CHAPTER 6 DISCUSSION

In this chapter, we discuss a real-time and low calculation tone recognition on recognition accuracy compared with autocorrelation method, vowel segmentation, and recognition accuracy of automatic tonal speech recognizer with tone and without tone.

6.1 Recognition Accuracy

This thesis compares our vowel magnitude different function (V_{MDF}) tone recognition performance with autocorrelation method (AC).

When we use the entire syllable as an input signal, the fundamental frequency (F_0) is sometimes affected by the neighboring syllable. Figure 6.1 shows F_0 which is affected by the preceding syllable. In order to reduce error rates of neighboring syllable effect, this thesis therefore estimates F_0 by only vowel signals.

However, there are still some F_0 errors due to double pitch. The double pitch error occurs when we detect very low pitch point value compared with the other frames. The very low pitch point is shown in Figure 6.2. To reduce double pitch error, we apply mesh 5 for both V_{MDF} and AC methods. We realize that when we apply mesh 5 to those methods V_{MDF} provides better F_0 trajectory while AC still shows double pitch errors. This results in tone recognition accuracy reduction using AC method. Figure 6.2 shows V_{MDF} and AC after mesh 5.

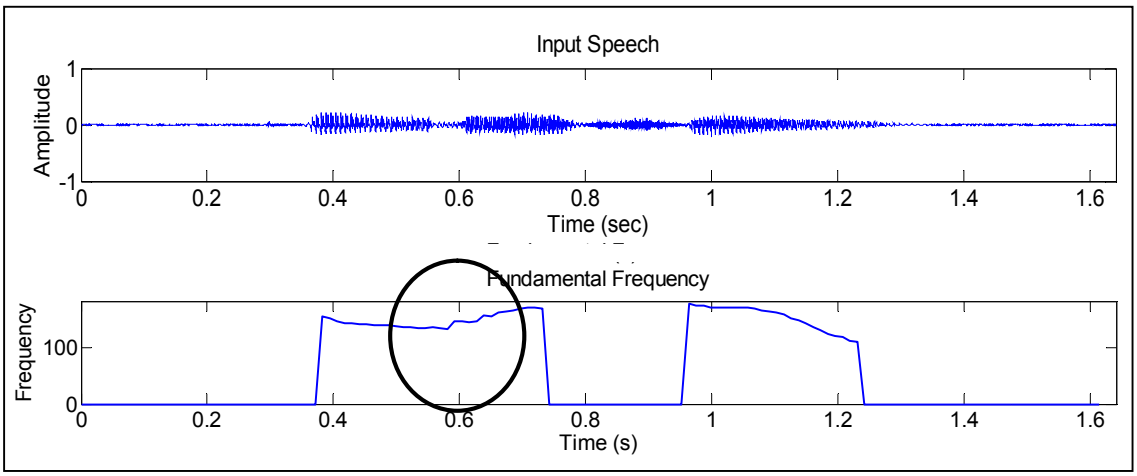

Figure 6.1 F_0 affected by the preceding syllable

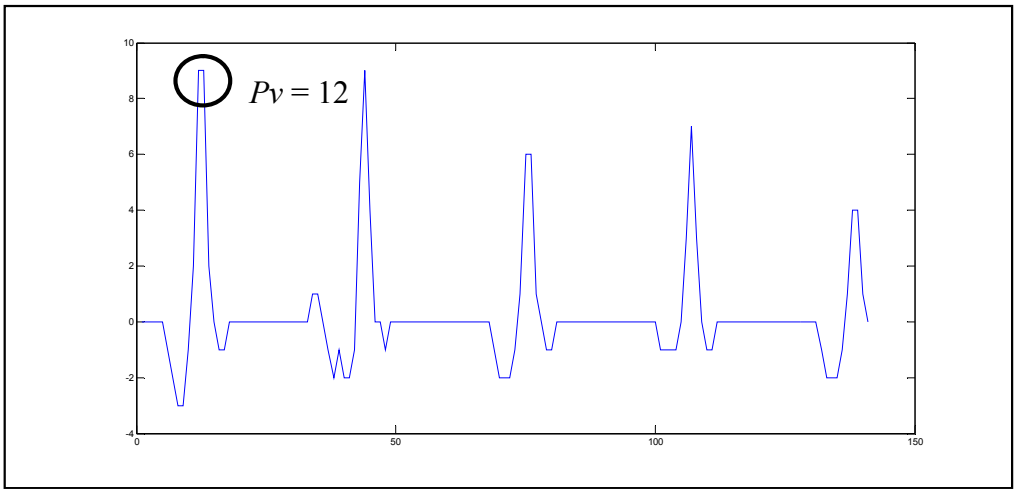

Figure 6.2 A very low pitch point value detected by feature extraction process

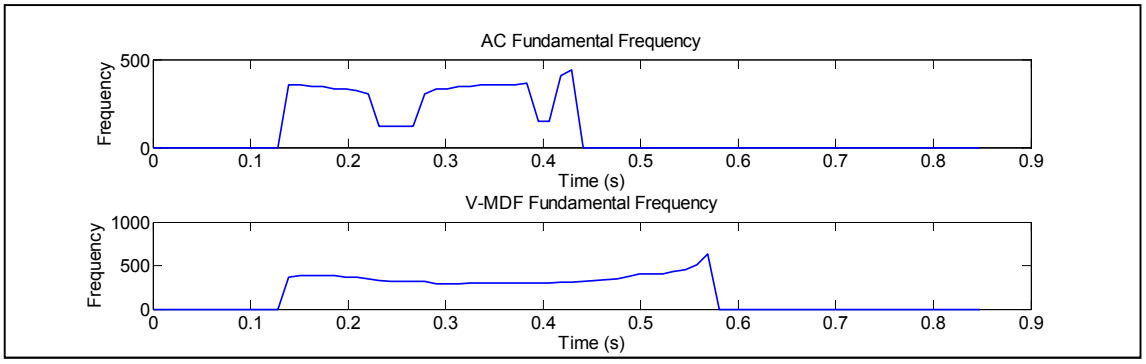

Figure 6.3 V_{MDF} and AC after mesh 5

This thesis therefore proposes vowel-magnitude different function (V_{MDF}) to reduce F_0 error from neighboring syllable influence. V_{MDF} provides nucleus part of V_{MDF} without tone modeling process. In addition, V_{MDF} requires the lower number of frames due to input vowel signals. This makes V_{MDF} can process the tone recognition faster than AC.

6.2 Vowel-MDF Segmentation

V_{MDF} achieves time processing performance. However, the tone recognition accuracy can be dropped when Vowel-MDF segmentation detects the incorrect vowel regions.

Figure 6.4 illustrates the segmentation error. There is no vowel region of syllable 1. This makes tone result incorrect. Figure 6.5 shows vowel segmentation error in syllable 3. Due to incorrect vowel region, the initial F_0 part of syllable 3 provides high frequency.

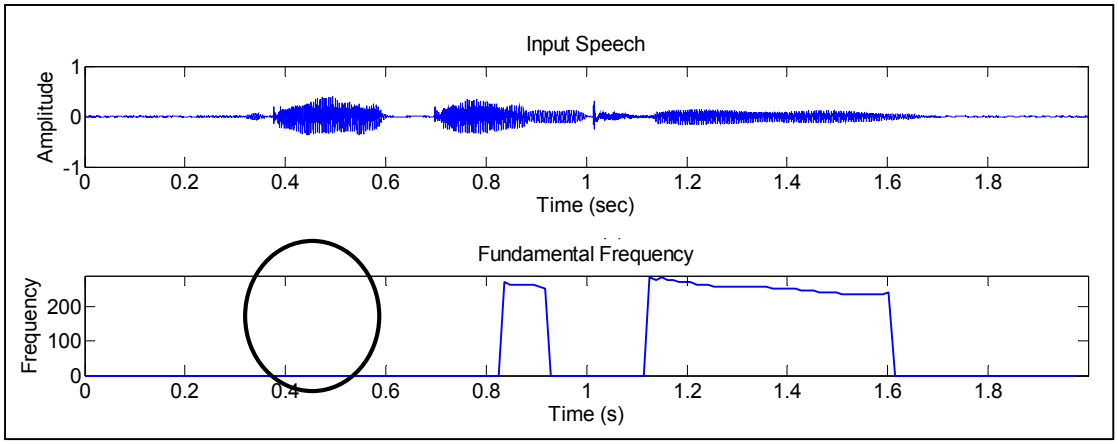

Figure 6.4 Vowel region error of syllable 1

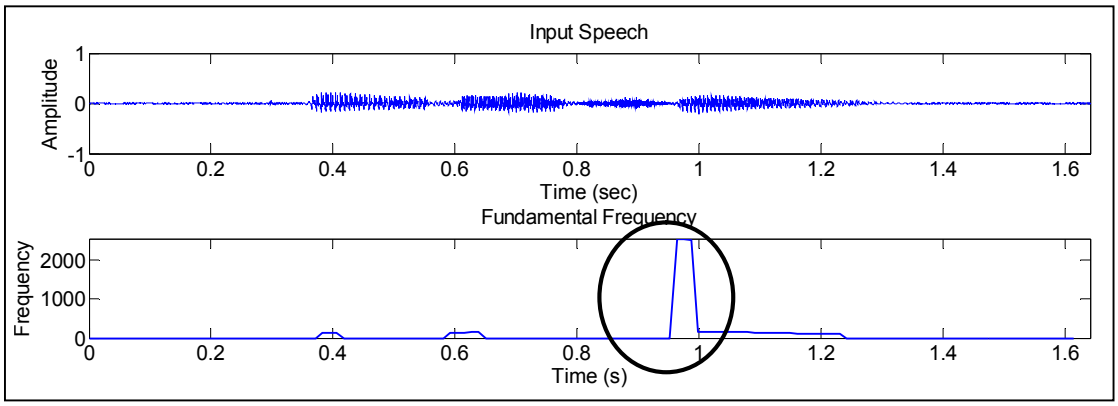

Figure 6.5 Vowel segmentation error of syllable 3

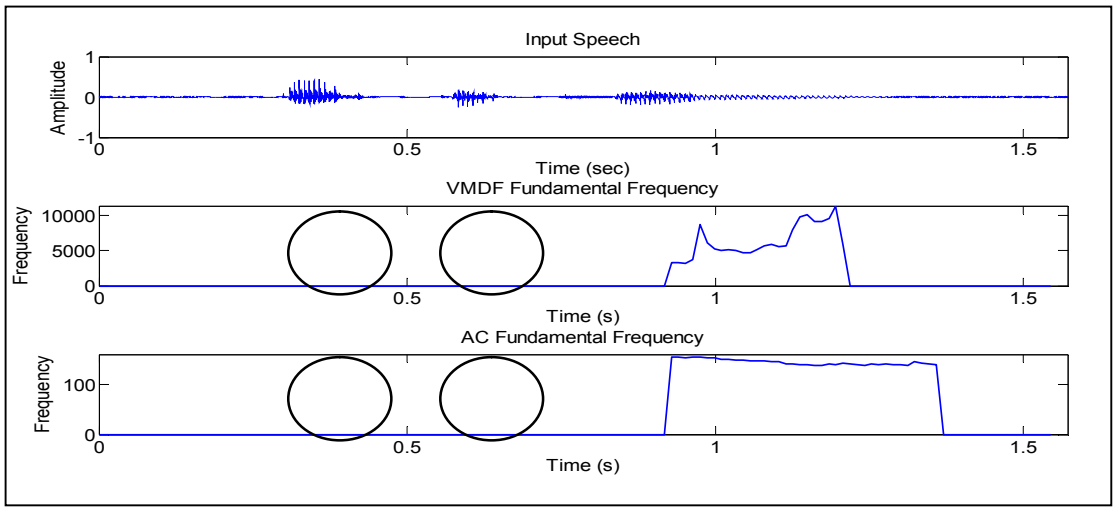

Figure 6.6 Vowel segmentation errors

Figure 6.6 indicates the vowel segmentation errors. The errors occur when the zero-crossing rate of vowel signal is higher than T_z. shown in Figure 6.7 and therefore the there are no F_0 trajectories the circles. In addition, V_{MDF} is sensitive to low amplitude input signal. This makes its F_0 trajectory incorrect.

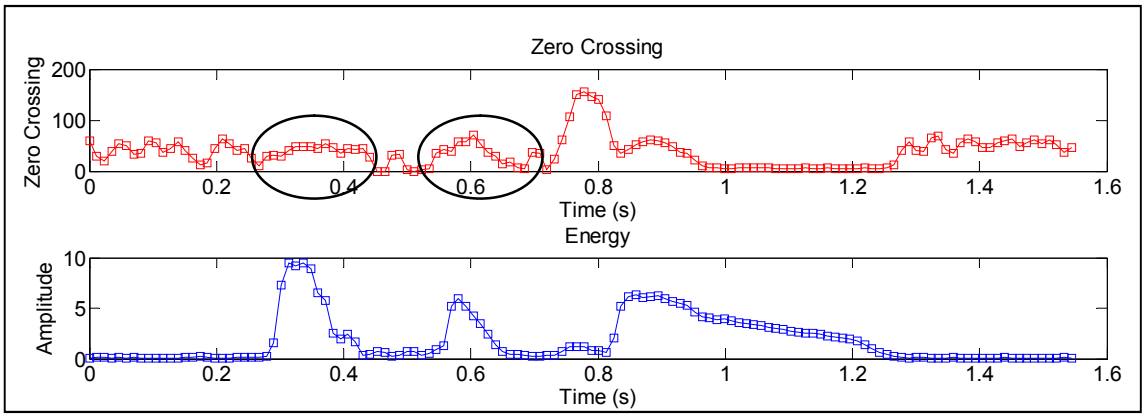

Figure 6.7 Zero-crossing rate errors

6.3 ATSR Without Tone Recognition

This thesis proposes ATSR with tone recognition system. The system improves recognition accuracy especially in case of words having same phoneme but different tones. However, if either tone recognition or word recognition provides incorrect result, the recognition result is wrong.

We state this problem in Figure 6.8. For instance, the word "*kāa*" is and input speech. If word recognition recognizes the correct word "*kaa*" but the tone recognition recognizes the wrong "low" tone, the result will provide wrong result "*ka`a*". On the other hand, if the tone recognition recognizes the incorrect tone while the word recognition provides correct result, the final result is still incorrect.

We therefore require very high recognition accuracy in both systems.

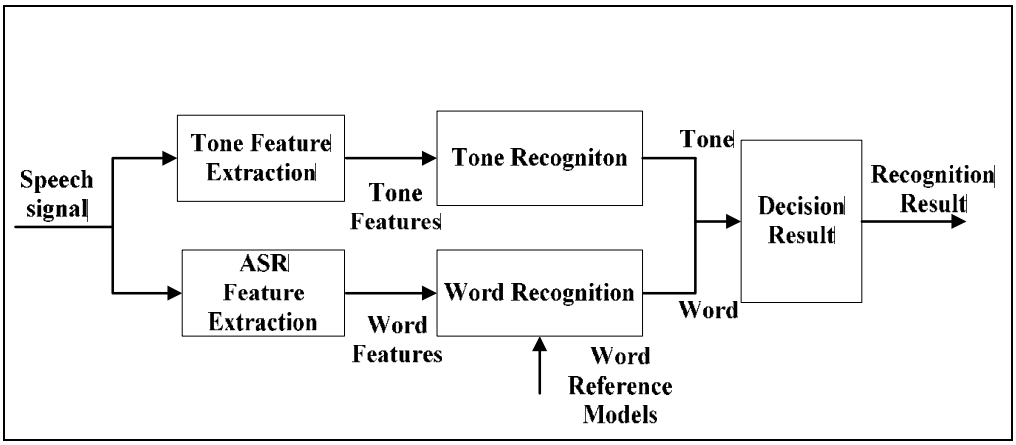

Figure 6.8 Overview of ATSR