

第四章 實驗數據與分析

在 3 章我們針對前人的方法提出分析與新的改進，而在這裡我們將會針對我們實驗過後的數據與結果提出分析探討。

4.1 精準度

在精準度的方面，我們運用下列四種誤差來表示浮點數的值與整數的值產生誤差的情形：

$$\text{Maximum absolute error} : \max_{n=0}^{N-1} \{ |dataF[n] - dataI[n]| \}$$

$$\text{Maximum relative error} : \max_{n=0}^{N-1} \{ |dataF[n] - dataI[n]| / dataF[n] \}$$

$$\text{Average absolute error} : \left(\sum_{n=0}^{N-1} |dataF[n] - dataI[n]| \right) / N$$

$$\text{Average relative error} : \left(\sum_{n=0}^{N-1} |dataF[n] - dataI[n]| \right) / \sum_{n=0}^{N-1} |dataF[n]|$$

$dataF$: Floating Data , $dataI$: Integer Data

4.1.1 Cos and Sin Table 之誤差比較：

表格 4-1、表格 4-2 為我們的方法與先前建立 Cos 及 Sin 表的方法誤差值的比較，很明顯的，我們的誤差值降低相當的多，而表格 4-3 為 Cos 及 Sin 表所佔記憶體大小比較。

表格 4-1 Cos Table 的誤差

Cos Table	最大誤差	最大相對誤差	平均誤差	平均相對誤差
[2]、[3]的方法	2	468.35 %	0.826027	129.791 %
我們的方法	5.81E-05	100 %	2.43E-05	0.00381 %

表格 4-2 Sin Table 的誤差

Cos Table	最大誤差	最大相對誤差	平均誤差	平均相對誤差
[2]、[3]的方法	9.7E-04	100 %	4.7E-04	0.0752 %
我們的方法	5.64E-05	100 %	2.45E-05	0.00546 %

表格 4-3 Cos 與 Sin Table 佔記憶體大小比較

	Size of Cos Table	Size of Sin Table
[2]、[3]的方法	4Kbytes	4Kbytes
我們的方法	1.22Kbytes	0.12Kbytes

4.1.2 Logarithm and Square Root Table 之誤差比較

表格 4-4 為[2]、[3]與我們所建 Log 表的誤差比較，由表可知，我們的誤差值相較先前的方法下降相當多，而表格 4-5，則是我們的開方根表與標準函式庫中的開方根函數之誤差值。

表格 4-4 Logarithm Table 的誤差

Log Table	最大誤差	最大相對誤差	平均誤差	平均相對誤差
[2]、[3]的方法	9.001	100 %	7.71	94.012 %
我們的方法	0.2857	+ ∞	0.0167	0.0204 %

表格 4-5 Square Root Table 的誤差

Square Root Table	最大誤差	最大相對誤差	平均誤差	平均相對誤差
我們的方法	0.49874	29.829 %	0.24996	0.375 %

4.1.3 MFCC 各個步驟誤差的比較

我們使用 TCC300 及 TIMIT 語料各 100 句，來測試誤差情形，每一句語料都是以 20ms 為一個單位的音框大小。表格 4-6 為[2]、[3]的方法在 TCC300 語料下 MFCC 各個步驟中所產生出的四種誤差值；表格 4-7 則為我們改進過後，

在 TCC300 語料環境中之誤差；表格 4-8 為[2]、[3]的方法在 TIMIT 語料下之誤差；而表格 4-9 為我們的方法在 TIMIT 語料下的誤差

表格 4-6 [2][3]的方法針對 TCC300 語料所產生的誤差

Errors Functions	Maximum		Average	
	Absolute error	Relative error	Absolute error	Relative error
Energy	1.224397	3125.52717 %	0.779034	122.442522 %
Pre-emphasis	1.777343	110.756283 %	0.502014	0.240768 %
Windowing	6.097656	100%	0.792114	0.702443 %
FFT	745597.1875	9658170497.2 %	742.784652	33.543611 %
Triangular filter	2152004436	$+\infty$	1073298958.9	100.000817%
Logarithm	3.93088	408045871.145 %	0.1242663	1.337575 %
DCT	5.3998	404236090.82 %	0.18537	19.0496 %
Weight Cepstrum	21.020127	404236090.82 %	1.4163288341	25.166302 %
Delta	21.020127	404236090.82 %	0.8347596644	26.56967208 %
Acceleration	21.020127	408045871.145 %	0.6092158652	27.180834 %

表格 4-7 我們的方法針對 TCC300 語料所產生的誤差

Errors Functions	Maximum		Average	
	Absolute error	Relative error	Absolute error	Relative error
Energy	1.22439717	3125.52717 %	0.77903400	122.442522 %
Pre-emphasis	2.174194	110.7562 %	0.50223	0.24083 %
Windowing	6.09765	100%	0.792131861	0.70245213 %
FFT	646366.5	9432897797.4 %	582.55845871	25.96243 %
Triangular filter	2151875678	$+\infty$	1073298622.8	100.00078 %
Logarithm	4.54862	90.1297 %	0.05821	0.61552 %
DCT	4.18056	414368395.132 %	0.1166	12.005 %
Weight Cepstrum	27.3529	414202583.35 %	0.907	15.8155 %
Delta	27.3529	414202583.35 %	0.5213	16.279 %
Acceleration	27.3529	414202583.350 %	0.37671	16.4986 %

表格 4-8 [2][3]的方法針對 TIMIT 語料所產生的誤差

Errors Functions	Maximum		Average	
	Absolute error	Relative error	Absolute error	Relative error
Energy	1.22428749	122.89031 %	0.8654990	122.40858 %
Pre-emphasis	1.2502746	100 %	0.4941173	0.3746495 %
Windowing	2.123046875	100 %	0.77052	1.08196 %
FFT	126594.32	1274604075.9 %	362.66144	32.463512 %
Triangular filter	2148490350	$+\infty$	1073520286.7	100.0003 %
Logarithm	3.7583	215.54560 %	0.1113	1.2498 %
DCT	4.9989	6224711.201 %	0.1751	16.427 %
Weight Cepstrum	16.124829	6214868.90 %	1.32690822	22.7336514 %
Delta	16.124829	8700327.12 %	0.79291050	24.6836269 %
Acceleration	16.12482	11020803.92 %	0.5816822	25.501602 %

表格 4-9 我們的方法針對 TIMIT 語料所產生的誤差

Errors Functions	Maximum		Average	
	Absolute error	Relative error	Absolute error	Relative error
Energy	1.22428749	122.89031 %	0.8654990	122.40858 %
Pre-emphasis	1.250274	100 %	0.49419	0.374706 %
Windowing	2.1230468	100 %	0.770528	1.081972 %
FFT	122648.32	820230374.6 %	280.07471	25.101727 %
Triangular filter	2148460253	$+\infty$	1073520142.63	100.0003 %
Logarithm	2.53228	221.1282 %	0.042115	0.47255 %
DCT	4.43561	1354309.307 %	0.07678	7.20696 %
Weight Cepstrum	11.38171	1353576.937 %	0.63146	10.80133 %
Delta	11.38171	6859157.902 %	0.37397	11.6258 %
Acceleration	11.3817	6859157.902 %	0.2733	11.968 %

我們可以於表格 4-6、表格 4-7 歸納出如圖 4-1 在平均相對誤差的比較，在圖中我們可以看到在 Energy、Pre-emphasis、Hamming Window 中兩個方法的誤差相差不到 0.0001%，而 FFT 時我們的誤差值降低了約 7.581%，Logarithm 降低了約 0.722375%，DCT 及 Weight Cepstrum 各別下降了約 7.0437%、9.3%，Delta 中降低了約 10.31%，Acceleration (即 Overall Average Relative Error)降低了約 10.6841%。

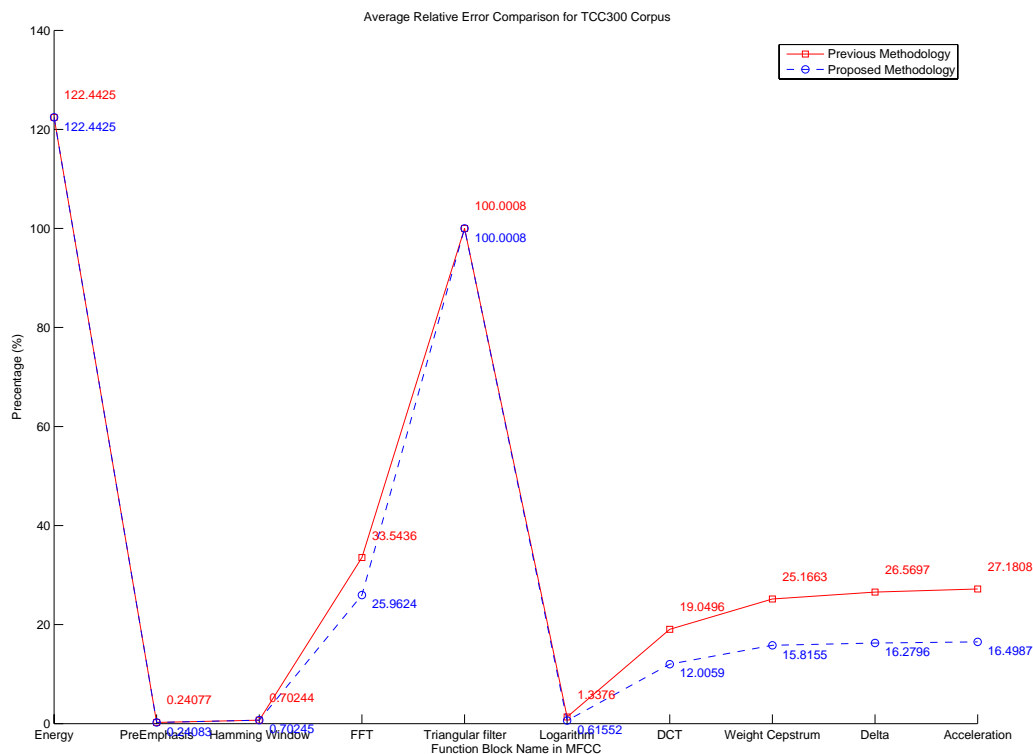


圖 4-1 Average Relative Error Comparison of TCC300 Corpus

而表格 4-8、表格 4-9 歸納出如圖 4-2 在平均相對誤差的比較，在圖中我們的方法在 FFT 中誤差約降低了 7.36%，Logarithm 約下降了 0.7773%，DCT 的誤差下降 7.043%，Weight Cepstrum 降低了 9.35%，Delta 約下降了約 9.29%，Acceleration (即 Overall Average Relative Error)約降低了 10.685%。

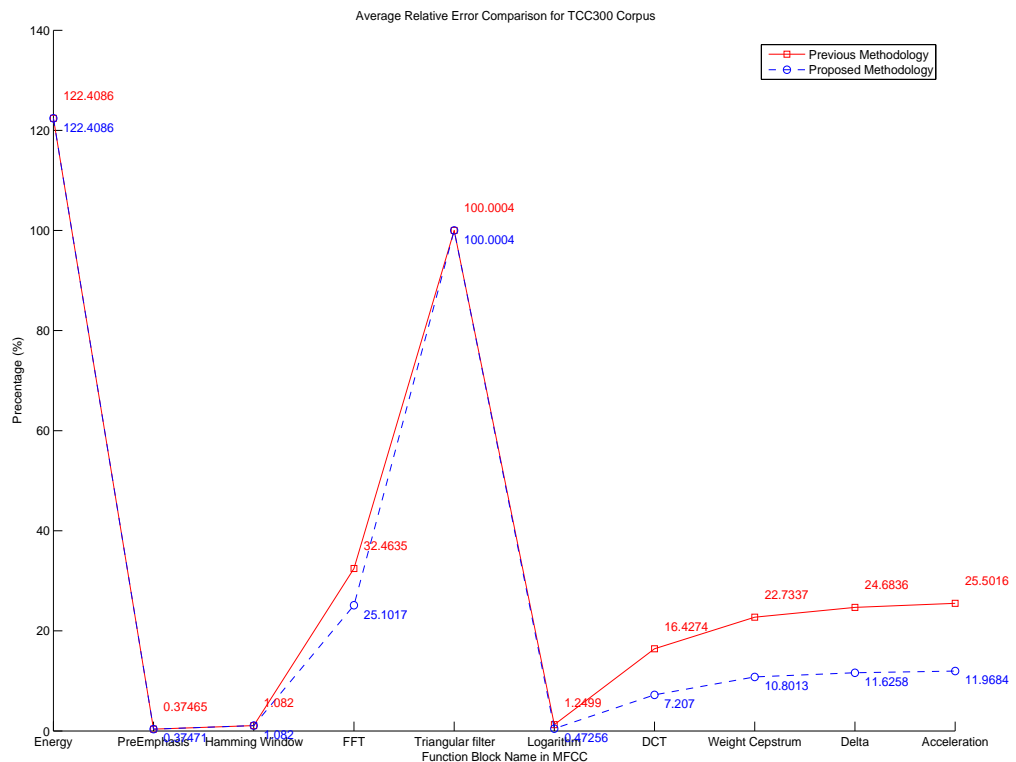


圖 4-2 Average Relative Error Comparison of TIMIT Corpus

4.1.4 辨識率比較

TCC300 測試語句資訊如下，而 TIMIT 則是用 Free Phone Decoding，表格

4-11、表格 4-12 則是 Tang Poem 辨識結果與 TIMIT Free Phone Decoding 之結果。

	Testing Data
Content	Tang Poem
Speakers	7 Males and 1 Female
Sample Rate	16 KHz
Bits Per-sample	16 bits
Total	2331 Files = 3.2 hours

表格 4-10 Tang Poem 資訊

Tang Poem	[2]、[3]的方法	我們的方法
辨識率	88.37 %	89.45 %

表格 4-11 TCC300 辨識率比較

TIMIT	[2]、[3]的方法	我們的方法
辨識率	48.66 %	47.41 %

表格 4-12 TIMIT 辨識率比較

4.1.5 執行時間比較

我們以唐詩”二十四橋明月夜語句”，取樣頻率 16Khz、解析度 16bits，錄音長度五秒，來測試[2][3]及我們所提出之 MFCC 特徵擷取方法於各個步驟在 Intel Pentium X86-機器(CPU 時脈：2.4GHz)上執行時所需的時間。

	[2][3]的方法	我們的方法
計算能量	5139 ns	5159 ns
預強調	2019 ns	2014 ns
漢明窗	2737 ns	2769 ns
快速傅立葉轉換	52834 ns	53935 ns
三角帶通濾波器	15662 ns	80342 ns
對數計算	3619 ns	10696 ns
離算餘弦轉換	12275 ns	4515 ns
Weight Cepstrum	1219 ns	549 ns
Delta and Acceleration	904 ns	894 ns

表格 4-13 [2][3]及我們的方法所需計算時間的比較

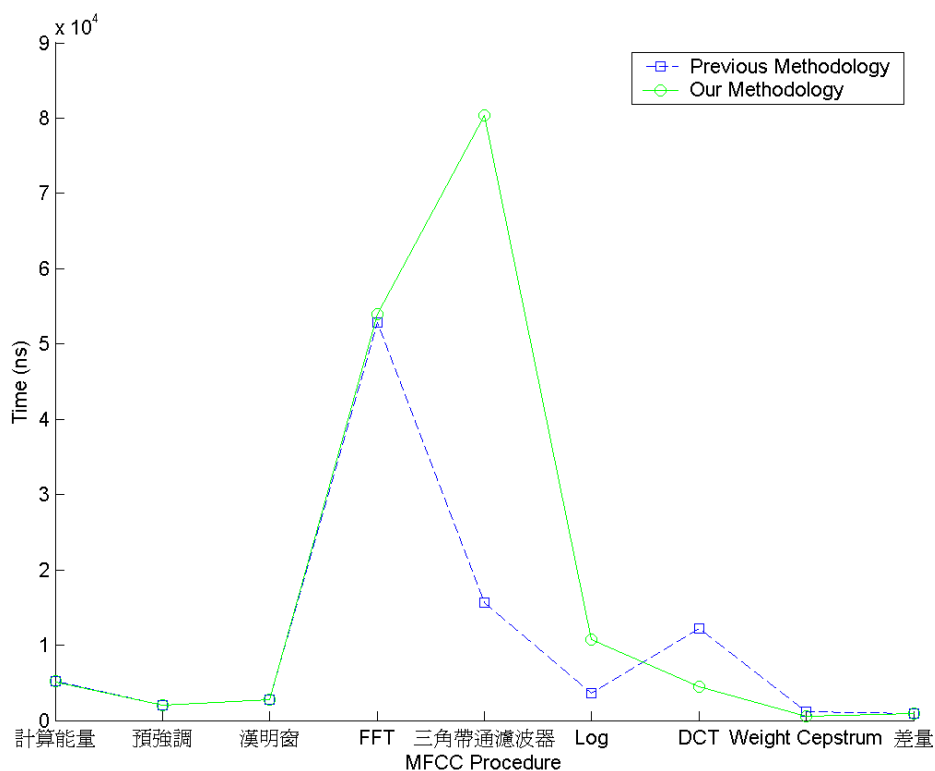


圖 4-3 [2][3]與我們的方法所需時間比較

由圖 4-3 我們可以得知，因為我們採用新的查表方式，所以造成在三角帶通濾波器及計算對數的步驟中會比先前的方法較耗時。

4.2 錯誤分析

依照我們改進的方法，實驗的結果很明顯的較先前的做法在誤差方面降低了許多，而先前我們預期如果誤差值降低，有助於提升辨識率，也在實驗過後，有小幅度較原先的方法提升了約接近 1%，符合我們想將誤差降低以期提升辨識率的想法。不過在這邊 TIMIT 的辨識率還是不盡理想，應該與 Free Phone Decoding 有相當的關聯，不過 FFT 及 Triangular Band-pass Filters 這兩個步驟的誤差相對於其他步驟來的大，因此在這兩個步驟，還需要提出更好的調整參數來改進之。