We used human exon and intron datasets downloaded at HTUhttp://www.ncbi.nlm.nih.gov/UTH for testing. In our first result, we chose exons and introns with a length approximately equal to 1500. Fig. 1 shows $\tilde{X}[k]$ in the FT approach and the spectrum for $s_2[n]$. It can be seen that both approaches can detect the 3-periodicity in the coding regions as peaks are observed at *k*=501 which corresponds to the 2Π/3 frequency. Fig. 2 shows results for another exon with GenBank accession number "AX136319". In the FT approach, the peak cannot be easily identified. In contrast, the spectrum of $s_2[n]$ clearly shows the peak at 2Π/3 (*k*=411). As discussed in Section IV, this is due to the fact that the periodicity assumption is made with respect to the biological property embedded in the DNA sequence.

Recognizing human exons is sometimes a very challenging problem as human exons can be very short in length (137 bp in average). Exon sequences with GenBank accession numbers "AB061839" and "AB050050" are chosen for testing. The first sequence has 123 bp while the second sequence has 127 bp. Results for these sequences are shown respectively in Fig. 3 and Fig. 4. Due to the short length of the exon sequences, peaks are not observed at 2Π/3 for both sequences for $\tilde{X}[k]$. In contrast, peaks are observed at 2Π/3 (k=41 and k=43 respectively for Fig. 3 and Fig. (4) for the spectrum of the modified sequences $s_i[n]$.

In a further experiment, we extracted two different datasets [7]. These datasets consist of 6000 Yeast ORFs and 6000 Yeast No Feature sequences, 1500 human exons and 1500 introns whose length is less than 140bp. We performed classification experiments for both the FT approach and our proposed approach. In the FT approach, the value at 2Π/3 is extracted as the feature. In our proposed approach, the three power spectra of $s_0[n], s_1[n]$ and $s_2[n]$ are firstly obtained. Then the features for classification are the values at 2Π/3 and the DC values for these three spectra.

Classification experiments using these selected features were then performed using the k-nearest-neighbor classifier as in Wu *et al.,*. **[8]** Table I summarizes the results. Note that sensitivity is defined as the proportion of coding sequences that have been correctly classified as coding while specificity is the proportion of non-coding sequences that have been correctly classified as non-coding. From Table 1, we see that for human sequences, low specificity is observed for the FT approach. This implies that many non-coding sequences are wrongly classified as coding sequences. However, using the proposed approach, the performance is greatly improved. For Yeast sequences, both sensitivity and specificity are increased by using the proposed features.