CS353 Coursework #1

This assessment is worth 30% of your overall mark for the module.

The purpose of this assessment is to test your understanding of the Python programming concepts taught in the first part of this course. This primarily covers structured programming in Python, use of basic programming constructs such as selection and iteration, string handling, basic data manipulation using NumPy, file handling and working with JSON files and objects in Python.

It also tests your ability to read and interpret a moderately complex scenario/problem and translate that scenario/problem into a fully working Python program that takes account of all of the criteria detailed in the scenario.

Furthermore you are tested on your ability to independently research Python testing utilities and implement them to demonstrate the integrity of your code.

Your task is to build a small application in Python which imports, collates and reports on the results of subjective video quality testing.

Background & Rationale for Data Collection

In modern telecommunications systems video and multimedia applications now account for the vast majority of network traffic. However, despite many recent advances in networking and communications technologies (e.g. 5G mobile networks), it is still not always possible to meet a user's expectations with regards to video playback quality whenever network saturation or other adverse events occur. In order to design mitigation/intervention schemes and algorithms researchers need to understand the impact of adverse network events on the users' perception of video playback quality. This is called measuring the user's Quality of Experience (QoE).

By understanding how different adverse network events impact on a user's QoE, network engineers can design algorithms that will automatically intervene to manage adverse network events in an optimal manner with respect to maintaining the user's QoE.

The main principle of subjective video quality evaluation experiments (measuring QoE) is that a user should be shown a series of different video clips under controlled conditions. Although there are numerous ways of conducting such experiments the most common way is to show the user different versions of each clip. Some of which were delivered over a network link without network impairments and others with a range of different impairments applied to the network link. Statistical analysis is then used to quantify the impact of each type of impairment on the user's QoE.

Data Collection Process and Data Explained

Human Subjects

In these subjective video quality experiments a number of human subjects named "Subject A", "Subject B", etc. were asked to view a series of pairs of short video clips and rate their

perception of the difference in quality of the video playback between each of the videos in a pair. The Absolute Category Rating (ACR) scale shown below was used to map scores to user perception of video quality.

Score	Description
5	No difference
4	Difference perceptible but not annoying
3	Slightly Annoying
2	Annoying
1	Very Annoying

Video Clips

Each video in the set of video clips used was uniquely identified by a single character ID (note lower case and upper case represent distinct video clips so a clip with ID “R” is not the same clip as that with ID “r”). All video clips were encoded using variable bitrate encoding which means that the bandwidth required to successfully deliver the video changes over time in line with the content being presented to the user (e.g. scenes with high levels of motion or detail require a greater bandwidth than those with limited motion or levels of detail).

Network Impairments

The network impairments applied in these experiments were simple bandwidth throttling impairments (i.e. reducing the bandwidth available on the link without changing any other parameters). Each time a clip was shown it have had one of the following network conditions applied to the network link.

Code	Description
MAX	The maximum bandwidth requirement for the video clip, this should always ensure the video is correctly delivered without any loss of perceptual quality.
AV	The average bandwidth required over the duration of the video clip, this will often be lower than the MAX bandwidth requirement and may lead to some minor impairment in scenes of high motion or detail.
-3	The network link is throttled to 3% less than the average bandwidth requirement
-5	The network link is throttled to 5% less than the average bandwidth requirement
-10	The network link is throttled to 10% less than the average bandwidth requirement

Test Number

Each test was individually numbered and consisted of the viewing of two versions of the same video clip, one played locally with no impairment and the other played over a network link under a given network constraint (e.g. video clip “A” under network condition -3). The first clip was the reference clip and the second (impaired clip) was the condition under test. The

data set for each such individual test shows the subject ID and score registered of each human subject who viewed this particular video clip/bandwidth combination. This score represents the human subject's perception of the quality of the impaired clip in comparison to the reference clip using the ACR scale. It is important to note that these combinations were randomly allocated to human subjects. Therefore not all subjects will have participated in every test and some subjects may have viewed the same video clip/bandwidth combination more than once.

A small subset of the data collected from these experiments can be found in the ***subjective_results.json*** file on MyPlace. This is the data input file you will use in this assessment. Fig. 1 shows an extract from this file.

Figure 1. An extract from the data input file showing a single test record.

The Assessment Task

Your task is to build a small application in Python which imports, collates and reports on the results of subjective video quality testing.

The system should have a menu similar to that shown below:

Subjective Quality Evaluation System

Select an option:

1. Import Test Data From File
2. Print Data Summary Report
3. Output Summary Report to File
4. View Details of an Individual Test
5. View Details of an Individual Video Clip
6. Quit

Menu Option 1

This should prompt the user for a file name. The system should only accept JSON files, should always perform a check that the user entered filename exists and prompt the user to re-enter a correct filename where appropriate. The data from the file should be loaded into the system. After importing the data the user should be returned to the menu.

Menu Option 2

Selecting this menu option should generate an appropriately formatted summary report containing all of the data shown below. This report shows the average, minimum and maximum ACR score recorded against each test together with details of the clip and bandwidth setting used. If data has not previously been loaded from file the system must provide appropriate error messaging to the user.

Test No	Clip Id	Bandwidth Setting	Avg. Score	Min Score	Max Score
1	X	-10	3	1	4
2	Y	MAX	5	4	5
.					
.					
60	A	-5	4	2	5

The user should then be given a prompt to return to the main menu.

Menu Option 3

Selecting this menu option should create a JSON file named ***summary_data.json***. The file should be a correctly formatted JSON file containing the same data as produced by menu option 2.

Your application should also produce a second JSON file named ***scaled_data.json***. This file should again be correctly formatted and only contain the test number and average score for each test. N.B to produce the data for this file the average score should be scaled from a range of 1 to 5 to a range of 1 to 20 using an appropriate Numpy arithmetic operation. N.B. You MUST use Numpy for this scaling operation.

The user should then be given a prompt to return to the main menu.

Menu Option 4

Selecting this option should prompt the user to enter a valid test ID number. (You should validate as required). The application should then display an appropriately formatted full set of data (as provided by the input file) for the chosen test number. The user should then be given a prompt to return to the main menu.

Menu Option 5

Selecting this menu option should prompt for a valid Clip ID and produce a report similar to that shown below.

Summary Data for Test Clip X					
Total Number of tests conducted 10					
Bandwidth Setting	MAX	AV	-3	-5	-10
Mean Opinion Score	5	4	3	3	2
Count (N)	14	14	28	14	28

Mean Opinion Score (MOS) is defined as the arithmetic mean of the scores recorded by individual human subjects for each test condition (combination of test clip and bandwidth setting). When R is the rating (ACR score) given by an individual human subject and N is the number of subjects who rated this particular video clip/bandwidth combination, the MOS is calculated as follows:

$$MOS = \frac{\sum_{n=0}^{N} R_n}{N}$$

After presenting the report the user should then be given a prompt to return to the main menu.

Testing

You must independently research testing in Python and implement an appropriate set of tests to ensure that your application is robust and meets the requirements of the scenario described in this document.

Submission

Your Python file, tests and example output files should be zipped in a folder and uploaded to the submission link.

Something to Consider

This exercise is only a basic starting point for the exploration of the data contained in the data set provided in the file. As you learn more in this course you will be able to understand how to analyse the data to properly address the stated aims of the experiments as outlined above. The full data set for these experiments contains many other fields, for example each of the video clips also has a score which reflects the complexity of the video in terms of the amount of motion and changes in texture in the video. This impacts on the peaks and troughs in variable bitrate video streaming as complex scenes with high motion require a higher bandwidth. It could be that you do some clustering of the different video clips based on this value!