Universidad Politécnica de Madrid

Escuela Técnica Superior de Ingenieros Informáticos

Information retrieval, extraction and integration

Assignment 2: Non-textual data extraction assignment

Authors:

José Antonio Ruiz Heredia

Joseph Tartivel

Álvaro Honrubia

Teacher:

Raul Alonso Calvo

Date:
February 26, 2025

Contents

1 Motivation

On platforms like Milanuncios, sellers often list cars with incomplete, ambiguous or wrong information. Sometimes listing with only the base model name without key details that properly classify the vehicle (e.g., "Renault clio" instead of "Renault Clio IV phase 1", and all the specification related to the model). This lack of specificity can make it difficult for buyers to find what they are looking for and creates inconsistencies in how car listings are organized, leading to less sells and customers, a bad experience for both sellers and buyers.

With this project, we aim to explore a way to improve the accuracy and completeness of car listings using a Content-Based Image Retrieval (CBIR) approach. The idea is to suggest the most relevant car details based on the image provided by the seller, reducing the need for manual input and improving search precision. Our method relies on a dataset of pre-classified car images, each linked to complete metadata, and attempts to match query images to the closest known model.

To achieve this, we use a two-step matching process: ORB (Oriented FAST and Rotated BRIEF) Matching: We chose ORB because it provides a good balance between speed and accuracy when identifying key structural features of a car. More advanced techniques like SIFT or SURF could be more precise but are significantly slower and computationally expensive, making them less practical for a large-scale platform. ORB allows us to quickly find images with similar shapes and contours, helping us identify posible matches.

When multiple cars have similar shapes, structure alone is not enough to distinguish between models. To refine our results, we incorporate a double histogram-based color matching technique that combines Conventional Colour Histogram (CCH) and Stacked Colour Histogram (SCH). This approach enhances retrieval by capturing both global color distribution and texture details, improving robustness against variations in lighting and perspective. While color isn't a defining characteristic for model identification, this method acts as a useful tie-breaker by filtering out unlikely matches and prioritizing images with a more similar overall appearance.

Our approach faces several challenges, including variations in seller-uploaded images (lighting, angles, and resolution), the similarity of many car models, and the limitations of color matching due to reflections or edits. Ensuring computational efficiency at scale is another key concern. Despite these obstacles, exploring a combination of structural analysis and color-based refinement could offer a more automated and accurate way to enhance online car listings. By addressing these challenges, we aim to contribute to a more efficient and user-friendly vehicle classification system.

2 State of Art

2.1 ORB

Oriented FAST and Rotated BRIEF (ORB) was introduced in 2011 by Ethan Rublee et al. as an efficient alternative to Scale-Invariant Feature Transform (SIFT) and Speeded-Up Robust Features (SURF). It was designed to provide fast, robust and scale and rotation invariant feature detection and description for real-time applications. It quickly became a preferred method for fast and efficient feature matching, especially in resource-limited environments. ORB improves two existing feature detection and description techniques, Oriented FAST (oFAST) feature detection by adding orientation information and Rotated BRIEF (rBRIEF) feature description by making it rotation-invariant. While ORB is fast and robust, it still has some limitations, particularly in image retrieval applications.[2]

Hybrid solutions are applied to solve this, for example Hybrid ORB Algorithm proposed in 2024 by Thakur & Gondkar.[3] Helping to focus on important edges, reducing noise and false feature matches, smoothing images for improved feature consistency and accuracy and improving retrieval accuracy by combining ORB with preprocessing techniques.

2.2 Histogram-Based Color Matching

The paper ”Effective Image Representation using Double Colour Histogram for Content-Based Image Retrieval” presents an advanced method for improving image retrieval by combining Conventional Colour Histogram (CCH) and Stacked Colour Histogram (SCH). This approach enhances color-based matching by capturing both global color distribution and texture information, making it highly effective for distinguishing images with similar colors but different patterns. In our car image matching system, this methodology is particularly useful as it ensures robust color matching even under variations in lighting, angles, and minor transformations. By integrating both CCH and SCH, our system can more accurately compare vehicle images, improving the reliability of identifying similar cars based on color and texture cues. [1]

3 Implementation

For this work we selected the Stanford Cars Dataset for its 16,185 images and diverse range of vehicle models, under multiple angle, making it a reliable reference for our system.[4] While our approach focuses on visual matching using structural and color-based features, the dataset's detailed metadata could provide additional classification insights if we were to extend our project further.

This code implements a hybrid image matching system that integrates ORB feature matching with color-based analysis to identify similar images within a dataset given query image.

3.1 Orb Matcher

The *OrbMatcher* class utilizes ORB to detect keypoints and extract descriptors from images, computing a similarity distance metric to identify the closest matching images. First, the dataset images are processed, and their descriptors are precomputed and stored to enhance the efficiency of the actual CBIR process. This preprocessing step significantly reduces computation time during retrieval. Then, *OrbMatcher* compares the query image descriptors against those stored in the precomputed dataset. Matching is performed using a brute-force matcher (*cv2.BFMatcher*), which finds correspondences between keypoints. The final similarity score is determined based on the number of ORB keypoint matches between the query and the dataset descriptors. For better insight into the inference process, the final output also visualizes the keypoints detected in the images. This allows for a performance evaluation of the algorithm and helps identify any irregularities in the matching procedure.

3.2 Color Matcher

The *ColorMatcher* class computes two types of color-based features for images, the Conventional Color Histogram (CCH) and the Stacked Color Histogram (SCH). CCH is based on standard RGB color histograms, while SCH applies blur and averaging over several iterations. These features are combined into a single vector and compared to find the most similar images based on color similarity.

3.3 Hybrid Matcher

The *hybrid_matching* method integrates ORB and color-based results to provide a combined ranking. First, it filters images using the *OrbMatcher* to reduce the search space based on the *orb_matches_count* metric, then it computes the *ColorMatcher* to obtain the *color_distance* metric of the reduced list. Finally, a combined score is computed, weighing the *orb_matches_count* more heavily than the *color_distance*, and the top results are displayed along with their corresponding metrics and combined score. This method allows the system to benefit from both the geometric features (ORB) and visual color features (CCH + SCH), providing a robust image matching solution.

3.4 Other Experiments

During implementation, we conducted several experiments that did not yield the anticipated results. Initially, we pursued background removal techniques after discovering that background elements adversely affected our color matcher algorithm. The matcher was unintentionally incorporating background colors into its analysis, while ORB feature detection frequently identified points in the background rather than on the target subject, thus corrupting our matching model's performance. However, the background removal method *BackgroundSubtractorMOG2* lacked sufficient sophistication to address the extensive diversity of backgrounds present in our dataset. Consequently, this approach introduced additional complexity to our model while producing negative performance outcomes.

We also evaluated alternatives to our brute-force ORB matcher. Our first alternative, Fast Library for Approximate Nearest Neighbors (FLANN), implemented a nearest neighbor search approach that increased processing speed but resulted in reduced precision compared to our original method. Subsequently, we tested Hierarchical Navigable Small World (HNSW), which utilizes a graph-based approach. Similarly to FLANN, this method enhanced system velocity but at the cost of decreased accuracy. After comprehensive evaluation of these performance trade-offs, we determined that accuracy was the paramount concern for our specific application. As the processing time achieved with the brute force ORB matcher was deemed acceptable for our requirements, we ultimately decided to maintain this approach in our final implementation.

4 Future Work

Our image processing requires several critical improvements to improve accuracy and efficiency. We need to implement a robust segmentation technique to completely remove background elements from images while simultaneously evaluating various algorithms to determine optimal performance across different conditions.

To address inconsistencies in our data, we want to develop standardized procedures that normalize lighting variations, ensuring consistent vehicle positioning through calibration protocols, and correct for environmental biases that currently skew our results. In addition, as an alternative to our current ORB matcher, we have to test SIFT, SURF, and AKAZE algorithms, measuring their comparative performance in terms of accuracy, processing speed, and overall robustness.

Finally, we are moving beyond traditional color matching by investigating texture-based recognition methods, exploring the potential of deep learning-based feature extraction, and researching hybrid approaches that combine multiple recognition strategies to create a more comprehensive and efficient system.

References

[1] Martey, Ezekiel Mensah, Lei, Hang, Li, Xiaoyu, and Appiah, Obed. "Effective image representation using double colour histograms for content-based image retrieval." *Informatica*. Available: https://www.informatica.si/index.php/informatica/article/view/3715/1663

[2] Rublee, Ethan, Bradski, Gary, Konolige, Kurt, and Rabaud, Vincent. "ORB: an efficient alternative to SIFT or SURF." Available: https://www.researchgate.net/publication/221111151_ORB_an_efficient_alternative_to_SIFT_or_SURF

[3] Manish Kumar Thakur and Raju Ramakrishna Gondkar. "Enhanced Image Retrieval using Hybrid ORB Algorithm." *International Journal of Intelligent Systems and Applications in Engineering (IJISAE)*. Available: https://www.ijisae.org/index.php/IJISAE/article/view/7190/6159

[4] Li, Jessica. "Stanford Cars Dataset" *Kaggle*. Available: https://www.kaggle.com/datasets/jessicali9530/stanford-cars-dataset