

POLITECNICO DI MILANO
Scuola di Ingegneria dell'Informazione



POLO TERRITORIALE DI COMO

Corso di Laurea Magistrale in Ingegneria Informatica

**ALGORITMI ADATTATIVI
PER LA CANCELLAZIONE DEL DISTURBO
INTRODOTTO DAGLI AVVISATORI ACUSTICI
NELLE COMUNICAZIONI VIA RADIO
DEL CORPO NAZIONALE DEI VIGILI DEL FUOCO**

Relatore: Prof. Luigi Piroddi

Tesi di Laurea di: Nicola Stalliti, matr. 764406

Anno Accademico 2013-2014

POLITECNICO DI MILANO
Scuola di Ingegneria dell'Informazione



POLO TERRITORIALE DI COMO
Master of Science in Computer Engineering

**ADAPTIVE ALGORITHMS FOR
THE ATTENUATION
OF SIREN NOISE
IN RADIO COMMUNICATIONS OF THE
ITALIAN NATIONAL FIRE CORPS**

Supervisor: Prof. Luigi Piroddi

Master Graduation Thesis by: Nicola Stalliti, ID 764406

Academic Year 2013-2014

ad Arianna, per l'affetto e il sostegno
a Marco, Matteo e Paola, per il piacere di aver condiviso con voi quest'esperienza
alla mia famiglia, per tutto

Sommario

Gli avvisatori acustici installati sugli automezzi del Corpo Nazionale dei Vigili del Fuoco svolgono l'importante funzione di allertare i pedoni, e chi è alla guida di un autoveicolo, del passaggio del mezzo di emergenza. Il suono di sirena che essi producono porta con sé l'inconveniente di interferire con le comunicazioni radio degradando notevolmente la comprensione del parlato.

In passato sono stati proposti alcuni metodi per la cancellazione del suono di sirena che necessitano di microfoni o di collegamenti accessori per funzionare correttamente.

In questa tesi proponiamo due metodi per attenuare il suono di sirena che disturba le comunicazioni radio, utilizzando esclusivamente il segnale proveniente dal microfono della radio veicolare. Per fare ciò, abbiamo utilizzato due sistemi adattativi differenti che hanno entrambi la caratteristica di riuscire a tracciare nel tempo la frequenza che descrive l'andamento della sirena e, tramite questa informazione, applicano al segnale un adeguato filtraggio. I risultati ottenuti attraverso le simulazioni in ambiente virtuale sono stati ampiamente soddisfacenti.

Nello studio che abbiamo affrontato abbiamo scelto di basare la realizzazione e il *testing* dei nostri sistemi sul caso particolare delle sirene in dotazione al Corpo Nazionale dei Vigili del Fuoco. In fase di studio successivo, però, abbiamo rilevato che, con le opportune modifiche, questi metodi possono trovare applicazione anche per le sirene di tipologia diversa che sono in dotazione ad altri servizi di emergenza.

Abstract

The siren alarms installed on the vehicles used by the Italian National Fire Corps play the vital role in alerting pedestrian and drivers of the incoming of the emergency vehicle. The sound produced by such devices unfortunately interferes with the radio communications and degrades significantly the comprehension of the speech.

The methods proposed in the past for canceling the sound of the siren noises need microphones or additional connections to work correctly.

In this work we are going to present two methods that to damp the sound that interferes with the radio communications only use the signal coming from the microphone of vehicle's radio.

In order to do that, we have used two different adaptive systems, both with the ability to track in time the frequency that describes the trend of the siren and - thanks to this information - apply to the signal an appropriate filtering. The results obtained through the simulation in the virtual environment are fully satisfying.

In the research conducted we choose to set up the creation and the testing of our systems on the specific type of alarm used by the Italian National Fire Corps. Nevertheless, in the following step of the study, we also found out that, making the appropriate changes, those methods can find work also for different types of sirens used by other emergency corps.

Ringraziamenti

Desidero innanzitutto ringraziare il Professor Piroddi. Il suo aiuto preciso e puntuale è stato determinante per la realizzazione di questo lavoro.

Voglio poi porgere un sentito ringraziamento ai colleghi del Comando Provinciale dei Vigili del Fuoco di Como per avermi aiutato, non senza curiosità ed entusiasmo, a effettuare le registrazioni e le misurazioni sugli automezzi.

Infine ringrazio i colleghi del Comando Provinciale dei Vigili del Fuoco di Milano per avermi supportato nei momenti in cui, tra un intervento e l'altro, mi svincolavo dai lavori di caserma per dedicarmi alla stesura della tesi.

Indice

Sommario	V
Abstract	VII
1. Introduzione	1
1.1 Struttura della tesi	2
2. Presentazione della problematica e del contesto	3
2.1 Specifiche tecniche delle sirene in dotazione al C.N.VV.F	3
2.2 Misura e analisi delle sirene installate sugli automezzi	4
2.3 Il suono della sirena come disturbo nelle conversazioni radio	8
2.3.1 Estensione frequenziale e ampiezza della voce	8
2.3.2 Estensione frequenziale e ampiezza della sirena	8
3. Metodi documentati in letteratura	11
3.1 Cancellazione con utilizzo del suono di sirena come segnale di riferimento	11
3.1.1 Cancellazione con aggiornamento di centro-banda dei filtri <i>notch</i>	11
3.1.2 Cancellazione con ritardo e filtro adattativo LMS a doppio canale	14
3.1.3 Cancellazione con ritardo, filtro <i>anti-speech</i> e filtro adattativo NLMS a doppio canale	17
3.2 Cancellazione senza l'utilizzo del suono di sirena come segnale di riferimento	23
3.2.1 Cancellazione con ritardo temporale e filtro adattativo LMS a singolo canale	23
3.2.2 Cancellazione con ritardo temporale, filtro <i>anti-speech</i> e filtro adattativo NLMS a singolo canale	25
4. Metodo 1: con utilizzo dell' <i>Average Magnitude Difference Function</i>	29
4.1 Panoramica d'insieme del sistema proposto	29
4.2 Estrazione del <i>pitch</i> tramite <i>Average Magnitude Difference Function</i>	30
4.2.1 La Funzione di Autocorrelazione	31
4.2.2 La funzione <i>Average Magnitude Difference Function</i>	32
4.2.3 Estrazione del <i>pitch</i>	34
4.3 Filtro Passa-Alto	34
4.4 Filtri <i>notch</i>	37
4.5 Buffer	39
4.6 Linea di ritardo	39
4.7 Il metodo applicato	41
5. Metodo 2: con utilizzo dell' <i>IIR Adaptive Notch Filtering</i>	43
5.1 Panoramica d'insieme del sistema proposto	43
5.2 Estrazione del <i>pitch</i> tramite l' <i>Adaptive Notch Filtering</i>	44
5.2.1 Filtro <i>notch</i> in forma <i>lattice</i>	45
5.2.2 Algoritmo con Discesa del Gradiente	47
5.2.3 Algoritmo <i>Lattice Semplificato</i>	49

5.3 Filtro Passa-Banda	54
5.4 Il metodo applicato	55
6. Valutazione e confronto dei risultati sperimentali	57
6.1 Cancellazione per mezzo dell' <i>Average Magnitude Difference Function</i>	57
6.1.1 AMDF applicata a suono di sirena simulato	57
6.1.2 AMDF applicata a suono di sirena simulato immerso in rumore bianco	61
6.1.3 Filtro Passa-Alto	62
6.1.4 Banco di filtri <i>notch</i>	64
6.1.5 Ritardo Temporale	65
6.1.6 Simulazione del sistema completo con segnali reali	66
6.2 Cancellazione per mezzo dell' <i>IIR Adaptive Notch Filter</i>	67
6.2.1 ANF applicato a sinusoidi pure	67
6.2.2 ANF applicato a sinusoidi immerse in rumore bianco	69
6.2.3 Filtri passa-banda	70
6.2.4 Simulazione del sistema completo con segnali reali	72
6.3 Confronto tra i due metodi	73
6.4 Applicazione dei sistemi ad altre tipologie di sirene	73
7. Conclusioni e sviluppi futuri	75
7.1 Possibili sviluppi futuri	76
Appendice A	77
Bibliografia	80

Indice delle figure

2.1	Evoluzione nel tempo della frequenza fondamentale di un ciclo di sirena	4
2.2	Spettrogramma della voce umana	8
2.3	Spettrogramma del suono di sirena registrato all'interno di un'APS Iveco EuroCity	9
3.1	Variazione, rispetto al tempo, della frequenza fondamentale del suono di sirena di tipo <i>Wail</i>	12
3.2	Variazione, rispetto al tempo, della frequenza fondamentale del suono di sirena di tipo <i>Yelp</i>	12
3.3	Filtro per la cancellazione del suono di sirena con tono ascendente, risultante dalla sovrapposizione di tre filtri <i>notch</i> aventi frequenza di centro-banda leggermente differente	13
3.4	Filtro per la cancellazione del suono di sirena con tono discendente, risultante dalla sovrapposizione di tre filtri <i>notch</i> aventi frequenza di centro-banda leggermente differente	14
3.5	Diagramma a blocchi del filtro adattativo LMS con suono sirena di riferimento e ritardo temporale	15
3.6	Diagramma a blocchi del filtro adattativo NLMS con microfono di riferimento, filtro <i>anti-speech</i> e ritardo temporale	18
3.7	Risposta in frequenza dell'ampiezza di un segnale contenente la somma della voce e del suono di sirena di tipo <i>Yelp</i> , prima e dopo il filtraggio con filtri <i>anti-speech</i> di Yule-Walker	19
3.8	Risposta in frequenza dell'ampiezza di un segnale contenente la somma della voce e del suono di sirena di tipo <i>Hi-lo</i> , prima e dopo il filtraggio con filtro <i>anti-speech</i> Yule-Walker	20
3.9	Spettrogramma di un segnale contenente la somma della voce e del suono di sirena di tipo <i>Yelp</i> , prima (a) e dopo (b) il filtraggio con filtro <i>anti-speech</i> Yule-Walker	20
3.10	Il sistema di acquisizione dei segnali composto da due microfoni	21
3.11	Posizionamento dei microfoni utilizzato per la sperimentazione. (Mic.1) rappresenta il microfono della radio veicolare, (Mic.2.) quello di riferimento	23
3.12	Diagramma a blocchi del filtro adattativo LMS a singolo canale che utilizza come segnale di riferimento lo stesso segnale in ingresso ritardato nel tempo	24
3.13	Diagramma a blocchi del filtro adattativo NLMS che utilizza filtro <i>anti-speech</i> e ritardo temporale	25
4.1	Schema a blocchi del sistema per la cancellazione della sirena che utilizza l'AMDF e un banco di filtri <i>notch</i>	29
4.2	<i>Frame</i> di 800 campioni di segnale audio avente componenti periodiche e sua relativa (<i>Short-Time</i>) <i>Autocorrelation Function</i>	32
4.3	<i>Frame</i> di 800 campioni di segnale audio non periodico e sua relativa (<i>Short-Time</i>) <i>Autocorrelation Function</i> .	32
4.4	<i>Average Magnitude Difference Function</i> applicata ad un segnale avente componenti	33
4.5	Filtro Passa-Alto ideale	34
4.6	Specifiche necessarie per la realizzazione di un filtro passa-alto attraverso la funzione di MATLAB "fdesign.highpass"	36
4.7	Risposta in frequenza dell'ampiezza del filtro passa-alto applicato nel sistema di Figura 4.1	37
4.8	Posizionamento, sulla circonferenza unitaria, dei <i>poli</i> e degli <i>zeri</i> di un filtro <i>notch</i>	38

4.9	Risposta in frequenza dell'ampiezza di uno dei k filtri <i>notch</i> applicati nel blocco di filtraggio di sistema di Figura 4.1. Nell'esempio la frequenza di centro-banda è settata sul valore di 392 Hz e la banda passante (calcolata a -3 dB) è ampia circa 8 Hz (fattore di qualità $Q=45$).	39
5.1	Schema a blocchi del sistema per la cancellazione della sirena che utilizza l' <i>Adaptive Notch Filtering</i> e un banco di filtri <i>notch</i>	43
5.2	Realizzazione di una coppia complementare di filtri passa-banda/elimina-banda a partire da un filtro passa-tutto in forma <i>lattice</i>	45
5.3	Blocco base θk della struttura <i>lattice</i>	46
5.4	Generazione dei segnali regressori filtrati necessari all'algoritmo adattativo descritto nella (5.7)	48
5.5	Ampiezza della Funzione Costo al variare della larghezza della banda del filtro <i>notch</i> con frequenza di centro-banda fissa	49
5.6	Stima della frequenza fondamentale del segnale in ingresso con a) filtro con Algoritmo con Discesa del Gradiente e b) filtro con Algoritmo <i>Lattice</i> Semplificato nel caso di ampia larghezza della relativa banda passante	52
5.7	Stima della frequenza fondamentale del segnale in ingresso con a) filtro con Algoritmo con Discesa del Gradiente e b) filtro con Algoritmo <i>Lattice</i> Semplificato nel caso della relativa banda passante stretta	53
5.8	Filtro Passa-Banda ideale	54
5.9	Risposta in frequenza dell'ampiezza del filtro passa-banda utilizzato nel sistema di Figura 5.1 per la nota bassa	55
5.10	Risposta in frequenza dell'ampiezza del filtro passa-banda utilizzato nel sistema di Figura 5.1 per la nota alta	55
6.1	Spettrogramma di un ciclo di sirena bitonale generato con MATLAB	59
6.2	<i>Pitch</i> rilevato dalla funzione AMDF applicata al suono di sirena "artificiale" creato attraverso lo <i>Script</i>	60
6.3	<i>Pitch</i> rilevato dalla funzione AMDF applicata al suono "sirena+rumore" con 14 dB di differenza di ampiezza tra la sirena e il rumore bianco	62
6.4	<i>Pitch</i> rilevato dalla funzione AMDF applicata al suono "sirena+rumore" con sirena e rumore bianco aventi stessa ampiezza	62
6.5	Spettrogramma del segnale "voce+sirena"	63
6.6	Spettrogramma del segnale "voce+sirena" filtrato con filtro passa-alto avente frequenza di taglio 1.8kHz	64
6.7	Spettrogramma del segnale "voce+sirena" filtrato con sistema non adattativo basato su due blocchi di 30 filtri <i>notch</i> in cascata	65
6.8	Spettrogramma del segnale "voce+sirena" prima del filtraggio	67
6.9	Spettrogramma del segnale "voce+sirena" dopo il filtraggio	67
6.10	Variazione della frequenza della sinusoide del tempo	69

6.11	Frequenza tracciata dall'ANF applicato al segnale di Figura 6.10	69
6.12	Frequenza tracciata dall'ANF applicato al segnale di Figura 6.10 a cui è sommato un rumore bianco di ampiezza inferiore ad esso di 25 dB	70
6.13	Frequenza tracciata dall'ANF applicato al segnale di Figura 6.10 a cui è sommato un rumore bianco di ampiezza inferiore ad esso di 5 dB	70
6.14	Spettrogramma del segnale "voce+sirena" rappresentato in Figura 6.8 filtrato con i filtri passa-banda creati dallo <i>Script</i> 6.8	71
6.15	Spettrogramma del segnale "voce+sirena" rappresentato in Figura 6.8, dopo la cancellazione della sirena effettuata con il sistema che utilizza l'ANF	72

Indice delle tabelle

2.1 Misura della frequenza fondamentale delle note emesse dalle sirene installate su mezzi diversi	5
2.2 Misurazione del periodo di evoluzione “nota bassa/nota alta” delle sirene installate su mezzi diversi	6
2.3 Misurazione della pressione sonora emessa dalle sirene installate su diversi mezzi	7
2.4 Pressione sonora esercitata dalla voce umana misurata a varie distanze dalla bocca	9
3.1 Dati sperimentali ricavati applicando il metodo esposto nel Paragrafo 3.1.2	16
3.2 Dati sperimentali ottenuti applicando il metodo esposto nel Paragrafo 3.1.2 modificato	17
3.3 Incremento del Rapporto Segnale-Rumore ottenuto applicando il sistema di cancellazione con ritardo, filtro <i>anti-speech</i> e filtro adattativo NLMS a doppio canale	22
3.4 Confronto tra le prestazioni di sistemi a canale singolo con algoritmi differenti per sirena tipo <i>Yelp</i>	27
3.5 Risultati ottenuti applicando il sistema adattativo a singolo canale con filtro <i>anti-speech</i> e ritardo temporale a più tipi di sirena	27
4.1 Valori di ingresso alla funzione “ <code>fdesign.highpass</code> ” per la realizzazione del filtro passa-alto nel sistema di Figura 4.1	36
5.1 Valori di ingresso alla funzione “ <code>fdesign.bandpass</code> ” per la realizzazione dei filtri passa-banda nel sistema di Figura 5.1	55
6.1 Valori dei ritardi introdotti nell’esecuzione, con MATLAB, dei vari blocchi di elaborazione del segnale	66

Indice degli script MATLAB

6.1 Generazione di un ciclo di sirena bitonale completo	58
6.2 Funzione che estrae il <i>pitch</i> del segnale in ingresso attraverso AMDF	59
6.3 Generazione del segnale “sirena+rumore bianco” ed estrazione del <i>pitch</i> tramite AMDF	61
6.4 Creazione del filtro passa-alto e filtraggio del segnale “sirena+voce”	63
6.5 Banco di filtri <i>notch</i>	64
6.6 Funzione per l’applicazione di un ritardo temporale L al segnale in ingresso	66
6.7 <i>Simplified Lattice Adaptive IIR Notch Filter</i>	68
6.8 Filtri passa-banda	71

Capitolo 1

Introduzione

Quando un operatore presente all'interno di un mezzo in dotazione al Corpo Nazionale dei Vigili del Fuoco, durante il tragitto verso il luogo dove la squadra deve compiere l'intervento, comunica via radio con un operatore presente su un altro mezzo o con la Sala Operativa, colui che ascolta riceve il segnale di chi trasmette fortemente disturbato dal suono della sirena che, all'interno dell'abitacolo, viene catturato dal microfono assieme alla voce.

Il *range* di frequenze in cui si sviluppa il suono delle sirene (300 Hz - 10000 Hz) si va a sommare alla banda coperta dalla voce (100Hz - 3000 Hz). Inoltre i volumi dei due suoni, circa 75 dB per il suono di sirena registrato all'interno dell'abitacolo con finestrini chiusi e circa 90 dB per la voce emessa da un operatore che parla a 1 cm dal microfono, risultano di intensità comparabile. Questo costituisce un problema, spesso risolto brutalmente spegnendo temporaneamente la sirena quando vi è la necessità di comunicare via radio. È evidente come questa pratica risulti altamente pericolosa privando di una segnalazione efficace il passaggio del mezzo in servizio di emergenza.

In letteratura sono documentati alcuni sistemi capaci di cancellare il suono di sirena senza eliminare con esso anche il suono della voce. Tali sistemi, tuttavia, forniscono in uscita un suono vocale distorto e, funzionando con algoritmi *Least Mean Square* oppure attraverso l'analisi del *pitch* della sirena tramite *zero-crossing*, necessitano non solo del segnale rilevato dal microfono della radio, ma anche dell'installazione di microfoni ausiliari (o del cablaggio di cavi elettrici) al fine di reperire un segnale di sirena "puro" da utilizzare come riferimento.

L'obiettivo dichiarato di questa tesi è quello di proporre, sviluppare e testare dei metodi adattativi che attuino la cancellazione della sirena utilizzando il solo segnale proveniente dal microfono della radio veicolare e che diano in uscita un segnale di voce non distorto e altamente comprensibile per chi lo riceve.

Partendo da questi presupposti abbiamo sviluppato due sistemi differenti per ottenere questo obiettivo. Entrambi, attraverso il continuo riconoscimento nel tempo della frequenza fondamentale che descrive l'andamento della sirena, riescono a impostare e modificare in continuazione un banco di filtri *notch* elimina-banda settati sulla frequenza fondamentale della sirena e sulle relative armoniche.

Per effettuare quest'operazione, il primo dei due utilizza l'*Average Magnitude Difference Function* applicata a brevi "finestre" estratte con continuità dal segnale in ingresso.

Il secondo metodo traccia istante per istante la frequenza fondamentale della sirena, attraverso l'utilizzo di un *IIR Adaptive Notch Filter*.

Il secondo metodo traccia istante per istante la frequenza fondamentale della sirena, attraverso l'utilizzo di un *IIR Adaptive Notch Filter*.

Con il supporto di adeguate misurazioni e registrazioni effettuate direttamente sui mezzi in dotazione al Corpo Nazionale dei Vigili del Fuoco, abbiamo implementato entrambi i sistemi in un ambiente virtuale di simulazione e siamo effettivamente riusciti a ottenere una buona cancellazione del suono di sirena senza che al contempo si verificasse una distorsione del segnale vocale.

In fase successiva abbiamo appurato che, apportando alcune modifiche in relazione al tipo di sirena da cancellare, è possibile estendere la validità dei due metodi anche al caso di avvisatori acustici di genere diverso.

1.1 Struttura della tesi

La tesi è impostata come segue:

- nel Capitolo 2 verrà presentata la problematica del disturbo introdotto dalle sirene nelle conversazioni radio e il contesto in cui essa si verifica. Inoltre, verranno analizzate le specifiche tecniche delle sirene in dotazione al Corpo Nazionale dei Vigili del Fuoco e verranno illustrati i risultati delle misure acustiche relativamente ai suoni di voce e di sirena all'interno degli automezzi in dotazione al Corpo;
- nel Capitolo 3 verranno esposte le metodologie presenti in letteratura per la cancellazione dei suoni di sirena, se ne valuterà l'efficacia e se ne discuteranno i limiti;
- nel Capitolo 4 presenteremo un primo metodo per l'eliminazione del suono emesso dagli avvisatori acustici, basato sull'*Average Magnitude Difference Function*, per l'estrazione del *pitch* del suono di sirena e che, attraverso questa informazione, imposta un banco di filtri *notch* che effettua la cancellazione;
- nel Capitolo 5 mostreremo un secondo metodo, per l'eliminazione del suono emesso dagli avvisatori acustici, che utilizza l'*IIR Adaptive Notch Filter* per tracciare la frequenza istantanea del suono di sirena e che, utilizzando questa informazione, imposta un banco di filtri *notch* che effettua la cancellazione;
- nel Capitolo 6 presenteremo i risultati ottenuti dalla simulazione in ambiente virtuale dei metodi proposti valutandoli e confrontandoli tra loro;
- nel Capitolo 7 concluderemo il lavoro riassumendo i risultati ottenuti e proponendo possibili sviluppi futuri.

Capitolo 2

Presentazione della problematica e del contesto

In questo capitolo ci occuperemo di analizzare le caratteristiche acustiche del suono emesso dalle sirene in dotazione al Corpo Nazionale dei Vigili del Fuoco (C.N.VV.F.). In primo luogo ne andremo a descrivere le specifiche di legge, successivamente mostreremo lo studio effettuato sui reali dispositivi acustici installati sugli automezzi. In ultimo esporremo la problematica inerente al disturbo che tale suono genera nelle conversazioni radio tra il personale presente sui mezzi di soccorso e quello presente in Sala Operativa.

2.1 Specifiche tecniche delle sirene in dotazione al C.N.VV.F.

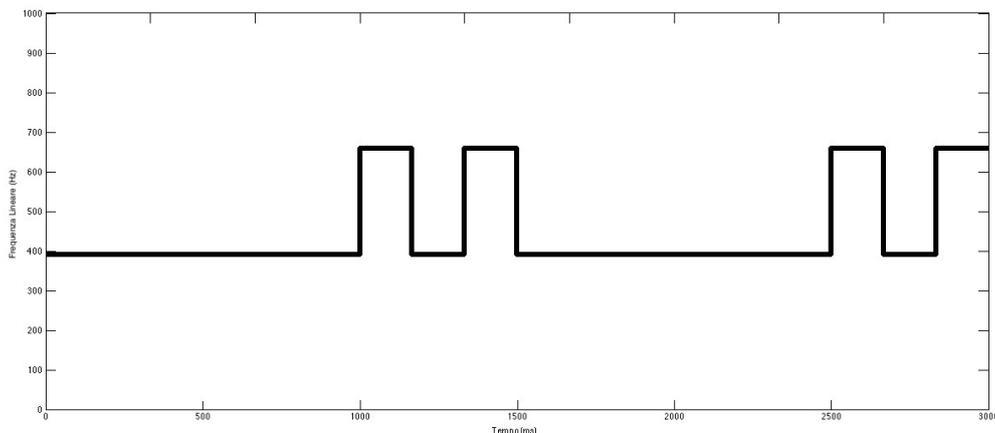
I dispositivi acustici di allarme in dotazione ai mezzi di soccorso del Corpo Nazionale dei Vigili del Fuoco, unitamente a quelli in dotazione alle autoambulanze, sono regolamentati in maniera dettagliata dal Decreto Ministeriale del 17/10/1980 (G.U. n.310 del 12/11/1980)¹. Qui troviamo una descrizione esatta del suono che devono generare tali dispositivi: la tonalità, l'andamento nel tempo e il livello di pressione sonora.

La sirena è, per legge, di tipo bitonale: ovvero è composta da due soli toni²: *L* e *H* che si susseguono nel tempo secondo caratteristiche ben definite. Il tono *L* con frequenza fondamentale pari a 392 Hz (Sol 4 Naturale o G4 nella notazione anglosassone) e il tono *H* con frequenza fondamentale pari a 660 Hz (Mi 5 Naturale o E5 nella notazione anglosassone). Su tali valori è consentita una tolleranza del $\pm 5\%$. Il livello di pressione sonora di questi suoni, a una distanza di 2 metri dal dispositivo di segnalazione, deve essere compreso tra i 115 dB e i 125 dB e il loro contenuto armonico, nella banda frequenziale da 1800 Hz a 4500 Hz, non deve essere inferiore a 105 dB.

Un ciclo completo di evoluzione nel tempo dei due toni è rappresentato in Figura 2.1. Esso, secondo le specifiche, deve durare nel complesso 3 secondi (con

¹ il testo completo del Decreto Ministeriale è riportato in Appendice A

² in questo contesto per "tono" si intende il tono musicale (la nota) e non "tono" inteso come singola forma d'onda sinusoidale

Figura 2.1: Evoluzione nel tempo della frequenza fondamentale di un ciclo di sirena.

Un ciclo completo di evoluzione nel tempo dei due toni è rappresentato in Figura 2.1. Esso, secondo le specifiche, deve durare nel complesso 3 secondi (con una tolleranza di ± 0.5 s) ripetendosi sempre uguale nel tempo e seguendo questa sotto-suddivisione.

2.2 Misura e analisi delle sirene installate sugli automezzi

Attraverso la registrazione dei suoni emessi da diversi dispositivi acustici di allarme presenti su tipologie diverse di automezzi attualmente in dotazione al C.N.VV.F., abbiamo voluto testare l'adeguatezza alle specifiche imposte dal Decreto Ministeriale delle sirene realmente installate e in funzione sugli stessi. Sono state interessate dallo studio quattro AutoPompa-Serbatoio (APS), due AutoScala (AS), due AutoBotte-Pompa (ABP), una Campagnola (CA) e una Autogrù (AG). La ricerca si è basata su cinque tipi di misurazione:

- misurazione della tonalità dei suoni emessi dalle sirene;
- misurazione del periodo di evoluzione “nota bassa/nota alta” (media calcolata su un intervallo di 3 cicli completi);
- misurazione della pressione sonora percepita all'esterno del mezzo a 2 metri di distanza dalla sirena montata sul veicolo (media calcolata su valori rilevati ogni 1.5 s su un intervallo di 3 cicli completi);
- misurazione della pressione sonora percepita all'interno del veicolo, lato passeggero con finestrini aperti (media calcolata su valori rilevati ogni 1.5 s su un intervallo di 3 cicli completi);

- misurazione della pressione sonora percepita all'interno del veicolo, lato passeggero con finestrini chiusi (media calcolata su valori rilevati ogni 1.5 s su un intervallo di 3 cicli completi);

La scelta di effettuare quest'ultima misurazione nella posizione del passeggero non è casuale: è, infatti, dove si siede il CapoSquadra che, nella quasi totalità dei casi, è l'unica figura che interagisce via radio con la Sala Operativa o con il personale presente su altri mezzi durante lo spostamento verso il luogo in cui si porta soccorso.

I valori registrati per le prime due tipologie di misurazione sono riportati rispettivamente nella Tabella 2.1 e nella Tabella 2.2 mentre quelli registrati per le ultime tre misurazioni sono riassunti assieme nella Tabella 2.3.

Dai dati esposti nella Tabella 2.1 si evince come, nella realtà, la tonalità dei suoni emessi dalle sirene non si differenzi troppo da quella imposta dalle specifiche ministeriali. Seppure in nessun caso si abbia una esatta collocazione a 392 Hz e a 660 Hz dei due toni, nove volte su dieci non ci si allontana di più del 3% da questi valori. Solo nel caso dell'AS Iveco Zeta, si ha uno scostamento maggiore rispetto ai valori indicati dalle specifiche. Lo scostamento non supera comunque il 5% e risulta quindi in linea con la disposizione ministeriale. È importante sottolineare che questo mezzo è il più vecchio tra tutti quelli presi in considerazione e ha installata, come la CA Land Rover Defender, una sirena di tipo elettropneumatico. Questa tipologia di sirena necessita di una periodica manutenzione che, se non effettuata correttamente, fa sì che la sirena si "scordi" nel tempo e inoltre, anche se in perfetto stato di mantenimento, risulta essere molto meno precisa di quelle di tipo elettronico montate su tutti gli altri automezzi presi in considerazione.

Dai dati rilevati, esposti nella Tabella 2.2, notiamo come, anche in questa circostanza, i dispositivi acustici di allarme presenti sui mezzi presi in esame

Tabella 2.1: Misura della frequenza fondamentale delle note emesse dalle sirene installate su mezzi diversi.

Mezzo	Frequenza Nota Bassa	Frequenza Nota Alta
APS Mercedes Actros	394 Hz	661 Hz
APS Iveco Stralis	387 Hz	658 Hz
APS Iveco EuroCity	389 Hz	657 Hz
APS Volvo fl 280	393 Hz	665 Hz
AS Iveco Magirus	392 Hz	661 Hz
AS Iveco Zeta	374 Hz	635 Hz
ABP Mercedes Actros	390 Hz	658 Hz
ABP Iveco EuroCargo	386 Hz	652 Hz
CA Land Rover Defender	401 Hz	678 Hz
AG Cormach Astra	396 Hz	664 Hz

Tabella 2.2: Misurazione del periodo di evoluzione “nota bassa/nota alta” delle sirene installate su mezzi diversi.

Mezzo	1° Ciclo	2° Ciclo	3° Ciclo	4° Ciclo	5° Ciclo	6° Ciclo	7° Ciclo	8° Ciclo	Totale
APS Mercedes Actros	1.03s	0.17s	0.17s	0.16s	1.03s	0.17s	0.17s	0.16s	3.06s
APS Iveco Stralis	1.02s	0.16s	0.17s	0.17s	1.03s	0.17s	0.17s	0.16s	3.06s
APS Iveco EuroCity	1.03s	0.17s	0.17s	0.17s	1.04s	0.17s	0.17s	0.16s	3.05s
APS Volvo fl 280	0.99s	0.15s	0.15s	0.16s	1.00s	0.16s	0.16s	0.15s	2.92s
AS Iveco Magirus	0.99s	0.16s	0.16s	0.16s	0.98s	0.16s	0.17s	0.16s	2.94s
AS Iveco Zeta	0.97s	0.16s	0.15s	0.15s	0.98s	0.15s	0.15s	0.16s	2.87s
ABP Mercedes Actros	1.03s	0.17s	0.16s	0.16s	1.03s	0.17s	0.16s	0.16s	3.04s
ABP Iveco EuroCargo	1.02s	0.17s	0.17s	0.17s	1.03s	0.16s	0.17s	0.17s	3.06s
CA Land Rover Defender	1.04s	0.15s	0.17s	0.17s	1.03s	0.17s	0.17s	0.17s	3.07s
AG Cormach Astra	0.99s	0.17s	0.17s	0.17s	0.99s	0.17s	0.17s	0.17s	2.97s

tolleranza di ± 0.5 s. Ancora una volta osserviamo che la performance peggiore è realizzata dalla sirena di tipo elettropneumatico installata sull'AS Iveco Zeta.

La prima colonna della Tabella 2.3 ci mostra come, per tutti i dispositivi acustici di allarme installati sui differenti mezzi, il volume del suono di sirena misurato in ambiente esterno rispecchia pienamente le specifiche ministeriali. Per tutti i casi, infatti, la media calcolata sui 6 valori ottenuti rilevando ogni 1.5 s il valore istantaneo del volume del suono emesso dai diversi dispositivi, è sempre compreso tra i 115 e i 125 dB. La seconda colonna ci mostra invece come nella posizione, del passeggero, il suono della sirena entri nella cabina del mezzo con finestrini aperti ad un volume notevole. Tale volume, per tutti casi presi in esame, si attesta ad un valore di circa 40 dB in meno del valore misurato all'esterno in posizione frontale e a 2 metri di distanza dal diffusore della sirena. Questo spontaneo abbattimento di volume è dovuto principalmente a due fattori:

Tabella 2.3: Misurazione della pressione sonora emessa dalle sirene installate su diversi mezzi, percepita rispettivamente: all'esterno a 2 m di distanza dalle sirene in direzione frontale; all'interno lato passeggero con finestrini aperti; all'interno lato passeggero con finestrini chiusi. Tutti valori sono espressi in dB SPL (riferimento a 20 μ Pa).

Mezzo	Press. Sonora Est.	Press. Sonora Int. Finestrino Aperto	Press. Sonora Int. Finestrino Chiuso
APS Mercedes Actros	120 dB	83dB	74 dB
APS Iveco Stralis	121 dB	83 dB	73 dB
APS Iveco EuroCity	117 dB	80 dB	72 dB
APS Volvo fl 280	119 dB	81 dB	73 dB
AS Iveco Magirus	118 dB	81 dB	74 dB
AS Iveco Zeta	122 dB	84 dB	79 dB
ABP Mercedes Actros	121 dB	81 dB	74 dB
ABP Iveco EuroCargo	119 dB	79 dB	73 dB
CA Land Rover Defender	122 dB	85 dB	79 dB
AG Cormach Astra	121 dB	82 dB	75 dB

- la direzionalità molto spiccata dei diffusori delle sirene che crea per sua stessa natura una zona a volume ridotto nella posizione opposta a quella della direzione della tromba di diffusione;
- la schermatura dovuta alla struttura stessa della cabina: la carrozzeria in metallo, il parabrezza in vetro, rivestimenti in plastica, stoffa e gomma.

La terza colonna ci mostra come questo abbattimento di volume, dovuto ai materiali che compongono la cabina e l'abitacolo, risulta ulteriormente intensificato dalla chiusura dei finestrini. In tutti gli automezzi presi in esame, a eccezione dei due più vecchi (AS Iveco Zeta e CA Land Rover Defender), la chiusura dei finestrini contribuisce all'abbattimento di altri circa 8 dB facendo così attestare il valore registrato nell'abitacolo lato passeggero su un valore di 74 dB circa.

2.3 Il suono della sirena come disturbo nelle conversazioni radio

2.3.1 Estensione frequenziale e ampiezza della voce

Come è noto, lo spettro della voce umana si estende in una banda che va da circa 50 Hz a circa 5000 Hz [1] e il contenuto armonico strettamente necessario alla compressione del parlato è considerato, per convenzione, compreso tra i 300 Hz e i 3000 Hz [2]. Per avere un riscontro di questa estensione in frequenza, in Figura 2.2 mostriamo lo spettrogramma dell'evoluzione nel tempo di una breve sequenza di parlato.

Notiamo come effettivamente, nell'esempio preso in considerazione, lo spettro vocale si estenda fino a oltre 5000 Hz e come il grosso del suo contenuto armonico sia compreso tra i 100 Hz e i 3500 Hz.

Dal punto di vista dell'ampiezza, invece, il valore della pressione sonora esercitata dalla voce umana a 1 metro di distanza dalla bocca, si estende da un minimo di circa 30 dB SPL per il sussurro ad un massimo di circa 90 dB SPL per l'urlo [3]. Nella Tabella 2.4 sono schematizzati alcuni valori statistici che abbiamo ottenuto facendo una media dei valori misurati registrando voci di donna e di uomo di diversa età posizionando il microfono a differenti distanze dalla bocca. La tabella mostra esplicitamente quanto sia ampia la dinamica vocale umana. È importante focalizzarsi sui valori riportati per la distanza di 1 centimetro della bocca dal microfono in quanto questa è la situazione tipica che si crea nelle conversazioni radio oggetto del nostro studio.

2.3.2 Estensione frequenziale e ampiezza della sirena

Nei paragrafi 2.1 e 2.2 abbiamo osservato come la sirena bitonale in dotazione al C.N.VV.F. sia composta da due toni posti uno a circa 392 Hz e l'altro a circa 660 Hz. Questi due toni però non risultano essere due sinusoidi pure ma, al contrario, hanno una ricca colorazione armonica. A dimostrazione di ciò, nella Figura 2.3 è

Figura 2.2: Spettrogramma della voce umana

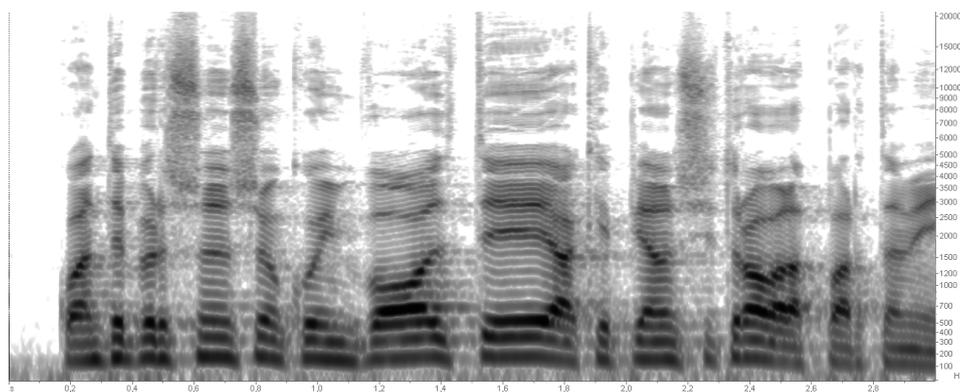


Tabella 2.4: Pressione sonora esercitata dalla voce umana misurata a varie distanze dalla bocca.

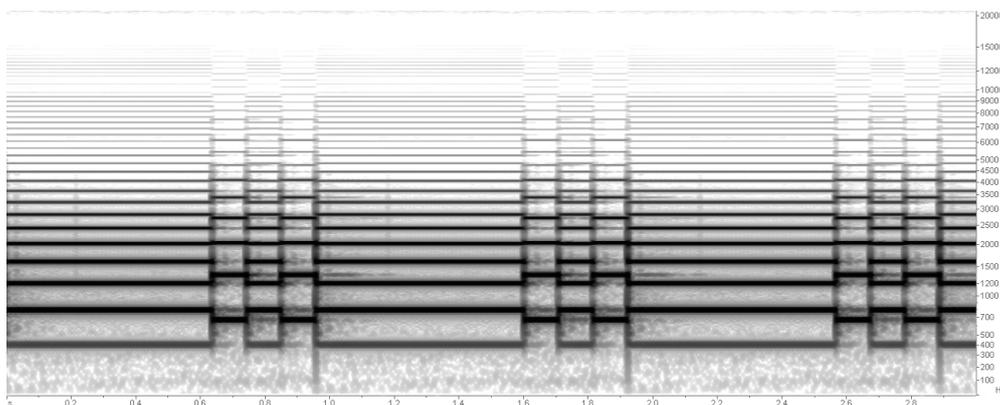
Distanza dalla bocca	Sussurro	Parlato normale	Urlo
0.01 m	53 dB	91 dB	118 dB
0.30 m	40 dB	83 dB	100 dB
1.00 m	33 dB	74 dB	89 dB
5.00 m	22 dB	65 dB	81 dB
10.00 m	9 dB	52 dB	70 dB

raffigurato lo spettrogramma di una sirena che compie circa tre cicli completi registrata all'interno dell'abitacolo di un'autopompa APS Iveco EuroCity. Si noti come, sia per il suono a 392 Hz che per il suono a 660 Hz, siano presenti più di 20 armoniche ciascuno.

2.3.3 Il segnale risultante dalla somma della voce e della sirena

Durante una conversazione radio tra il personale presente su un automezzo che si muove a sirene spiegate e la Sala Operativa, nel segnale audio che viene trasdotto dal microfono presente nella cabina dell'automezzo sono presenti contemporaneamente sia la voce di chi parla sia il suono di sirena che entra dentro l'abitacolo. Questi suoni, come mostrato dalla Figura 2.2 e dalla Figura 2.3 sono assolutamente compatibili dal punto di vista della banda frequenziale coperta dal loro spettro e perciò si trovano a miselarsi e a confondersi tra loro.

Inoltre, alla luce di quanto emerso dalle misurazioni esposte nel Paragrafo 2.2 e riassunte nella Tabella 2.3, e per quanto esposto nel Paragrafo 2.3.1 e nella Tabella 2.4, anche dal punto di vista del volume i due suoni risultano essere di intensità comparabili. Il suono di voce risulta, sì, generalmente di volume più alto - poiché nella normale conversazione radio la bocca di chi parla si trova a circa 1 cm dal microfono della radio veicolare producendo quindi tra gli 85 dB e i 95 dB di pressione sonora - ma soltanto di circa 10 dB in più della sirena che, come abbiamo visto in precedenza, all'interno dell'abitacolo si attesta sempre tra i 75 dB e gli 85 dB.

Figura 2.3: Spettrogramma del suono di sirena registrato all'interno di un'APS Iveco EuroCity

Da queste osservazioni risulta chiaro che il suono di sirena, che entra all'interno delle cabine degli automezzi e viene trasdotto dal microfono insieme alla voce, rende molto difficoltosa la comprensione delle comunicazioni formulate via radio. Per questo motivo si rende necessario sviluppare un sistema che riesca a cancellarlo o che ne limiti notevolmente l'ampiezza.

Capitolo 3

Metodi documentati in letteratura

In questo capitolo verrà fornita una panoramica dei metodi di cancellazione del suono di sirena dalle comunicazioni radio dei servizi di emergenza già presenti in letteratura.

Per chiarezza suddivideremo l'esposizione in due parti. La prima (Paragrafo 3.1) in cui illustreremo i metodi che, per effettuare questa cancellazione, utilizzano due segnali: uno di riferimento, che contiene il solo suono diretto della sirena, l'altro, proveniente dal microfono della radio veicolare, che contiene la somma della voce e del suono della sirena; nella seconda parte (Paragrafo 3.2) descriveremo dei metodi più recenti che utilizzano esclusivamente il segnale proveniente dal microfono della radio veicolare, contenente la somma del suono della sirena più quello della voce.

3.1 Cancellazione con utilizzo del suono di sirena come segnale di riferimento

3.1.1 Cancellazione con aggiornamento di centro-banda dei filtri *notch*

Il metodo proposto da Horvei, Ottesen e Svean [4] si prefigge l'obiettivo di ridurre di almeno 20 dB il suono di sirena che entra nelle comunicazioni radio sui mezzi di emergenza. I tipi di sirena che prende in considerazione sono quelli in dotazione alle forze di polizia norvegesi:

- il segnale di tipo *Wail* che è strutturato come un'onda quadra che varia la sua frequenza fondamentale nel tempo e che, come mostrato in Figura 3.1, in un ciclo completo della durata di 5.45 s, si sposta linearmente da 500 Hz a 1800 Hz e da 1800 Hz a 500 Hz;
- il segnale di tipo *Yelp* che si presenta anch'esso come un'onda quadra che varia la sua frequenza fondamentale nel tempo linearmente da 500 Hz a 1800 Hz e da 1800 Hz a 500 Hz ma che, come mostrato in Figura 3.2., compie un ciclo completo in appena 0.38 s.

Per ottenere l'eliminazione del suono di sirena dal segnale che contiene sia quest'ultimo che il suono della voce, questo metodo utilizza dodici filtri elimina banda a banda stretta - comunemente noti come filtri *notch* - che vengono posizionati di volta in volta, a gruppi di tre, sulla frequenza fondamentale del segnale e sulle prime tre relative armoniche. La scelta di applicarne tre, e non soltanto uno, nasce dall'esigenza di aumentare l'attenuazione in dB della banda

Figura 3.1: Variazione, rispetto al tempo, della frequenza fondamentale del suono di sirena tipo *Wail*.

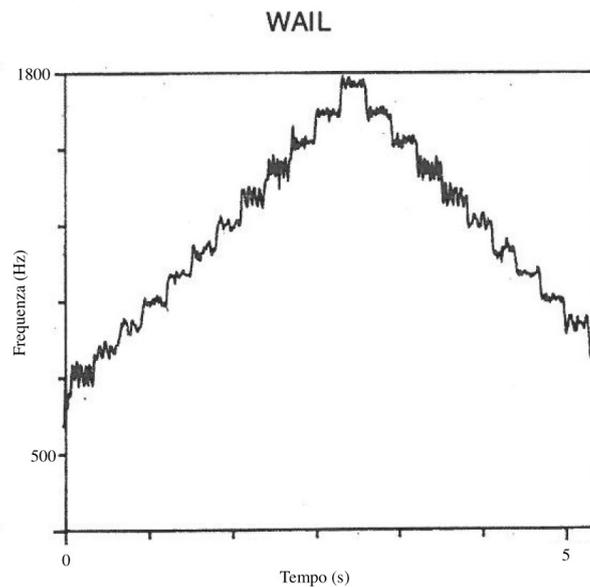
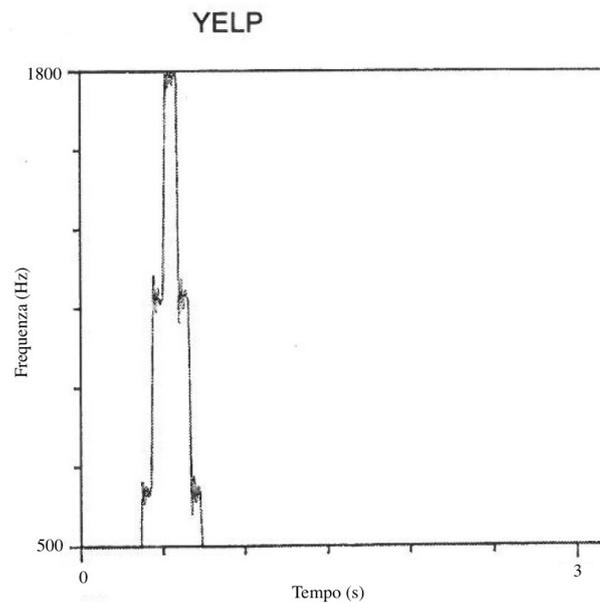


Figura 3.2: Variazione, rispetto al tempo, della frequenza fondamentale del suono di sirena tipo *Yelp*.



eliminata. Per poter continuamente aggiornare la frequenza di centro-banda dei filtri, il sistema utilizza, come segnale di riferimento, il segnale preso direttamente in uscita dal generatore elettrico della sirena. Attraverso l'osservazione dei punti in cui il segnale elettrico attraversa, nella sua frequenza fondamentale, il valore 0 (*zero-crossing*), il procedimento riesce ad identificare il semi-periodo e quindi la frequenza istantanea della sirena. Una volta ottenuta questa informazione, esso applica al segnale trasdotto dal microfono della radio veicolare, contenente sia la voce che la sirena, tre filtri *notch* per la frequenza fondamentale, tre per la prima

armonica, tre per la seconda armonica e tre per la terza armonica. I coefficienti necessari alla corretta costruzione dei filtri *notch* digitali sono ottenuti attraverso l'osservazione di una tabella pre-registrata nel sistema (*look-up table*).

I tre filtri presenti in ognuna delle quattro sezioni non sono, però, perfettamente accordati sulla stessa frequenza di centro-banda ma sono leggermente differenziati tra loro per poter coprire una banda maggiore. Questa accortezza nasce dall'esigenza di volere ovviare all'effetto dovuto alla riverberazione di circa 100 ms, naturalmente presente nell'abitacolo dei mezzi della polizia presi in esame. Infatti, osservando il segnale in ingresso al microfono ci si accorge che, dato un preciso istante temporale, non è presente solo il segnale emesso dalla sirena in quell'istante, ma anche suoni di sirena emessi ad istanti precedenti e quindi a frequenze fondamentali leggermente diverse. Il filtro totale risultante dalla somma di questi tre filtri *notch* leggermente diversi tra loro è visualizzabile nella Figura 3.3 per quanto riguarda la fase ascendente della frequenza fondamentale della sirena e nella Figura 3.4 per la fase discendente. Nel loro studio, Horvei, Ottesen e Svean riportano di aver ottenuto, con il metodo da loro proposto, un'attenuazione media di 28 dB per la sirena di tipo *Wail* e un'attenuazione media di 22 dB per la sirena di tipo *Yelp*. Essi riferiscono inoltre che, durante numerose sperimentazioni con radio installate sulle autovetture della polizia norvegese da loro prese in considerazione, al momento dell'accensione del sistema, non si manifestava alcuna degradazione della voce per effetto del filtraggio ma, al contrario, si verificava un notevole incremento della comprensibilità del parlato dovuto alla forte soppressione del suono di sirena.

Tuttavia, questo sistema possiede tre importanti limiti:

Figura 3.3: Filtro per la cancellazione del suono di sirena con tono ascendente, risultante dalla sovrapposizione di tre filtri *notch* aventi frequenza di centro-banda leggermente differente.

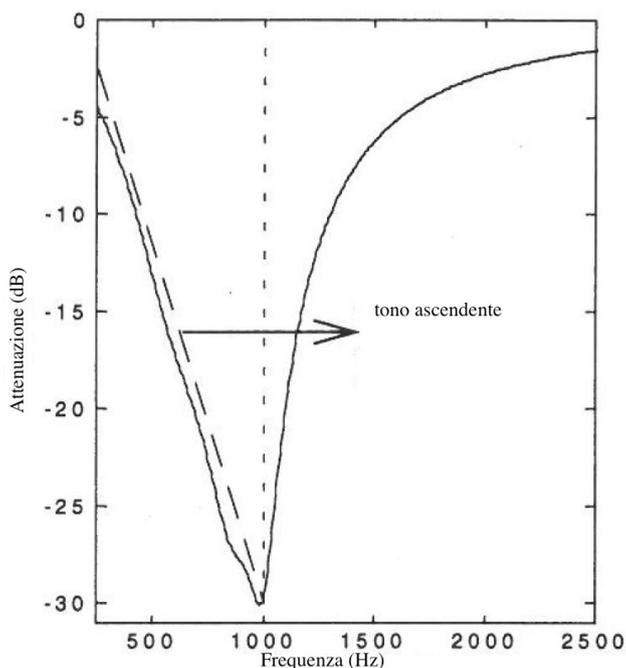
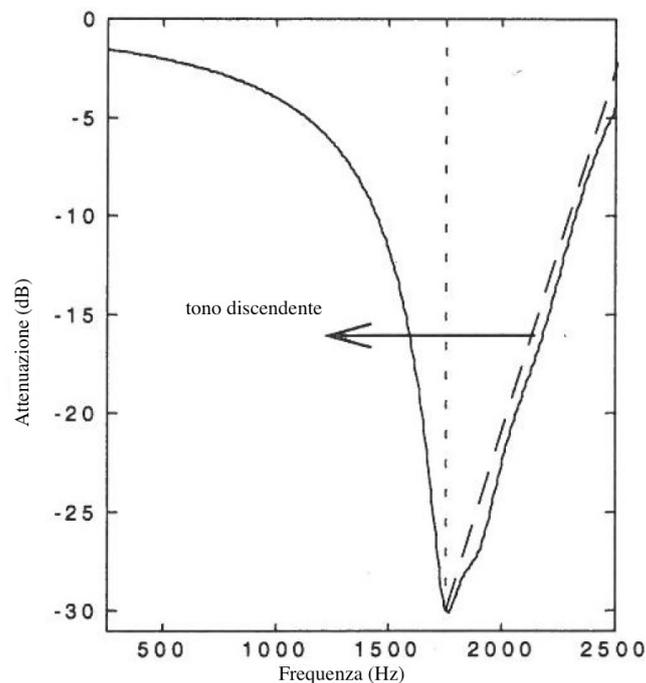


Figura 3.4: Filtro per la cancellazione del suono di sirena con tono discendente, risultante dalla sovrapposizione di tre filtri *notch* aventi frequenza di centro-banda leggermente differente.

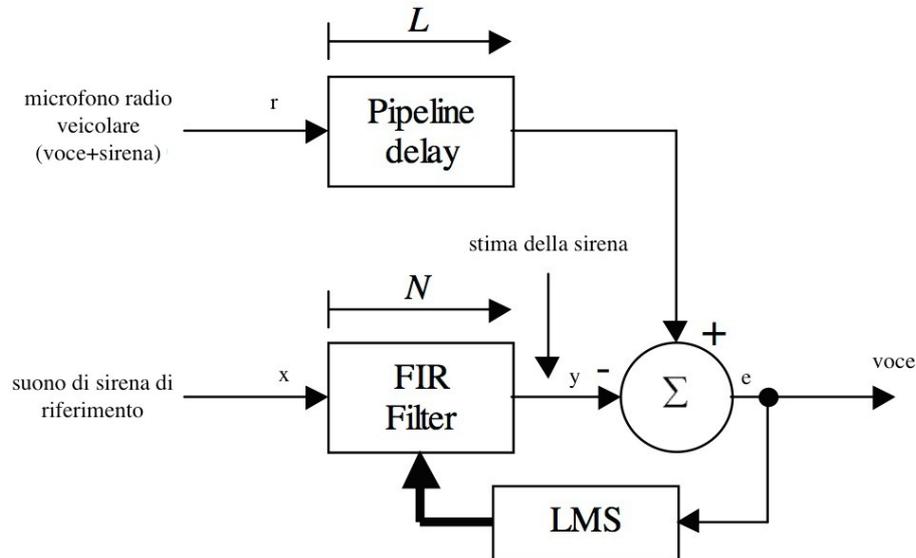


- latenza: nel tempo necessario al sistema per identificare la frequenza della sirena e impostare correttamente i filtri, la sirena ha già aumentato o diminuito la sua frequenza fondamentale;
- scarso numero di armoniche rimosse: il sistema cancella solo la fondamentale e le tre armoniche che compongono l'onda quadra in uscita al generatore elettronico di sirena quando in realtà il numero di armoniche presenti nel segnale audio in uscita è maggiore. Il sistema non è poi stato testato su sirene di tipo bitonale (*Hi-Lo*) che sono quelle che, rispetto alle altre, contengono il più largo spettro armonico;
- numero finito di coefficienti (relativi al numero finito di filtri) immagazzinati nella *look-up table*: la frequenza istantanea della sirena rilevata attraverso lo *zero-crossing* deve essere arrotondata per poter essere un valido valore di ingresso alla *look-up table*.

3.1.2 Cancellazione con ritardo e filtro adattativo LMS a doppio canale

Nel metodo proposto da Sherratt R.S., Townsend D. M., Guy C.G. [5], si utilizza un filtro di tipo adattativo per effettuare la cancellazione del suono di sirena dalle conversazioni radio. Il filtro proposto è rappresentato in Figura 3.5 ed è composto da quattro blocchi:

Figura 3.5: Diagramma a blocchi del filtro adattativo LMS con suono sirena di riferimento e ritardo temporale.



- un ritardo (*pipeline delay*) che genera in uscita il segnale in ingresso ritardato di L istanti temporali;
- un filtro digitale FIR a risposta finita all'impulso (*Finite Impulse Response*) che restituisce in uscita il segnale in ingresso filtrato con un filtro digitale di ordine N ;
- un blocco che tramite l'algoritmo LMS (*Least Mean Square*) minimizza l'errore quadratico medio tramite successive iterazioni;
- un sommatore che restituisce istante per istante la somma (in questo caso, la differenza) dei segnali in ingresso.

Indicando, con n il valore dell'istante temporale corrente, con x il suono di sirena di riferimento in ingresso al sistema e con y il segnale in uscita al filtro FIR possiamo scrivere l'equazione:

$$y(n) = \sum_{k=0}^{N-1} W(k)x(n-k) \quad (3.1)$$

dove \mathbf{W} è il vettore dei coefficienti del filtro ed è controllato dall'algoritmo LMS standard:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \mu e(n)\mathbf{X}(n) \quad (3.2)$$

dove μ rappresenta il coefficiente che determina la grandezza del passo di avanzamento dell'algoritmo LMS (*step-size*) e dove e rappresenta il segnale "errore" dato dalla differenza tra il segnale r (che contiene la somma della voce più il suono di sirena modificato dal sistema-ambiente in cui è immerso) ritardato di L istanti, e il segnale y in uscita al filtro FIR:

$$e(n) = r(n - L) - y(n). \quad (3.3)$$

Il sistema, una volta a regime, fa sì che il segnale e contenga quasi esclusivamente il segnale voce. Questo perché nella differenza espressa dall'equazione (3) si andrà a sottrarre dal segnale r il segnale di sirena filtrato in modo tale da emulare la risposta dell'ambiente presente tra l'uscita del suono di sirena e il microfono. Per quanto riportato dai dati relativi ai test sperimentali effettuati che sono mostrati nella Tabella 3.1, si potrebbe sostenere che il sistema funzioni correttamente. Infatti, in tutti i casi riportati, l'attenuazione media delle sirene è sempre maggiore di 20 dB. Purtroppo, però, a detta degli stessi autori, il tempo necessario al sistema per convergere non è comparabile con la velocità di transizione dei tre tipi di sirena presi in considerazione. Per questo motivo, quindi, nonostante senza cambi di tono il metodo in questione si comporti egregiamente, quando questo cambio repentino si verifica - e ciò avviene in tutte le tipologie di sirena prese in considerazione -, si manifestano in uscita sgradevoli *click* in grado di corrompere in maniera non tollerabile le comunicazioni radio.

Inoltre, considerare aperiodico un segnale come quello prodotto dalla voce umana risulta essere una scelta poco esatta. Esso possiede infatti numerose manifestazioni di periodicità e queste, quando per similarità di volume e di ampiezza, vanno a incrociarsi con la periodicità tipica del suono di sirena, riescono a mandare in fallo il sistema.

Una leggera modifica al metodo esposto nel paragrafo precedente è stata proposta dallo stesso Sherratt R.S. [6], qualche anno dopo. Il sistema LMS è stato sostituito

Tabella 3.1: Dati sperimentali ricavati applicando il metodo esposto nel Paragrafo 3.1.2 a comunicazioni radio disturbate da tipi diversi di sirena.

Tipo di Sirena	Attenuazione Media
Sirena tipo Bitonale (<i>Hi-lo</i>) generata elettropneumaticamente	29 dB
Sirena tipo bitonale (<i>Hi-lo</i>) generata elettronicamente	34 dB
Sirena tipo <i>Wail</i> generata elettronicamente	29 dB
Sirena tipo <i>Yelp</i> generata elettronicamente	21 dB

con la sua versione normalizzata NLMS (*Normalized Least Mean Square*), ovvero l'equazione (3.2) è stata modificata in:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \frac{\mu e(n)\mathbf{X}(n)}{[\mathbf{X}^T(n)\mathbf{X}(n)]} \quad (3.4)$$

L'algoritmo NLMS ha il vantaggio di adattarsi più velocemente rispetto all'algoritmo LMS anche se ciò va a discapito di una maggiore capacità di elaborazione richiesta per il calcolo aggiuntivo di $\mathbf{X}^T(n)\mathbf{X}(n)$ e per l'esecuzione di una divisione ad ogni passaggio. Come mostrato in Tabella 3.2, con l'utilizzo di questo algoritmo, l'attenuazione media del suono di sirena, aumenta significativamente e, per quanto riportato dall'autore, anche gli effetti indesiderati riscontrati in precedenza con l'LMS, si riducono notevolmente.

3.1.3 Cancellazione con ritardo, filtro *anti-speech* e filtro adattativo NLMS a doppio canale

Ding H., Lu J., Qiu X., e Xu B. [7] propongono un metodo simile a quello appena trattato ma con una significativa variazione: essi applicano un complesso filtro elimina-voce (*anti-speech*) al segnale usato come riferimento. Ciò consente di acquisire tale segnale non necessariamente all'uscita diretta del generatore elettrico della sirena ma da posizioni differenti come davanti al diffusore della sirena oppure in diverse posizioni nell'abitacolo, lontano dalla diretta esposizione alla voce.

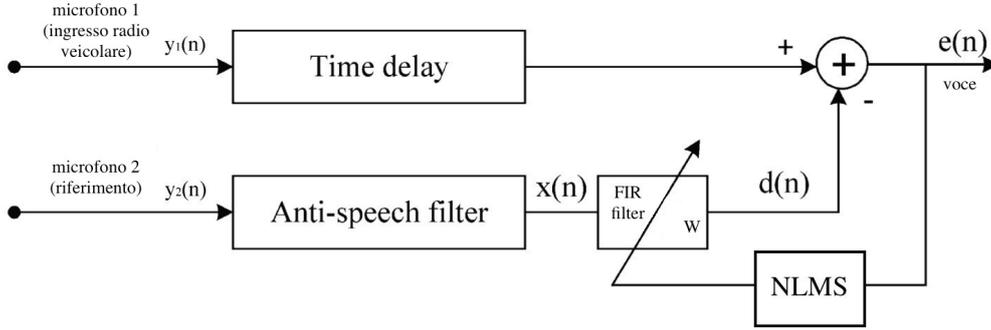
Lo schema a blocchi del sistema in questione è illustrato in Figura 3.6. Si noti come anche in questo caso venga utilizzato come algoritmo adattativo l'NLMS, ovvero la versione normalizzata del *Least Mean Squares Algorithm*.

Il filtro *anti-speech* è realizzato utilizzando un algoritmo di tipo Yule-Walker [8]. Esso adopera un processo di minimizzazione dei quadrati (*least-squares*) per ottenere i coefficienti B e A e l'ordine K , nella funzione di trasferimento del filtro,

Tabella 3.2: Dati sperimentali ottenuti applicando il metodo esposto nel Paragrafo 3.1.2 modificato (algoritmo LMS sostituito con algoritmo NLMS) a comunicazioni radio disturbate da diversi tipi di sirena.

Tipo di Sirena	Attenuazione Media
Sirena tipo Bitonale (<i>Hi-lo</i>) generata elettropneumaticamente	37 dB
Sirena tipo bitonale (<i>Hi-lo</i>) generata elettronicamente	42 dB
Sirena tipo <i>Wail</i> generata elettronicamente	37 dB
Sirena tipo <i>Yelp</i> generata elettronicamente	32 dB

Figura 3.6: Diagramma a blocchi del filtro adattativo NLMS con microfono di riferimento, *filtro anti-speech* e ritardo temporale.



mostrata nell'equazione (3.5), così che il medesimo risulti compatibile con la risposta in frequenza dell'ampiezza desiderata.

$$\frac{B(z)}{A(z)} = \frac{b(1) - b(2)z^{-1} + \dots + b(K)z^{-(K-1)}}{1 + a(1)z^{-1} + \dots + a(K)z^{-K}} \quad (3.5)$$

I coefficienti del denominatore, parametri auto-regressivi AR (*Auto-Regressive*), sono calcolati tramite le equazioni di Yule-Walker Modificate (MYW). Supponiamo che $\{y(n)\}$ sia un processo scalare di tipo ARMA (*Auto-Regressive Moving Average*) di ordine K , definito dalla seguente equazione alle differenze:

$$y(n) = -\sum_{i=1}^k a(i)y(n-i) + \sum_{i=1}^k a(i)w(n-i) \quad (3.6)$$

dove $w(n)$ è un rumore bianco a media nulla e a varianza unitaria. Moltiplicando entrambi i membri dell'equazione (3.5) per $y(n-k-l)$ e calcolandone il valore atteso, si ottiene la seguente equazione:

$$\rho(k+l) + a(1)\rho(k+l-1) + \dots + a(k)\rho(l), \quad l \geq 1 \quad (3.7)$$

dove $\rho(i)$ sono i coefficienti di auto-correlazione del processo $y(n)$:

$$\rho(i) = E\{y(n)y(n-i)\}, \quad \rho(-i) = \rho(i). \quad (3.8)$$

La risoluzione delle equazioni lineari espresse nella (3.6) consente di ottenere i coefficienti AR del denominatore.

Per ricavare, invece, i coefficienti a media mobile (*Moving Average*) del denominatore occorre attuare una procedura a quattro passi. Prima di tutto si calcola un polinomio del numeratore corrispondente alla decomposizione additiva della risposta in frequenza della potenza. Successivamente si valuta la completa risposta

in frequenza corrispondente ai polinomi del numeratore e del denominatore. Poi, attraverso una tecnica di fattorizzazione spettrale, si ottiene la risposta all'impulso del filtro. Infine si ricava il polinomio del numeratore attraverso la minimizzazione dei quadrati (*least squares*) dell'ampiezza di quest'ultima risposta all'impulso.

La Figura 3.7 e la Figura 3.8 riportano la risposta in frequenza dell'ampiezza di un segnale contenente rispettivamente il suono della voce sommato a quello di una sirena di tipo *Yelp* e il suono della voce sommato a quello di una sirena di tipo bitonale (*Hi-lo*), prima e dopo il filtraggio con filtro IIR *anti-speech* Yule-Walker.

Si noti come, secondo quanto riportato dagli autori e mostrato in queste due figure, e con più precisione dallo spettrogramma proposto dalla Figura 3.9, il grosso del segnale vocale viene cancellato dal filtro mentre il suono di sirena resta pressoché intatto.

Nel sistema qui proposto, il filtro *anti-speech* è applicato al segnale trasdotto dal microfono usato come riferimento. Il luogo di posizionamento di quest'ultimo risulta quindi essere più flessibile.

Come mostrato in Figura 3.10, i segnali $y_1(n)$ e $y_2(n)$ sono i segnali acquisiti rispettivamente dal Microfono 1 (della radio veicolare) e del Microfono 2 (di riferimento). Assumendo che la bocca dell'operatore sia molto più vicina al Microfono 1 che al Microfono 2, il segnale trasdotto dai due microfoni può essere descritto rispettivamente come:

$$y_1(n) = s(n) + \text{sirena}(n) * h_1(n) \quad (3.9)$$

Figura 3.7: Risposta in frequenza dell'ampiezza di un segnale contenente la somma della voce e del suono di sirena di tipo *Yelp*, prima e dopo il filtraggio con filtri *anti-speech* di Yule-Walker.

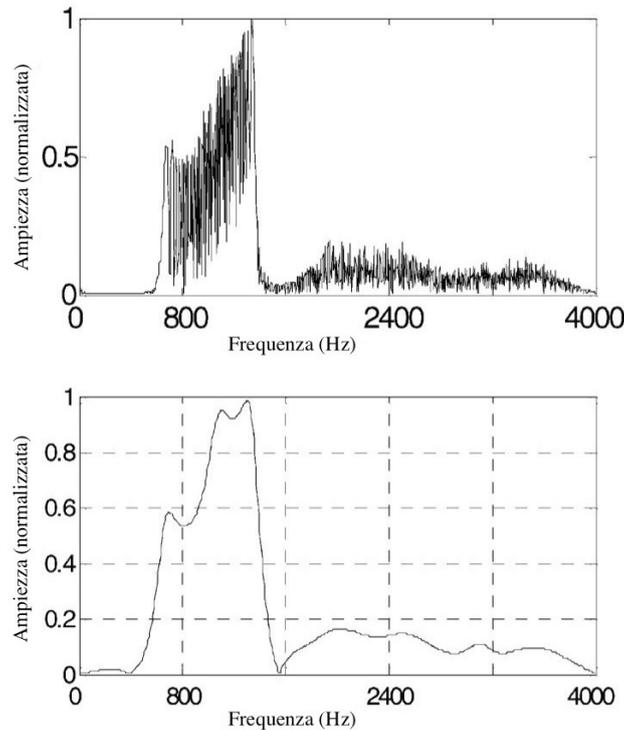


Figura 3.8: Risposta in frequenza dell'ampiezza di un segnale contenente la somma della voce e del suono di sirena di tipo *Hi-lo*, prima e dopo il filtraggio con filtro *anti-speech* Yule-Walker.

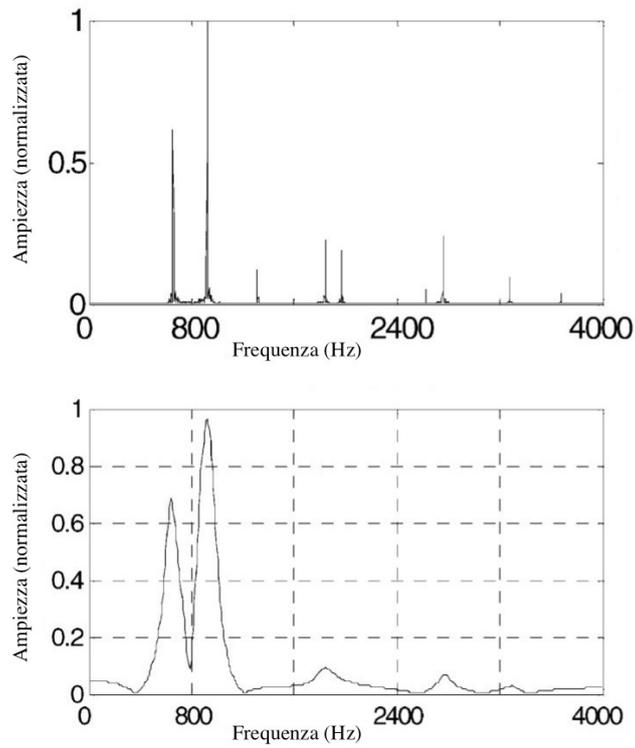
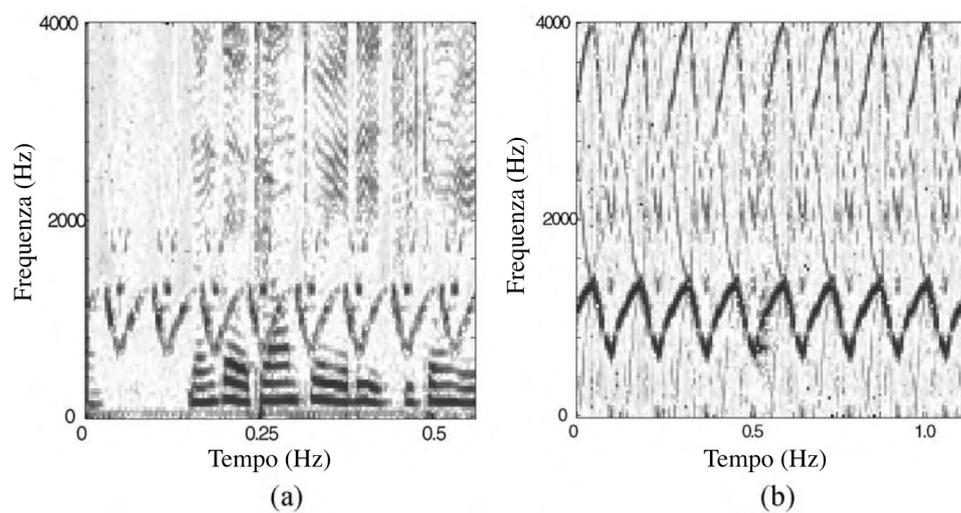


Figura 3.9: Spettrogramma di un segnale contenente la somma della voce e del suono di sirena di tipo *Yelp*, prima (a) e dopo (b) il filtraggio con filtro *anti-speech* Yule-Walker.



e

$$y_2(n) = s(n) * h_{12} + \text{sirena}(n) * h_2(n) \quad (3.10)$$

dove h_1 e h_2 sono le risposte all'impulso del sistema-ambiente presente tra il diffusore della sirena e i due rispettivi microfoni e dove h_{12} rappresenta la risposta all'impulso del sistema-ambiente presente tra la bocca dell'operatore e il Microfono 2.

Dopo un ritardo temporale di Ns campioni, il segnale $y_1(n)$ proveniente dal Microfono 1 (vedere Figura 3.6), diventa:

$$y_1'(n) = y_1(n + Ns). \quad (3.11)$$

Poiché, attraverso la cancellazione del suono della voce operata dal filtro *anti-speech*, il segnale prodotto dal Microfono 2 filtrato presenta quasi esclusivamente il suono di sirena, il segnale errore $e(n)$ può essere stimato attraverso l'equazione:

$$e(n) = y_1'(n) - \mathbf{W}^T(n)\mathbf{X}(n) \quad (3.12)$$

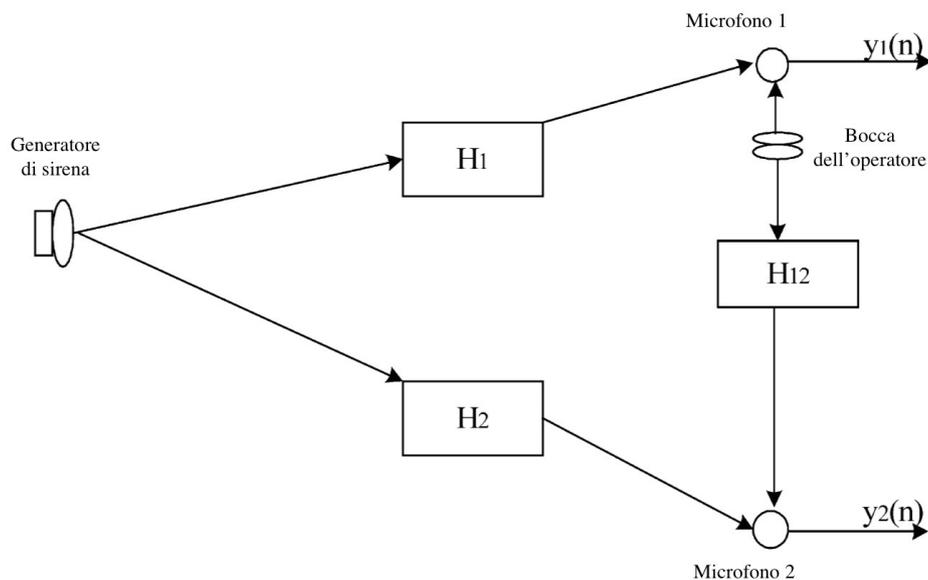
dove \mathbf{W}^T rappresenta il vettore dei coefficienti del filtro FIR:

$$\mathbf{W}^T = [w_1(n), w_2(n), \dots, w_N(n)], \quad (3.13)$$

e

$$\mathbf{X}^T = [x(n), x(n-1), \dots, x(n-N+1)]. \quad (3.14)$$

Figura 3.10: Il sistema di acquisizione dei segnali composto da due microfoni.



Ad ogni istante, l'algoritmo NLMS aggiusta il vettore $\mathbf{W}(n)$ secondo la seguente relazione:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \frac{\mu e(n)\mathbf{X}(n)}{[\mathbf{X}^T(n)\mathbf{X}(n)]} \quad (3.15)$$

dove μ è un fattore di convergenza costante che regola il passo di avanzamento (*step-size*).

Il Rapporto Segnale-Rumore (*Signal to Noise Ratio*) che si ha rispettivamente in ingresso al sistema e in uscita dallo stesso si esprime come:

$$SNR_{in} = 10 \log_{10} \left| \frac{\sum_{n=1}^N s(n)^2}{\sum_{n=1}^N [y(n) - s(n)]^2} \right| \quad e \quad SNR_{out} = 10 \log_{10} \left| \frac{\sum_{n=1}^N e'(n)^2}{\sum_{n=1}^N [e'(n) - s(n)]^2} \right|. \quad (3.16)$$

Nello studio, gli autori riportano di aver ottenuto un positivo incremento del Rapporto Segnale-Rumore, ovvero della differenza tra il suo valore in uscita e il suo valore in ingresso.

Nella Tabella 3.3 trascriviamo i dati ricavati dalle sperimentazioni da loro effettuate applicando il sistema ai due segnali provenienti dai due microfoni posizionati come illustrato in Figura 3.11 e campionati con frequenza di campionamento 8 kHz.

Per la sperimentazione è stato assegnato al ritardo temporale, applicato al segnale proveniente dal microfono della radio veicolare (Mic.1 in Figura 3.11), un valore di 64 campioni.

Se ci limitiamo a osservare i risultati conseguiti presentati nella Tabella 3.3 il metodo sembrerebbe essere valido, dato che i valori esposti sembrano essere soddisfacenti. Anche in questo caso, però, sussistono importanti inconvenienti:

Tabella 3.3: Incremento del Rapporto Segnale-Rumore ottenuto applicando il sistema di cancellazione della sirena con ritardo temporale, filtro *anti-speech* e filtro adattativo NLMS a doppio canale, posizionando il microfono di riferimento nelle tre posizioni diverse mostrate in Figura 3.11.

Incremento del Rapporto Segnale-Rumore	Microfono di riferimento in Posizione 1	Microfono di riferimento in Posizione 2	Microfono di riferimento in Posizione 3
Sirena tipo Yelp	28.56 dB	28.88 dB	29.99 dB
Sirena tipo Hi-Lo	27.56 dB	27.63 dB	27.63 dB
Sirena tipo Wail	31.12 dB	31.34 dB	31.56 dB

- per ogni tipo di sirena che si vuol cancellare occorre costruire un differente filtro *anti-speech*;
- nel caso di sirene di tipo bitonale (*Hi-lo*), nella fase di transito dal tono basso al tono alto e viceversa, per il periodo - di durata non irrilevante - necessario all'algoritmo NLMS per convergere, il suono di sirena non viene cancellato ed emergono, inoltre, fastidiosi *click* che deteriorano notevolmente la comprensibilità del parlato.

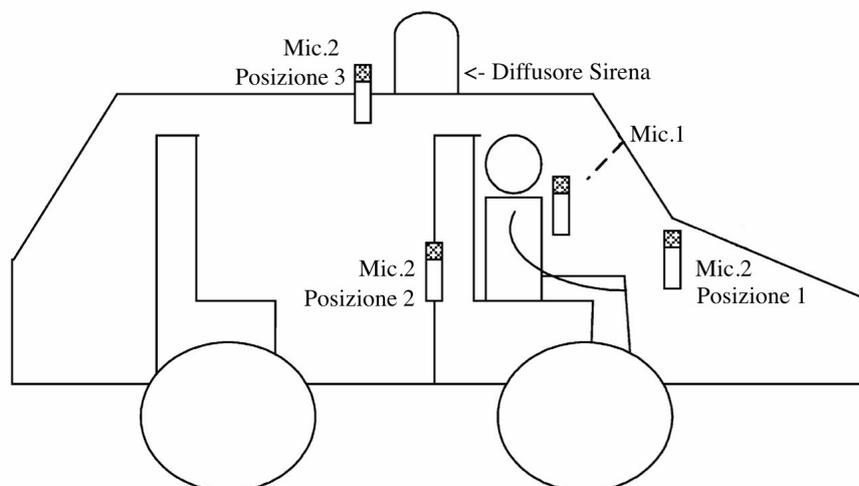
Nonostante, quindi, alcuni passi in avanti fatti rispetto al procedimento proposto nel Paragrafo 2.1.2, anche questa metodologia risulta lontana dall'essere definitiva e da potersi applicare per l'effettiva cancellazione del suono di sirena dalle comunicazioni radio. Discorso valido, soprattutto, per quanto riguarda le sirene di tipo bitonale come quelle in dotazione al Corpo Nazionale dei Vigili del Fuoco.

3.2 Cancellazione senza l'utilizzo del suono di sirena come segnale di riferimento

3.2.1 Cancellazione con ritardo temporale e filtro adattativo LMS a singolo canale

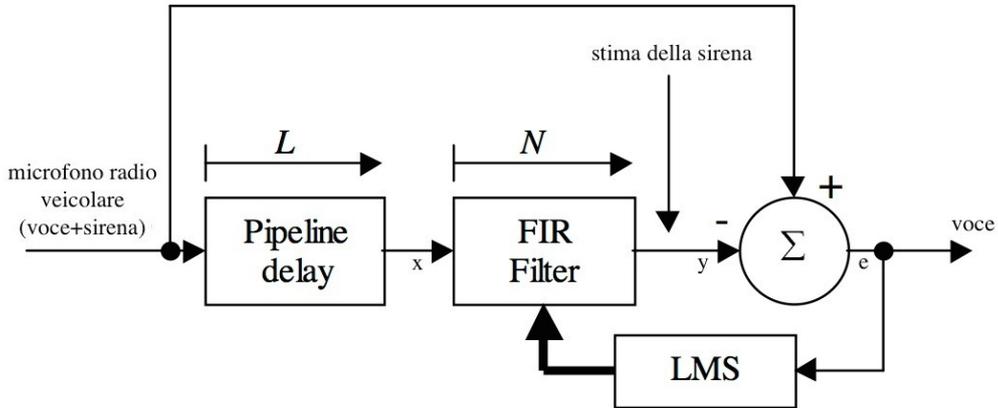
Oltre al sistema di cui abbiamo parlato in precedenza³, Sherratt R.S., Townsend D.M. e Guy C.G. hanno proposto una versione simile che, però, per il suo funzionamento, utilizza esclusivamente il segnale proveniente dal solo microfono della radio veicolare [5]. Come abbiamo detto più volte, questo segnale contiene la

Figura 3.11: Posizionamento dei microfoni utilizzato per la sperimentazione. (Mic.1) rappresenta il microfono della radio veicolare, (Mic.2.) quello di riferimento.



³ vedi paragrafo 3.1.2

Figura 3.12: Diagramma a blocchi del filtro adattativo LMS a singolo canale che utilizza come segnale di riferimento lo stesso segnale in ingresso ritardato nel tempo.



somma del suono emesso dalla voce dell'operatore che effettua le comunicazioni radio e del segnale di sirena che entra all'interno dell'abitacolo e che costituisce la fonte di disturbo che vogliamo cancellare.

La versione modificata del sistema è rappresentata nello schema a blocchi di Figura 3.12. Come si vede da questa rappresentazione i blocchi che compongono il sistema sono gli stessi già descritti nel Paragrafo 3.1.2, ovvero: un ritardatore, un filtro FIR, un blocco di calcolo dell'algoritmo LMS e un sommatore.

Il Predittore Lineare (*Linear Predictor*), istruito dall'algoritmo *Least Mean Square* [9], non richiede necessariamente l'ausilio di un segnale aggiuntivo ma utilizza, come suo riferimento, una versione ritardata dello stesso segnale principale in ingresso.

Anche in questo caso, il filtro adottato nel procedimento, è un filtro FIR a N elementi e il valore del ritardo temporale, applicato con l'obiettivo di ottenere due segnali correlati nella loro componente sirena e decorrelati nella loro componente voce, è fissato a ad una lunghezza di L campioni. Il segnale y in uscita al filtro rappresenta la stima effettuata dal sistema del suono di sirena catturato dal microfono ed è descritta dalle stesse relazioni valide nel metodo descritto nel Paragrafo 3.1.2 - equazione (1), equazione (3.2) -. Chiamando *mic* il segnale rappresentato dalla seguente equazione:

$$mic(n) = x(n - N - 1 - L), \quad (3.17)$$

dove x rappresenta il segnale trasdotto dal microfono e n il valore dell'istante temporale corrente, una volta ottenuta la convergenza dell'algoritmo LMS che regola i coefficienti del filtro, possiamo considerare l'uscita errore (e) dal sommatore come stima del segnale *voce*, privo del suono di sirena:

$$voce(n) = e(n) = mic(n) - y(n). \quad (3.18)$$

Secondo i dati sperimentali riportati dagli stessi autori, il sistema, così realizzato, non riesce nell'intento sperato. Anche nella configurazione che, per i test eseguiti, risulta la migliore, ovvero con N settato con un valore di 100 campioni, L con un valore di 50 campioni e il coefficiente μ dell'algoritmo LMS - equazione (3.2) - con un valore di 0.01, il procedimento soffre di tutte le stesse problematiche riportate nel Paragrafo 3.1.2 che, addirittura, in questa circostanza si aggravano: esso non riesce a ridurre di valori superiori a 6 dB nessun suono di sirena in nessuna condizione. Nel caso specifico della sirena bitonale, poi, il sistema non sortisce alcun tipo di cancellazione o attenuazione.

3.2.2 Cancellazione con ritardo temporale, filtro anti-speech e filtro adattativo NLMS a singolo canale

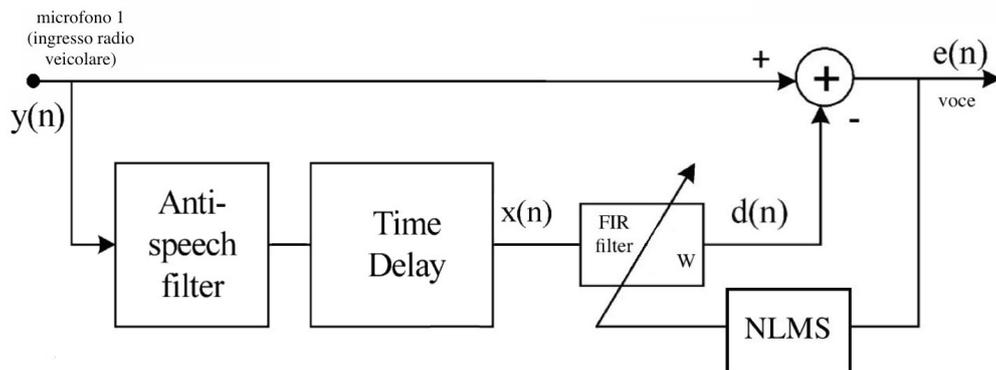
Ding H., Lu J., Qiu X., e Xu B., oltre al metodo già esposto nel Paragrafo 3.1.3, hanno proposto anche una versione dello stesso che non necessita di un segnale ausiliario di riferimento ma che lavora a singolo canale. Per apportare la cancellazione, esso utilizza in ingresso esclusivamente il segnale proveniente dal microfono della radio veicolare. Lo schema a blocchi di questo sistema è rappresentato in Figura 3.13. Dall'immagine si vede che, come nel caso a doppio canale, gli elementi che concorrono al funzionamento del processo di cancellazione sono un filtro *anti-speech*, un ritardo temporale (*time delay*) e un blocco di attuazione dell'algoritmo *Normalized Least Mean Squares*.

Assumendo che la bocca dell'operatore che sta effettuando la conversazione via radio sia molto vicina al microfono, il segnale in ingresso può essere espresso come:

$$y(n) = s(n) + sirena(n) * h(n), \tag{3.19}$$

dove $y(n)$ è il segnale captato dal microfono, $s(n)$ è il suono emesso dalla voce dell'operatore e $h(n)$ è la risposta all'impulso acustica della ambiente presente tra il diffusore della sirena e il microfono. Se la lunghezza del ritardo è N_s , il segnale usato come riferimento dall'algoritmo adattativo $x(n)$ può essere descritto come:

Figura 3.13: Diagramma a blocchi del filtro adattativo NLMS che utilizza *filtro anti-speech* e ritardo temporale.



$$x(n) = y(n + N_s). \quad (3.20)$$

Quando il sistema è in funzione, il segnale errore $e(n)$, con indice temporale n , conterrà il segnale vocale estratto con il suono di sirena residuo e può essere espresso come:

$$e(n) = y(n) - \mathbf{W}^T(n)\mathbf{X}(n), \quad (3.21)$$

dove \mathbf{W}^T è il vettore dei coefficienti del filtro FIR espresso come:

$$\mathbf{W}^T = [w_1(n), w_2(n), \dots, w_N(n)], \quad (3.22)$$

e

$$\mathbf{X}^T = [x(n), x(n-1), \dots, x(n-N+1)]. \quad (3.23)$$

Il vettore dei coefficienti \mathbf{W}^T è costantemente aggiornato dall'algoritmo NLMS dato da:

$$\mathbf{W}(n+1) = \mathbf{W}(n) + \frac{\mu e(n)\mathbf{X}(n)}{[\mathbf{X}^T(n)\mathbf{X}(n)]} \quad (3.24)$$

dove μ è la costante che regola il passo di adattamento dell'algoritmo NLMS.

Il ritardo temporale è scelto sufficientemente grande in modo che la componente di segnale vocale che, nonostante la cancellazione attuata del filtro *anti-speech*⁴ è ancora presente nel canale di riferimento diventi incorrelata rispetto a quella nel canale principale.

Per valutare le prestazioni di diversi sistemi a canale singolo che si avvalgono di questo metodo, sono stati considerati assieme il livello in uscita del SNR⁵ e la sensazione acustica soggettiva dell'ascoltatore.

Nella Tabella 3.3 si possono osservare i valori ottenuti applicando a una conversazione radio disturbata da una sirena di tipo *Yelp*, in ordine: un sistema privo del filtro *anti-speech* ma con blocco di ritardo temporale, un sistema con filtro *anti-speech* ma senza il blocco di ritardo temporale e, infine, un sistema con entrambe le componenti. In questo esperimento la bocca dell'operatore è posta molto vicina al microfono e pertanto l'influenza di altri rumori di sottofondo può essere ignorata. Essendo molto alto il volume di sirena all'interno della cabina, il rapporto segnale-rumore risulta essere minore di -20 dB e, a questo livello, è impossibile, per un ascoltatore, capire bene il significato delle parole comunicate. Si può notare come il

⁴ filtro dettagliatamente descritto nel Paragrafo 3.1.3

⁵ formalmente definito nella (16)

Tabella 3.4: Confronto tra le prestazioni di sistemi a canale singolo con algoritmi differenti applicati ad una conversazione radio disturbata da un suono di sirena di tipo *Yelp*.

Rapporto Segnale-Rumore	Sirena tipo <i>Yelp</i>	Sirena bitonale (<i>Hi-Lo</i>)	Sirena tipo <i>Wail</i>
SNR_{in}	-22.62 dB	-26.21 dB	-25.94 dB
SNR_{out}	4.41 dB	-0.42 dB	0.12 dB
Differenza SNR_{out}-SNR_{in}	27.03 dB	25.78 dB	26.05 dB

livello di attenuazione ottenuto con l'utilizzo del sistema senza filtro *anti-speech* sia piuttosto limitato e il SNR in uscita risulti ancora inferiore a 0 dB. In questa condizione, le parole comunicate via radio sono incomprensibili per la presenza di importanti interferenze di sottofondo. Introducendo il filtro *anti-speech*, specialmente in combinazione con il metodo del ritardo temporale, le prestazioni del sistema di cancellazione del rumore a singolo canale sono fortemente migliorate. I valori del SNR in uscita risultano essere maggiori di 0 dB e i livelli di comprensione del parlato sono accettabili, anche se disturbati, per il sistema privo di ritardo temporale e, addirittura, buoni per quello che implementa il ritardo temporale.

La Tabella 3.5 mostra i dati ottenuti applicando il sistema a singolo canale completo di entrambi i blocchi (*anti-speech* e ritardo temporale) ai tre tipi diversi di sirena. Per le sirene di tipo *Yelp* e *Wail* i risultati si possono ritenere soddisfacenti mentre per la sirena di tipo bitonale (*Hi-lo*) il valore di SNR in uscita minore di 0

Tabella 3.5: Risultati ottenuti applicando il sistema adattativo a singolo canale con filtro *anti-speech* e ritardo temporale a più tipi di sirena.

Rapporto Segnale Rumore	Senza <i>Anti-Speech</i> con Ritardo Temporale di 64 campioni	Con <i>Anti-Speech</i> Senza Ritardo Temporale	Con <i>Anti-Speech</i> e Ritardo Temporale di 64 campioni
SNR_{in}	-22.62 dB	-22.62 dB	-22.62 dB
SNR_{out}	-2.31 dB	1.27 dB	4.41 dB
Differenza SNR_{out}-SNR_{in}	20.32 dB	23.89 dB	27.03 dB
Valutazione qualitativa soggettiva	Interferenze molto forti	Interferenze presenti in sottofondo	Buona comprensione del parlato

dB indica che per questo tipo di segnalatore acustico, il sistema risulta ancora essere fallace e la cancellazione apportata non sufficiente per una buona comprensione del parlato.

Capitolo 4

Metodo 1: con utilizzo dell'*Average Magnitude Difference Function*

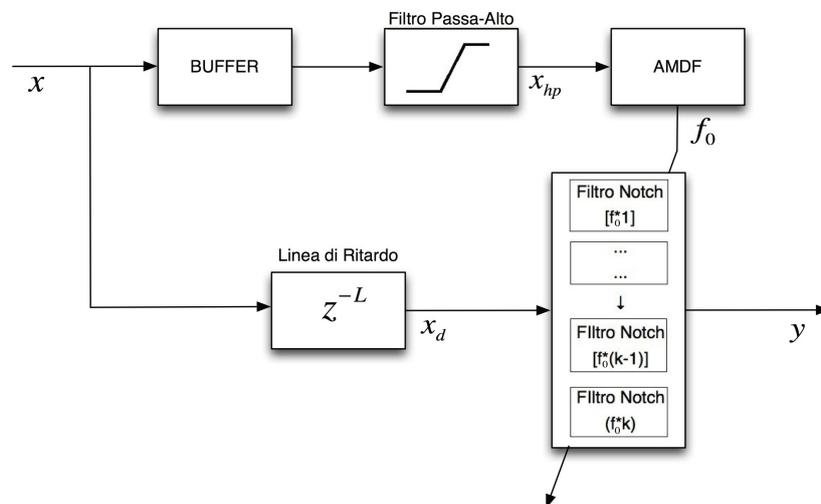
In questo capitolo verrà presentato il primo dei due metodi qui proposti per la cancellazione del suono di sirena. Inizieremo fornendo una visione d'insieme attraverso la rappresentazione e la descrizione dello schema a blocchi che raffigura l'interno sistema. Successivamente ci soffermeremo sull'analisi delle tecniche che caratterizzano i singoli blocchi. Infine riprenderemo il processo nella sua interezza per mostrarne il funzionamento applicato a un segnale vocale disturbato dalla presenza del suono della sirena di tipo bitonale installata sui mezzi del Corpo Nazionale dei Vigili del Fuoco.

4.1 Panoramica d'insieme del sistema proposto

Nella Figura 4.1 è rappresentato lo schema a blocchi che descrive il metodo che proponiamo in questo capitolo. I blocchi che lo compongono sono:

- un *buffer*, ovvero un registro intermedio, che memorizza il segnale in ingresso e dà in uscita il segnale diviso in *frame* di dimensione prestabilita ;

Figura 4.1: Schema a blocchi del sistema per la cancellazione della sirena che utilizza l'AMDF e un banco di filtri *notch*.



- un filtro passa-alto;
- una linea di ritardo
- un blocco che applica l'*Average Magnitude Difference Function* (AMDF) al segnale in ingresso e che invia in uscita la frequenza fondamentale (*pitch*) che caratterizza lo stesso;
- un banco di k filtri *notch*;

Nel sistema, il segnale digitale in ingresso, proveniente da precedente campionatura del segnale trasdotto dal microfono della radio veicolare (denominato x), si incanala su due percorsi differenti.

Nel primo di questi, esso viene memorizzato in un *buffer*⁶ e, a blocchi di lunghezza predefinita (*frame*⁷), passa attraverso un filtro passa-alto⁸. Una volta privato delle sue componenti in bassa frequenza (x_{hp}), si trasferisce al blocco dove viene eseguita l'AMDF⁹. Questa funzione dà in uscita il valore della frequenza fondamentale (*pitch*) che descrive il tono del segnale presente nel *frame* in ingresso e lo invia al banco di filtri *notch*.

Il secondo percorso inizia ritardando il segnale in ingresso di un tempo L^{10} , in modo da compensare il ritardo che si crea nelle elaborazioni presenti sul percorso parallelo. Successivamente il segnale ritardato (x_d) viene filtrato da una serie di k filtri *notch*¹¹. Questi sono filtri elimina-banda, a banda molto stretta, e sono impostati con frequenza di centro-banda pari alla frequenza (f_0) in uscita dal blocco AMDF e alle relative k armoniche [f_0*2 ; ... $f_0*(k-1)$; f_0*k]. Una volta filtrato dalla serie di filtri *notch*, il segnale ottenuto (y) è inviato in uscita e trasmesso via radio.

4.2 Estrazione del *pitch* tramite *Average Magnitude Difference Function*

⁶ la necessità di inserire uno stadio di *buffering* e quale sia il suo funzionamento sono spiegati nel Paragrafo 4.5.1

⁷ il giusto dimensionamento dei *frame* è ottenuto dalle verifiche sperimentali esposte nel Capitolo 6

⁸ la struttura e la frequenza di taglio del filtro passa-alto sono oggetto del Paragrafo 4.4

⁹ l'*Average Magnitude Difference Function* è trattata nel Paragrafo 4.2

¹⁰ la dimensione del ritardo viene calcolata nel Paragrafo 4.6

¹¹ una dettagliata analisi dei filtri *notch* è affrontata nel Paragrafo 4.3

L'AMDF è una funzione che permette di estrarre informazioni da un segnale periodico o che contenga delle periodicità. Nel nostro caso specifico la funzione serve a estrarre, a intervalli regolari, l'altezza (*pitch*) che assume il suono di sirena presente nel segnale che vogliamo filtrare. Essa trae origine dalla Funzione di Autocorrelazione (ACF, *Auto-Correlation Function*). Per questo, in primo luogo, ci soffermeremo ad analizzare quest'ultima.

4.2.1 La Funzione di Autocorrelazione

Per definizione l'ACF di un segnale è:

$$\phi(k) = \sum_{m=-\infty}^{\infty} s(m)s(m+k). \quad (4.1)$$

Essa conserva le informazioni che riguardano le frequenze fondamentali del segnale e delle relative armoniche. Dall'equazione (4.1) si vede che, in un certo modo, $\phi(k)$ misura la somiglianza del segnale con la sua versione traslata. Questa avrà quindi valori più grandi in corrispondenza dei ritardi k per cui $s(m)$ e $s(m+k)$ hanno forme d'onda simili. Oltre a queste, di seguito, altre sue importanti caratteristiche:

- è una funzione pari, ovvero: $\phi(k) = \phi(-k)$;
- assume il suo valore massimo per $k=0$;
- quando $k=0$, $\phi(0)$ corrisponde al valore dell'energia del segnale e, nel caso di segnali periodici non deterministici, anche al valore della sua potenza media;
- se il segnale è periodico con periodo P , anche l'autocorrelazione risulta periodica con lo stesso periodo, ovvero $\phi(k) = \phi(k+P)$.

Applicando una finestra temporale $w(n)$ alla funzione espressa nella (4.1) otteniamo la (*Short-Time*) *Autocorrelation Function* (ACF). E, modificando per convenienza la variabile m nella sua espressione come $m = n + m$, otteniamo:

$$R_n(k) = \sum_{m=-\infty}^{\infty} [s(n+m)w(-n)][s(n+m+k)w(-k-m)]. \quad (4.2)$$

Ponendo $w(-n) = w'(n)$, abbiamo, infine:

$$R_n(k) = \sum_{m=0}^{N-1-k} [s(n+m)w'(m)][s(n+m+k)w'(k+m)]. \quad (4.3)$$

In Figura 4.2 è mostrato un segnale campionato con frequenza di campionamento 8 kHz di lunghezza 800 campioni che presenta evidenti periodicità e, subito sotto, la sua ACF. Nella Figura 4.3 sono rappresentati, invece, un segnale che rappresenta un rumore privo di periodicità e, di nuovo, la sua relativa ACF. In questi due esempi

Figura 4.2: Frame di 800 campioni di segnale audio avente componenti periodiche e sua relativa (*Short-Time*) Autocorrelation Function.

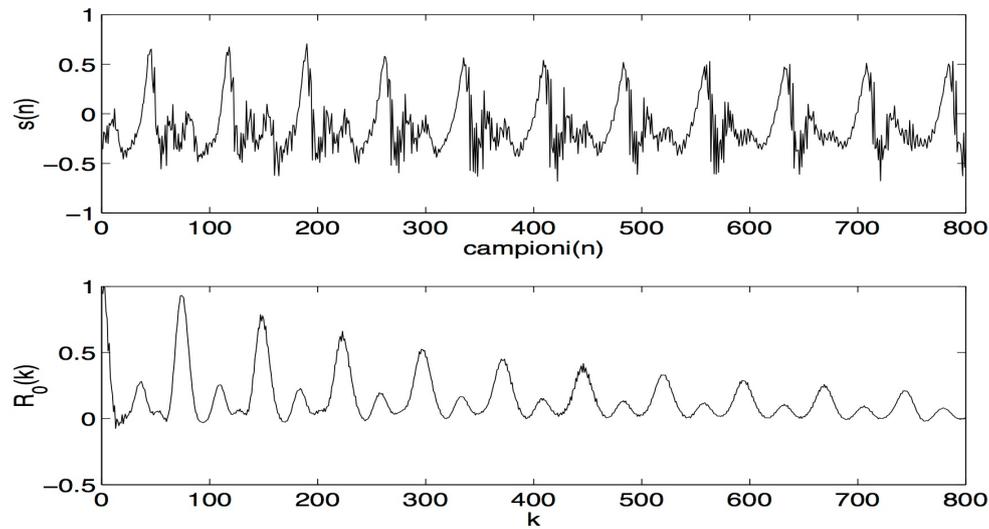
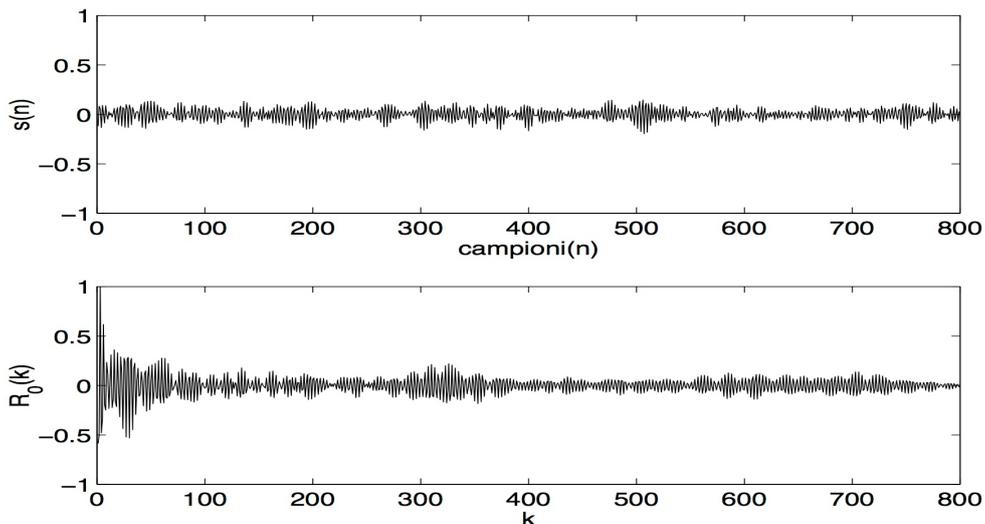


Figura 4.3: Frame di 800 campioni di segnale audio non periodico e sua relativa (*Short-Time*) Autocorrelation Function.



possiamo verificare la validità delle caratteristiche proprie della funzione sopra riportate; in particolar modo, appare visibile come le periodicità presenti nel primo *frame* emergano anche nella sua ACF mentre, nel secondo, anche quest'ultima risulti priva di ripetizioni periodiche di parti di segnale.

4.2.2 La funzione *Average Magnitude Difference Function*

Per un segnale periodico che possiede periodicità P sappiamo che la successione:

$$d(n) = s(n) - s(n - k) \quad (4.4)$$

si annulla per $k = 0, \pm P, \pm 2P, \dots$. Risulta, dunque, sensato utilizzare questa differenza per sviluppare una funzione che ci possa dare informazioni sulle periodicità presenti in un segnale. Se, a differenza di quanto fatto per il calcolo dell'Autocorrelazione nella (4.1), invece di moltiplicare $s(m)$ per $s(m-k)$, procediamo con la sottrazione descritta dalla (4.4) presa nel suo valore assoluto, la (4.2) diventa:

$$\gamma_n(k) = \sum_{m=-\infty}^{\infty} |s(n+m)w(m) - s(n+m-k)w(m-k)| \quad (4.5)$$

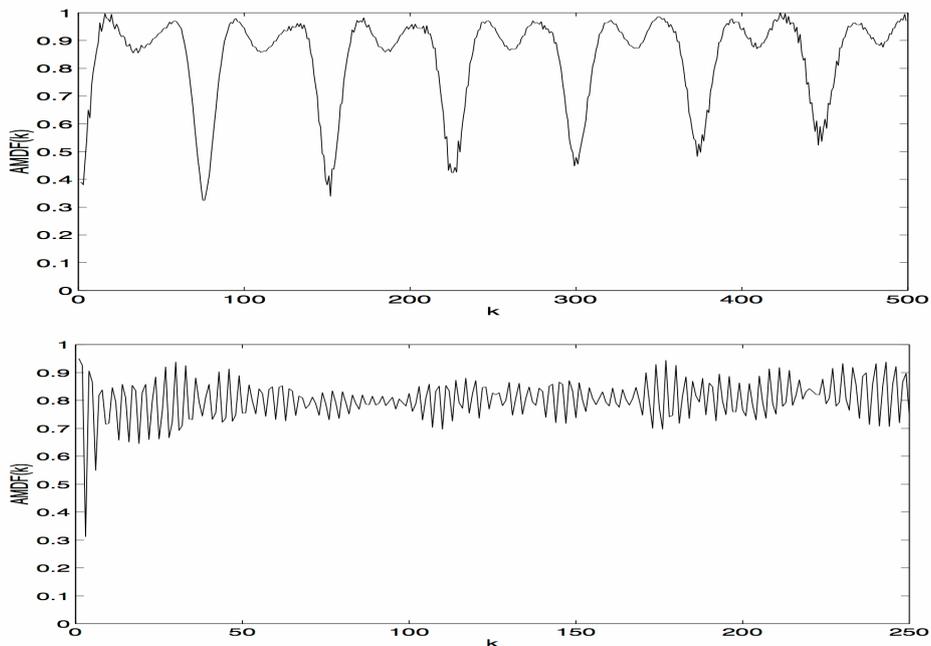
e da questa equazione, prendendo una finestra temporale $w(n)$ di tipo rettangolare e di durata N , ricaviamo la (*Short-Time*) *Average Magnitude Difference Function* [11]:

$$AMDF(k) = \sum_{m=k}^{N-1} |s(m) - s(m-k)|. \quad (4.6)$$

In Figura 4.4 sono riportate le AMDF applicate agli stessi segnale trattati in Figura (4.1) e in Figura (4.2), ovvero a un segnale che contiene delle componenti periodiche e a un segnale totalmente aperiodico. Dal grafico più in alto si evince come, dalla spaziatura presente tra i valori minimi dell'AMDF del segnale con periodicità, si possa ricavare il reciproco della frequenza fondamentale del segnale¹².

Dal segnale sottostante notiamo, invece, come ad un segnale aperiodico sia effettivamente associata un'AMDF aperiodica.

Figura 4.4: *Average Magnitude Difference Function* applicata ad un segnale avente componenti periodiche (sopra) e ad un segnale aperiodico (sotto).



¹² per l'Autocorrelazione la stessa operazione è eseguibile analizzando la spaziatura tra i suoi massimi

4.2.3 Estrazione del *pitch*

Come abbiamo visto nel paragrafo precedente, è possibile misurare la distanza tra i primi due valori minimi che troviamo nella sua AMDF (o tra i primi due valori massimi presenti nella sua ACF) per risalire alla frequenza fondamentale che caratterizza un suono periodico. Indicando questa distanza come k_M , possiamo ricavare la frequenza fondamentale f_0 come:

$$f_0 = \frac{F_s}{k_M} \quad (4.7)$$

dove F_s è la frequenza di campionamento del segnale.

Nell'esempio di Figura 4.4, rileviamo che il primo minimo, dopo il minimo presente in 0, si verifica per $k_M=66$ per cui la frequenza fondamentale risulta essere:

$$f_0 = \frac{8 \text{ kHz}}{66} \approx 120 \text{ Hz}$$

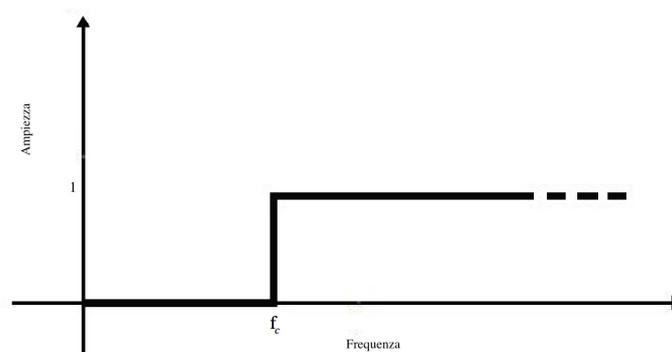
Nel nostro sistema, metteremo in atto questa operazione per poter valutare l'altezza che il suono della sirena bitonale assume coll'evolvere del tempo. L'AMDF, però, non sarà applicata a tutto il segnale che proviene dal microfono della radio veicolare ma sarà calcolata su un segnale opportunamente filtrato con un filtro passa-alto per limitare la presenza del parlato al suo interno. Questo perché, per sua stessa natura, la voce umana presenta delle notevoli caratteristiche di periodicità che potrebbero portare a un'errata valutazione della frequenza fondamentale da parte dell'AMDF.

Il funzionamento e le caratteristiche di questo filtro passa-alto saranno oggetto di studio del prossimo paragrafo.

4.3 Filtro Passa-Alto

Come mostrato in Figura 4.5, un filtro passa-alto ideale è un filtro che ha risposta in frequenza nulla per valori al di sotto di una certa frequenza - detta frequenza di

Figura 4.5: Filtro Passa-Alto ideale.



taglio - e che, per valori maggiori o uguali a questa, ha risposta in frequenza pari a uno. Ovviamente, nella realtà, risulta impossibile ottenere una risposta in frequenza identica a quella ideale. Esistono, però, diverse tecniche che permettono di ottenere risposte molto simili.

Per quanto riguarda il filtro utilizzato nel nostro sistema per poter ridurre la presenza della voce nel suono che deve essere analizzato attraverso l'AMDF, abbiamo utilizzato un filtro numerico con risposta infinita all'impulso IIR sul modello dei filtri di Butterworth.

Il filtri IIR sono rappresentabili nel dominio della frequenza con una funzione di trasferimento razionale in z^{-1} così espressa [12]:

$$H(z) = \frac{\sum_{k=0}^{M-1} b_k z^{-k}}{\sum_{k=0}^{L-1} a_k z^{-k}} \quad (4.8)$$

dove b_k sono i coefficienti che denotano gli *zeri* della funzione e a_k sono i coefficienti che caratterizzano i *poli*.

I filtri di Butterworth sono caratterizzati da una risposta in frequenza dall'ampiezza *massimamente* piatta nella banda passante e da una risposta della fase distorta, anche se non eccessivamente. Per questo motivo, essendo nel caso specifico non interessati alla fase ma interessati a un'alta fedeltà in ampiezza nella banda passante, la nostra scelta è ricaduta su questa tipologia di filtri.

La forma generale del modulo di un filtro di Butterworth è del tipo:

$$\left| B_N \left(\frac{\omega}{\omega_c} \right) \right| = \sqrt{1 + \left(\frac{\omega}{\omega_c} \right)^{2N}} \quad (4.09)$$

dove N rappresenta l'ordine del filtro, ω_c la frequenza di taglio e B_N l' N -esimo polinomio di Butterworth. Di seguito, alcune caratteristiche salienti di questa famiglia di polinomi:

- hanno coefficienti reali;
- hanno tutti gli zeri posti su di una circonferenza di raggio unitario e centrata nell'origine;
- se N è pari, le radici del polinomio sono radici $2N$ -esime di -1 a parte reale negativa;

- se N è dispari, le radici del polinomio sono radici $2N$ -esime di 1 a parte reale negativa.

Nel nostro caso abbiamo ricavato i coefficienti del filtro che vogliamo implementare nel sistema, utilizzando una specifica funzione presente in MATLAB chiamata “fdesign.highpass”. Questa funzione ci permette di ottenere questi coefficienti dando in ingresso le specifiche del filtro secondo quanto mostrato in Figura 4.6: con F_{stop} si indica l’ultima frequenza fino alla quale si estende la banda di stop; con F_{pass} si assegna il valore della frequenza da dove deve iniziare la banda passante; con A_{stop} si definisce il valore in dB dell’attenuazione applicata dal filtro nella banda di stop e, infine, con A_{pass} si configura l’ampiezza massima consentita, in dB, dell’ondulazione (*ripple*) presente in banda passante.

La frequenza di taglio che ci permette di filtrare il segnale “voce+sirena”, in modo tale da ottenere il miglior compromesso tra riduzione della presenza della voce e permanenza di un’ampiezza adeguata delle armoniche della sirena presenti in banda passante, si attesta sul valore di 1.8 kHz¹³. Alla luce di ciò, i valori che abbiamo immesso nella funzione di MATLAB, per la realizzazione del filtro passa-alto, sono quelli mostrati nella Tabella 4.1. La risposta in frequenza dell’ampiezza che abbiamo ottenuto è quella mostrata in Figura 4.7.

Figura 4.6: Specifiche necessarie per la realizzazione di un filtro passa-alto attraverso la funzione di MATLAB “fdesign.highpass”.

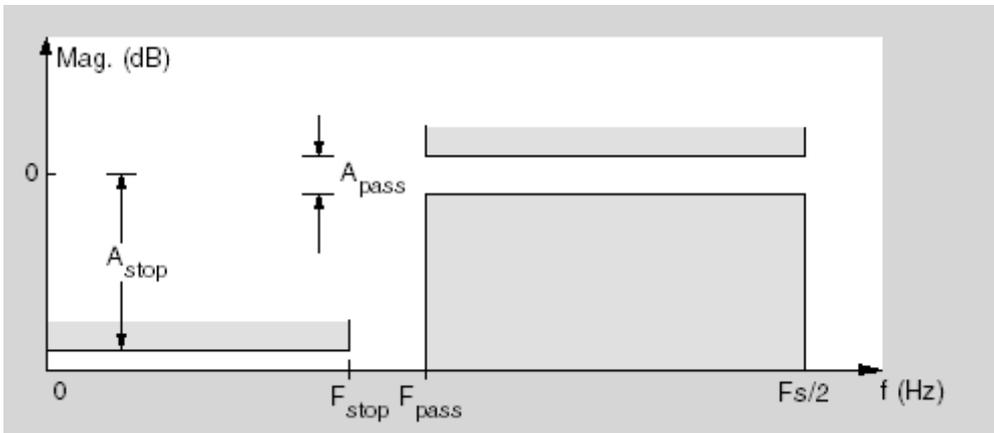
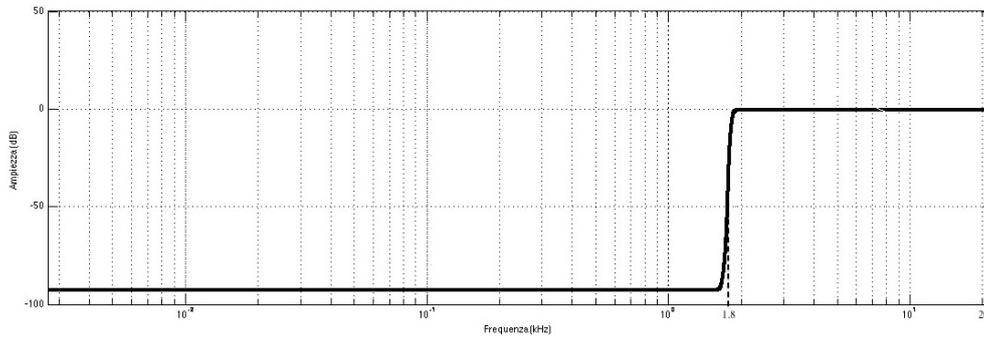


Tabella 4.1: Valori di ingresso alla funzione “fdesign.highpass” per la realizzazione del filtro passa-alto applicato nel sistema di Figura 4.1.

F_{stop}	F_{pass}	A_{stop}	A_{pass}
1.7 kHz	1.9 kHz	80 dB	1 dB

¹³ vedi Capitolo 6

Figura 4.7: Risposta in frequenza dell'ampiezza del filtro passa-alto applicato nel sistema di Figura 4.1



4.4 Filtri notch

Come abbiamo visto in fase di presentazione preliminare del sistema, il segnale “voce+sirena”, per rimuovere la componente associata alla sirena, viene filtrato da un banco di filtri elimina-banda. I k filtri *notch* presenti nel blocco di filtraggio sono tutti filtri IIR del secondo ordine e hanno frequenza di centro-banda impostata sul valore della frequenza fondamentale rilevata dal blocco che esegue l'AMDF e sulle sue relative k armoniche.

Il filtro *notch* ideale ha risposta in frequenza [13]:

$$H(e^{j\omega}) = \begin{cases} 0, & \omega_0 \\ 1, & \text{altrimenti} \end{cases} \quad (4.10)$$

dove ω_0 rappresenta la frequenza di centro-banda.

Nella realtà non può esistere un filtro che possa soddisfare la risposta espressa dalla (4.10); esistono però soluzioni che riescono ad approssimarla molto. Una di queste è fornita dal filtro IIR del secondo ordine descritto dalla seguente funzione di trasferimento in z :

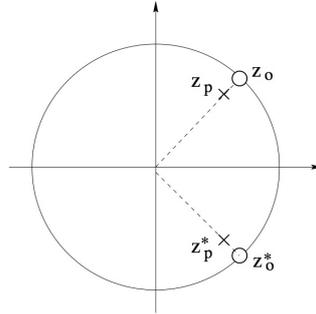
$$H(z) = \frac{(z - e^{j\omega_0})(z - e^{-j\omega_0})}{(z - re^{j\omega_0})(z - re^{-j\omega_0})} \quad (4.11)$$

in cui $e^{j\omega_0} = z_0$ rappresenta lo *zero* della funzione, $re^{j\omega_0} = z_p$ il polo ed $0 < r < 1$.

In Figura 4.8 si mostra come la caratteristica elimina-banda, per questa funzione di trasferimento, possa essere ottenuta posizionando lo *zero* sulla circonferenza unitaria e collocando il *polo* molto “vicino” ad esso (valore di r prossimo ad 1) all'interno del cerchio di raggio unitario in modo che il filtro risulti stabile. Inoltre, poiché i coefficienti della risposta all'impulso risultino numeri reali, si inseriscono il relativo *polo* e il relativo *zero* coniugati.

Facendo muovere ω tra 0 e π possiamo osservare il comportamento del filtro e ricavare, così, la sua funzione guadagno:

Figura 4.8: Posizionamento, sulla circonferenza unitaria, dei *poli* e degli *zeri* di un filtro *notch*.



- quando ω è molto “vicino” a ω_0 , $e^{j\omega} \approx e^{j\omega_0}$ e quindi: $|H(e^{j\omega})| \approx |H(e^{j\omega_0})| = 0$;
- quando ω è invece “lontano” da ω_0 , $e^{j\omega} - e^{j\omega_0} \gg 0$, ricordando che $z_0 \approx z_p$

e che $z_0^* \approx z_p^*$, abbiamo:
$$|H(e^{j\omega})| = \left| \frac{e^{j\omega} - z_0}{e^{j\omega} - z_p} \right| \cdot \left| \frac{e^{j\omega} - z_0^*}{e^{j\omega} - z_p^*} \right| \approx 1 \cdot 1 = 1.$$

Dall’equazione (4.11), procedendo con la sostituzione $e^{jx} = \cos(x) + j \sin(x)$, otteniamo:

$$H(z) = \frac{1 - 2 \cos \omega_0 z^{-1} + z^{-2}}{1 - 2r \cos \omega_0 z^{-1} + r^2 z^{-2}} \quad (4.12)$$

e la relativa equazione alle differenze:

$$y(n) = x(n) - 2 \cos(\omega_0)x(n-1) + x(n-2) + 2r \cos(\omega_0)y(n-1) - r^2y(n-2). \quad (4.13)$$

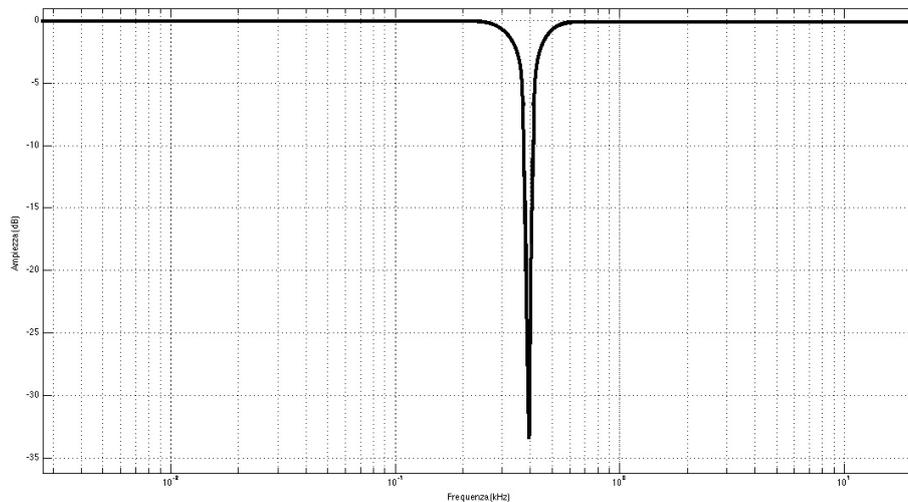
In Figura 4.9 è mostrata la risposta in frequenza dell’ampiezza di un filtro costruito attenendosi a queste due equazioni. Esse sono esattamente quelle che abbiamo utilizzato per la realizzazione dei filtri *notch* nella nostra simulazione con MATLAB.

Nei filtri, la frequenza di centro-banda si imposta assegnando direttamente il valore di ω_0 mentre la larghezza della banda eliminata si ottiene variando il valore di r : più r si avvicina ad 1, più la banda si restringe, più si allontana, più la banda si allarga.

I k filtri¹⁴ presenti nel blocco di filtraggio del sistema sono posti in modo tale che, a partire dal primo fino ad arrivare al k -esimo, l’uscita di uno sia l’ingresso del successivo. In questo modo, esso riesce nell’intento di eliminare il suono di sirena dal segnale in ingresso: il primo *notch* del gruppo filtra la frequenza fondamentale $\omega_0 = 2\pi f_0$ ricavata dal blocco che applica l’AMDF e, di passo in passo, gli altri *notch* presenti in cascata filtrano le armoniche fino alla k -esima, ovvero quella con frequenza $k\omega_0 = 2\pi k f_0$.

¹⁴ k , dalle sperimentazioni effettuate nel Capitolo 6, risulterà assumere il valore 12

Figura 4.9: Risposta in frequenza dell'ampiezza di uno dei k filtri *notch* applicati nel blocco di filtraggio di sistema di Figura 4.1. Nell'esempio la frequenza di centro-banda è settata sul valore di 392 Hz e la banda passante (calcolata a -3 dB) è ampia circa 8 Hz (fattore di qualità $Q=45$).



4.5 Buffer

Come abbiamo visto nel paragrafo 4.2, l'AMDF, per poter estrarre il *pitch* del segnale a cui è applicata, necessita di un *frame* del segnale di dimensione maggiore o uguale al valore massimo del periodo relativo alla frequenza più bassa che si intende tracciare. La frequenza più bassa raggiunta dal suono di sirena è 392 Hz meno il suo 5%, ovvero 372 Hz. In questo caso il periodo (l'inverso della frequenza) risulta quindi essere di 0.003s. Per un segnale campionato a 44100 campioni al secondo, come quello che abbiamo preso in considerazione in tutto il nostro lavoro, il numero di campioni che compongono il *frame* risulta quindi essere 132. Occorre, perciò, che prima che il segnale venga inviato alla fase di filtraggio passa-alto e a quella del riconoscimento del tono della sirena, ci sia un sistema che memorizzi i campioni che compongono il *frame*. Questa operazione è, per l'appunto, l'operazione eseguita dal *buffer* [14].

Una volta che questo segmento di segnale è acquisito nella sua interezza e inviato al blocco di calcolo che esegue l'AMDF, il *buffer* può aggiornare i suoi registri procedendo con l'acquisizione dei campioni che compongono il *frame* successivo.

4.6 Linea di ritardo

Come abbiamo visto nei capitoli precedenti, affinché il sistema possa cancellare correttamente il suono di sirena presente nel segnale in ingresso, occorre che:

- memorizzi i campioni in ingresso che compongono un *frame*;
- calcoli l'AMDF;
- ne estragga il relativo *pitch*;
- setti correttamente i k filtri notch

Queste operazioni necessitano, ovviamente, di un certo lasso di tempo per poter essere eseguite. Per questo motivo, prima della fase di filtraggio del segnale, abbiamo inserito un ritardo volto a compensare questo scostamento temporale.

Indicando con:

- N_{frame} il numero di campioni presenti in un *frame*;
- N_{hp} il ritardo introdotto dal filtro passa-alto;
- N_{AMDF} il tempo (espresso in campioni) necessario al calcolo dell'AMDF;
- N_{pitch} l'intervallo temporale (sempre in campioni) utile all'estrazione del *pitch*;
- N_{k_notch} il tempo che serve perché tutti i k filtri *notch* si posizionino sulla esatta frequenza di centro-banda;

possiamo ottenere il ritardo complessivo del sistema nel suo “ramo superiore”:

$$N_{TOT} = N_{frame} + N_{hp} + N_{AMDF} + N_{pitch} + N_{k_notch} \quad (4.14)$$

Ponendo il valore del ritardo temporale L , come $L = N_{TOT}$, abbiamo che, in Figura 4.1: $x_d(n) = x(n - L)$. Grazie a questa accortezza, nel momento in cui il segnale ritardato arriva al blocco di filtraggio, la frequenza di centro-banda dei *notch* risulta appena aggiornata e di conseguenza il primo e i successivi $(N_{frame} - 1)$ campioni del segnale in ingresso al banco di filtri possono essere filtrati correttamente.

Questa operazione comporta evidentemente il generarsi di una latenza tra il segnale in ingresso al sistema e il suo corrispettivo filtrato in uscita. Il valore della latenza totale può essere misurato come:

$$Latency = L + kN_{filter} \quad (4.15)$$

Dove N_{filter} rappresenta la latenza che introduce ogni singolo filtro durante il filtraggio attuato dal blocco di filtri *notch*.

Il latenza complessiva, però, vista la potenza di calcolo dei processori attualmente presenti sul mercato, si attesta sempre su valori accettabili¹⁵ e risulta che:

$$N_{frame} > N_{hp} + N_{AMDF} + N_{pitch} + N_{k_notch} \quad (4.16)$$

ovvero, il valore che in influenza maggiormente il ritardo generale del sistema è N_{tor} : la dimensione della finestra necessaria al calcolo dell'AMDF.

Per una comunicazione “ad una sola via”, come per le radio installate sui mezzi di emergenza (*one-way radio*) - dove un operatore, per poter parlare, deve attendere che il suo interlocutore abbia chiuso la trasmissione e viceversa - sono tollerabili valori di latenza addirittura fino a 0.5s, di gran lunga più grandi del valore della dimensione del frame di 0.003s che utilizziamo nel nostro sistema.

4.7 Il metodo applicato

Prendendo in considerazione quanto esposto nei paragrafi precedenti, possiamo analizzare cosa succede quando in ingresso al sistema proposto troviamo un segnale che contiene la voce e, insieme ad essa, il disturbo generato dal suono di sirena bitonale. Superato lo stadio di *buffering*, il *frame* del segnale, che secondo quanto riportato nel Paragrafo 4.5, consta di 132 campioni, viene filtrato dal filtro passa-alto con frequenza di taglio 1.8 kHz. In uscita al filtro troviamo, quindi, un segnale che presenta ancora una buona componente di voce ma dove le armoniche generate dal suono di sirena hanno valori in ampiezza largamente maggiori di essa. Per questo motivo, attraverso l'*Average Magnitude Difference Function*, risulterà certo che la frequenza fondamentale ricavata sia riferita al suono di sirena e non ad elementi periodici presenti nel suono della voce. Questo ci dà la sicurezza che di volta in volta, i k filtri *notch* preposti alla cancellazione della sirena, siano “accordati” sulla nota corretta. La larghezza di banda di questi filtri, determina da una parte la profondità della cancellazione effettuata - più la banda è larga, più siamo certi che la frequenza che si vuol cancellare sia effettivamente eliminata, andando anche a correggere piccoli errori di valutazione della funzione AMDF - e dall'altra la qualità della voce rimanente nel segnale dopo il filtraggio. Infatti, in caso di larghezza di banda troppo elevata dei filtri *notch*, la quantità di segnale vocale eliminato involontariamente assieme alla sirena risulterebbe tale da comprometterne fortemente la qualità durante l'ascolto. In fase sperimentale occorrerà quindi trovare un compromesso tra questi due effetti: applicare ai filtri *notch* una larghezza di banda tale che i filtri svolgano adeguatamente il compito di cancellazione della sirena senza che, allo stesso tempo, distorcano troppo il segnale vocale. Una volta presa questa accortezza, il segnale y in uscita al sistema si rivelerà essere il solo suono della voce, privo del disturbo della sirena. Esso risulterà non eccessivamente

¹⁵ vedi stime riportate nel Capitolo 6

degradato (e quindi comprensibile per un'ascoltatore) e pronto per la trasmissione via radio.

Capitolo 5

Metodo 2: con utilizzo dell'IIR *Adaptive Notch* Filtering

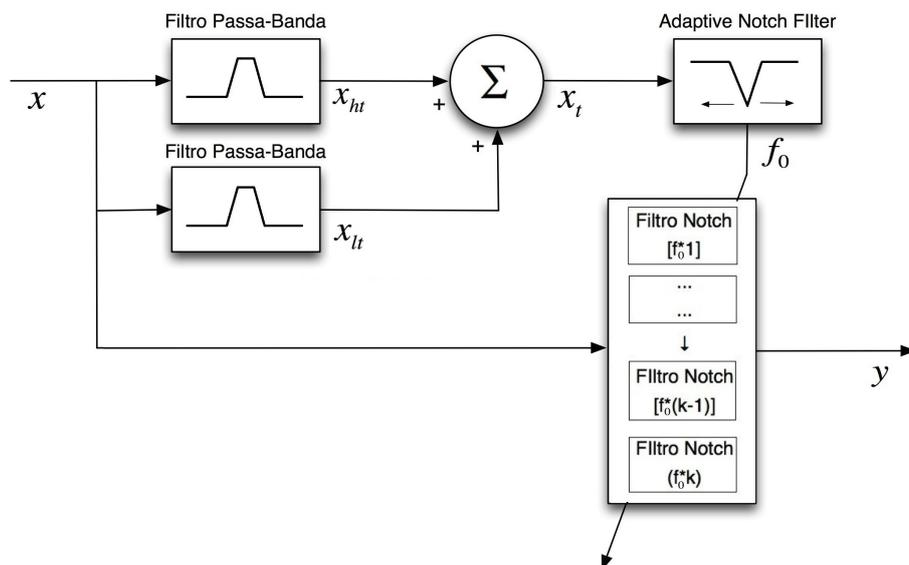
In questo capitolo descriveremo il secondo dei due metodi che proponiamo per la cancellazione del suono di sirena. Come per il primo metodo, innanzitutto forniremo una visione complessiva del sistema attraverso la sua rappresentazione tramite schema a blocchi. Successivamente analizzeremo approfonditamente ogni singolo blocco (ad esclusione di quelli già trattati nel capitolo precedente), per poi applicare il metodo completo a un segnale vocale disturbato dalla presenza del suono di sirena di tipo bitonale - come quelle installate sui mezzi del Corpo Nazionale dei Vigili del Fuoco -, e osservarne il comportamento complessivo.

5.1 Panoramica d'insieme del sistema proposto

Nella Figura 5.1 è mostrato lo schema a blocchi che descrive il funzionamento del metodo che proponiamo in questo capitolo. I blocchi che lo compongono sono:

- due filtri passa-banda;

Figura 5.1: Schema a blocchi del sistema per la cancellazione della sirena che utilizza l'*Adaptive Notch Filtering* e un banco di filtri *notch*.



- un sommatore;
- un blocco che, tramite filtraggio adattativo con un filtro IIR *notch* (*Adaptive Notch Filtering*), estrae la frequenza fondamentale che caratterizza il segnale in ingresso (*pitch*) e la invia in uscita;
- un banco di k filtri *notch*.

Seguendo quanto riportato dallo schema a blocchi proposto in Figura 5.1, il segnale digitale in ingresso, campionato dal segnale trasdotto dal microfono della radio veicolare, prende tre percorsi diversi. Nei primi due è previsto che al segnale siano applicati due filtri passa-banda¹⁶. In uscita ai due filtri, troviamo, quindi, due segnali (x_l e x_m) contenenti, ciascuno, una sola banda passante. Questi due segnali vengono successivamente sommati tra loro da un sommatore e vanno a comporre nuovamente un solo segnale (x_i) che, a questo punto, conterrà soltanto due bande rispetto al segnale originario. Questo è quello che va in ingresso all'*Adaptive Notch Filter*,¹⁷ il quale dà in uscita la frequenza fondamentale (f_0) che caratterizza il suo tono (*pitch*) e la invia al banco di filtri *notch*¹⁸.

Il terzo percorso è quello che porta all'effettiva cancellazione del suono di sirena dal segnale "voce+sirena". Il segnale in ingresso viene, infatti, filtrato da una serie di k filtri *notch*. Questi, come nel metodo esposto nel capitolo precedente, sono filtri elimina-banda, a banda molto stretta, e sono impostati con frequenza di centro-banda pari alla frequenza (f_0) - in uscita dal blocco dove è effettuato il filtraggio *notch* adattativo - e alle relative k armoniche [f_0*2 ; ... $f_0*(k-1)$; f_0*k].

5.2 Estrazione del *pitch* tramite l'*Adaptive Notch Filtering*

L'*Adaptive Notch Filtering* è una tecnica che permette di tracciare la frequenza fondamentale che caratterizza una sinusoidale immersa in un rumore. Nel nostro caso ci permette di individuare, istante per istante, il *pitch* del suono di sirena che entra nel sistema. L'algoritmo specifico che abbiamo utilizzato per svolgere questa funzione è l'Algoritmo *Lattice Semplificato* (*Simplified Lattice Algorithm*) proposto per primo da Phillip A. Regalia [15]. Esso rappresenta una variante semplificata dell'Algoritmo di Discesa del Gradiente applicato a un filtro *notch* scritto nella forma *lattice*. Pertanto, prima di fornire l'algoritmo definitivo utilizzato nel nostro sistema, introdurremo questi due concetti.

¹⁶ la struttura e la banda passante del filtro passa-banda sono oggetto del Paragrafo 5.3

¹⁷ l'*Adaptive Notch Filtering* è trattato nel Paragrafo 5.2

¹⁸ una dettagliata analisi dei filtri *notch* è affrontata nel Capitolo 4, Paragrafo 4.3

5.2.1 Filtro notch in forma lattice

Una modalità differente, rispetto a quanto esposto nel Paragrafo 4.4, per la rappresentazione di un filtro notch è quella che è costruita intorno alla funzione passa-tutto del secondo ordine scritta nella forma lattice che, in formule, è descritta dalla seguente espressione:

$$V(z) = \frac{\sin \theta_2 + \sin \theta_1(1 + \sin \theta_2)z + z^2}{1 + \sin \theta_1(1 + \sin \theta_2)z + \sin \theta_2 z^2}. \quad (5.1)$$

Da questa espressione possiamo ricavare, rispettivamente, un filtro elimina-banda ($\hat{H}(z)$) e il suo complementare passa-banda ($\hat{G}(z)$) operando direttamente sulla struttura del filtro passa-tutto, come mostrato in Figura 5.2, ottenendo:

$$\begin{aligned} \hat{H}(z) &= \frac{1}{2}[1 + V(z)] \\ \hat{G}(z) &= \frac{1}{2}[1 - V(z)]. \end{aligned} \quad (5.2)$$

I blocchi θ_1 e θ_2 , in figura, rappresentano i blocchi standard della struttura lattice (Figura 5.3), e, quindi, possiamo scrivere le (5.2) come:

$$\begin{aligned} \hat{H}(z) &= \frac{1 + \sin \theta_2}{2} \frac{1 + 2 \sin \theta_1 z + z^2}{1 + \sin \theta_1(1 + \sin \theta_2)z + \sin \theta_2 z^2} \\ \hat{G}(z) &= \frac{1 - \sin \theta_2}{2} \frac{1 - z^2}{1 + \sin \theta_1(1 + \sin \theta_2)z + \sin \theta_2 z^2}. \end{aligned} \quad (5.3)$$

Figura 5.2: Realizzazione di una coppia complementare di filtri passa-banda/elimina-banda a partire da un filtro passa-tutto in forma lattice.

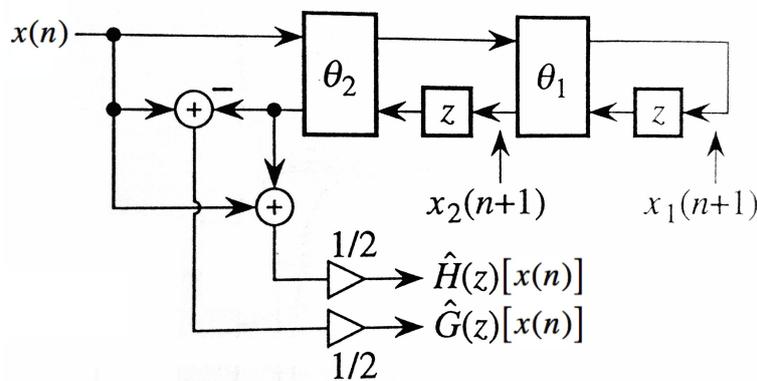
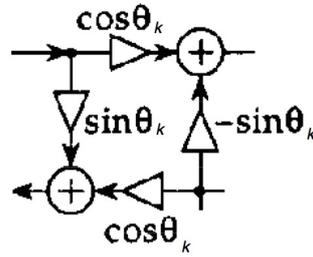


Figura 5.3: Blocco base θ_k della struttura *lattice*.



Poiché $|V(e^{i\omega})| = 1$ per ogni ω , possiamo scrivere $|V(e^{i\omega})| = e^{i\phi(\omega)}$, dove $\phi(\omega)$ è la risposta in frequenza della fase. Dato che $V(z)$ è una funzione passa-tutto causale, quando $|\sin\theta_1| < 1$ e $|\sin\theta_2| < 1$, $\phi(\omega)$ risulta monotonicamente crescente e soddisfa le seguenti condizioni al contorno: $\phi(0) = 0$, $\phi(\pi) = 2\pi$.

Applicando queste condizioni al filtro elimina-banda abbiamo:

$$\hat{H}(e^{i0}) = \frac{1}{2}[1 + e^{i\phi(0)}] = 1$$

$$\hat{H}(e^{i\pi}) = \frac{1}{2}[1 + e^{i\phi(\pi)}] = 1$$
(5.4)

e, per il valore ω_0 per cui $\phi(\omega_0) = \pi$, si ha che $\hat{H}(e^{i\omega_0}) = 0$.

ω_0 risulta quindi essere la frequenza che viene eliminata dal filtro *notch* e può essere ottenuta risolvendo la seguente equazione:

$$e^{i\pi} = V(e^{i\omega_0}) = \frac{\sin\theta_1 + \sin\theta_1(1 + \sin\theta_2)e^{i\omega_0} + e^{i2\omega_0}}{1 + \sin\theta_1(1 + \sin\theta_2)e^{i\omega_0} + \sin\theta_2 e^{i2\omega_0}}$$
(5.5)

che ha come soluzione $\omega_0 = \theta_1 + \frac{\pi}{2}$ per $|\theta_1| < \frac{\pi}{2}$.

La larghezza della banda eliminata (valutata a -3dB) si ottiene, invece, estraendo il valore di B da:

$$\sin\theta_2 = \frac{1 - \tan(B/2)}{1 + \tan(B/2)}$$

È importante osservare che la frequenza di centro-banda del filtro *notch* e la relativa larghezza di banda sono rispettivamente (ed esclusivamente) dipendenti da θ_1 e da θ_2 . Pertanto i due valori possono essere settati indipendentemente l'uno dall'altro fornendo notevole libertà nella realizzazione del filtro.

5.2.2 Algoritmo con Discesa del Gradiente

Supponendo di poter considerare il segnale in ingresso $x(n)$ come la somma di un segnale utile $u(n)$, dal quale ricavare la frequenza fondamentale della sinusoide che lo caratterizza, e di un rumore bianco $\zeta(n)$, possiamo scrivere il valore in uscita dal filtro elimina-banda sopra descritto come:

$$\hat{y}(n) = \hat{H}(z)[u(n) + \zeta(n)] \quad (5.6)$$

Prendendo la Funzione Costo di questo segnale ($E[\hat{y}^2(n)]$) e osservando che essa ha un singolo punto di minimo, possiamo indicare il seguente come algoritmo adattativo volto alla sua minimizzazione:

$$\theta_1(n+1) = \theta_1(n) - \mu \hat{y}(n) \frac{\partial \hat{y}(n)}{\partial \theta_1} \quad (5.7)$$

dove μ rappresenta il guadagno dell'algoritmo adattativo (*stepsize*).

Calcolando la derivata al secondo membro della (5.7), abbiamo:

$$\frac{\partial \hat{y}(n)}{\partial \theta_1} = \left[\frac{z \cos \theta_1 \cos^2 \theta_2}{2} \frac{1 - z^2}{(1 + \sin \theta_1 (1 + \sin \theta_2) z + \sin \theta_2 z^2)^2} \right] [u(n) + \zeta(n)]$$

Ponendo $D(z) = (1 + \sin \theta_1 (1 + \sin \theta_2) z + \sin \theta_2 z^2)^2$, risulta:

$$\frac{\partial \hat{y}(n)}{\partial \theta_1} = \left[\frac{z \cos \theta_1 \cos \theta_2}{D(z)} \frac{1 - z^2}{D(z)} \frac{1 - \sin \theta_2}{2} \frac{1 + \sin \theta_2}{\cos \theta_2} \right] [u(n) + \zeta(n)]$$

da cui, sostituendo

$$G(z) = \frac{1 - z^2}{D(z)} \frac{1 - \sin \theta_2}{2}$$

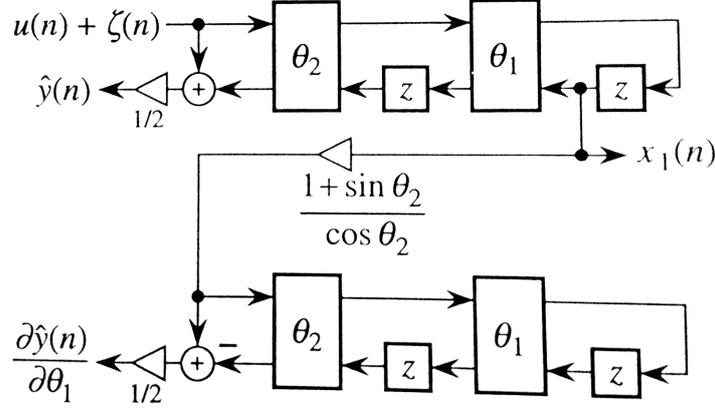
e osservando dalla Figura 5.4 che

$$x_1(n) = \frac{z \cos \theta_1 \cos \theta_2}{D(z)} [u(n) + \zeta(n)],$$

si ottiene:

$$\frac{\partial \hat{y}(n)}{\partial \theta_1} = \frac{1 + \sin \theta_2}{\cos \theta_2} G(z) x_1(n). \quad (5.8)$$

Figura 5.4: Generazione dei segnali regressori filtrati necessari all'algorithm adattativo descritto nella (5.7)



Il fattore di scala $(1 + \sin\theta_1)/\cos\theta_2$ presente nella (5.8) non varia con il parametro che determina la frequenza di centro banda del filtro notch, ovvero θ_1 , e può quindi essere assorbito nel fattore di guadagno (*stepsiz*e) μ .

L'algorithm che, dalle precedenti considerazioni, risulta in grado di adattare nel tempo il valore di θ_1 , affinché questo converga verso la frequenza fondamentale che caratterizza il segnale $u(n)$ in ingresso, è riassunto come segue:

$$\begin{aligned} \begin{bmatrix} g_1 \\ w_1 \end{bmatrix} &= \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix} \begin{bmatrix} x(n) \\ x_2(n) \end{bmatrix} \\ \hat{y}(n) &= \frac{1}{2} [x(n) + w_1] \\ \begin{bmatrix} g_2 \\ w_2 \end{bmatrix} &= \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix} \begin{bmatrix} x_1(n) \\ \xi_2(n) \end{bmatrix} \\ \nabla_{\theta_1} &= \frac{1}{2} [x_1(n) - w] \end{aligned} \quad (5.9)$$

$$\theta_1(n+1) = \theta_1(n) - \mu \hat{y}(n) \nabla_{\theta_1}$$

$$\begin{bmatrix} x_1(n+1) \\ x_2(n+1) \end{bmatrix} = \begin{bmatrix} \cos\theta_1(n+1) & -\sin\theta_1(n+1) \\ \sin\theta_1(n+1) & \cos\theta_1(n+1) \end{bmatrix} \begin{bmatrix} g_1 \\ x_1(n) \end{bmatrix}$$

$$\begin{bmatrix} \xi_1(n+1) \\ \xi_2(n+1) \end{bmatrix} = \begin{bmatrix} \cos\theta_1(n+1) & -\sin\theta_1(n+1) \\ \sin\theta_1(n+1) & \cos\theta_1(n+1) \end{bmatrix} \begin{bmatrix} g_1 \\ \xi_1(n) \end{bmatrix}$$

Dove, all'istante n , risultano disponibili:

- la frequenza di centro-banda del filtro *notch* $\theta_1(n)$;

- gli stati del filtro $x_1(n)$, $x_2(n)$, $\xi_1(n)$ e $\xi_2(n)$;
- il segnale di ingresso $x(n) = u(n) + \zeta(n)$.

5.2.3 Algoritmo *Lattice Semplificato*

Per l'algoritmo con Discesa del Gradiente, la Funzione Costo è una funzione di tipo non-quadratico del parametro θ_1 , ovvero della frequenza di centro-banda del filtro *notch*. In Figura 5.5 è mostrato il grafico che rappresenta tale funzione, al variare di θ_1 , con tre differenti valori di θ_2 (e, di conseguenza, tre larghezze di banda diverse del filtro *notch*) dato un valore di centro banda fisso. Notiamo come ogni curva somigli a una funzione di tipo *notch* che diventa più sottile quando la larghezza di banda diminuisce. Per questo motivo il gradiente

$$\partial E[\hat{y}(n)]/\partial \theta_1$$

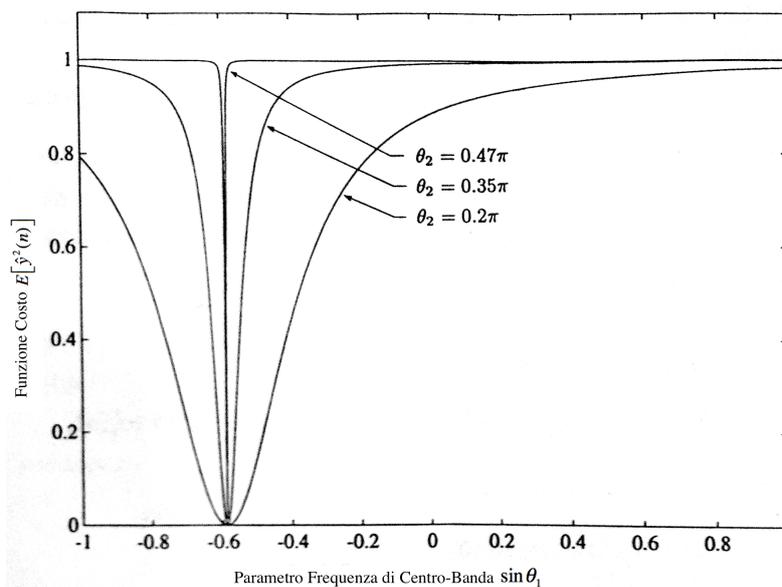
diventa poco significativo se la frequenza di centro-banda risulta lontana dalla frequenza fondamentale presente in $u(n)$ e perciò la convergenza dell'algoritmo può risultare molto lenta. La versione semplificata, proposta in questo paragrafo, ha la proprietà di aumentarne la velocità.

L'aggiornamento dell'algoritmo (5.7), in questa versione semplificata, diventa:

$$\theta_1(n+1) = \theta_1(n) - \mu \hat{y}(n) x_1(n) \quad (5.10)$$

che, a ogni iterazione, non utilizza il gradiente ma direttamente lo stato interno del filtro *lattice* $x_1(n)$.

Figura 5.5: Ampiezza della Funzione Costo al variare della larghezza della banda del filtro *notch* con frequenza di centro-banda fissa.



Per analizzare le proprietà di convergenza dell'algoritmo di aggiornamento (5.10), possiamo studiare il comportamento dell'equazione differenziale ordinaria:

$$\frac{d\theta_1}{dt} = -E[\hat{y}(n)x_1(n)] \quad (5.11)$$

i punti di convergenza (in media) della (5.10) corrisponderanno, infatti, ai “punti di attrazione” di questa equazione differenziale.

Per un segnale di ingresso del tipo:

$$x(n) = u(n) + \zeta(n) = p_1 \sin(\omega_1 n + \phi) + \zeta(n)$$

otteniamo la soluzione:

$$\frac{d\theta_1}{dt} = -p_1^2 \frac{\cos\theta_1 \cos\theta_2}{|D(e^{j\omega_1})|^2} (\cos\omega_1 - \sin\theta_1).$$

Possiamo affermare che questa espressione è globalmente asintoticamente stabile per un punto stazionario che soddisfi l'uguaglianza:

$$\sin\theta_1 = \cos\omega_1.$$

Ciò implica che la frequenza di centro-banda del filtro *notch* ($\omega_0 = \theta_1 + \pi/2$) coincide con la frequenza fondamentale ω_1 del segnale utile $u(n)$.

L'Algoritmo *Lattice* Semplificato, nella sua interezza, è descritto dalle seguenti computazioni:

$$\begin{aligned} \begin{bmatrix} g_1 \\ w \end{bmatrix} &= \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix} \begin{bmatrix} x(n) \\ x_2(n) \end{bmatrix} \\ \hat{y}(n) &= \frac{1}{2}[x(n) + w] \\ \theta_1(n+1) &= \theta_1(n) - \mu \hat{y}(n)x_1(n) \\ \begin{bmatrix} x_1(n+1) \\ x_2(n+1) \end{bmatrix} &= \begin{bmatrix} \cos\theta_1(n+1) & -\sin\theta_1(n+1) \\ \sin\theta_1(n+1) & \cos\theta_1(n+1) \end{bmatrix} \begin{bmatrix} g_1 \\ x_1(n) \end{bmatrix} \end{aligned} \quad (5.12)$$

Dove, all'istante n , sono disponibili:

- la frequenza di centro-banda del filtro *notch* $\theta_1(n)$;

- gli stati del filtro $x_1(n), x_2(n)$;
- il segnale di ingresso $x(n) = u(n) + \zeta(n)$.

In Figura 5.6 mostriamo rispettivamente il comportamento del filtro *notch* adattativo con Algoritmo di Discesa del Gradiente e del filtro con Algoritmo *Lattice* Semplificato quando in ingresso ai sistemi è posto un segnale sinusoidale a frequenza variabile immerso in un rumore bianco e nel caso in cui entrambi i filtri abbiano un'ampia larghezza della banda passante.

In Figura 5.7, invece, proponiamo lo stesso esperimento ma nel caso in cui entrambi i filtri presi in esame abbiano una banda passante molto stretta.

Dai primi due grafici si può notare come i filtri abbiano tempi di convergenza simili ma anche come il filtro con Algoritmo con Discesa del Gradiente risulti leggermente più preciso nel tracciare la frequenza del segnale in ingresso rispetto al filtro con Algoritmo *Lattice* Semplificato.

Nei secondo due, invece, è facile osservare come il filtro con Algoritmo con Discesa del Gradiente perda pesantemente la capacità di tracciare la frequenza fondamentale del segnale in ingresso, mentre il filtro con Algoritmo *Lattice* Semplificato, seppur con tempi di convergenza maggiori rispetto al caso con filtro ad ampia banda passante, continua a riuscire efficacemente nell'operazione.

Per questo motivo e per una minore complessità di calcolo per lo svolgimento del (5.12) rispetto al (5.9), nel nostro metodo è stato scelto un filtraggio *notch* adattativo con Algoritmo *Lattice* Semplificato.

È importante sottolineare che gli algoritmi che abbiamo illustrato in questo paragrafo riescono nell'intento di tracciare la frequenza fondamentale del segnale in ingresso quando questo è composto da sinusoidi pure immerse in un rumore bianco. Questa, ovviamente, non è la condizione che si verifica quando si vuol tracciare il *pitch* del suono di sirena che disturba una conversazione radio. Filtrando, però, il segnale "voce+sirena" con due filtri passa-banda prima che lo stesso sia inviato al filtro *notch* adattativo, riusciamo ad avvicinarsi fortemente alla condizione richiesta. Come vedremo nel prossimo paragrafo, per poter settare correttamente la larghezza e la frequenza centrale della banda passante dei due filtri, utilizzeremo le informazioni pregresse forniteci dalle specifiche tecniche dei suoni di sirena.

Figura 5.6: Stima della frequenza fondamentale del segnale in ingresso con:

- a) filtro con Algoritmo con Discesa del Gradiente
- b) filtro con Algoritmo *Lattice* Semplificato

entrambi nel caso di ampia larghezza della relativa banda passante.

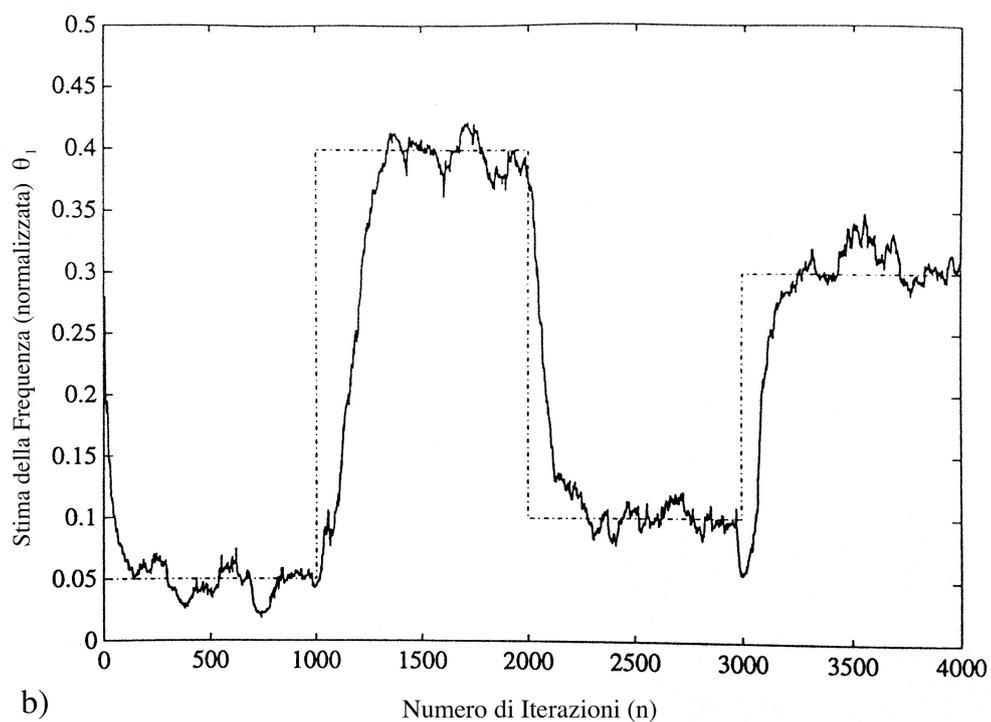
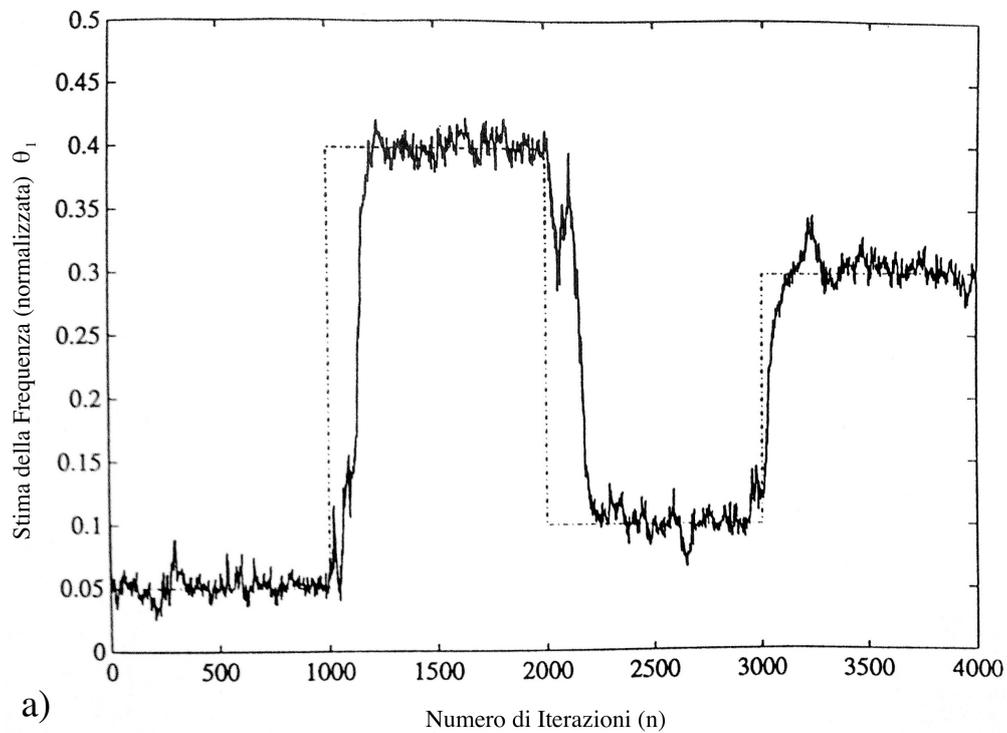
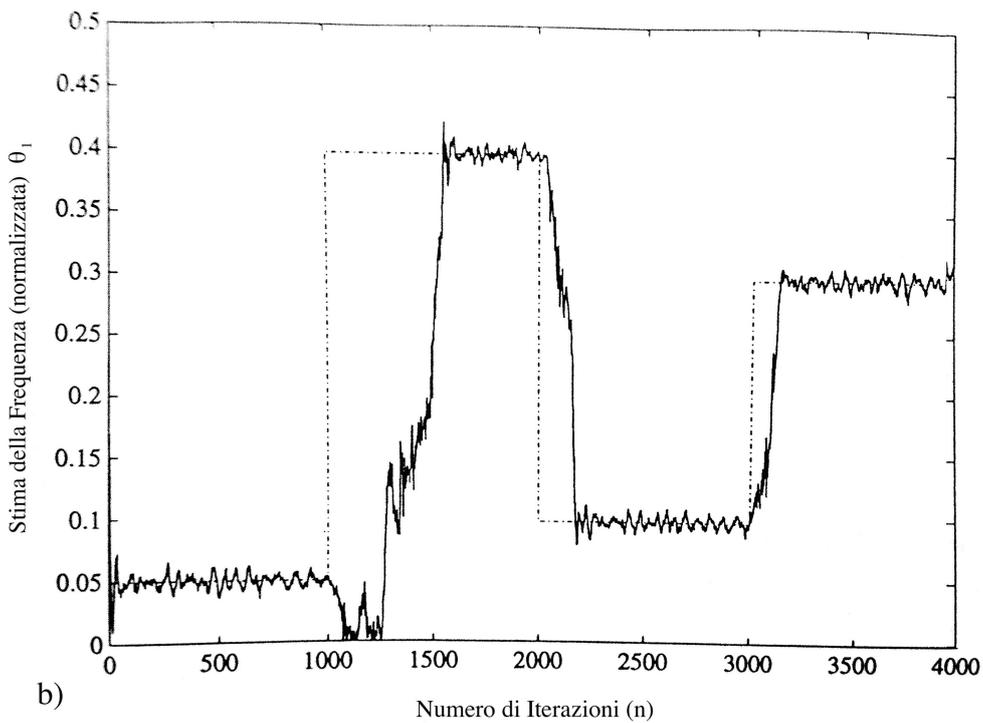
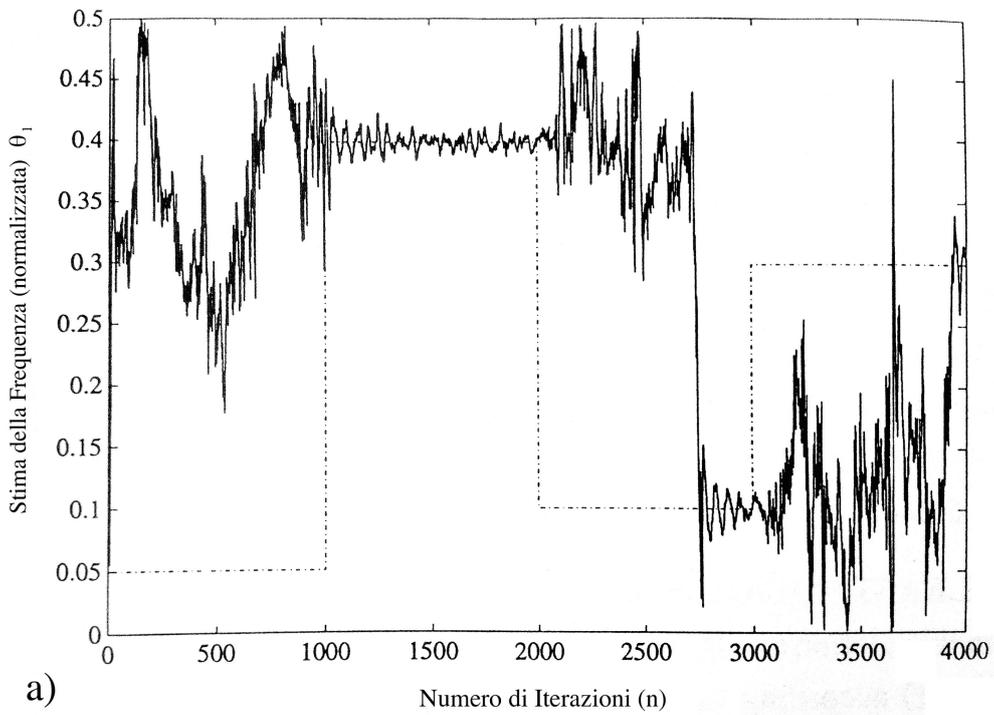


Figura 5.7: Stima della frequenza fondamentale del segnale in ingresso con:

- a) filtro con Algoritmo con Discesa del Gradiente
- b) filtro con Algoritmo *Lattice* Semplificato

entrambi nel caso della relativa banda passante stretta.



5.3 Filtro Passa-Banda

In Figura 5.8 è rappresentato un filtro passa-banda ideale. Come per il filtro passa-alto descritto nel Paragrafo 4.3, per realizzare un filtro digitale che approssimi la risposta in frequenza descritta in figura, utilizzeremo un filtro IIR di Butterworth. Anche in questo caso, infatti, non siamo interessati alla risposta in frequenza della fase ma desideriamo soltanto che la risposta dell'ampiezza sia la più "piatta" possibile nella banda passante; proprietà tipica dei filtri di questa tipologia.

Nella simulazione del sistema rappresentato in Figura 5.1, abbiamo realizzato i due filtri passa-banda attraverso la funzione di MATLAB "fdesign.bandpass". Fornendo in ingresso i valori delle frequenze di inizio e di fine della banda passante (f_L e f_H) - calcolate con ampiezza a -3 dB - e l'ordine del filtro, essa ci dà in uscita la funzione di Butterworth desiderata.

Nel caso specifico abbiamo utilizzato come valori di inizio e di fine della banda passante, i valori che si discostano di $\pm 5\%$ dalle due note della sirena bitonale poste a 392 Hz e 660 Hz. In questo modo siamo sicuri di riuscire a coprire tutte le possibili frequenze che si discostano dalle frequenze regolamentari ma che rimangono entro i limiti tollerati dal decreto esposto nel Paragrafo 2.1.

I valori che abbiamo immesso nella funzione MATLAB, per la realizzazione dei due filtri passa-alto utilizzati nel sistema in Figura 5.1, sono riassunti nella Tabella 5.1

Le risposte in frequenza dell'ampiezza che abbiamo, così, ottenuto sono quelle mostrate in Figura 5.9 e in Figura 5.10.

Figura 5.8: Filtro Passa-Banda ideale.

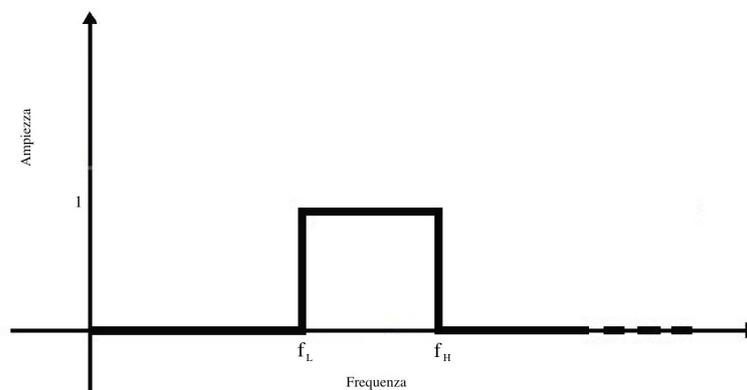


Tabella 5.1: Valori di ingresso alla funzione “fdesign.bandpass” per la realizzazione dei filtri passa-banda applicati nel sistema di Figura 5.1.

Filtro Passa-Banda	f_L	f_H	Ordine del Filtro N
Per la Nota Bassa (≈ 392 Hz)	372 Hz	412 Hz	10
Per la Nota Alta (≈ 660 Hz)	627 Hz	693 Hz	10

Figura 5.9: Risposta in frequenza dell’ampiezza del filtro passa-banda utilizzato nel sistema di Figura 5.1 per la nota bassa

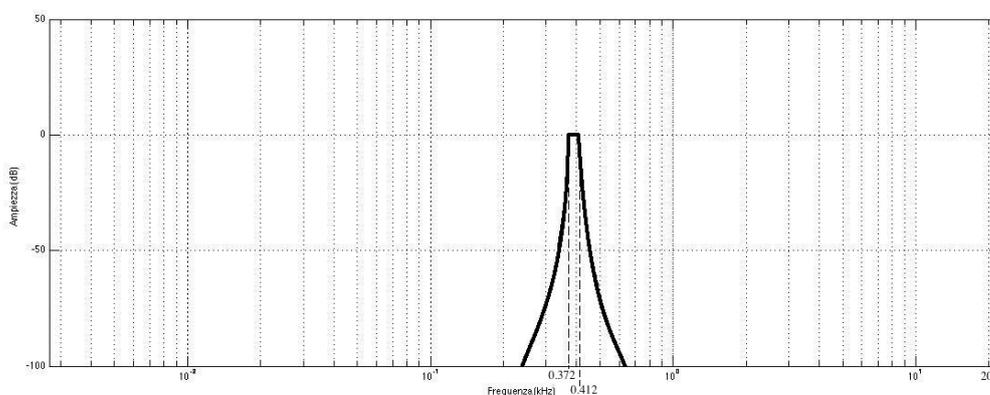
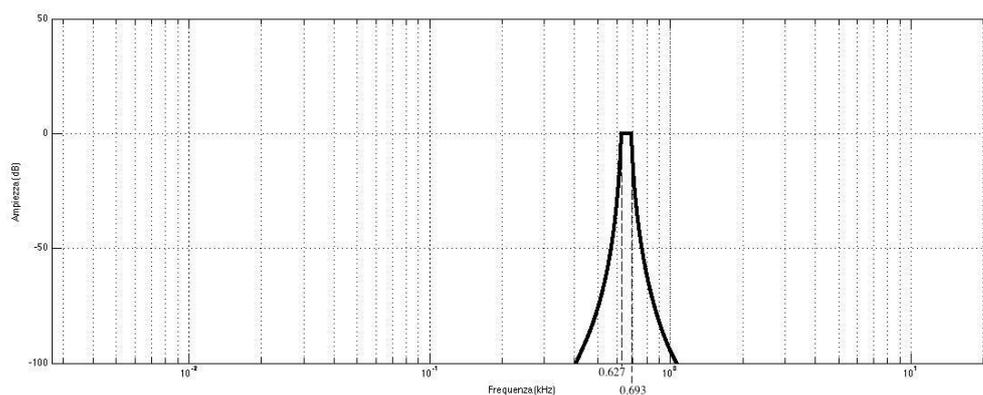


Figura 5.10: Risposta in frequenza dell’ampiezza del filtro passa-banda utilizzato nel sistema di Figura 5.1 per la nota alta



5.4 Il metodo applicato

Una volta descritti i vari blocchi che compongono lo schema di Figura 5.1, possiamo descrivere il funzionamento del sistema nella sua interezza quando in

ingresso allo stesso troviamo un segnale che contiene la voce e, insieme ad essa, il disturbo creato dal suono della sirena bitonale.

Il segnale in ingresso $x(n)$ viene filtrato dai due filtri passa-banda, dei quali uno è centrato sulla frequenza fondamentale della nota bassa del sirena, l'altro sulla frequenza della nota alta. In uscita ai due filtri troviamo rispettivamente:

- un segnale contenente solo la banda che va dai 372 Hz ai 412 Hz (x_{lt});
- un segnale contenente solo la banda che va dai 627 Hz ai 693 Hz (x_{ht}).

I due segnali vengono successivamente sommati tra loro e il segnale somma risultante (x_t) è inviato al filtro *notch* adattativo.

La funzione di questo filtro è quella di riuscire a tracciare il valore della frequenza fondamentale che caratterizza la sirena, muovendosi all'interno delle due piccole bande passanti. La parte del suono contenente la voce che, assieme alla sirena, passa attraverso i due filtraggi passa-banda - essendo la differenza di ampiezza tra i due suoni nelle bande passanti nettamente a vantaggio del suono di sirena -, è assimilata da esso come rumore di fondo. Pertanto, in uscita all'*Adaptive Notch Filter*, abbiamo il valore di *pitch* (f_0) che, istante per istante, caratterizza la sirena. Attraverso questo valore si settano le frequenze di centro-banda dei k filtri *notch*. Questi, applicati in cascata al segnale in ingresso $x(n)$, fanno sì che il segnale in uscita al sistema $y(n)$ contenga il solo suono della voce.

Come per il metodo proposto nel Capitolo 4, è opportuno che, in fase sperimentale, si trovi un compromesso tra una larghezza di banda dei filtri *notch* troppo grande - che riuscirebbe perfettamente nella cancellazione della sirena ma degraderebbe troppo la voce - e una banda troppo stretta - che lascerebbe quasi intatta la qualità della voce ma non riuscirebbe a cancellare perfettamente il suono della sirena -.

Capitolo 6

Valutazione e Confronto dei risultati sperimentali

In questo capitolo studieremo il comportamento dei due sistemi di cancellazione del suono di sirena proposti in precedenza, attraverso la loro simulazione con il software MATLAB. Per entrambi i casi inizieremo con l'implementazione delle rispettive componenti fondamentali, ovvero l'*Average Magnitude Difference Function* e l'*Adaptive Notch Filter*, impiegandole su suoni creati "artificialmente". Solo successivamente esamineremo il funzionamento dei metodi nella loro forma completa applicandoli a segnali realmente acquisiti dai microfoni delle radio veicolari presenti sugli automezzi dei Vigili del Fuoco. In seguito, confronteremo tra loro i risultati ottenuti e, infine, applicheremo, con opportune modifiche, gli stessi sistemi ad avvisatori acustici di tipologia diversa da quella utilizzata dal Corpo Nazionale dei Vigili del Fuoco. In questo modo potremo testare se i sistemi studiati per il caso particolare possono anche avere valore di carattere generale.

6.1 Cancellazione per mezzo dell' *Average Magnitude Difference Function*

6.1.1 AMDF applicata a suono di sirena simulato

Seguendo le specifiche tecniche dettate dal Decreto Ministeriale del 17/10/80¹⁹ risulta facile poter realizzare un suono di sirena regolamentare attraverso MATLAB. Per costruirlo creiamo due sinusoidi con frequenza a 392 Hz e a 660 Hz e sommiamo ad entrambe una relativa serie di armoniche in numero tale da coprire lo spettro audio fino a 20 kHz. Per la sinusoide a frequenza bassa andremo ad aggiungere 51 armoniche (fino a 392 Hz x 51 = 19.992 kHz) mentre, per quella a frequenza più alta ne aggiungeremo 30 (fino a 660 Hz x 30 = 19.800 kHz). Per rendere più verosimili i suoni ottenuti in questo modo, applicheremo un fattore di smorzamento di $1/k^2$ ad ogni k^{esima} armonica. Una volta ottenute le due note base, costruiremo un ciclo completo del suono della sirena affiancando tra di loro i periodi di nota alta e quelli di nota bassa secondo quanto riportato nel Paragrafo 2.1.

Di seguito mostriamo lo *script* che genera questo suono e, successivamente (Figura 6.1), il relativo spettrogramma.

¹⁹ vedi Appendice A

Script 6.1: Generazione di un ciclo di sirena bitonale completo.

```

%frequenza di campionamento
Fs=44100;
%durata del ciclo in campioni
ciclo=3*Fs;
%vettore tempo periodo lungo
t1=linspace(0,ciclo/(3*Fs),44100);
%vettore tempo periodo corto
t2=linspace(0,ciclo/(18*Fs),7350);
%durata del periodo lungo in campioni
T1=ciclo/3;
%durata del periodo corto in campioni
T2=ciclo/18;

%sinusoide fondamentale periodo lungo nota bassa
lb=sin(392*2*pi*t1)/2;
%sinusoide fondamentale periodo corto nota alta
ca=sin(660*2*pi*t2)/2;
%sinusoide fondamentale periodo corto nota bassa
cb=sin(392*2*pi*t2)/2;

%armoniche periodo lungo nota bassa
for k=2:51
    arm=((1/(k^2))*(sin(392*k*2*pi*t1)));
    lb=lb+arm;
end

%armoniche periodo corto nota alta
for k=2:30
    arm=((1/(k^2))*(sin(660*k*2*pi*t2)));
    ca=ca+arm;
end

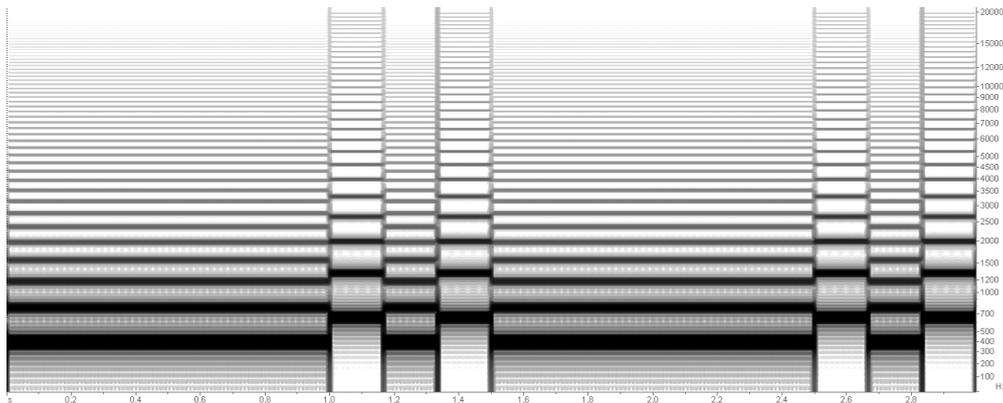
%armoniche periodo corto nota bassa
for k=2:51
    arm=((1/(k^2))*(sin(392*k*2*pi*t2)));
    cb=cb+arm;
end

%costruzione ciclo completo
% primo periodo
x(1:T1)=lb;
%secondo periodo
x(T1+1:T1+T2)=ca;
%terzo periodo
x(T1+1+T2:T1+2*T2)=cb;
%quarto periodo
x(T1+2*T2+1:T1+3*T2)=ca;
%quinto periodo
x(T1+3*T2+1:2*T1+3*T2)=lb;
%sesto periodo
x(2*T1+3*T2+1:2*T1+4*T2)=ca;
%settimo periodo
x(2*T1+4*T2+1:2*T1+5*T2)=cb;
%ottavo ciclo
x(2*T1+5*T2+1:2*T1+6*T2)=ca;

%normalizzazione
x=x/max(x);
Fs=44100;

%scrittura file audio
audiowrite('sirena.wav',x,Fs);

```

Figura 6.1: Spettrogramma di un ciclo di sirena bitonale generato con MATLAB.

Per applicare l'*Average Magnitude Difference Function* al suono di sirena generato con MATLAB, così da poterne estrarre il *pitch*, occorre implementare, su finestre temporali mobili, le funzioni (4.7) e (4.8) che qui, per comodità, riportiamo:

$$AMDF(k) = \sum_{m=k}^{N-1} |s(m) - s(m-k)|$$

$$f_0 = \frac{F_S}{k_M},$$

dove $s(m)$ è il segnale sirena in ingresso e k_M è la distanza tra i primi due valori minimi dell'AMDF.

Lo *script* che esegue queste due funzioni è il seguente:

Script 6.2: Funzione che estrae il *pitch* del segnale in ingresso attraverso l'*Average Magnitude Difference Function*.

```
function amdf(x,Ts)

% Grandezza Finestra di Analisi e dimensione dello scostamento tra una
% finestra e la sua successiva.
xlen = length(x);
wlen = 256;
wlag = 64;

% Zero-pad del segnale per renderlo divisibile in finestre temporali.
xpad = wlag-rem(xlen-wlen-1,wlag)-1;
x = [x(:); zeros(xpad,1)];
L = (xlen+xpad-wlen)/wlag+1;
L = L-wlen/wlag/2;

% Calcolo dell'AMDF e dei periodi di pitch.
for k1 = 1:L
    k2 = (k1-1)*wlag;
    for k3 = 1:wlen/2-1, c(k3) = sum(abs(x(k2+(1:wlen))-
x(k2+k3+(1:wlen))));
    end
% Valori minimi
n = trovapicchi(-c);
```

```

if ~isempty(n), n(c(n) > mean(c(n))-sqrt(var(c(n)))) = [];
end
if ~isempty(n), t0 = n(1);
else t0 = [];
end
if ~isempty(t0)
u1 = x(k2+t0+(1:wlen))-x(k2+(1:wlen));
u2 = x(k2+t0+(1:wlen))-x(k2+t0+1+(1:wlen));
t1 = sum(u1.*u2)/sum(u2.*u2);
p(k1) = 1/Ts/(t0+t1); %valore del pitch in Hz
else
p(k1) = NaN;
end
end

% Disegno delle note su scala musicale.
keys = {'C','C#','D','D#','E','F','F#','G','G#','A','A#','B'};
key0 = 12*log2(p/440)+57; %trasformazione da pitch a nota
u1 = max(0,min(floor(key0)));
u2 = min(96,max(ceil(key0)));
key1 = mod((u1:u2),12)+1;
key2 = floor((u1:u2)/12);
for k = 1:u2-u1+1, tick(k) = strcat(keys(key1(k)),num2str(key2(k)));
end

figure, plot((0.5+(0:L-1))*wlag*Ts,key0,'b. ');
set(gca,'FontSize',8,'XLim',[0,L*wlag]*Ts,'YLim',[u1,u2],'YTick',
(u1:u2),'YTickLabel',tick);
xlabel('Time'); ylabel('Key');

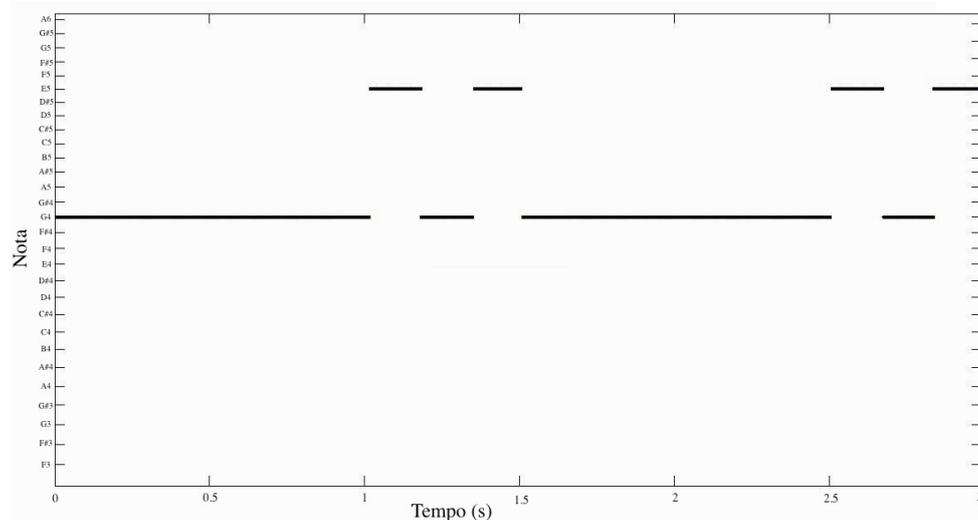
% funzione per rilevare i picchi
function n = trovapicchi(x)

n = find(diff(diff(x) > 0) < 0);
u = find(x(n+1) > x(n));
n(u) = n(u)+1;

```

Per facilitare la visualizzazione del *pitch* estratto con l'AMDF abbiamo inserito nella funzione MATLAB la generazione di un grafico che riporta, nel tempo, il valore della nota rilevata su scala musicale invece che il valore della frequenza in Hz.

Figura 6.2: *Pitch* rilevato dalla funzione AMDF applicata al suono di sirena “artificiale” creato attraverso lo Script 6.1.



La grandezza della finestra temporale è stata scelta di dimensione tale per cui l'AMDF possa riuscire a rilevare frequenze maggiori di 300 Hz²⁰ e il valore dello scostamento tra una finestra e l'altra è stato scelto di grandezza pari ad 1/4 della dimensione della finestra stessa.

Applicando la funzione descritta nello *Script* 6.2 alla sirena generata “artificialmente”, otteniamo il grafico di Figura 6.2. Confrontando questo risultato con la Figura 6.1, si nota come le note tracciate dalla funzione AMDF corrispondono esattamente con quelle della sirena in ingresso al sistema, infatti G4 (Sol quarta ottava) ha frequenza 392 Hz e E5 (Mi quinta ottava) ha frequenza 660 Hz. Questo dimostra che il sistema di rilevazione del *pitch*, nel caso di suono di sirena “puro” e creato “artificialmente”, funziona con estrema precisione.

6.1.2 AMDF applicata a suono di sirena simulato immerso in rumore bianco

Per testare il funzionamento in condizioni meno ideali, abbiamo sommato al suono di sirena, creato nel paragrafo precedente, un segnale della stessa lunghezza contenente rumore bianco. Attraverso la funzione di MATLAB “*wgn*” (*White Gaussian Noise*) riusciamo facilmente a creare il rumore desiderato, semplicemente impostandone la lunghezza e la potenza in dB (relativi ad 1 Watt).

Di seguito lo *script* che crea questo segnale, che lo somma ad un ciclo di suono di sirena e che, successivamente, analizza il segnale risultante con AMDF:

Script 6.3: Generazione del segnale “sirena+rumore bianco” ed estrazione del pitch tramite AMDF.

```
% caricamento del segnale sirena
[u,Fs]=audioread('sirena.wav');

% generazione del rumore bianco (stessa lunghezza del
% segnale sirena, ampiezza variabile)
L=length(x);
w=wgn(L,1,-15);

% somma dei due segnali
x=u+w;

% estrazione e grafico del pitch tramite l'AMDF (Script 6.2)
amdf(x,1/Fs);
```

Assegnando di volta in volta valori di ampiezza diversi per il rumore bianco, abbiamo eseguito lo *Script* 6.3 e abbiamo osservato il comportamento della funzione AMDF. Dagli esperimenti svolti è risultato che, per differenze di ampiezza media tra il suono di sirena e il rumore superiori ai 15 dB, il comportamento dell'AMDF non subisce variazioni e la valutazione del *pitch* collima precisamente con la frequenza

²⁰ nel caso della rilevazione del *pitch* di sirene bitonali in uso al Corpo Nazionale dei Vigili del Fuoco, non dovremmo mai avere frequenze < 372 Hz. Per dare maggiore elasticità al sistema, di contro, abbiamo scelto di fare in modo che esso possa rilevare frequenze fino a 300 Hz. La dimensione della finestra scelta risulta, quindi, essere di $1/(300 \text{ Hz}) = 0.003333 \text{ s}$. Ovvero, in campioni: $0.003333 \times 44100 = 147$. Inoltre, per facilitare le operazioni di calcolo, abbiamo approssimato questo valore per eccesso alla vicina potenza di 2, vale a dire: 256 campioni.

fondamentale della sirena. Quando, invece, la differenza di ampiezza tra i due segnali si riduce a valori inferiori ai 15 dB, l'AMDF comincia a perdere di precisione fino al punto che, con 0 dB di differenza, essa perde del tutto la capacità di tracciare la frequenza fondamentale. La responsabilità di ciò è dovuta alla natura stessa del rumore bianco. In esso, infatti, sono contenute, ad ampiezza costante, frequenze su tutto lo spettro audio e, perciò, quando quest'ultime hanno ampiezza comparabile con le frequenze fondamentali della sirena, l'AMDF risulta priva di picchi di massimo e di minimo e non è quindi possibile estrarre un valore di *pitch* valido.

In Figura 6.3 e in Figura 6.4, sono riportate valutazioni del *pitch* su segnali con differenza tra ampiezza della sirena e ampiezza del rumore rispettivamente di 14 dB e di 0 dB.

Figura 6.3: *Pitch* rilevato dalla funzione AMDF applicata al suono “sirena+rumore” con 14 dB di differenza di ampiezza tra la sirena e il rumore bianco.

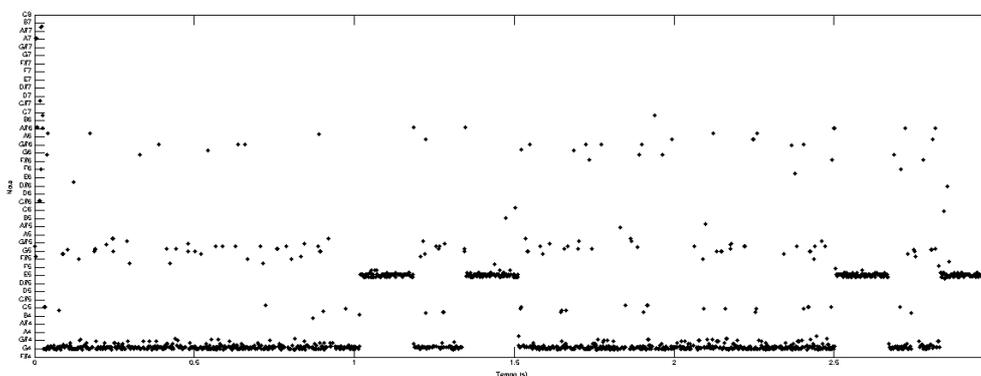
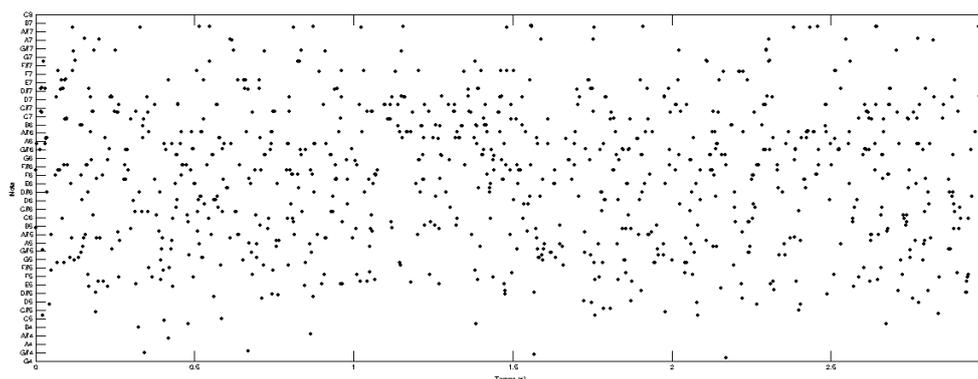


Figura 6.4: *Pitch* rilevato dalla funzione AMDF applicata al suono “sirena+rumore” con sirena e rumore bianco aventi stessa ampiezza.



6.1.3 Filtro Passa-Alto

Da quanto emerso nel paragrafo precedente, al fine di evitare errori nell'applicazione dell'AMDF, occorre mettersi nella condizione per cui il suono di sirena abbia ampiezza maggiore, rispetto alla voce, di almeno 15 dB. Per questo motivo, nel sistema descritto in Figura 4.1 è stato introdotto un filtro passa-alto.

Infatti, mentre per frequenze inferiori ad 1 kHz la differenza di volume delle armoniche della sirena e di quelle della voce risulta quasi nulla, a frequenze più alte le prime cominciano ad avere volume maggiori delle seconde. Per individuare la corretta frequenza di taglio del filtro passa-alto abbiamo scelto, tra una numerosa quantità di file “reali” contenenti il segnale “voce+sirena”, quelli in cui la sirena aveva volume più basso così da poterci mettere nella condizione peggiore. Su questi file abbiamo applicato un filtro passa-alto del quale, a partire da 1 kHz, abbiamo di volta in volta innalzato la frequenza di taglio fino a quando non si riusciva ad ottenere un segnale in cui, la differenza tra l’ampiezza della sirena e l’ampiezza della voce fosse superiore a 15 dB. La frequenza di taglio che abbiamo rilevato da questa sperimentazione/misurazione è 1.8 kHz. Applicando per controprova il filtro a tutti i file abbiamo verificato che, nei segnali in uscita, il volume delle armoniche della sirena si attestassero effettivamente ad un valore di almeno 15 dB sopra il volume della voce.

Di seguito mostriamo lo script necessario alla creazione del filtro passa-alto con MATLAB e in Figura 6.5 e in Figura 6.6 rispettivamente lo spettrogramma di un segnale “voce+sirena” prima e dopo il filtraggio.

Script 6.4: Creazione del filtro passa-alto e filtraggio del segnale “sirena+voce”.

```
% valori filtro passa alto %
Fs = 44100;
Fstop = 1700;          % Frequenza di stop
Fpass = 1900;         % Frequenza di apertura
Astop = 80;           % Attenuazione (dB)
Apass = 1;            % Ripple del PassaBanda (dB)
match = 'stopband';  % Banda che deve coincidere

% Design come object del filtro PassaAlto usando filtro Butter %
h = fdesign.highpass(Fstop, Fpass, Astop, Apass, Fs);
Hd1 = design(h, 'butter', 'MatchExactly', match);
fvtool(Hd1);

%caricamento del file, filtrario e registrazione del segnale % in
uscita al filtro
[x,Fs] = audioread('voce_sirena.wav');
y=filter(Hd1,x);
audiowrite('voce_sirena_hp.wav',y,Fs);
```

Figura 6.5: Spettrogramma del segnale “voce+sirena”.

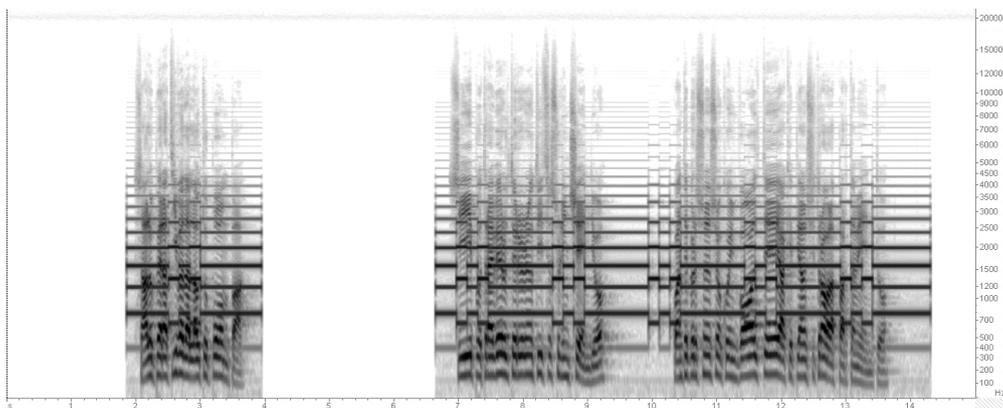
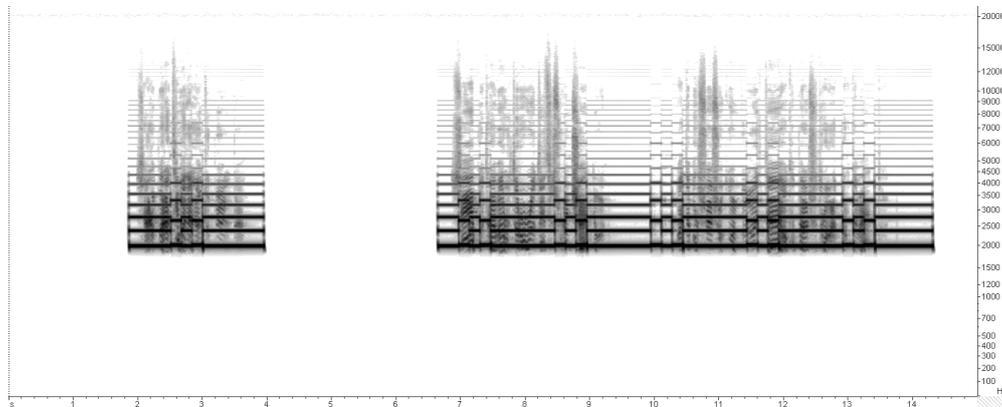


Figura 6.6: Spettrogramma del segnale “voce+sirena” filtrato con filtro passa-alto avente frequenza di taglio 1.8kHz.



6.1.4 Banco di filtri *notch*

Il banco di filtri *notch* è la componente del sistema che attua la vera e propria cancellazione del suono di sirena.

In fase di sperimentazione abbiamo innanzitutto creato una funzione MATLAB che, ricevendo in ingresso la frequenza di campionamento, il segnale da filtrare, la frequenza di centro banda f_0 del primo filtro e il fattore di qualità Q^{21} dei filtri, genera k filtri *notch* centrati su f_0 e sulle relative k armoniche e, attraverso questi, dà in uscita il segnale filtrato. Di seguito si mostra lo *script* che genera tale funzione:

Script 6.5: Banco di filtri *notch*.

```
function [ y ] = notch(x,Fs,f,Q)

%Filtro Notch che elimina f e le relative k armoniche
y = x;
% settaggio del valore di k
k = 30;

for i = 1:k
    wo = f*i/(Fs/2); bw = wo/Q;
    [b,a] = iirnotch(wo,bw,-20);
    y=filtfilt(b,a,y);
end
end
```

Successivamente, per verificare l’oggettiva necessità di utilizzare un filtro adattativo per la cancellazione, abbiamo provato quello che succederebbe muovendosi utilizzando le sole informazioni disponibili a priori, ossia: le frequenze fondamentali ufficiali delle note che compongono la sirena bitonale e il margine di tolleranza concesso su queste frequenze. Con questi soli elementi a disposizione l’unica operazione praticabile è quella del filtraggio contemporaneo del segnale con

²¹ il fattore di qualità Q è definito come il rapporto tra la frequenza di centro banda e la banda eliminata dei filtri *notch*, in formule $Q = f_0 / B$

due banchi di filtri *notch* centrati contemporaneamente su 392 Hz e 660 Hz e impostando il fattore di qualità Q a 10^{22} . Applicando i filtri siffatti a un segnale “sirena+voce” otteniamo il risultato il cui spettrogramma è presentato in Figura 6.7. Si può notare come, insieme al suono di sirena, i filtri cancellino anche gran parte della voce presente. Questo perché l’eccessiva larghezza delle bande dei filtri e l’aver applicato i due blocchi contemporaneamente ha generato l’effetto di “unire” tutte le bande tra loro e creare una sorta di passa-basso. Si può osservare, infatti, che del segnale in ingresso rimangono solo le componenti al di sotto dei 392 Hz che, da un’analisi qualitativa ottenuta con l’ascolto, non sono assolutamente accettabili per essere utilizzate nelle comunicazioni radio in quanto le parole pronunciate si rivelano incomprensibili.

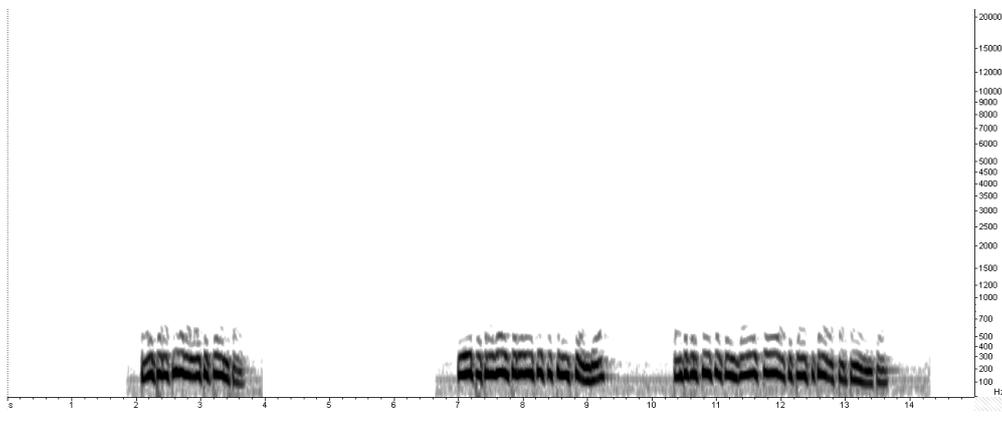
6.1.5 Ritardo Temporale

Il ritardo temporale L da applicare al segnale, prima del filtraggio con il banco di filtri *notch*, è quello che serve a compensare il ritardo introdotto dal filtro passa-alto, dal calcolo dell’AMDF, dall’estrazione del *pitch* e dall’aggiornamento dei k filtri *notch* ai quali va poi sommata la dimensione prescelta per la finestra in cui si calcola l’AMDF. Occorre, quindi, che $L = N_{TOT}$, dove:

$$N_{TOT} = N_{frame} + N_{hp} + N_{AMDF} + N_{pitch} + N_{k_notch}.$$

Osservando i tempi impiegati da MATLAB per svolgere le operazioni suddette, siamo riusciti ad avere una stima dei ritardi che, insieme, concorrono al ritardo complessivo. Questi valori, per una frequenza di campionamento di 44.1 kHz, sono riportati in Tabella 6.1. Prendendo la somma di questi ultimi e aggiungendo i 256 campioni che compongono la finestra che abbiamo scelto per l’analisi dell’AMDF, otteniamo $L = 440$ campioni. Ciò implica che in uscita al sistema avremo un segnale ritardato, rispetto a quello in ingresso, di circa 10 ms.

Figura 6.7: Spettrogramma del segnale “voce+sirena” filtrato con sistema non adattativo basato su due blocchi di 30 filtri *notch* in cascata.



²² infatti abbiamo che per una tolleranza del $\pm 5\%$,
 $Q = f / \{ [f + (f \times 5) / 100] - [f - (f \times 5) / 100] \} = 10$

Tabella 6.1: Valori dei ritardi introdotti nell'esecuzione, con MATLAB, dei vari blocchi di elaborazione del segnale, con frequenza di campionamento 44.1 kHz, $k=12$ e ordine del filtro passa alto 10.

	N_{hp}	N_{AMDF}	N_{pitch}	N_{k_notch}
Valore in campioni	10	120	18	36
Valore in millisecondi	0.18	2.72	0.36	0.91

Lo *script* di MATLAB che ci permette di creare un “filtro” che applichi il ritardo temporale desiderato è il seguente:

Script 6.6: Funzione per l'applicazione di un ritardo temporale L al segnale in ingresso.

```
function [x_d] = ritardo(x,L)

%Funzione che applica al segnale di ingresso il ritardo L
h = dfilt.delay(L);
x_d = filter(h,x);

end
```

6.1.6 Simulazione del sistema completo con segnali reali

Nel Paragrafo 6.1.4 abbiamo appurato la necessità che il filtro che effettua la cancellazione con i *notch* debba essere di tipo adattativo. Nel Paragrafo 6.1.2 abbiamo, invece, valutato quali sono le condizioni che rendono certa l'efficacia dell'utilizzo dell'AMDF per l'estrazione del *pitch* della sirena. A questo punto, possiamo applicare il sistema descritto in Figura 4.1 ad una serie di segnali “voce+sirena” provenienti dal mondo reale. Nei paragrafi precedenti abbiamo illustrato tutte le funzioni MATLAB che sono necessarie per effettuare questa operazione. Con la sperimentazione ci è rimasto soltanto da ricavare il valore di k che identifica il numero di *notch* presenti nel blocco di filtraggio e il fattore di qualità Q su cui questi debbano essere impostati. Il valore di k è stato scelto come il valore più basso che riesce a togliere sufficienti armoniche del suono di sirena in modo tale che, dal punto di vista qualitativo, esso non risulti più udibile. Il fattore Q è stato, invece, scelto come il valore più grande che riesce ad eliminare efficacemente la sirena senza distorcere eccessivamente la voce. Questi due valori sono stati identificati facendo una pura scelta qualitativa soggettiva ottenuta ascoltando i risultati del filtraggio adattativo. Per quanto riguarda k il valore è risultato essere 12, mentre per Q abbiamo individuato il valore 45. I risultati ottenuti imponendo k e Q , così ricavati, e applicando il sistema nella sua interezza, sono molto buoni: mentre la voce rimane altamente comprensibile, la sirena viene cancellata quasi del tutto.

In Figura 6.8 e in Figura 6.9 sono rappresentati, rispettivamente, lo spettrogramma di un segnale “voce+sirena” prima del filtraggio e lo spettrogramma dello stesso dopo il filtraggio. Dal confronto tra le due immagini notiamo come il

filtraggio abbia eliminato i suoni della sirena e le relative armoniche. Purtroppo, però, dalle stesse immagini emerge evidente un limite: nelle fasi di transizione tra una nota e l'altra il sistema non riesce a cancellare del tutto la sirena. Questo, in fase di ascolto, si manifesta con la percezione di brevi ma fastidiosi transitori in corrispondenza delle variazioni di nota. Il motivo per cui questo accade è dovuto da un lato alla presenza di frequenze “spurie”, emesse dalla sirena in fase di passaggio da una nota all'altra e difficilmente tracciabili, dall'altro dalla dimensione della finestra necessaria al calcolo dell'AMDF che risulta troppo elevata per poter seguire variazioni di frequenza così repentine.

Figura 6.8: Spettrogramma del segnale “voce+sirena” prima del filtraggio.

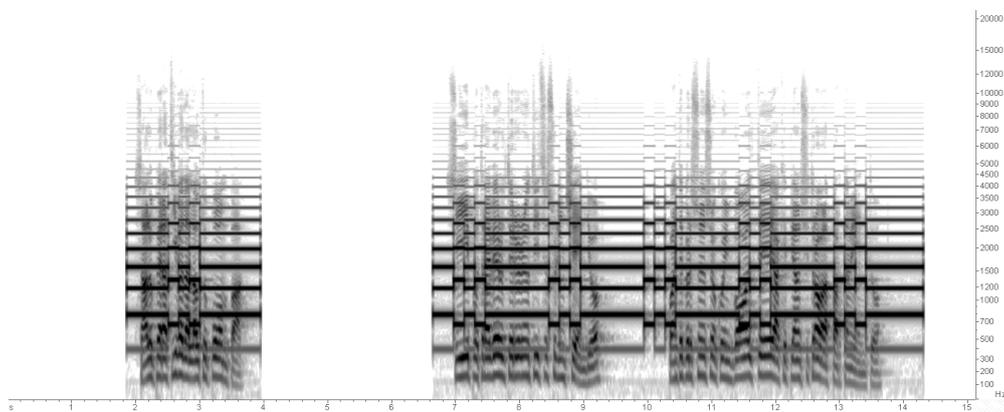
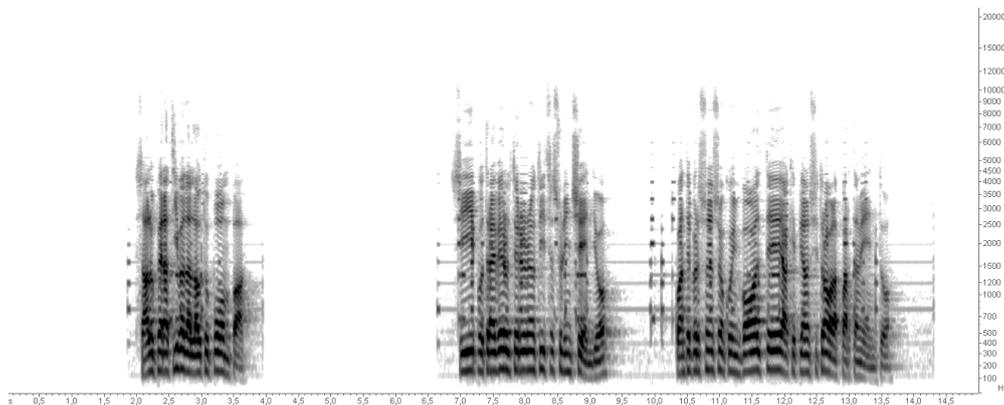


Figura 6.9: Spettrogramma del segnale “voce+sirena” dopo il filtraggio.



6.2 Cancellazione per mezzo dell'IIR *Adaptive Notch Filter*

6.2.1 ANF applicato a sinusoidi pure

Nel Paragrafo 5.2.3 abbiamo visto che l'algoritmo che descrive l'*Adaptive Notch Filter* nella versione *Lattice Semplificata* è il seguente:

$$\begin{bmatrix} g_1 \\ w \end{bmatrix} = \begin{bmatrix} \cos \theta_2 & -\sin \theta_2 \\ \sin \theta_2 & \cos \theta_2 \end{bmatrix} \begin{bmatrix} x(n) \\ x_2(n) \end{bmatrix}$$

$$\hat{y}(n) = \frac{1}{2} [x(n) + w]$$

$$\theta_1(n+1) = \theta_1(n) - \mu \hat{y}(n) x_1(n)$$

$$\begin{bmatrix} x_1(n+1) \\ x_2(n+1) \end{bmatrix} = \begin{bmatrix} \cos \theta_1(n+1) & -\sin \theta_1(n+1) \\ \sin \theta_1(n+1) & \cos \theta_1(n+1) \end{bmatrix} \begin{bmatrix} g_1 \\ x_1(n) \end{bmatrix}.$$

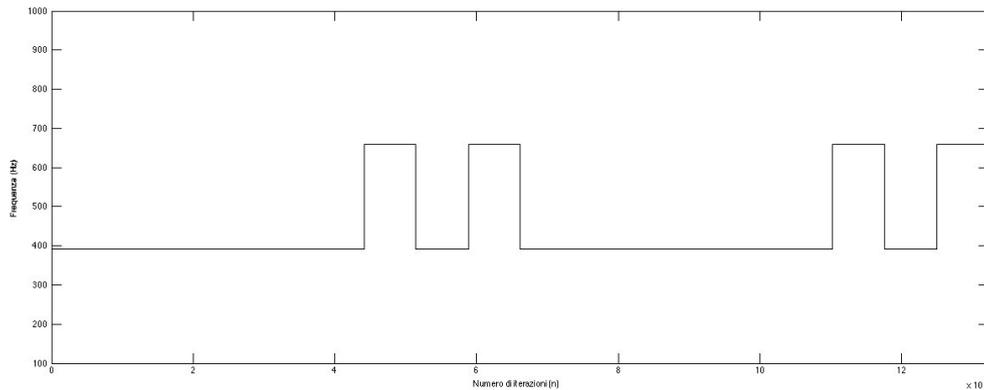
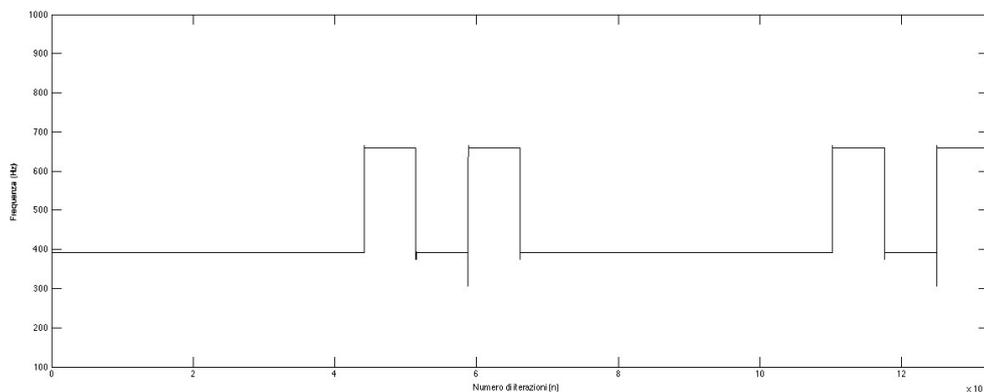
Da questo, risulta relativamente semplice ricavare un serie di comandi MATLAB che eseguano gli stessi passaggi.

Di seguito lo *script* che mette in atto questo algoritmo:

Script 6.7: Simplified Lattice Adaptive IIR Notch Filter.

```
% Simplified Lattice Adaptive IIR Notch Filter Algorithm.
%
% Inizializzazione dei parametri
%
% Frequenza di campionamento
Fs = 44100;
% Stati del filtro adattativo
xst = zeros(2,1);
% Segnali intermedi
temp = zeros(2,1);
pihalf = 0.5*pi;
% Valore iniziale per il parametro "frequenza di notch".
thetal = 0.5(392/Fs-pi);
sth = sin(thetal);
cth = cos(thetal);
bw = 0.20*pi;
% Parametro "larghezza di banda" del filtro notch (theta2)
sth2 = sin(bw);
cth2 = cos(bw);
% Step size del filtro adattativo
mu = 0.023;
%
% Filtro adattativo:
%
% Frequenza iniziale di notch
init_notch_freq=392;
freqold = init_notch_freq/Fs;
for kk=1:iter
    insig = insamples(kk);
    temp = [cth2 -sth2;sth2 cth2]*[insig; xst(2)];
    error = mu*(insig + temp(2));
    y(kk) = error;
    % frequenza normalizzata istantanea rilevata
    theta = theta - error*xst(1);
    % frequenza istantanea rilevata
    freqnew = 0.5*(thetal+pihalf)/pi;
    % Grafico della frequenza rilevata
    if kk>1
        plot([kk-1 kk], [freqold*Fs freqnew*Fs]);
    end
end
```

Applicando lo *script* ad un segnale sinusoidale che varia nel tempo come quello mostrato in Figura 6.10, ovvero che segue gli stessi periodi di evoluzione che

Figura 6.10: Variazione della frequenza della sinusoide del tempo.**Figura 6.11:** Frequenza tracciata dall'ANF applicato al segnale di Figura 6.10.

seguono le frequenze fondamentali in un ciclo di sirena bitonale, otteniamo una stima istantanea del filtro praticamente identica a quella effettiva (Figura 6.11).

6.2.2 ANF applicato a sinusoidi immerse in rumore bianco

Nelle situazioni reali in cui vogliamo applicare l'ANF non ci troveremo mai in condizioni ideali come quelle illustrate nel paragrafo precedente. Per questo, come abbiamo fatto per l'AMDF, in seconda fase di *testing* del filtraggio adattativo, abbiamo provato il suo funzionamento nel caso di una sinusoide che si muove sempre come quella di Figura 6.10 ma a cui abbiamo sommato un rumore bianco di intensità variabile. Questa sperimentazione ha evidenziato come, andando a sommare al suono sinusoidale un rumore a cui, a partire da zero, abbiamo aumentato man mano l'intensità, il sistema presenti sempre maggiore incertezza nel tracciare la frequenza fino ad arrivare al punto (quando sinusoide e rumore hanno lo stesso volume) in cui esso perde del tutto la capacità di indicare la frequenza istantanea del segnale a cui è applicato. Nel nostro sistema possiamo ritenere come valida un'incertezza sulla frequenza tracciata che sia al di sotto del $\pm 2\%$ e questo, secondo le diverse prove effettuate, si manifesta soltanto quando il rumore è inferiore al suono sinusoidale di almeno 25 dB.

Figura 6.12: Frequenza tracciata dall'ANF applicato al segnale di Figura 6.10 a cui è sommato un rumore bianco di ampiezza inferiore ad esso di 25 dB.

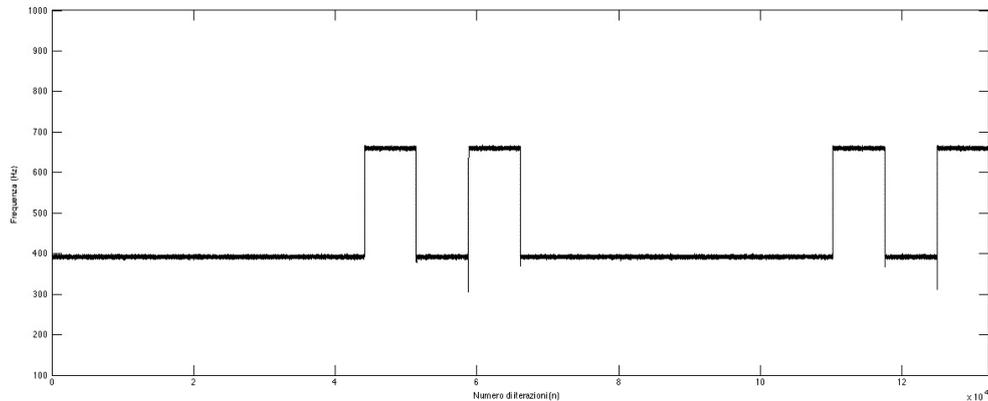
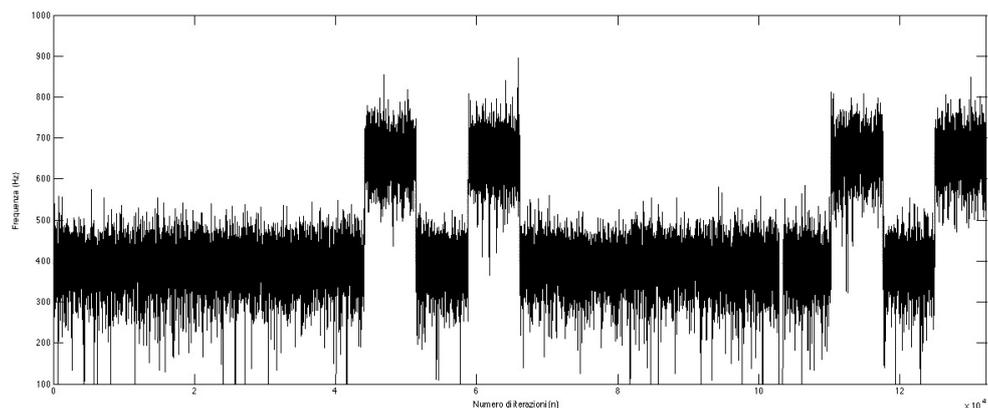


Figura 6.13: Frequenza tracciata dall'ANF applicato al segnale di Figura 6.10 a cui è sommato un rumore bianco di ampiezza inferiore ad esso di 5 dB.



In Figura 6.12 e in Figura 6.13 sono illustrati i grafici che mostrano la stima della frequenza istantanea dell'ANF quando in ingresso abbiamo un suono sinusoidale sommato ad un rumore di intensità inferiore ad esso, rispettivamente di 25 dB e 5 dB.

6.2.3 Filtri passa-banda

Per portarci il più vicino possibile alla condizione in cui abbiamo, in ingresso all'ANF, una sola sinusoide che varia la sua frequenza nel tempo, abbiamo applicato al segnale “sirena+voce” due filtri passa-banda i cui segnali in uscita sono stati successivamente sommati tra di loro. Per quanto trattato nel Paragrafo 5.3 questi filtri passa-banda devono essere settati su le seguenti bande passanti: 372-412 Hz e 627-693 Hz.

Di seguito si espone lo script MATLAB che, attraverso la funzione “fdesign.bandpass” genera i due filtri desiderati, li applica al segnale e somma tra loro i segnali in uscita.

Script 6.8: Filtri passa-banda.

```

% caricamento de segnale voce+sirena
[x,Fs] = audioread('voce_sirena.wav');

% valori filtro passa-banda nota bassa%
FL = 372; %frequenza di taglio basse frequenze
FH = 412; %frequenza di taglio alte frequenze
N = 10; %ordine del filtro
% Design come object del filtro PassaBanda usando filtro Butter %
h = fdesign.bandpass('N,F3dB1,F3dB2',N,FL,FH,Fs);
HdL = design(h,'butter');
x_bpL=filter(HdL,x);

% valori filtro passa-banda nota alta%
FL = 627; %frequenza di taglio basse frequenze
FH = 693; %frequenza di taglio alte frequenze
N = 10; %ordine del filtro
% Design come object del filtro PassaBanda usando filtro Butter %
h = fdesign.bandpass('N,F3dB1,F3dB2',10,FL,FH,Fs);
HdH = design(h,'butter');
x_bpH=filter(HdH,x);

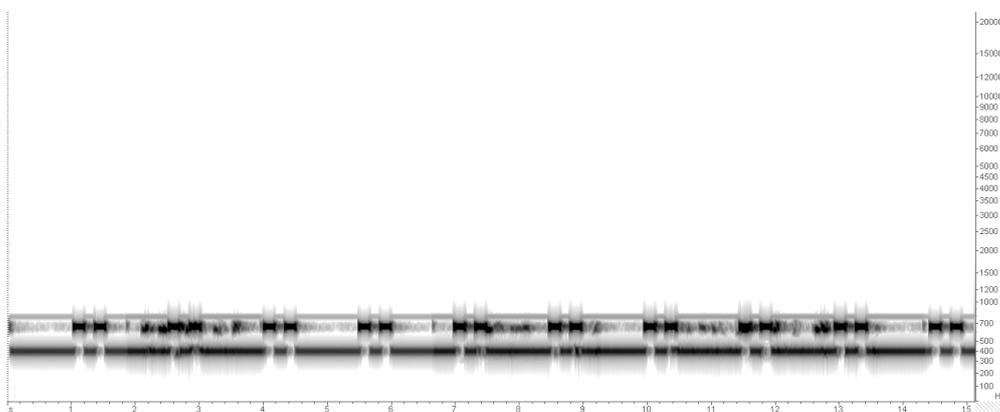
%Somma e normalizzazione
y=x_bpL+x_bpH;
y=0.99*y/max(y);

% salvataggio del segnale contenente la somma delle due bande passanti
audiowrite('voce sirena bp.wav',y,Fs);

```

Dallo spettrogramma di Figura 6.14 si può osservare il risultato ottenuto applicando questi due filtri allo stesso segnale “voce+sirena” è rappresentato in Figura 6.8. Notiamo come il segnale ricomposto dalla somma dei segnali in uscita ai due filtri passa-banda contenga, quasi esclusivamente, le frequenze fondamentali che descrivono l’andamento della sirena. Purtroppo in esso sono presenti anche delle “impurità” dovute ai passaggi discontinui del suono di voce su quelle frequenze. Nonostante ciò, la differenza di ampiezza tra sirena e voce rimane, comunque, sufficientemente buona affinché l’ANF riesca a tracciare il *pitch* con efficacia.

Figura 6.14: Spettrogramma del segnale “voce+sirena” rappresentato in Figura 6.8 filtrato con i filtri passa-banda creati dallo Script 6.8.



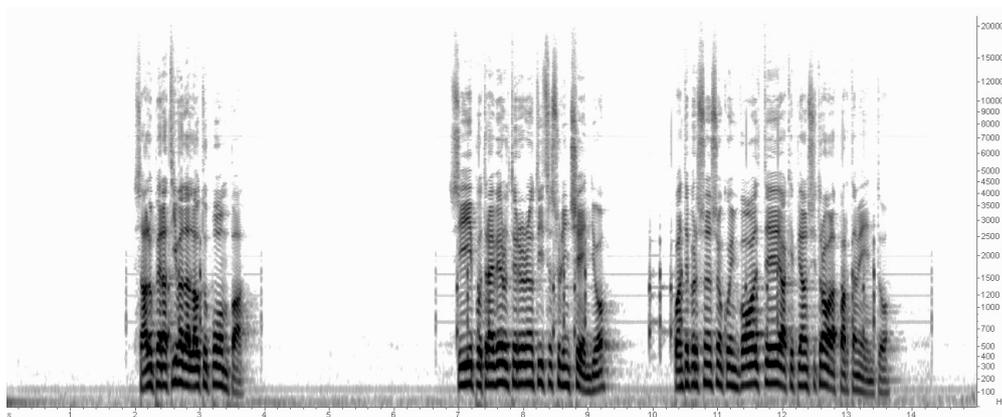
6.2.4 Simulazione del sistema completo con segnali reali

Avendo a disposizione le funzioni MATLAB che simulano il comportamento di tutti i blocchi che compongono il sistema di cancellazione di Figura 5.1, abbiamo potuto applicare il metodo nella sua interezza a segnali provenienti dal mondo reale. Per il banco di filtri *notch* abbiamo usato le stesse impostazioni emerse dalla sperimentazione esposta nel Paragrafo 6.1.4, ossia: $k = 12$; $Q = 45$.

Da un'analisi qualitativa effettuata ascoltando i segnali in uscita dal filtraggio adattativo, abbiamo verificato come, effettivamente, il suono di sirena venga cancellato e come la voce, invece, riesca a passare attraverso di esso senza subire eccessive distorsioni. Anche in questo caso permangono dei piccoli disturbi presenti in corrispondenza degli istanti di cambiamento tra una nota e l'altra ma risultano molto meno invadenti in confronto alla cancellazione effettuata per mezzo dell'AMDF. Questo miglioramento è dovuto alla "continuità" con cui l'ANF traccia, istante per istante il *pitch* del segnale a cui è applicato. Infatti, in fase di cambiamento tra una nota e l'altra - ricordando che in ingresso al filtro adattativo giungono solo due bande costruite intorno alla probabile frequenza fondamentale del suono di sirena -, esso, per effetto del tempo necessario alla sua convergenza, non passa istantaneamente da una frequenza bassa ad una alta (o viceversa) ma indica, nel passaggio, una serie di frequenze intermedie. La "lentezza" di convergenza dell'ANF, che normalmente rappresenta un fenomeno intrinseco indesiderato, in questo caso fa, invece, sì che il sistema si adatti in maniera più precisa al generatore reale del suono di sirena, incapace, evidentemente, di far variare istantaneamente il suono di sirena da una nota all'altra.

In Figura 6.15 mostriamo lo spettrogramma relativo allo stesso segnale rappresentato in Figura 6.8. Dall'immagine si può osservare come le armoniche della sirena siano state del tutto cancellate e come le piccole componenti indesiderate che sono ancora presenti, abbiano volume minore rispetto al sistema di cancellazione effettuato per mezzo dell'AMDF.

Figura 6.15: Spettrogramma del segnale "voce+sirena" rappresentato in Figura 6.8, dopo la cancellazione della sirena effettuata con il sistema che utilizza l'ANF.



6.3 Confronto tra i due metodi

Confrontando i risultati ottenuti applicando i due metodi ad una serie di segnali provenienti dal campionamento di differenti microfoni connessi alle radio dei veicoli in dotazione al Corpo Nazionale dei Vigili del Fuoco, abbiamo osservato che, in linea di massima, il sistema che utilizza l'ANF effettua una cancellazione della sirena migliore rispetto a quello che utilizza l'AMDF. Questo è dovuto, come spiegato nel Paragrafo 6.1.5, alla necessità di quest'ultima tecnica di “finestrare” il segnale per l'estrazione del *pitch*. Ci sono però alcune condizioni che fanno sì che questo sistema funzioni meglio di quello che utilizza l'ANF. Infatti, nel caso di un operatore che parli molto forte o che addirittura urli durante una conversazione radio, esso continua, con più facilità, ad essere in grado di attuare l'eliminazione mentre l'ANF perde del tutto la capacità di tracciare il *pitch* della sirena e, di conseguenza, il sistema ad esso connesso non riesce ad effettuare la cancellazione. In questo caso ci ritroviamo in una situazione paradossale in cui: minore è la presenza di suono di sirena, più difficile risulta la sua cancellazione. Ma, poiché, in questa circostanza appena descritta, il volume della voce risulta già di per sé ampiamente maggiore rispetto a quello della sirena, e risulta quindi non strettamente necessario un sistema che attui la cancellazione, possiamo comunque ritenere globalmente più valido il sistema che utilizza l'ANF.

6.4 Applicazione dei sistemi ad altre tipologie di sirene

Al fine di testare una possibile valenza generale dei metodi proposti, modellati per trattare il caso particolare delle sirene bitonali in dotazione al Corpo Nazionale dei Vigili del Fuoco, abbiamo provato ad applicare gli stessi ad altre tipologie di sirene reperendo i relativi file audio dal *web*. Dalle prove effettuate abbiamo osservato come, con le dovute modifiche, i sistemi riescano in alcuni casi a funzionare correttamente.

Nel caso di sirene di tipo bitonale, una volta a conoscenza delle relative specifiche tecniche a cui devono attenersi (in modo da poter settare correttamente i filtri passa-alto o passa-banda), i due sistemi funzionano come per le sirene per cui sono stati progettati. Nel caso, invece, di sirene di tipo *Yelp* o *Wail*, il metodo che utilizza l'ANF non riesce a tracciare il *pitch* della sirena e ad effettuare la sua cancellazione, non essendo applicabili dei filtri passa-banda adeguati. Viceversa, il metodo con AMDF, nonostante a tratti generi in uscita *click* e *glitch* dovuti all'inadeguatezza della dimensione della finestra di analisi rispetto alla variazione repentina della frequenza fondamentale della sirena, riesce comunque ad applicare una generica attenuazione.

In conclusione, possiamo affermare che:

- il metodo che utilizza l'AMDF è un metodo globalmente valido per l'attenuazione (che solo in alcuni casi risulta essere vera e propria cancellazione) delle sirene, di qualsiasi tipo esse siano;

- il metodo che utilizza l'ANF è un metodo globalmente valido per la cancellazione delle sirene di tipo bitonale quando siano note, a priori, le loro caratteristiche fondamentali.

Capitolo 7

Conclusioni e sviluppi Futuri

In questa tesi abbiamo discusso la problematica inerente al disturbo generato dagli avvisatori acustici, installati sugli automezzi del Corpo Nazionale dei Vigili del Fuoco, quando accade che il suono che essi emettono, per l'elevato volume con cui è diffuso, interferisce con le comunicazioni effettuate via radio, disturbandole notevolmente.

Nel nostro lavoro abbiamo proposto due metodologie che hanno come obiettivo quello di migliorare le prestazioni dei sistemi preesistenti e di ottenere la cancellazione del suono di sirena servendosi esclusivamente del segnale trasdotto dal microfono della radio. Dalle sperimentazioni effettuate è emerso come i due sistemi proposti riescano a raggiungere gli obiettivi prefissati ovvero l'attenuazione del disturbo provocato dalle sirene nelle comunicazioni radio e il miglioramento, rispetto alle metodologie passate, della qualità percepita del parlato.

Entrambi i metodi agiscono tracciando la frequenza fondamentale della sirena e utilizzano questa informazione per filtrare via il disturbo con opportuni filtri *notch*. Per fare ciò, un metodo esamina continue "finestre" di segnale, precedentemente filtrato con un filtro passa-alto (per eliminare la parte più significativa dello spettro della voce), l'altro applica un filtraggio adattativo a due bande di segnale poste intorno alle possibili frequenze fondamentali sulle quali si sviluppa la sirena. Il primo ha il vantaggio di poter tracciare il suono di sirena qualsiasi sia il suo sviluppo e, di contro, la lacuna di non procedere con la cancellazione in maniera continuativa ma ad intervalli proporzionati alla dimensione della finestra di analisi. Il secondo, che invece ha il beneficio di tracciare il *pitch* della sirena istante per istante, ha il difetto di dover sapere a priori quale sia l'andamento teorico della stessa (in modo da poter settare correttamente i filtri passa-banda a monte del filtraggio adattativo).

Nonostante lo sviluppo dei nostri metodi si sia concentrato nella risoluzione del problema riferendosi al caso specifico delle sirene del Corpo Nazionale dei Vigili del Fuoco, successivamente abbiamo effettuato ulteriori sperimentazioni su sirene in dotazione ad altri servizi di emergenza. Da queste prove è emerso che, a patto di apportare alcune modifiche relative ai tipi di sirena trattati, i metodi proposti hanno una validità di carattere più generale.

7.1 Possibili sviluppi futuri

I risultati incoraggianti che abbiamo ottenuto dalle simulazioni degli algoritmi in ambiente virtuale, denotano che la metodologia può essere implementata in un sistema reale che utilizzi un *Digital Signal Processor*. Il passo successivo si delinea come l'installazione del sistema e il *testing* dello stesso sugli automezzi del Corpo Nazionale dei Vigili del Fuoco. Dopo aver effettuato ciò si renderà necessario progettare la modalità in cui esso possa essere inserito all'interno degli apparati radio, a metà strada tra la campionatura del segnale proveniente dal microfono e la sua trasmissione via radio.

Riteniamo che una linea di sviluppo futuro possa consistere nello sviluppo di un *Adaptive Notch Filter* che riesca a tracciare non solamente una sinusoide, come per il filtro qui presentato ed utilizzato, ma più sinusoidi contemporaneamente, sempre in presenza di un rumore di sottofondo. Questo potrà dare ulteriore precisione e una valenza più ampia al sistema complessivo.

Appendice A

MINISTERO DEI TRASPORTI, Decreto Ministeriale 17/10/1980 (G.U. n. 310 del 12/11/1980)

“Modifiche sperimentali delle caratteristiche acustiche dei dispositivi supplementari di allarme da applicare ad autoveicoli e motoveicoli adibiti a servizi antincendi e ad autoambulanze”

Visti gli articoli 210 e 211 del regolamento per l'esecuzione del testo unico per l'esecuzione delle norme sulla disciplina della circolazione stradale approvato con decreto del Presidente della Repubblica 30 giugno 1959, n. 420, concernenti le caratteristiche dei dispositivi supplementari d'allarme;

Visto l'articolo 465 del regolamento suddetto, che demanda al Ministero dei Trasporti la facoltà di concedere dispensa dalla osservanza, tra l'altro, dei citati articoli 210 e 211, quando la dispensa sia ritenuta necessaria per l'applicazione dei dispositivi in esperimento;

Considerato che il Corpo nazionale vigili del fuoco ha chiesto la dispensa prevista dall'art. 465 citato per dotare i propri veicoli di una sirena bitonale da impiegare nei servizi di soccorso;

Visto il favorevole parere del Ministero dell'interno;

Considerato che l'Associazione volontaria di pronto soccorso ed assistenza (SOS), di Milano ha chiesto analoga dispensa;

Ritenuta la necessità di migliorare l'efficienza dei dispositivi supplementari di allarme, facilitando nel traffico la localizzazione della provenienza dei veicoli che li impiegano, e la necessità di contenere la potenza sonora di tali dispositivi, per ridurre il disturbo acustico arrecato;

Riconosciuta l'opportunità di differenziare il dispositivo supplementare di allarme dei veicoli adibiti a servizi di soccorso da quello dei veicoli adibiti a servizi di polizia;

Decreta

Art. 1: In deroga alle disposizioni degli articoli 210 e 211 del decreto Presidente della Repubblica 30 giugno 1959, n. 420, gli autoveicoli e motoveicoli adibiti a servizi antincendi e le autoambulanze di cui all'art. 126 del decreto Presidente della Repubblica 15 giugno 1959, n. 393, possono essere muniti, in via sperimentale, di un dispositivo supplementare di allarme secondo le prescrizioni degli articoli seguenti.

Art. 2: Il dispositivo supplementare di allarme di cui all'art. 1 deve presentare, in condizioni di campo libero, le caratteristiche acustiche seguenti:

1.

emissione di due suoni di frequenza rispettivamente pari a 392 Hz (Sol naturale) e 660 Hz (Mi naturale). Sui valori citati è consentita una tolleranza in più o in meno del 5%;

2.

il livello di pressione sonora globale per ciascuna nota, misurata in condizioni di campo libero ed alla tensione nominale misurata ai capi del dispositivo, deve essere massimo sull'asse principale dell'apparecchio e compreso tra i 115 e i 125 dB a due metri di distanza da esso;

3.

lo spettro acustico di ciascuno dei due suoni, rilevato nelle condizioni di cui al punto 2) con filtri a terze di ottava, deve risultare, su tutta la gamma di frequenze, almeno 4 dB inferiore al livello di pressione sonora lineare. Nella banda da 1800 a 4500 Hz, nelle stesse condizioni di misura, il livello di pressione sonora deve risultare non inferiore a 105 dB;

4.

un ciclo acustico completo comprende un suono a 392 Hz per una durata di 1/3 della durata totale del ciclo, un suono a 660 Hz per la durata di 1/18 della durata totale del ciclo, un suono a 392 Hz per una durata di 1/18 della durata totale del ciclo, un suono a 660 Hz per la durata di 1/18 della durata totale del ciclo, un suono a 392 Hz per una durata di 1/3 della durata totale del ciclo, un suono a 660 Hz per la durata di 1/18 della durata totale del ciclo, un suono a 392 Hz per una durata di 1/18 della durata totale del ciclo e, infine un suono a 660 Hz per la durata di 1/18 della durata totale del ciclo. Sulle frazioni di tempo succitate è ammessa una tolleranza in più o in meno del 5%. I suoni devono susseguirsi senza interruzioni apprezzabili e senza sovrapposizioni. La durata totale del ciclo completo sopradescritto deve essere di 3 sec. \pm 0,5 sec. Fra un ciclo acustico completo ed il successivo può sussistere un intervallo la cui durata massima non deve comunque superare i 0,2 sec. Le condizioni di cui sopra devono essere soddisfatte anche quando la tensione ai morsetti del segnalatore varia da -10% a +10% rispetto alla tensione nominale;

5.

una volta azionato per mezzo di un impulso di azionamento di durata inferiore al ciclo acustico descritto al punto 4), questo deve arrestarsi e completarsi automaticamente;

6.

il ciclo acustico descritto al punto 4) deve potersi ripetere senza danneggiamenti o arresti anormali del dispositivo per 1 ora consecutiva, alimentando il dispositivo alla tensione nominale e ad una temperatura di $20 \pm 2^{\circ}\text{C}$. Al termine della prova il dispositivo dovrà arrestarsi e riavviarsi.

Art. 3: Il dispositivo supplementare di allarme di cui all'art. 1, applicato su un veicolo nelle condizioni normali di montaggio, alimentato dalla batteria carica, deve dare un livello sonoro soggettivo, misurato sull'asse del veicolo, a 30 metri davanti ad esso, non inferiore a 90 dB.

Art. 4: Il dispositivo deve essere riconosciuto conforme alle prescrizioni stabilite agli articoli 2 e 3 dal Ministero dei trasporti – Direzione generale della motorizzazione civile e dei trasporti in concessione. Ogni esemplare deve essere conforme al prototipo approvato e deve recare un marchio costituito dalle lettere KSP seguite dal numero di riferimento relativo al riconoscimento.

Art. 5: Il presente decreto sarà pubblicato nella Gazzetta Ufficiale della Repubblica italiana.

Roma, addì 17 ottobre 1980
Il Ministro FORMICA (10148)

Bibliografia

- [1] J. D. Gibson. *The Communications Handbook*. CRC Press, 2002.
- [2] R. J. Baken. *Clinical Measurement of Speech and Voice*. Taylor and Francis, London, 1987.
- [3] B. Boren, A. Roginska, B. Gill. *Maximum Averaged and Peak Levels of Vocal Sound Pressure*. In Proc. 135th AES Convention, New York, USA, 2013.
- [4] B. Horvei, G.E. Ottesen, J. Svean. *Cancellation of siren sound in emergency vehicle communication equipment*. In Proc. NORSIG, Tromsø, Norway, 1997.
- [5] C. G. Guy, R. S. Sherratt, D. M. Townsend. *Cancellation of Siren Noise From Two Way Voice Communications Inside Emergency Vehicles*. Computing & Control Engineering Journal, 2002.
- [6] R. S. Sherratt. *Results of a DSP Based Adaptive Cancellation Filter System For The Chiro, Pulsar and Localiser Sirens*. School of Computer Science, Cybernetics and Electronic Engineering, The University of Reading, UK, 2001.
- [7] H. Ding, J. Lu, X. Qiu, B. Xu. *An Adaptive Speech Enhancement Method for Siren Noise Cancellation*. Applied Acoustics, 65:385-399, 2004.
- [8] B. Friedlander, B. Porat. *The Modified Yule-Walker method of ARMA spectral estimation*. IEEE Transaction on Aerospace Electronic Systems, 20(2):58-73, 1984.
- [9] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, R. C. Goodlin. *Adaptive Noise Cancelling: Principles and Applications*. In Proc. IEEE, pp. 1692-1717, 1975.
- [10] M. Luise, G. M. Vitetta. *Teoria dei segnali*. McGraw-Hill, III ed., 2009.
- [11] M. J. Ross, H. L. Shaffer, A. Cohen, R. Freudberg, H. J. Manley. *Average Magnitude Difference Function Pitch Extractor*. IEEE Trans. on Acoustics, Speech and Signal Processing, 22(5):353-362, 1974.
- [12] J. G. Proakis, D. K. Manolakis. *Digital Signal Processing*. Pearson, 2006.

[13] A. V. Oppenheim, R. W. Schaffer, J. A. Buck. *Discrete-time signal processing*. Prentice Hall, pp. 468–471, 1999.

[14] J. R. Ziegler. *Time-sharing Data Processing Systems*. Prentice Hall, 1967.

[15] P. Regalia. *Adaptive IIR filtering in signal processing and control*. Marcel Dekker, New York, USA, 1995.