 2006.

[7] W. Weber and E. Weber, "Mechanik der menschlichen Gehwerkzeuge: eine anatomisch-physiologische Untersuchung," vol. 1, 1836.

[8] M. A. Cappozzo, "Tosi V. Marey and Muybridge: How modern biolocomotion analysis started," 1992.

[9] R. Bongaardt and O. G. Meijer, "Bernstein's Theory of Movement Behavior: Historical Development and Contemporary Relevance," *Journal of Motor Behavior,* vol. 32, no. 1, pp. 57-71, 2000.

[10] H. Elftman, "The measurement of the external force in walking," *Science,* vol. 88, no. 2276, pp. 152-153, 1938.

[11] D. A. Winter, "Overall principle of lower limb support during stance phase of gait," *Journal of Biomechanics,* vol. 13, no. 11, pp. 923-927, 1980.

[12] G. Bronwyn, "The Biomechanics of Equine Locomotion," in *The Athletic Horse (2nd Edition)*, ELSEVIER, 2014, pp. 266-281.

[13] A. Belli, P. Bui, A. Berger, A. Geyssant and J. Lacour, "A treadmill ergometer for three-dimensional ground reaction forces measurement during walking," *Journal of Biomechanics,*

vol. 34, pp. 105-112, 2001.

[14] D. Sutherland, "The Evolution of Clinical Gait Analysis Part I: Kinesiological EMG," *Gait and Posture,* vol. 14, pp. 61-70, 2001.

[15] M. Zanfir, M. Leordeanu and C. Sminchisescu, "The Moving Pose: An Efficient 3D Kinematics Descriptor," in *ICCV*, Sydney, 2013.

[16] B. Bilney, M. Morris and K. Webster, "Concurrent related validity of the GAITRite® walkway system for quantification of the spatial and temporal parameters of gait," *Gait & posture,* vol. 17, no. 1, pp. 68-74, 2003.

[17] J. H. Viitasalo and P. V. Komi, "Signal characteristics of EMG with special reference to reproducibility of measurements," *Acta Physiol Scand,* vol. 93, no. 4, pp. 531-9, 1975.

[18] A. De Stefano, J. H. Burridge, V. T. Yule and R. Allen, "Effect of gait cycle selection on EMG analysis during walking in adults and children with gait pathology," *Gait Posture,* vol. 20, no. 1, pp. 92-101, 2004.

[19] S. Bronner and A. Noah, "Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait," *Journal of Medical Engineering & Technology,* pp. 1-7, 2014.

[20] T. Owings and M. Grabiner, "Measuring step kinematic variability on an instrumented treadmill: how many steps are enough?," *Journal of Biomechanics,* vol. 36, no. 8, pp. 1215-1218, 2003.

[21] W. Herzog, B. Nigg, L. Read and E. Olsson, "Asymmetries in Ground Reaction Force Patterns in Normal Human Gait," *Medicine and Science in Sports and Exercise,* vol. 21, no. 1, pp. 110-114, 1989.

[22] J. Bigouette, J. Simon, K. Liu and C. Docherty, "Altered Vertical Ground Reaction Forces in Participants With Chronic Ankle Instability While Running," *J Athl Train,* vol. 51, no. 9, pp. 682-687, 2016.

[23] D. Padua and L. DiStefano, "Sagittal Plane Knee Biomechanics and Vertical Ground Reaction Forces Are Modified Following ACL Injury Prevention Programs," *Sports Health,* vol. 1, no. 2, pp. 165-173, 2009.

[24] A. Coetsee, "ANALYSIS OF THE VERTICAL GROUND REACTION FORCES IN SPORTS PARTICIPANTS WITH ADDUCTOR-RELATED GROIN PAIN: A COMPARISON STUDY," Stellenbosch University, South Africa, 2016.

[25] A. Abbasi, H. Sadeghi and M. Khaleghi, "GROUND REACTION FORCES ATTENUATION IN

SUPINATED AND PRONATED FOOT," in *ISBS Conference 2008*, Seoul, Korea, 2008.

[26] A. Tonazzini, "Blind Source Separation Techniques for Detecting Hidden Texts and Textures in Document Images," in *International Conference Image Analysis and Recognition*, Porto, Portugal, 2004.

[27] K. Prakash, "Blind Source Separation for Speech Music and Speech Speech Mixtures," *International Journal of Computer Applications,* vol. 110, no. 12, pp. 40-43, 2015.

[28] E. Visser, "Speech enhancement using blind source separation and two-channel energy based speaker detection," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Hong Kong, China, 2003.

[29] T. Heittola, "Musical Instrument Recognition in Polyphonic Audio Using Source-Filter Model for Sound Separation," in *10th International Society for Music Information Retrieval Conference*, Kobe, Japan, 2009 .

[30] C.-Y. Yu, "Blind source separation based x-ray image denoising from an image sequence," *Review of Scientific Instruments,* vol. 86, 2015.

[31] P. Duhamel, "Blind multivariable equalization," *Proc. DSP,* pp. 13-16, 1997.

[32] J. R. Treichler, "Practical blind demodulators for high-order QAM signals," *Proc. IEEE,* vol. 86, no. 10, p. 1907–1926, 1998.

[33] U. Madhow, "Blind adaptive interference suppression for Direct-Sequence CDMA," *Proc. IEEE,* vol. 86, no. 10, p. 2049–2069, 1998.

[34] S. Haykin, Unsupervised Adaptive Filtering, John Wiley&Sons, 2000, p. 13–112.

[35] T. Lee, Independent Component Analysis: Theory and Applications, Kluwer Academic Publ., 1998.

[36] A. K. Barros, A. Mansour and N. Ohnishi, "Removing artifacts from electrocardiographic signals using independent components analysis," *Neurocomputing,* vol. 22, no. 1-3, pp. 173-186, 1998.

[37] S. Dodel, J. M. Herrmann and T. Geisel, "Localization of brain activity – Blind separation of fMRI data," *Neurocomputing,* vol. 32, pp. 71-708, 2000.

[38] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Processing Magazine,* May 1996.

[39] U. R. Abeyratne, A. P. Petropulu and J. M. Reid, "Higher-order spectra based deconvolution of ultrasound images," *IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control,* vol. 42, no. 6, pp. 1064-1095, 1995.

[40] R. H. T. Bates, "Astronomical speckle imaging," *Physics Reports,* vol. 90, no. 4, p. 203–297, 1982.

[41] -, "Edges are the 'independent components' of natural scenes," *Advances in Neural Information Processing Systems 9,* vol. 9, 1996.

[42] A. Hyvarinen, E. Oja, P. Hoyer and J. Hurri, "Image feature extraction by sparse coding and independent component analysis," *Proc. ICPR,* p. 1268–1273, 1998.

[43] -, "Localisation et identification par la quadricovariance," *Traitement du Signal,* vol. 7, no. 5, p. 397–406, 1990.

[44] -, "Higher-order narrow-band array processing," in *Proc. Int'l Sig. Proc. Wkshp on HigherOrder Statistics*, 1991.

[45] J. L. Lacoume, P. O. Amblard and P. Comon, "Statistiques d' Ordre Superieur Pour le Traitement du Signal," Paris, 1997.

[46] R. Huez, D. Nuzillard and A. Billat, "Denoising using blind source separation for pyroelectric sensors," *EURASIP J. Appl. Signal Processing,* vol. 2001, no. 1, pp. 53-65, 2001.

[47] L. Tong, R. Liu, V. C. Soon and Y. F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits and Systems,* vol. 38, no. 5, pp. 499-509, 1991.

[48] C. Jutten and J. Karhunen, "Advances in Nonlinear Blind Source Separation," *Proc. of the 4th Int. Symp. on Independent Component Analysis and Blind Signal Separation,* vol. 41, no. 2, 2003.

[49] K. Zhang and A. Hyvarinen, "On the Identifiability of the Post-Nonlinear Causal Model," *UAI,* pp. 647-655, 2009.

[50] Y. Deville, "Matrix Factorization for Bilinear Blind Source Separation: Methods, Separability and Conditioning," in *23rd European Signal Processing Conference (EUSIPCO)*, Nice, 2015.

[51] A. Ziehe, M. Kawanabe, H. Stefan and K.-R. Muller, "Blind Separation of Post-nonlinear Mixtures using Linearizing Transformations and Temporal Decorrelation," *Journal of Machine Learning Research,* vol. 4, pp. 1319-1338, 2003.

[52] Y. Tan, Y. J. Wang and J. M. Zurada, "Nonlinear blind source separation using a radial basis

function network," *IEEE Transactions on Neural Networks,* vol. 12, no. 1, pp. 124-134, 2001.

[53] Z. Koldovsky, J. Malek and P. Tichavski, "Blind Speech Separation in Time-Domain Using Block-Toeplitz Structure of Reconstructed Signal Matrices," in *INTERSPEECH*, Florence, 2011.

[54] V. Capdevielle, C. Servire and J. L. Lacoume, "Blind separation of wide-band sources in the frequency domain," in *ICASSP95*, Detroit, 1995.

[55] C. Jutten and J. Herault, "Blind separation of Sources, part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing,* vol. 24, no. 1, pp. 1-10, 1991.

[56] J. Xi and J. P. Reilly, "Blind separation and restoration of signals mixed in convolutive environment," in *ICASSP'97*, Munich, 1997.

[57] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso and E. Moulines, "A blind source separation technique using second-order statistics," *IEEE Trans. Sig. Proc.,* vol. 45, no. 2, pp. 434-444, 1997.

[58] E. Weinstein, M. Feder and A. Oppenheim, "Multichannel signal separation by decorrelation," *IEEE Trans. Speech Audio Proc.,* vol. 1, no. 4, pp. 405-413, 1993.

[59] S. Shamsunder and G. B. Giannakis, "Multichannel blind signal separation and reconstruction," *IEEE Trans. Speech, Audio Proc.,* vol. 5, no. 6, pp. 515-528, 1997.

[60] A. Mansour, C. Jutten and P. Loubaton, "Subspace method for blind separation of sources and for a convolutive mixture model," *Signal Processing VIII, Theories and Applications,* pp. 2081-2084, 1996.

[61] R. K. Prasad, H. Saruwatari and K. Shikano, "Problems in Blind Separation of Convolutive Speech Mixtures by Negentropy Maximization," in *IWAENC'03*, Kyoto, 2003.

[62] C. Jutten, L. Nguyen Thi, E. Dijkstra, E. Vittoz and C. J., "Blind Separation of Sources: An Algorithm for Separation of Convolutive Mixtures," in *International Signal Processing Workshop*, 1992.

[63] C. Jutten and J. Herault, "Blind Separation of Sources, part I: An adaptive Algorithm based on neuromimetic architecture," *Signal Processing,* vol. 24, no. 1, pp. 1-10, 1991.

[64] A. J. Bell and T. J. Sejnowski, "An information mzximization approach to blind separation and blind deconvolution," *Neural Computation,* vol. 7, no. 6, pp. 1129-1159, 1995.

[65] L. Parra, C. Spence and B. de Vries, "Convolutive Source Separation with ML," in *ISIS'97*, 1997.

[66] J. F. Cardoso, "Blind Signal Separation: Statistical Principles," *Proc. IEEE,* vol. 9, no. 10, pp. 2009-2025, 1998.

[67] L. Parra, C. Spence and B. de Vries, "Convolutive source separation and signal modeling with ML," in *ISIS'97*, 1997.

[68] R. H. Lambert and C. L. Nikias, "Polynomial matrix whitening and application to the multichannel blind deconvolution problem," *MILCOM,* vol. 3, pp. 988-992, 1995.

[69] K. Torkkola, "Blind separation for audio signals - are we there yet?," in *ICA'99*, 1999.

[70] L. A. Lindgren and H. Broman, "Source separation using a criterion based on second-order statistics," *IEEE Trans. Sig. Proc. ,* vol. 46, no. 7, pp. 1837-1850, 1998.

[71] B. Yin and P. Sommen, "Adaptive blind signal separation using a new simplified mixing model," in *ProRISC'99*, Mierlo, 1999.

[72] L. Parra and C. Spence, "On-line convolutive source separation of non-stationary signals," *IEEE J. VLSI Sig. Proc,* vol. 26, no. 1-2, pp. 39-46, 2000.

[73] P. Sommen and B. Yin, "A new convolutive blind signal separation algorithm based on second order statistics using a simplified mixing model," in *EUSIPCO'00*, 2000.

[74] S. Shamsunder and G. Giannakkis, "Multichannel blind signal separation and reconstruction," *IEEE Trans. Speech. Audio Proc. ,* vol. 5, no. 6, pp. 515-528, 1997.

[75] K. Sabri, M. El Badaoui, F. Guillet and J.-B. Morin, "Cyclostationary modeling of ground reaction force signals," *Signal Processing,* vol. 90, no. 4, pp. 1146-1152, 2010.

[76] K. Sabri, "Cyclosparsity: A New Concept for Sparse Deconvolution," *Global Journal of Computer Science and Technology: F Graphics & Vision ,* vol. 14, no. 4, 2014.

[77] T. Mei and F. Yin, "Blind separation of convolutive mixtures by decorrelation," *Signal Processing,* vol. 84, no. 12, pp. 2297-2213, 2004.

[78] D. Burnett, N. H. Campbell-Kyureghyan, P. Ceritto and P. Quesada, "Symmetry of ground reaction forces and muscle activity in asymptomatic subjects during walking, sit-to-stand, and stand-to-sit tasks," *Journal of Electromyography and Kinesiology,* vol. 21, no. 4, pp. 610-615, 2011.

[79] A. Singh and A. Thakur, "Human gait analysis using wavelet denoising and total variation filtering," in *ICGCIoT*, Delhi, 2015.

[80] G.-J. Jang and T.-w. Lee, "A probabilistic approach to single channel blind signal separation," *Advances in neural information processing,* 2003.

[81] T. Hermle, C. Schwarz and M. Bogdan, "Employing ICA and SOM for spike sorting of multielectrode recordings from CNS," *Journal of Physiology,* 2004.

[82] R. G. Baraniuk and D. L. Jones, "A signal-dependent time-frequency representation: optimal kernel design," *IEEE Transactions on Signal Processing,* vol. 41, no. 4, pp. 1589-1602, 1993.

[83] M. Emresoy and P. Loughlin, "Weighted least squares implementation of Cohen-Posch time-frequency distributions," *IEEE Transactions on Signal Processing,* vol. 46, no. 3, pp. 753-757, 1998.

[84] SAS, 2013. [Online]. Available: http://support.sas.com/documentation/cdl/en/imlug/64248/HTML/default/viewer.htm#imlug_waveletanalysis_sect008.htm.

[85] S. Krishnan, R. M. Rangayyan, B. G. D and F. C. B, "Adaptive time-frequency analysis of knee joint vibroarthrographic signals for noninvasive screening of articular cartilage pathology," *IEEE Transactions on Biomedical Engineering,* vol. 47, no. 6, pp. 773-783, 2000.

[86] Z. M. Hussain and B. Boashash, "Adaptive Instantaneous Frequency Estimation of Multicomponent FM Signals Using Quadratic Time–Frequency Distributions," *IEEE Transactions on Signal Processing ,* vol. 50, no. 8, pp. 1866-1876, 2002.

[87] F. Berthommier and S. Choi, "Evaluation of CASA and BSS models for subband cocktail-party speech separation," *Proc. ICA,* 2011.

[88] B. Gao, W. Woo and S. Dlay, "Single-channel source separation using EMD-subband variable regularized sparse features," *IEEE Transactions on Audio, Speech and Language Processing,* vol. 19, no. 4, pp. 961-976, 2011.

[89] B. Gao, "Single channel blind source separation," Newcastle University, Newcastle, 2011.

[90] L. Pang and X. Deng, " A SCBSS methodology for time-frequency overlapped signals using non-negative matrix factorisation," *International Journal of Electronics,* vol. 104, no. 4, pp. 624-634, 2016.

[91] W. Xu, X. Liu and Y. Gong, "Document Clustering based on non-negative matrix factorization," in *SIGIR'03*, Toronto, 2003.

[92] J. Sivic, B. C. Russell, A. A. Efros, Z. A. and W. T. Freeman, "Discovering objects and their location in images," in *ICCV'05*, 2005.

[93] D. Soukup and I. Bajla, "Robust Object Recognition under Partial Occlusions Using NMF," *Computational Intelligence and Neuroscience,* vol. 2008, p. 14, 2008.

[94] M. Kalayeh, H. Idrees and M. Shah, "NMF-KNN: Image Annotation using Weighted Multi-view Non-negative Matrix Factrorization," in *CVPR'14*, 2014.

[95] J. Marial, F. Bach, J. Ponce and G. Sapiro, "Online Learning for Matrix Factorization and Sparse Coding," *JMLR,* pp. 19-60, 2010.

[96] R. Sandler and M. Lindenbaum, "Nonnegative Matrix Factorization with Earth Mover's Distance Metric for Image Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 33, no. 8, pp. 1590-1602, 2011.

[97] M. Berry, M. Browne, A. Langville, V. Pauca and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational Statistics and Data Analysis,* vol. 52, no. 1, pp. 155-173, 2007.

[98] V. Monga and M. Mihcak, "Robust and Secure Image Hashing via Non-Negative Matrix Factorizations," *IEEE Transactions on Information Forensics and Security,* vol. 2, no. 3, pp. 376-390, 2007.

[99] A. Cichocki et al., "Noninvasive BCIs: Multiway Signal-Processing Array Decompositions," *Computer,* vol. 41, no. 10, pp. 34-42, 2008.

[100] C. Damon, A. Liutkus, A. Gramfort and S. Essid, "Non-negative matrix factorization for single-channel EEG artifact rejection," in *ICASSP'13*, Vancouver, 2013.

[101] N. Mohammadiha, P. Smaragdis and A. Leijon, "Supervised and Unsupervised Speech Enhancement Using Nonnegative Matrix Factorization," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 21, no. 10, pp. 2140-2151, 2013.

[102] A. Ozerov and C. Fevotte, "Multichannel Nonnegative Matrix Factorization in Convolutive Mixtures for Audio Source Separation," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 18, no. 3, pp. 550-563, 2010.

[103] D. L. Sun and R. Mazumder, "Non-negative matrix completion for bandwidth extension: A convex optimization approach," in *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Southampton, 2013.

[104] A. Ozerov, C. Févotte, R. Blouet and J. L. Durrieu, "Multichannel nonnegative tensor factorization with structured constraints for user-guided audio source separation," in *ICASSP'11*, Prague, 2011.

[105] N. Bertin, R. Badeau and E. Vincent, "Enforcing Harmonicity and Smoothness in Bayesian Non-Negative Matrix Factorization Applied to Polyphonic Music Transcription," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 18, no. 3, pp. 538-549, 2010.

[106] C. Fevotte, N. Bertin and J.-L. Durrieu, "Nonnegative Matrix Factorization with the Itakura-Saito Divergence: With Application to Music Analysis," *Neural Computation,* vol. 21, no. 3, pp. 793-830, 2009.

[107] A. T. Cemgil, "Bayesian Inference for Nonnegative Matrix Factorisation Models," *Computational Intelligence and Neuroscience,* vol. 2009, p. 17, 2009.

[108] P. Hoyer, "Non-negative Matrix Factorization with Sparseness Constraints," *JMLR,* pp. 1457-1469, 2004.

[109] D. El Badawy, N. Q. K. Duong and A. Ozerov, "On-the-fly audio source separation," in *IEEE International Workshop on Machine Learning for Signal Processing*, Reims, 2014.

[110] N. Seichepine, S. Essid, C. Févotte and O. Cappé, "Soft Nonnegative Matrix Co-Factorization," *IEEE Transactions on Signal Processing,* vol. 62, no. 22, pp. 5940-5949, 2014.

[111] D. Cai, X. He, J. Han and T. S. Huang, "Graph Regularized Nonnegative Matrix Factorization for Data Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 33, no. 8, pp. 1548-1560, 2011.

[112] H. Lee and S. Choi, "CUR+NMF for learning spectral features from large data matrix," in *IEEE International Joint Conference on Neural Networks*, Hong Kong, 2008.

[113] C. Ding, T. Li and W. Peng, "Nonnegative Matrix Factorization and Probabilistic Latent Semantic Indexing: Equivalence, Chi-square Statistic, and a Hybrid Method," in *21st National Conference on Artificial Intelligence and the 18th Innovative Applications of Artificial Intelligence Conference*, Boston, MA, 2006.

[114] N. E. Huang, "Hilbert-Huang transform and its applications," *World Scientific,* vol. Vol. 16, 2014.

[115] H. Pengju and C. Xiaomeng, "A method for extracting fetal ECG based on EMD-NMF single channel blind source separation algorithm," *Technology and Health Care,* vol. 24, no. Recent Innovations on Biomedical Engineering, pp. 17-26, 2016.

[116] M. M. K. I. and K. Hirose, "Single-Mixture Audio Source Separation by Subspace Decomposition of Hilbert Spectrum," *IEEE Transactions on Audio, Speech, and Language Processing,* vol. 15, no. 3, pp. 893-900, 2007.

[117] A. Hayashi, H. Kameoka, T. Matsubayashi and H. Sawada, "Non-negative periodic component analysis for music source separation," in *016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Jeju, 2016.