MIR Homework 3 Report

Question 1.

Fourier Tempogram

Genre	P-score	ALOTC
ChaCha	0.3375	0.7658
Jive	0.4554	0.8333
Quickstep	0.4683	0.9024
Rumba	0.1908	0.4286
Samba	0.1105	0.2442
Tango	0.4340	0.8488
Viennese waltz	0.4841	0.9077
Waltz	0.3134	0.6818

Autocorrelation Tempogram

Genre	P-score	ALOTC
ChaCha	0.5136	0.9910
Jive	0.4523	0.9333
Quickstep	0.4181	0.8537
Rumba	0.4432	0.8980
Samba	0.3701	0.7326
Tango	0.5170	0.9651
Viennese waltz	0.4795	0.8923
Waltz	0.3402	0.6545

The results obtained by using the autocorrelation tempogram shows a significant improvement from fourier tempogram. Especially on the genre ‘Rumba’, ‘Samba’, both musica often have complex rhythmic patterns with multiple layers of percussion instruments. The autocorrelation method may be better suited for capturing the intricate rhythmic structures and detecting the periodicities within the music, leading to higher accuracy in tempo estimation.

And ChaCha also shows a notable improvement, in my opinion, the ChaCha rhythm may have strong periodic patterns that are better captured by the autocorrelation method. Autocorrelation is particularly effective for detecting repetitive patterns in music, which could explain the higher accuracy in tempo estimation for this genre.

Question 2.

4 seconds fourier

Genre	P-score	ALOTC
ChaCha	0.1299	0.2973
Jive	0.4443	0.8000
Quickstep	0.4885	0.8902
Rumba	0.2484	0.5306
Samba	0.0551	0.1163
Tango	0.4932	0.9419
Viennese waltz	0.4834	0.9231
Waltz	0.2059	0.4455

4 seconds autocorrelation

Genre	P-score	ALOTC
ChaCha	0.5056	0.9820
Jive	0.4533	0.9333
Quickstep	0.4480	0.8780
Rumba	0.4429	0.8980
Samba	0.3658	0.7209
Tango	0.5180	0.9651
Viennese waltz	0.5182	0.9231
Waltz	0.2398	0.4455

8 seconds fourier

Genre	P-score	ALOTC
ChaCha	0.3375	0.7658
Jive	0.4554	0.8333
Quickstep	0.4683	0.9024
Rumba	0.1908	0.4286
Samba	0.1105	0.2442
Tango	0.4340	0.8488
Viennese waltz	0.4841	0.9077
Waltz	0.3134	0.6818

8 seconds autocorrelation

Genre	P-score	ALOTC
ChaCha	0.5136	0.9910
Jive	0.4523	0.9333
Quickstep	0.4181	0.8537
Rumba	0.4432	0.8980
Samba	0.3701	0.7326
Tango	0.5170	0.9651
Viennese waltz	0.4795	0.8923
Waltz	0.3402	0.6545

12 seconds fourier

Genre	P-score	ALOTC
ChaCha	0.2482	0.5946
Jive	0.4822	0.8833
Quickstep	0.4393	0.8537
Rumba	0.2182	0.4898
Samba	0.1234	0.2558
Tango	0.4883	0.9419
Viennese waltz	0.4535	0.8769
Waltz	0.2532	0.5636

12 seconds autocorrelation

Genre	P-score	ALOTC
ChaCha	0.5097	0.9820
Jive	0.4457	0.9167
Quickstep	0.3941	0.8049
Rumba	0.4481	0.9082
Samba	0.3764	0.7442
Tango	0.5171	0.9651
Viennese waltz	0.4793	0.8923
Waltz	0.3510	0.6636

It is interesting that Walts seem to have better performance when the window size is larger, while QuickStep seem to have better performance when the window size is small. So I think the slower songs will be better detected by using larger window size.

Question 3.

Genre	P-score	ALOTC
ChaCha	0.4954	0.9910
Jive	0.5227	0.9667
Quickstep	0.3971	0.8293
Rumba	0.4094	0.8878
Samba	0.3642	0.7326
Tango	0.6072	0.9884
Viennese waltz	0.4433	0.8000
Waltz	0.4578	0.8182

The proposed method computes the tempogram by element-wise multiplication of the Fourier tempogram and the transformed autocorrelation tempogram. The scales of the autocorrelation tempogram are pooled to match the frequency bins of the Fourier tempogram. The maximum values of the pooled autocorrelation tempogram within each bin are extracted and combined with the Fourier tempogram.

From the result, we can see that the badly performed genre back in question 1 & 2 are improved, and now have the similar or better performance like the result from another type of tempogram. For example, Samba has 0.11 from fourier tempogram and 0.37 from autocorrelation tempogram, now it has 0.36.

Question 4.

Genre	Precision	Recall	F-scores
ChaCha	0.8991	0.8886	0.8938
Jive	0.8302	0.5393	0.6538
Quickstep	0.8421	0.4319	0.5710
Rumba	0.7554	0.7979	0.7761
Samba	0.5492	0.6121	0.5790
Tango	0.8404	0.7836	0.8110
Viennese waltz	0.8857	0.6747	0.7659
Waltz	0.5308	0.6754	0.5944

I think the result is affected by the music style, for example, ChaCha's beats are super easy to follow while Walts usually emphasize on the first beat of a bar which leads to unclear down beat.