EEE3032

Computer Vision and Pattern Recognition Coursework Assignment

Visual Search of an Image Collection

Kongkea Ouch

Academic Year: 2020-2021

Abstract

The objective of this assignment is to create a basic visual search program that has the capacity to accept an input image and search through an image collection to return the top 10 related image results that are ranked based on their similarity. This program incorporates the use of global colour histogram which is one of the most popular techniques for visual search as the core image descriptor and the Microsoft Research (MSVC-v2) dataset of 591 images (20 classes). Principal component analysis is then introduced to project the global colour histogram descriptors to lower dimensions, and multiple distance measures are tested interchangeably to find out difference in computing image similarity. Each experiment follows by a rigid evaluation methodology that produces precision recall statistics which are essential in assessing how one experiment compares to one another as well as the overall performance of the visual search system.

Table of Content

Description

Content-based image retrieval

A visual search, also known as content-based image retrieval (CBIR) [1], is a computer vision technique that searches for image in a database through the use of visual characteristics such as color, shape, texture or any other information that the image may contain. It is an automated alternative that is conducted on an electronic system which is far more efficient and advanced than traditionally having each image in a database manually fingerprinted by human.

Global Colour Histogram

Colour [2] is one of the most important features that is widely used in CBIR system to distinguish and recognize visual information in an image because it is easy to extract colour features from an image and it is quite fast in indexing and searching through a database.

In this visual search program, only colour attribute is used to determine the similarity between query image and the candidate images in the dataset. Specifically, a colour histogram is implemented on each image by quantizing the RGP space and mapping every pixel's color of the image mapped to a certain bin on the histogram. This colour histogram encapsulates the overall distribution of colour and it points to a position in feature space in which distance to another image can be calculated using many distance measurement metrics such as Euclidean distance, Manhattan distance and Mahalanobis distance.

However, colour histogram [3] has its major disadvantage as it solely depends on the colour of the query image and does not take into account its shape and texture. Colour histogram performs poorly in identifying images with entirely different colours but similar partial information. Yet, images with similar overall colour but different spatial content are very likely to get picked up as positive by colour histogram. For instance, a colour histogram approach may result in an orange cat being matched with an orange fruit while a green car is paired with an area of green grass. To solve this issue, others image descriptors that take into consideration not just colour but shape or texture such as Edge Orientation Histogram (EOH), Scale-invariant feature transform (SIFT) may be used.

Principal component analysis

The entire set of colour histogram descriptors can be readily used for querying related images. But, principal component analysis (PCA) [4] should be used to reduce the number of their elements in the feature space since there are correlations between descriptors which can be represented by lesser principal components due to redundancy.

Result

Evaluation methodology

As this visual search system is heavily based on global color histogram, the most suitable approach for evaluating such low-level visual features is to incrementally determine precision (P) and recall (R) values of each experiment across a range of image outputs. For a robust and consistent image retrieval, every experiment takes 20 handpicked images which are representative of all the 20 image categories from the dataset and query each of them against all 591 images of the dataset by using different descriptors and distance measurements.

The result images of each query can be marked as relevant or not by comparing their categories to the query image's category, and the precision recall (PR) statistics over the 591 files is calculated based on the relevance of these retrieved images. Average precision (AP) which is an overall indicator of how well an image category does in each query is then used to find the mean average precision (MAP) which can be used as a benchmark of an experiment to compare with other experiments. The PR statistics of the 20 images are plotted on the PR curve with depicting of important information such as ranking precision performance of image categories and the MAP of the entire queries.

The aforementioned evaluation methodology is applied to four main experiments which are related to different levels of RGB quantization, global colour histogram, principal component analysis (PCA) and varying distance measurements.

Different levels of RGB quantization

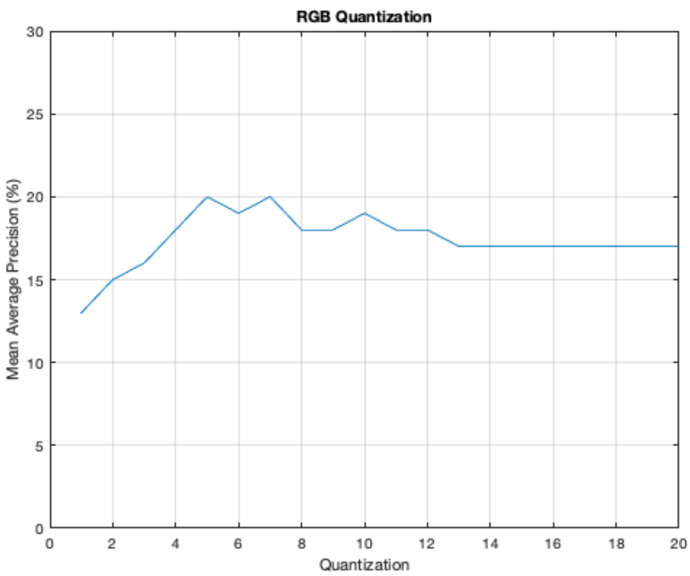

Figure 1: RGB Quantization

The RGB quantization followed the said evaluation methodology and it is experimented using global colour histogram as image descriptor and Euclidean distance measurement for ranking candidate images in feature space. As can be seen from Figure 1, the quantization starts from 1 to 20 incrementing by one and the most optimal values that yield the highest mean average precision are between 4 and 12 having MAP between 18-20% and peaks at 5 and 7 with 20% MAP. It can be noticed that applying higher quantization values is totally unnecessary and inefficient because it does not result in any better precision than the range of 4 to 12 and computing descriptor gets slower and more resource-intensive as the quantization value increases. Hence, the value of 5 is selected as the standard quantization value for other experiments due to the fact that it offers the fastest computing time for descriptor and the highest precision in querying images.

Global colour histogram

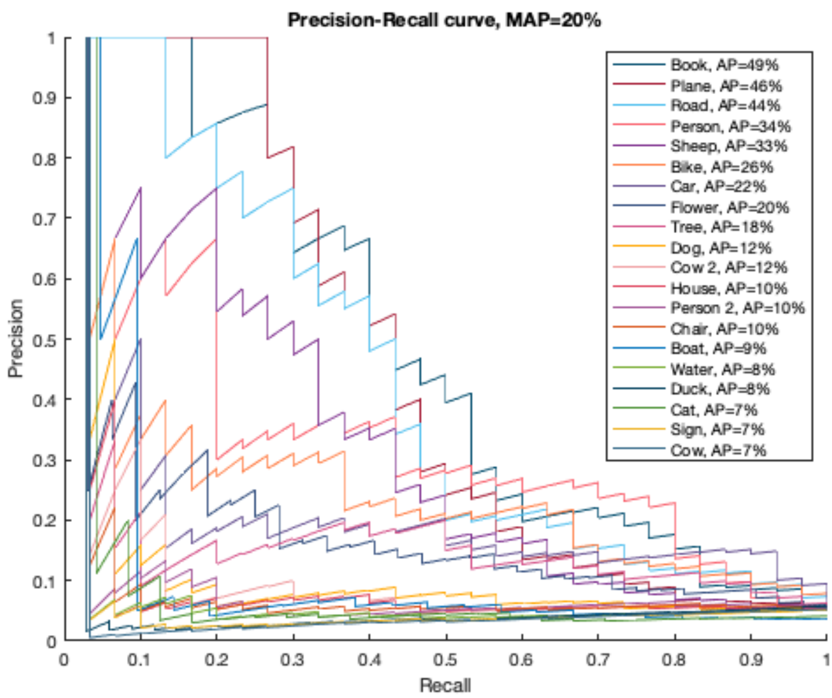

Figure 2: Main global colour histogram

The highest precision rate of 20% is achieved in querying according to Figure 5 with the application of global colour histogram, Euclidean distance and most optimal quantization value of 5. The global colour histogram descriptor focuses only on overall distribution of colors and ignores spatial properties of image, and the low precision output from the graph proves this very idea as it generally signifies a very lackluster image querying performance for most query images with related category's images varying in shape, texture or spatial location of the colors. It can be observed from the graph that the top 3 best performing categories are book

(49%), plane (46%) and road (44%) and the bottom 3 categories are cat, sign and cow which have the same precision of 7%. The top performing images likely contain related images in their categories with high relevance in similar overall colour while the bottom performing images have related images with different overall colour representation.

Figure 3: Top 10 results of book

For instance, the book image performs best because its colorful distribution is significantly shared by many other images in book category as seen in Figure 3 even though there are some faulty results such as bike and dog as they contain similar overall color to book.

Figure 4: Top 10 results of cat

On the other hand, the cat image results in precision according to Figure 4 because very few of its category images have the same greyish overall color like the original input image which instead gets confused and resonates with other similar greyish images such as water and duck.

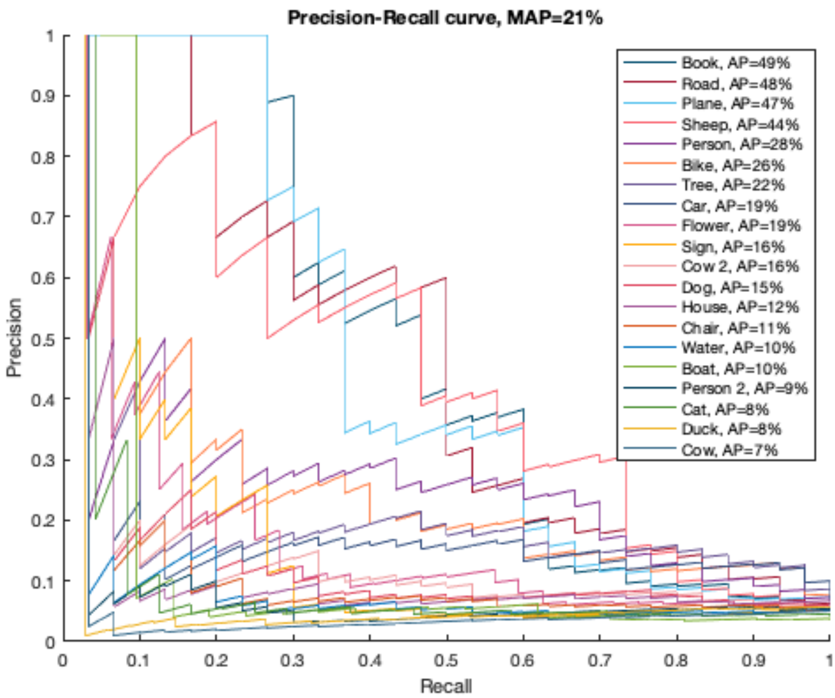

Figure 5: Global color histogram with Manhattan distance

A marginal increase of 1% in overall precision is seen after conducting the same experiment with Manhattan distance instead of Euclidean distance. Comparing between Figure 2 and Figure 5, there are also some noticeable drops and rises among some categories.

Principal component analysis (PCA)

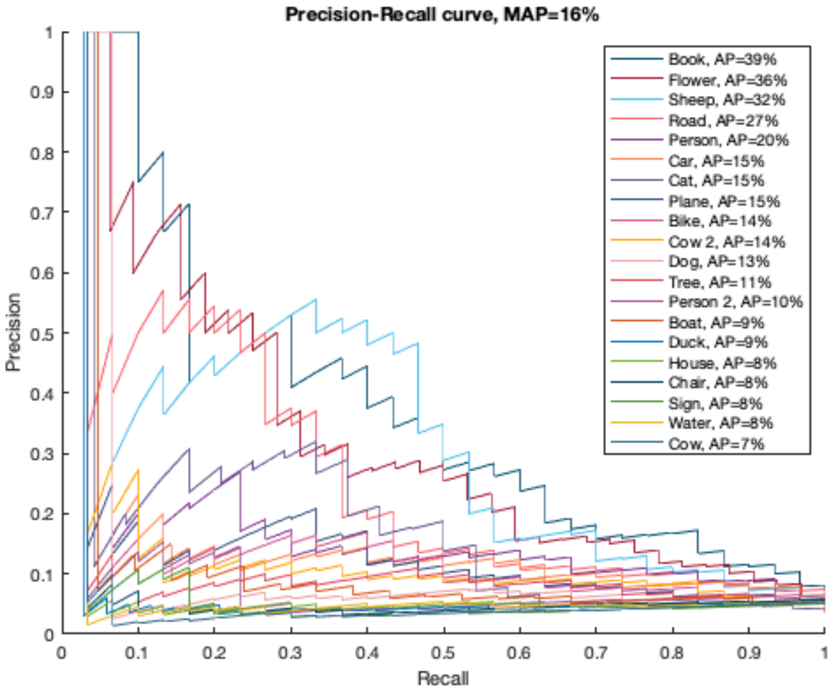

Figure 6: Principal component analysis

Principle component analysis is used to significantly reduce the number of dimensions of global histogram descriptors from 591x125 to 591x3 making it smaller and faster while still retaining the most significant information. As the PCA descriptors are based on color histogram descriptors that result in Figure 5, the ranking of image categories still remains the

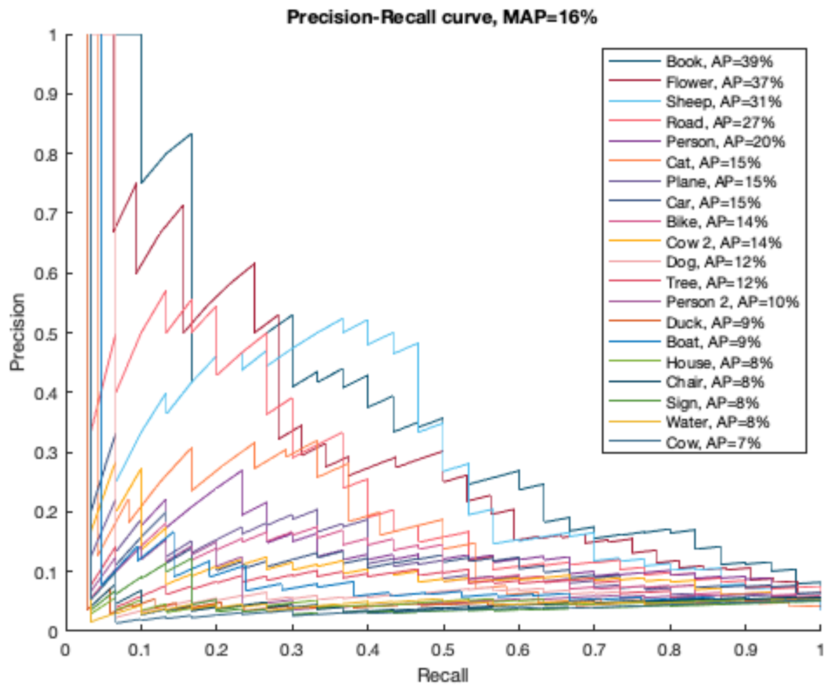

Figure 7: PCA with Mahalanobis

same despite decrease in each category's precision due to use of PCA. Comparing with the overall precision result of Figure 2, there is a 4% drop in mean average precision in Figure 6 as a result of trade-off between feature reduction and accuracy through PCA. When Mahalanobis or Manhattan distance is used to replace the Euclidean distance used in PCA descriptors of Figure 6, there is no noticeable overall improvement in average precision performance except minor variances in some images as exemplified in testing with one of the distances in Figure 7.

Conclusion

This visual assignment makes use of global colour histogram to count the occurrence of colors in image and searches for the most similar images that are ranked highly in termed of the composition of colors. Despite its strength in matching images with shared overall color, the color histogram suffers in performance when it is put to test of querying for images with similar spatial information. The quantization experiment shows that higher quantization level does not improve system's precision but rather regresses it. The most optimal quantization values which provide fast computing time for descriptors are between 4 to 12 with 5 and 7 being the top performing ones. The PCA experiment keeps the global color histogram descriptors at lower dimensions while trading off a degree of precision. It can also be inferred from the distance measurement experiment that the Manhattan distance has a slight edge in performance over the Euclidean distance and that changing measurement types for PCA experiment does not result in any boost in precision. Last but not least, other image descriptors such as SIFT and EOH should be used to identify both color and spatial information which may lead to better precision of the visual search system.

Bibliography

[1] M. Sharma and A. Batra, "Analysis of Distance Measures in Content Based Image Retrieval," *Global Journal of Computer Science and Technology: G Interdisciplinary ,* vol. 14, no. 2, 2014.

[2] G. Khosla, N. Rajpal and J. Singh, "Evaluation of Euclidean and Manhanttan Metrics In Content Based Image Retrieval System," *International Journal of Engineering Research and Applications,* vol. 4, no. 9, 2014.

[3] K. Meskaldji, S. Boucherkha and S. Chikhi, "Color Quantization and its Impact on Color Histogram Based Image Retrieval," in *Proceedings of the 2nd Conférence Internationale sur l'Informatique et ses Applications (CIIA'09)*, Saida, 2009.

[4] B. Malinga, D. Raicu and J. Furst, "Local vs. Global Histogram-Based Color Image Clustering," 2006.