HW1

105072123 黃海茵

- **Task1**

dataset	genre	Q1 ACC	Q2 ACC Γ = 1	Q2 ACC Γ = 10	Q2 ACC Γ = 100	Q2 ACC Γ = 1000	Q3 ACC
GiantSteps	Overall	27.88%	26.71%	24.21%	23.71%	23.37%	33.86%
GTZAN	Blues	7.14%	7.14%	4.08%	4.08%	5.1%	18.57%
	Country	32.32%	34.34%	31.31%	31.31%	31.31%	33.64%
	Disco	31.63%	32.65%	28.57%	28.57%	27.55%	34.69%
	Hiphop	13.58%	13.58%	14.81%	9.88%	8.64%	18.40%
	Jazz	16.46%	16.46%	11.39%	13.92%	13.92%	22.41%
	Metal	24.73%	21.51%	19.35%	19.35%	19.35%	30.54%
	Pop	41.49%	40.43%	39.36%	34.04%	32.98%	45.43%
	Reggae	32.99%	31.96%	28.87%	24.74%	25.77%	35.46%
	Rock	34.69%	33.67%	30.61%	27.55%	28.57%	41.12%
	Overall	26.52%	26.16%	23.54%	21.86%	21.86%	31.53%

Q1: Which genres have lower accuracy and can you guess why?

A1: Blues 和 Hiphop 的 ACC 較低，我認為是因為 Blues 主要使用的是 Hexatonic Blues Scale 以及 Heptatonic Blues Scale，而 Hiphop 不是使用傳統的七聲音階編曲技巧。

Q2: How this factor is related to the result?

A2: 加入 Γ 參數後，做了非線性的調整。觀察結果發現，除了 Hiphop 外，其餘都在 Γ=1 時有最佳的 ACC。而大部分的 genre 都是隨著 Γ 越大 ACC 越低，除了少部分有些微回升的情形。

Q3: Discuss the result.

A3: 相較於原先的方法，ACC 有顯著提升。因為這個方法，Perfect fifth, Relative major/minor, Parallel major/minor 也都能得到分數，不像之前只有 Same 才能得到分數，提升的原因顯而易見。

- **Task2**

dataset	genre	Q1 ACC	Q2 ACC Γ = 1	Q2 ACC Γ = 10	Q2 ACC Γ = 100	Q2 ACC Γ = 1000	Q3 ACC
GiantSteps	Overall	42.57%	42.41%	39.73%	37.73%	37.90%	46.63%
GTZAN	Blues	19.39%	18.37%	19.39%	20.41%	20.41%	29.49%
	Country	49.49%	48.48%	44.44%	39.39%	38.38%	51.51%
	Disco	34.69%	37.76%	35.71%	31.63%	31.63%	42.04%
	Hiphop	18.52%	17.28%	17.28%	17.28%	17.28%	24.44%
	Jazz	29.11%	26.58%	24.05%	18.99%	18.99%	35.57%
	Metal	34.41%	33.33%	27.96%	27.96%	25.81%	41.40%
	Pop	58.51%	55.32%	53.19%	53.19%	53.19%	62.02%
	Reggae	57.73%	55.67%	49.48%	39.18%	39.18%	59.07%
	Rock	42.86%	41.84%	39.80%	35.71%	34.69%	47.65%
	Overall	38.83%	37.75%	35.13%	32.02%	31.54%	44.17%

Q4-1: Which feature is better?

A4-1: 使用 K-S method 較好，因為它是透過人體感知實驗所找出的 template，而非太過絕對的 binary template。

Q4-2: Is there any limitation of these method?

A4-2: The algorithm uses an input vector which is weighted by duration of the pitch classes in the piece. It requires a list of notes with ontimes and offtimes. However, in the audio domain, overlap of harmonic components of individual notes in real-world musical recordings would make it a difficult task to determine the actual notes or their duration. A large number of notes are detected in the frequency analysis. Hence the algorithm cannot be directly applied.

Q4-3: Is there any limitation of using GTZAN dataset for key finding?

A4-3: The main limitations of GTZAN is the legality of the dataset, the small size, no complete meta-data regarding artist names and song titles, and no additional meta-data like ratings.

- **Task3**

dataset	KS method ACC	KS method ACC (MIREX)
BPS-FH	43.08%	46.44%
A-MAPS	35.84%	43.31%

Q5: Based on Task 1 and Task 2 的結果，我發現 ACC 最高的是使用 KS method 的 Task2，所以 Task3 我也決定使用這個演算法，並將 segment size 設為 30 秒，得到了還不錯的結果。