

## 第五章 實驗結果與討論

本論文中所做的實驗是為了驗證所提出的動態調整辦法可以有效的改善語音品質，讓語音在時間延遲的情形下，能夠讓語音更平順的播出。因為網路狀況不是我們所能決定的，所以所需要的網路延遲數據，是透過 OPNET 來模擬網路環境得到，模擬的網路將在 5.1 節作介紹。

在這一章將探討不同的緩衝區調整辦法與封包遺失率、以及緩衝個數的關係，並以主觀之人耳聽覺測試來評斷不同緩衝區調整方式對語音的影響，與客觀之 SSSNR 值評估方式。

### 5.1 VoIP 網路模擬環境

本論文的網路模擬環境，是使用 Winnt 上的網路模擬軟體（OPNET）來模擬實際網路可能發生的情形，而本論文模擬的網路環境架構為兩個區域網路之間的語音傳送，並且在兩個區域網路中分別設定  $n$  台電腦，讓兩個區域網路的電腦彼此溝通，兩個區域網路的每台電腦都是以 G.729 為語音壓縮標準並且每台電腦可以設定在不同的時間點傳送 G.729 的語音資料，然後藉由 OPNET 得到在這種網路環境下每個語音封包時間延遲數據，然後把網路延遲數據套用到論文所提出的不同的動態緩衝區調整方式，藉由 SSSNR 來比較語音在透

過動態緩衝區調整之後是否能達到良好的語音品質。

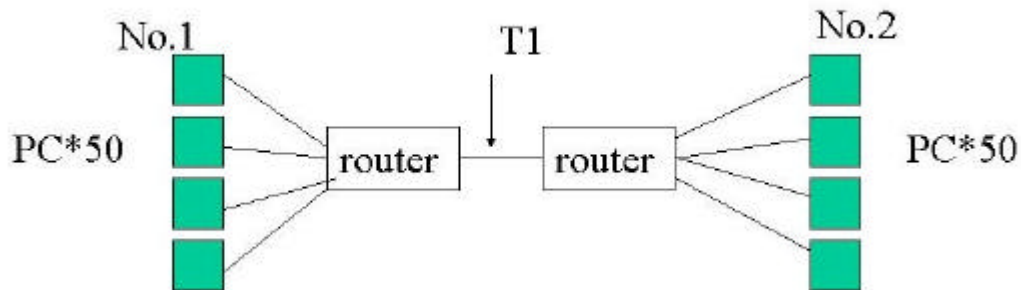


圖 5.1 VoIP 網路模擬環境

圖5.1是本論文所模擬的網路架構，而所量測的網路延遲數據，只有量測 No.1 和 No.2 之間兩台電腦之間的網路延遲數據，在兩個區域網路之間是以兩台路由器（Router）做連接，中間是以T1（1.544Mbps）連接兩台路由器。並藉由所得到的時間延遲數據來分析本論文所建立的動態調整辦法是否可以有效的改善語音的品質，在這裡的網路延遲數據是經由100次的實驗數據之後的平均值。

透過網路模擬軟體 OPNET，模擬所要的網路架構，可以得到每個封包經過 G.729 語音壓縮後，經過網路傳送到接收端所得到的網路延遲（Network end-to-end delay）分布圖（圖 5.2），從圖形分布中可以得知，在網路傳送的過程中，在一開始是處於少數語音資料傳送，

經過一段時間之後，突然有多部電腦（PC）同時在網路上傳送 G.729 的語音資料，造成網路時間延遲瞬間上升（圖 5.3），並且這一段時間是處於巨大時間延遲，經過一段時間之後網路又呈現快速下降，回復到剛開始時的網路狀況（圖 5.4）。

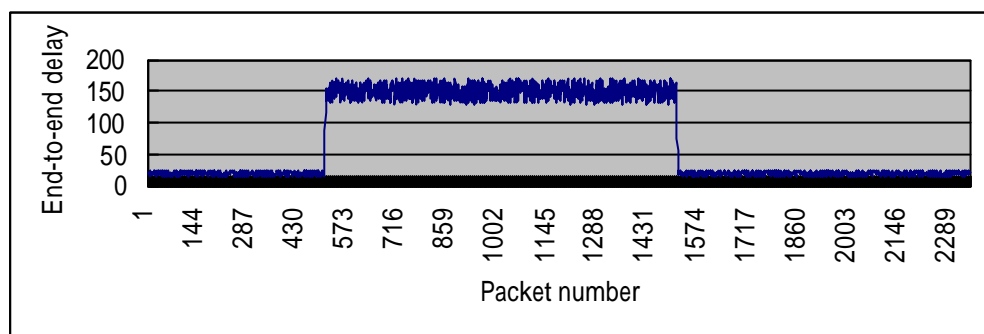


圖 5.2 VoIP 網路模擬時間延遲

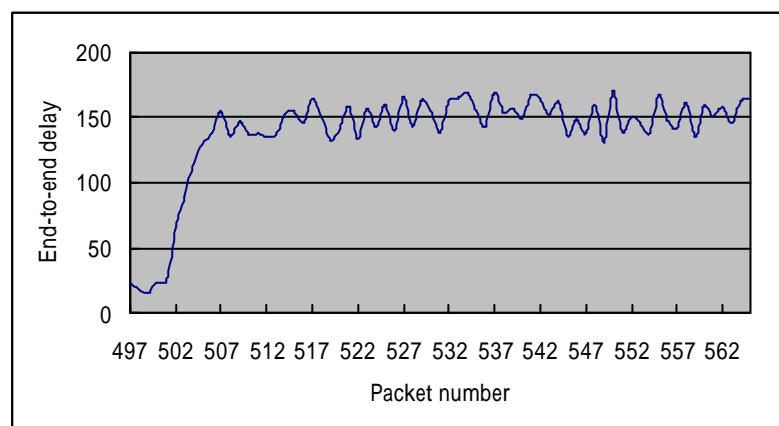


圖 5.3 網路模擬時間延遲快速上升

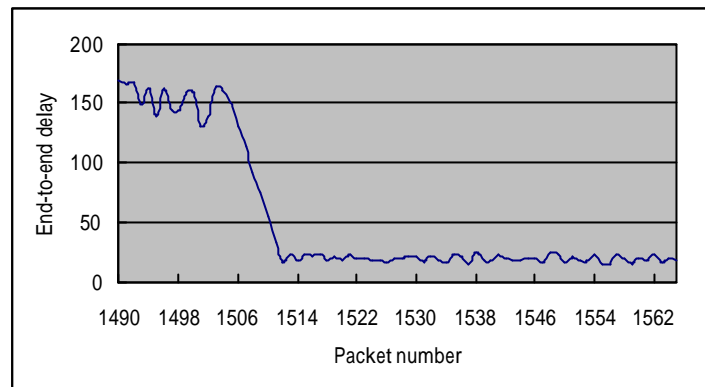


圖 5.4 網路模擬時間延遲快速下降

## 5.2 原始語音分布

在做不同緩衝區調整之前，必須對要測試的語音有基本的認識，包含語音有多少意義封包、靜音封包（圖 5.5）的分布和此段語音共多少個語音段落等，這對資料對接下來的實驗將有重要的影響，圖 5.6 為此語音的語音波形，這個語音檔案包含六段不同的語音，其中有中文發音、英文發音、男生發音、女生發音。

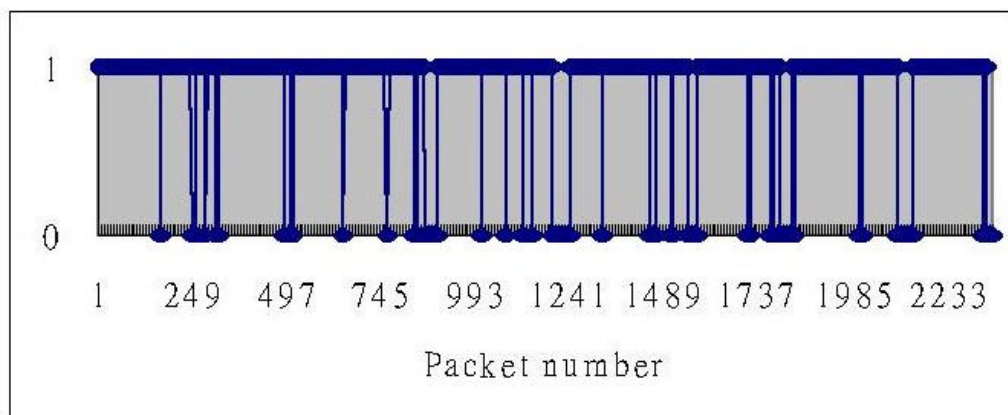


圖 5.5 原始語音封包分布

透過 G.729 語音編碼和語音段落 (Talk Spurt) 的計算可得到表 5.1, 在圖 5.5 中, 1 代表這個封包為有意義的封包, 0 表示此封包為靜音封包。

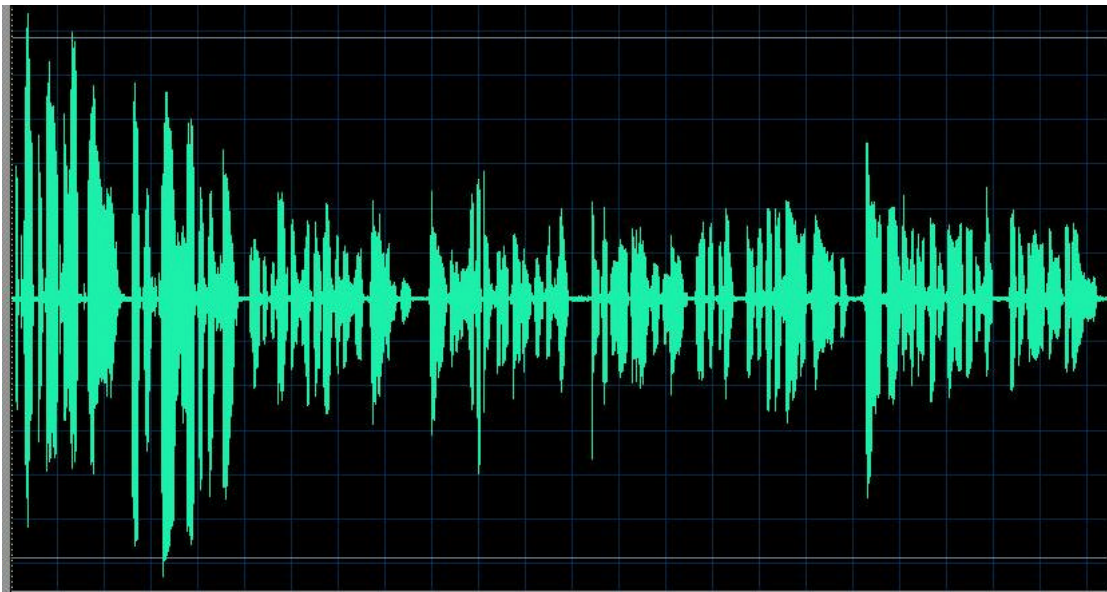


圖 5.6 原始語音波形

Total frame	2344
Total silence	282
Talkspurt	21

表 5.1 原始語音封包表

### 5.3 晚到封包丟棄 (Drop) 驗證

在語音封包該播放時此封包卻晚到而來不及播放時，採用封包丟棄 (Drop) 的方式來驗證不同的緩衝區調整方式，看其調整效果是否有改善。表 5.2 是在封包丟棄的方式下，不同緩衝區調整方式和封包遺失率 (Loss rate) 的比較。

no	buffer	buffer	buffer	buffer	buffer	buffer	buff	late-rate	max-jitt
buffer	1	2	3	4	5	er 6			er
247	0.3036	0	0	0	0	0	0	0	0
488	0.6058	0	0	0	0	0	0	0	0
757	0.7955	0.5018	0.2973	0.0446	0.0037	0.0037	0	0.0446	0
1068	0.5241	0.3247	0.1414	0.0032	0	0	0	0.0032	0.1414
1324	0.4291	0.3122	0.1215	0	0	0	0	0	0
1549	0.5422	0.4133	0.2177	0.0311	0	0	0	0.2177	0
1765	0.4768	0	0	0	0	0	0	0	0
2000	0.4042	0	0	0	0	0	0	0	0
2328	0.3164	0	0	0	0	0	0	0	0

表 5.2 在封包丟棄下,不同緩衝區調整方式和封包遺失率  
的關係

從表中可以知道語音封包透過網路延遲後，會造成大量的封包遺失，讓語音品質下降，從表 5.2 的第一項 (No buffer)中，便可以知道沒有在解碼端前配置任何緩衝區的後果，是會讓語音封包遺失率增

大，透過定量的緩衝個數 (Buffer 1~Buffer 6) 可以明顯看到，語音封包遺失率正慢慢的減少，直到定量緩衝 6 個封包之後，就沒有封包遺失率產生，但是增加大量的緩衝個數是不智的，相對的也增加語音播放的時間。再比較 Late-rate 和 Max-jitter 緩衝區調整方式，知道這兩種調整方式也可以得到較少的封包遺失率，單單看封包遺失率是不太客觀，所以請參考圖 5.7 緩衝區調整方式和緩衝個數的比較，這樣較能看出每種緩衝區的調整效果。

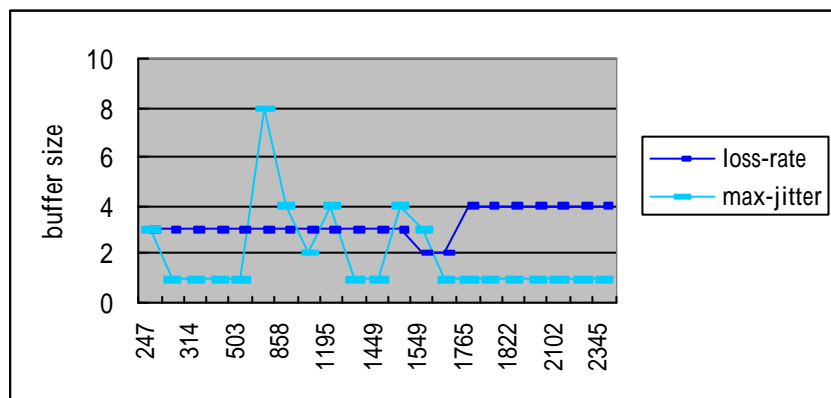


圖 5.7 不同緩衝區調整方式和緩衝個數的比較

從圖 5.7 可以知道每一個語音段落所緩衝的個數，在這裡值得注意的地方是，以 Late-rate 的方式調整緩衝區，雖能有效的降低封包遺失率，但是卻不能及時反應網路狀況，所以在第 15 個語音段落其緩衝區仍往上提升，這對語音而言無疑是增加多餘的延遲時間，反觀利用 Max-jitter 來調整的情形，在第 14 個語音段落之後，每一個語音段

落只需 1 個緩衝個數，便可以達到 0 封包遺失率，所以在緩衝個數和封包遺失率的考慮下 max-jitter 是較好的緩衝區調整方式，至於定量的調整方式，雖然在這也有不錯的效果，但這種方式完全不能有效反應網路狀況，一般都不採用這種方式來調整緩衝區。

為更進一步了解緩衝區調整的效果，利用封包遺失率與網路延遲分布的關係加以分析，請參考圖 5.8。

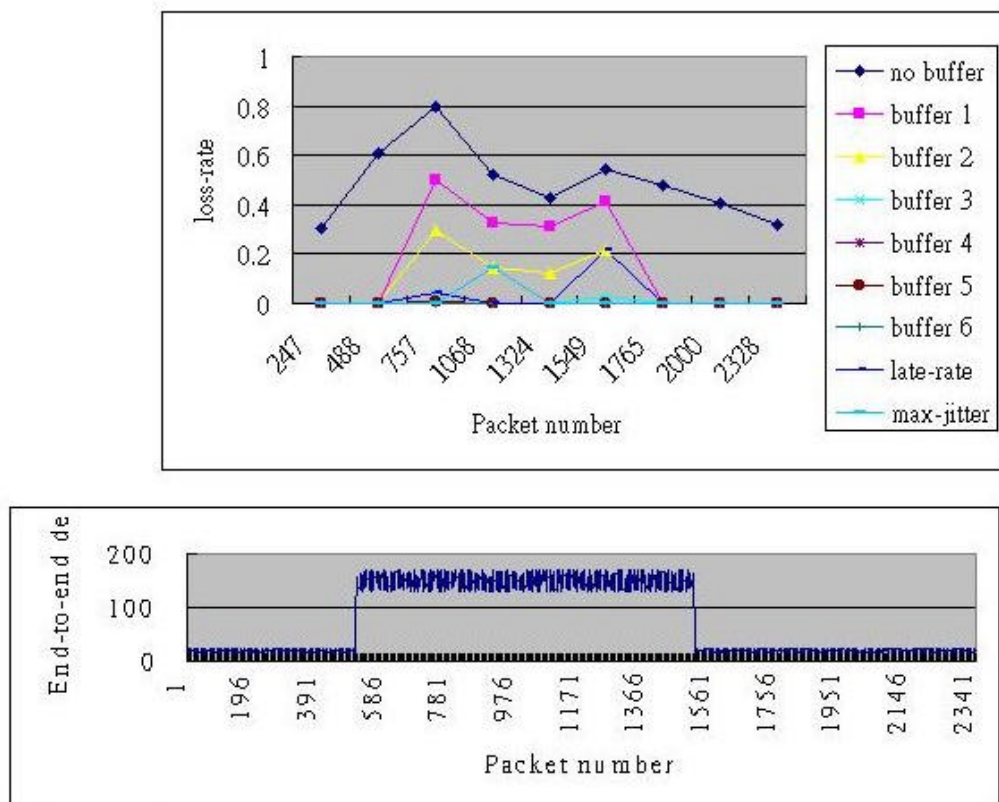


圖 5.8 封包遺失率與網路延遲分布

從上圖 5.8 中可以知道開始發生封包遺失率是起源於網路延遲分

布突然上升時，面對突然快速上升的網路延遲時，定量的靜態調整方式是不能反應過來，所以在定量緩衝 1~3 個封包數，仍然有較大的封包遺失率，但是動態的調整方式面對突然上升的網路延遲確可以有效的降低封包遺失率。

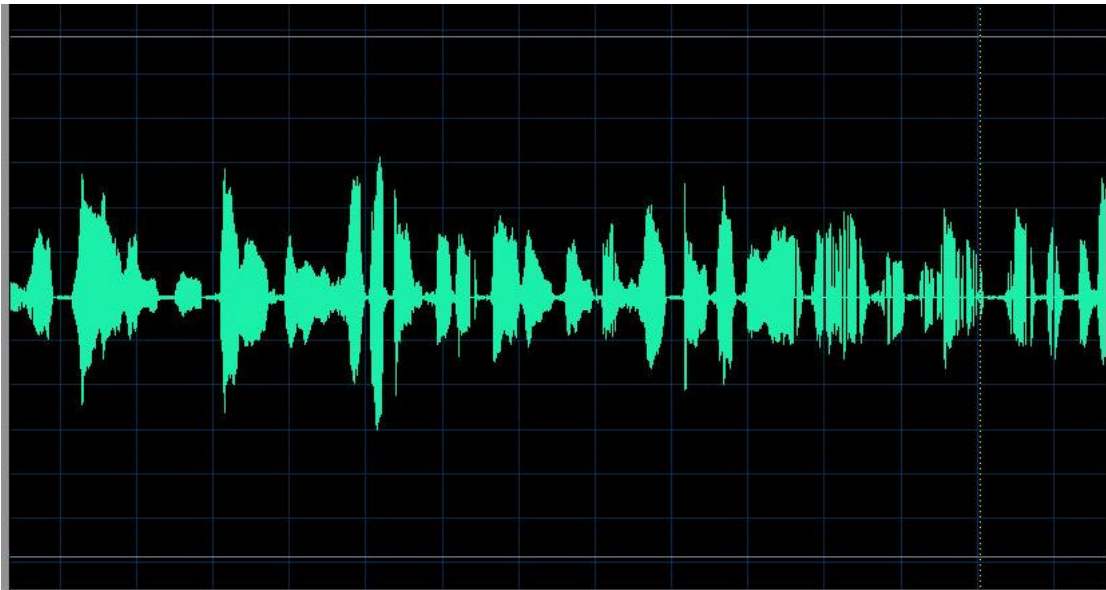


圖 5.9 在丟棄方式下，以 Late-rate 動態調整緩衝區  
之後的語音波形

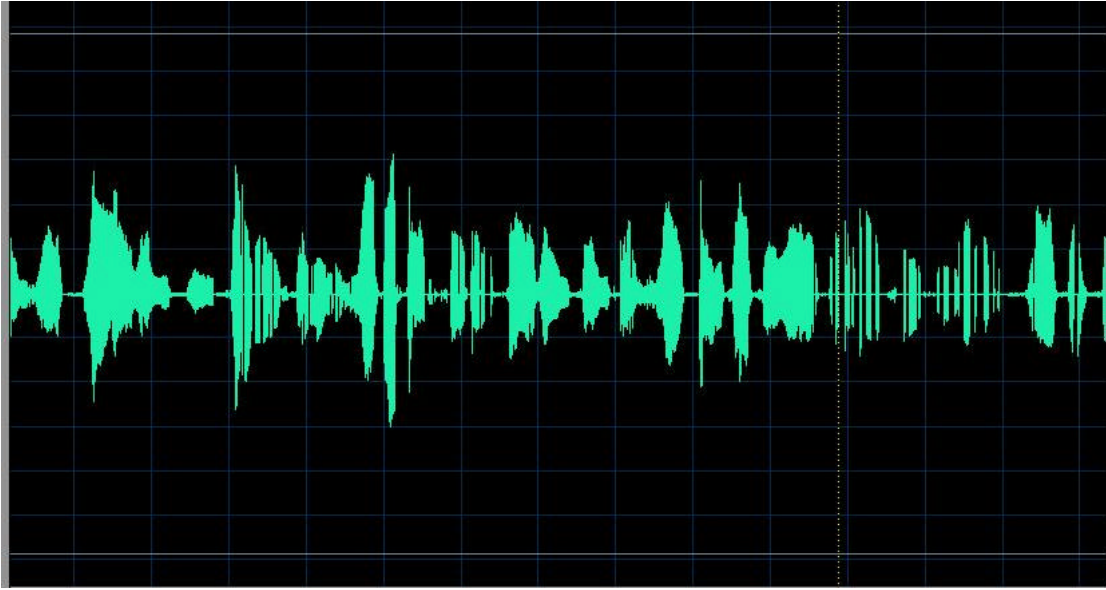


圖 5.10 在丟棄方式下，以 Max-jitter 動態調整緩衝區  
之後的語音波形

#### 5.4 晚到封包延遲（Delay）驗證

針對晚到的語音封包，不管如何接受端都要讓這個封包順利播放，所以就等待此封包到達使其能夠順利播放，採用這種延遲方式（Delay）來驗證不同的緩衝區調整方式，看其調整效果是否有改善

在對晚到語音封包採用延遲方式處理時，必須先確定對晚到的語音封包最多能夠等待多少個語音封包，如果晚到的語音封包在過了等待時間之後，仍沒有到達，則這個語音封包將被丟棄。

表 5.3~5.6 分別是在不同的封包延遲下，不同緩衝區調整方式和封包遺失率（Loss rate）的比較。

	no buffer	buffer 1	buffer 2	buffer 3	buffer 4	buffer 5	late-rate	max-jitter
247	0	0	0	0	0	0	0	0
488	0	0	0	0	0	0	0	0.0333
757	0.5018	0.2973	0.0446	0.0037	0.0037	0	0.0037	0.0128
1068	0.3247	0.1414	0.0032	0	0	0	0	0.0023
1324	0.3151	0.1028	0	0	0	0	0	0.148
1549	0.4133	0.2177	0.0311	0	0	0	0.0311	0
1765	0	0	0	0	0	0	0	0
2000	0	0	0	0	0	0	0	0
2328	0	0	0	0	0	0	0	0

表 5.3 延遲一個語音封包，不同緩衝區調整方式和封包遺失率  
的關係

	no buffer	buffer 1	buffer 2	buffer 3	buffer 4	late-rate	max-jitter
247	0	0	0	0	0	0	0
488	0	0	0	0	0	0	0.0333
757	0.2973	0.0446	0.0037	0.0037	0	0.0037	0.0098
1068	0.1414	0.0032	0	0	0	0	0.0186
1324	0.1028	0	0	0	0	0	0
1549	0.2177	0.0311	0	0	0	0	0
1765	0	0	0	0	0	0	0
2000	0	0	0	0	0	0	0
2328	0	0	0	0	0	0	0

表 5.4 延遲兩個語音封包，不同緩衝區調整方式和封包遺失率  
的關係

	no buffer	buffer 1	buffer 2	buffer 3	late-rate	max-jitter
247	0	0	0	0	0	0
488	0	0	0	0	0	0.0333
757	0.0446	0.0037	0.0037	0	0	0.0098
1068	0.0032	0	0	0	0	0
1324	0	0	0	0	0	0
1549	0.0311	0	0	0	0	0
1765	0	0	0	0	0	0
2000	0	0	0	0	0	0
2328	0	0	0	0	0	0

表 5.5 延遲三個語音封包，不同緩衝區調整方式和封包遺失率  
的關係

	no buffer	buffer 1	buffer 2	late-rate	max-jitter
247	0	0	0	0	0
488	0	0	0	0	0
757	0.2974	0.0037	0	0.0037	0
1068	0.1414	0	0	0	0.016
1324	0	0	0	0	0.0273
1549	0.2177	0	0	0	0
1765	0	0	0	0	0
2000	0	0	0	0	0
2328	0	0	0	0	0

表 5.6 延遲四個語音封包，不同緩衝區調整方式和封包遺失率  
的關係

從表 5.3 ~ 表 5.6 中，可以看出，在等待晚到的封包過程中，等待的時間愈久，封包的遺失率也就愈低，並且所需的緩衝數也較少，但是從語音品質考量，語音會因為採用延遲等待語音封包的關係，讓語音聽起來變得較慢。

評斷在不同的時間延遲下，對於不同的緩衝區調整辦法，可以知道採用延遲方法下，Late-rate 的辦法比 max-jitter 更有較少的封包遺失率，主要是採用晚到封包可以延遲播放，這種動作讓每個語音段落內的語音封包都可以等待更久的時間，所以其封包遺失率便會大大降低，而 Max-jitter 緩衝區調整方式，是利用每個語音段落內語音封包之間的最大時間差來決定緩衝區大小，所以對 Max-jitter 而言，其調

整方式將不會改變。採用 Late-rate 的調整方式將優於 Max-jitter 的調整方式。

圖 5.11 是在延遲 2 個封包下，以 Late-rate 動態調整緩衝區之後的語音波形

圖 5.12 是在延遲 2 個封包下，以 Max-jitter 動態調整緩衝區之後的語音波形

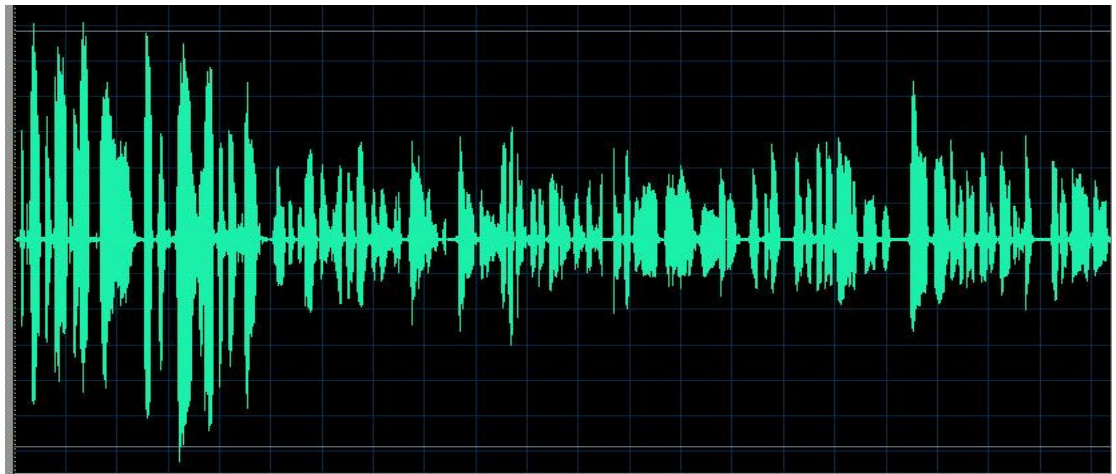


圖 5.11 在延遲 2 個封包下，以 Late-rate 動態調整緩衝區之後的語音  
波形

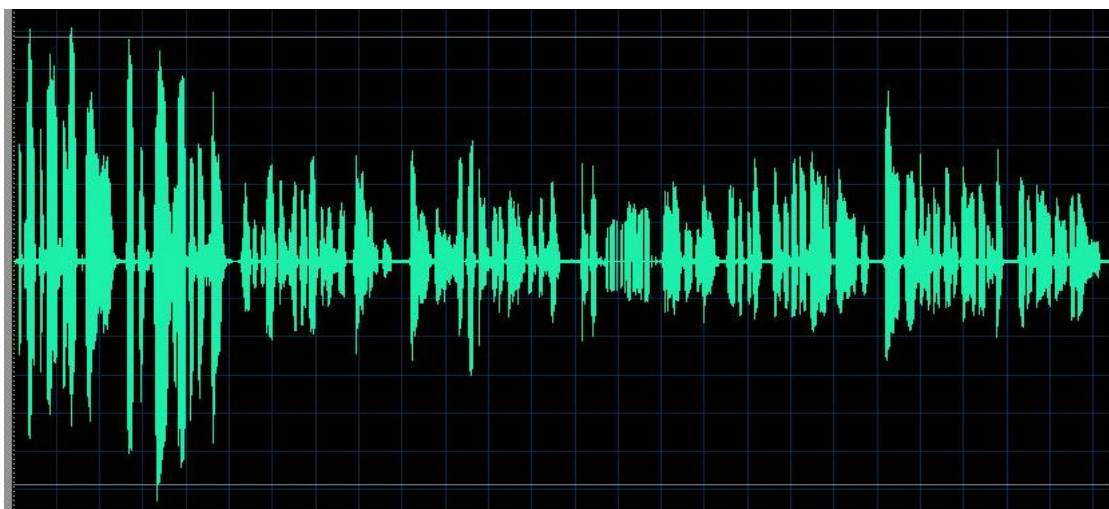


圖 5.12 在延遲 2 個封包下，以 Max-jitter 動態調整緩衝區之後的語音  
波形

### 5.5 不同緩衝區調整方式下的語音播放時間

不管對晚到封包採用丟棄（Drop）或者是延遲（Delay）播放的方式，對於經過緩衝區的語音都將應用語音時間校正，來減少語音的長度，在這裡將列出利用各種調整方式所得到的語音長度，及語音播形，並主觀的播放語音來評斷語音的好壞。

source	no buffer	buffer 1	buffer 2	buffer 3	buffer 4	buffer 5	late-rate	max-jitter
23.441	22.84	22.25	22.46	23.47	23.68	23.89	23.52	23.26

表 5.7 在封包丟棄方式下的語音長度

source	no buffer	buffer 1	buffer 2	buffer 3	buffer 4	buffer 5	late-rate	max-jitter
23.441	23.4	23.61	24.32	24.38	25.13	25.25	24.31	23.81

表 5.8 在封包延遲一個封包方式下的語音長度

source	no buffer	buffer 1	Buffer 2	buffer 3	late-rate	max-jitter
23.441	24.46	24.67	24.98	25.12	24.25	24.38

表 5.9 在封包延遲兩個封包方式下的語音長度

比較表 5.7 ~ 表 5.9 的語音播放長度數據之後，有透過時間校正的語音，其經過解碼之後，發現解碼後的語音長度和原始的語音長度相差不遠，這意味經過網路延遲的語音封包，再經過各種緩衝區調整方法之後，除了保有較少的封包遺失率之外，還能維持語音的長度，不至於讓語音經過長時間的緩衝之後，語音長度變長了。

### 5.6 不同緩衝區調整方式下的語音 SSSNR 品質評估

透過 3.4 節的語音評估方式，計算不同緩衝區調整方式下的語音 SSSNR 值，藉由這些 SSSNR 值作為客觀的評估方式，因為評估方式主要是考慮語音封包的遺失狀況，所以以在語音封包延遲的條件下，會得到較佳的語音品質，其中又以 Late-rate 緩衝區調整方式的品質更高。表 5.10 是不同緩衝區調整方式下的語音 SSSNR 與語音增長時間 (Extra delay)。

	Source	Late-rate (Drop)	Max-jitter (Drop)	Late-rate (Delay1)	Max-jitter (Delay1)	Late-rate (Delay2)	Max-jitter (Delay2)
SSSNR	8.17dB	7.35dB	7.41dB	7.62dB	7.51dB	7.88dB	7.57dB
Extra delay	0	280ms	420 ms	240 ms	420 ms	180 ms	420ms

表 5.10 不同緩衝區調整方式下的語音 SSSNR 與語音增長時間