Photo Tag Recommendation System

Multimedia Systems and Applications 4/M Coursework Submission

Magda Kowalska

2037342K

1. For this assignment I decided to use Java, because I’m the most familiar with I/O methods and data structures (especially Maps) in this language. When running the program, it creates two outputs. First is the coocurrencePhotoTags.csv file containing co-occurrence matrix, second is the standard output of Java Console containing top 5 recommendations with and without IDF for ‘water’, ‘people’ and ‘london’ tags.
   1. Tag Recommendation Strategy  
        
      The most popular tag recommendation technique gives you tags that co-occur the most with other tags within a given collection. However, this technique tends to recommend unrelated popular tags. The reason for this is because popular tags highly co-exist with almost all tags in a given collection.
      1. Pseudo-code (for my solution):

From a collection of photos extract all used tags

For every photo map co-existing tags

* 1. Tag Suggestion with popularity and significance

This technique takes into account inverse document frequency (IDF) to get rid of unrelated popular tags. The syntax is the same as for Tag Recommendation Strategy however values of co-occurrence are multiplied by IDF factor. IDF is calculated as follows , where I is the number of images in the collection and I(X) is number of images tagged with tag X. In fact, in my solution I used the same Map from Task 1 to calculate values for Task 3, only updating values with the IDF.

* + 1. Pseudo-code (for my solution):