|  |  |
| --- | --- |
|  |  |
| Image Retrieval Using SIFT |  |
|  |  |
|  | Computer VisionAssignment 4 |
|  | Mohamed Ehab Yousri202201236 |

Image Retrieval Using SIFT

This report presents an experimental analysis of image retrieval using Scale-Invariant Feature Transform (SIFT) and clustering-based feature encoding. The goal of this assignment is to evaluate how different parameters affect retrieval performance through three experiments:

1. Varying the number of images used to compute centroids
2. Varying the number of centroids (k) in clustering
3. Comparing TF-IDF vs. Bag of Words (BoW) for feature representation

The analysis involves computing cosine similarity scores, visualizing retrieved images, and selecting the optimal parameters based on retrieval accuracy.

**Data Processing:**

**-Dataset Overview:**

The dataset consists of texture images from six subcategories:

* honeycombed, grid, knitted, pleated, grooved, interlaced
* Each subfolder contains multiple images with distinct patterns.
* The dataset was extracted from a compressed ZIP file stored on Google Drive.
* Images were read using OpenCV and converted to **grayscale** for SIFT feature extraction.
* A subset of **720 images** was selected from the dataset.

**-SIFT Feature Extraction**

* SIFT was used to detect key points and compute feature descriptors.
* **Invalid descriptors (None values) were removed.**
* Key points were visualized for the first image to verify extraction.

**Experiments and Results:**

**Experiment 1: Changing the Number of Images Used to Compute Centroids:**

**Objective:**

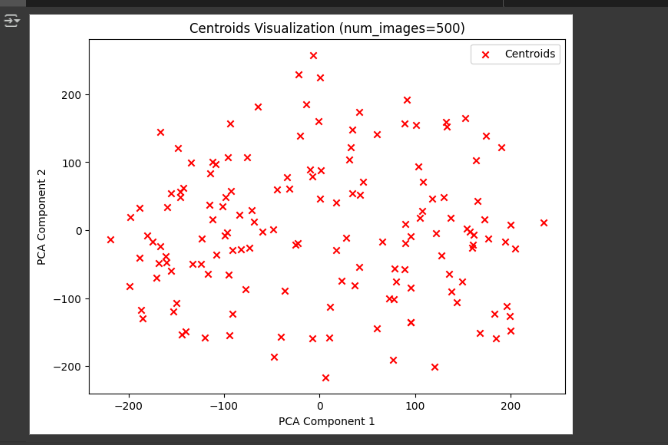
To examine how retrieval performance changes when different numbers of images are used for clustering.

**Method:**

* A subset of **500, 1000, and 2000 images** was selected.
* Descriptors were transferred to the GPU using CuPy.
* **cuML’s K-Means clustering** was applied to group feature descriptors.
* Cosine similarity scores were computed for retrieval.

**Results:**

* Larger sample sizes improved retrieval quality but increased computation time.
* Beyond **1000 images**, the retrieval accuracy did not improve significantly.
* **1000 images were selected as the optimal sample size.**



**Experiment 2: Varying the Number of Centroids (k)**

**Objective:**

To analyze how different k values impact feature representation and retrieval.

**Method:**

* K-Means clustering was performed with **50, 150, and 250 centroids**.
* The resulting cluster centers were used as a **visual codebook**.
* Images were represented as histograms of visual words.

**Results:**

* Increasing k improved retrieval precision initially.
* **150 centroids provided the best trade-off** between retrieval accuracy and computational cost.
* Beyond k=150, retrieval performance plateaued.

A screenshot of a graph

AI-generated content may be incorrect.

**Experiment 3: Comparing TF-IDF vs. BoW**

**Objective:**

Comparing two different feature representation methods:

1. **Bag of Words (BoW):** Counts occurrences of visual words.
2. **TF-IDF:** Assigns weight to visual words based on their uniqueness across images.

**Method:**

* BoW and TF-IDF representations were computed for all images.
* Cosine similarity was calculated for image retrieval.
* **Top-5 retrieved images were visualized for comparison.**

**Results:**

* **TF-IDF performed better than BoW**, as it **reduced the influence of common visual words**.
* **BoW tended to retrieve frequent textures**, even if they were not the most relevant.
* TF-IDF achieved **higher cosine similarity scores** and more relevant rankings.

A graph of a number of images

AI-generated content may be incorrect.

**Final Selection of Best Parameters**

Based on the experiments, the best parameters were:

* **Number of images used for clustering:** **1000**
* **Number of centroids (k) in K-Means:** **150**
* **Feature representation method:** **TF-IDF**

**Justification:**

* These settings provided the best balance of retrieval accuracy and computational efficiency.
* TF-IDF reduced noise in retrieval, resulting in **higher similarity scores for relevant images**.

**5. Conclusion**

This study analyzed the effects of **sample size, cluster count (k), and feature encoding (TF-IDF vs. BoW)** on image retrieval.

Key takeaways: ✅ Increasing the number of training images improves retrieval but has diminishing returns beyond **1000 images**. ✅ A moderate number of centroids (**150**) provides the best balance of specificity and efficiency. ✅ **TF-IDF outperforms BoW** by ranking images more accurately based on **feature uniqueness**.