Expanding Color Query Results via Image Recoloring

Sixing Hu,[1] Pierre-Yves Laffont,[2] Brian Price,[3] Scott Cohen,[3] Michael S. Brown[4]

[1] National University of Singapore, Singapore [2] ETH Zurich, Switzerland [3] Adobe Research, USA [4] York University, Canada

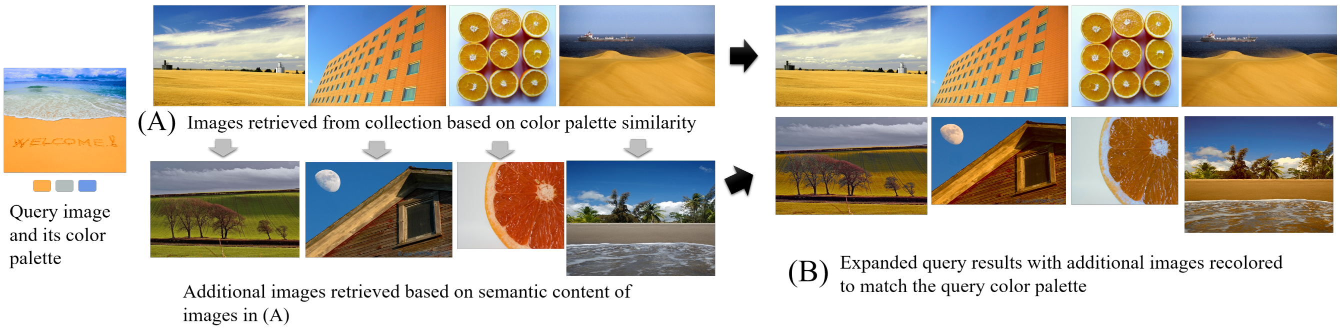


Figure 1: *An image's color palette is used to query an image collection. (A) shows typical results of this type of query where images in the collection with similar colors are found. (B) shows an example of the expanded results produced by the proposed framework. The results in (A) are used to find semantically similar images in the image collection. These images are then recolored to better match the query. This expansion by recoloring provides more results to the user and allows images in the collection to be utilized.*

Abstract

An image's color palette can be used to search an image collection to retrieve images that share similar colors to the query palette color. The results of this approach are naturally limited to only those images in the database that share similar colors to the query. The idea proposed in this paper is to expand the search results by finding additional images in the collection that can be computationally recolored to better match the query. This not only provides more results to the user but also helps to extend the usefulness of the image collection. We describe a prototype system to realize this idea, using a two-step procedure. First, images in the database with color palettes similar to the query are identified to produce a set of initial results. Then, additional images that are semantically similar to these initial results are found and are modified using palette-based recoloring such that they better match the color query. We demonstrate results from our prototype and discuss several challenges for developing such image search systems.

Categories and Subject Descriptors (according to ACM CCS): I.4.8 [IMAGE PROCESSING AND COMPUTER VISION]: Scene Analysis—Color

1. Introduction and Related Work

While stock photos and image collections are commonly searched by keywords or categories, another relevant search modality is by color. Query by color can be performed in a number of ways, one of which is to have the user provide an image whose color palette is used for the query. The collection is then searched to find images that have the most similar color palettes. Figure 1-(A) shows an example. The results of color-based queries are naturally limited to the small percentage of images in the collection that share similar colors. The goal of this work is to expand these results by finding images in the collection that can be dynamically recolored such that they better match the query colors. This system not only provides more results to the user but also has the added benefit of extending the usefulness of the image collection by allowing the images to be dynamically modified to satisfy queries they would previously not be suitable for. This functionality is desirable for commercial stock photo sites and for image content producers.

The challenge for this task is twofold. First, it is necessary to de-

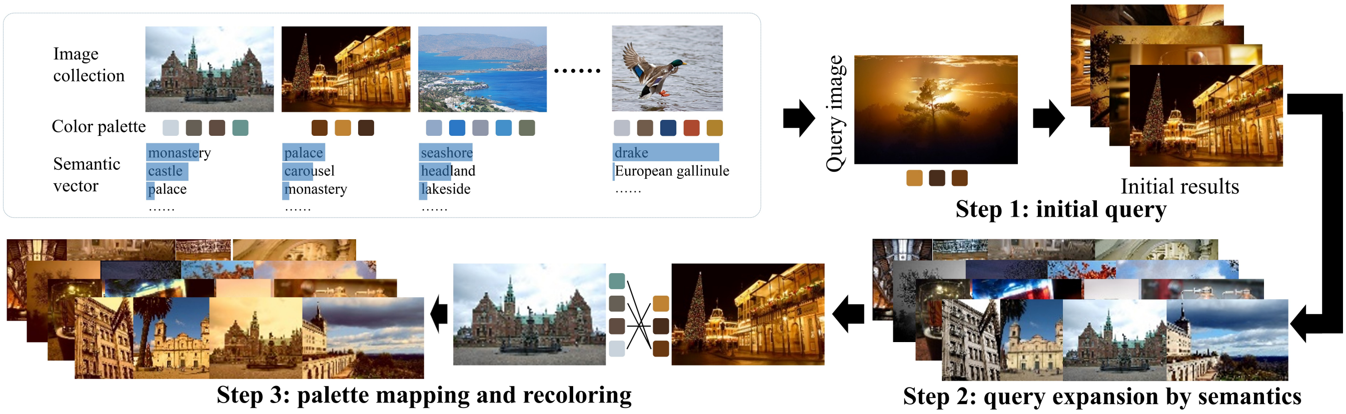


Figure 2: *Overview of the proposed framework. The image collection is preprocessed to extract color palettes and semantic information of each image. When a query image comes, a set of initial images is first retrieved based on color similarity (step 1). Then, additional images are retrieved based on semantic content (step 2). These additional images are recolored to match the query color (step 3) by a palette-based color transfer method.*

termine which images in the collection are suitable to be recolored for a given query. Finding a subset of images suitable for recoloring is necessary because it would be computationally prohibitive to recolor the entire image collection, and not all images in the collection may be suitable to be recolored by the query colors. The second challenge involves applying a suitable recoloring method that provides realistic and natural results.

We have developed an initial prototype system to address this problem. Our approach uses a two-step procedure that starts by finding images in the collection that are similar to the query image's color palette. Specifically, we use the earth mover's distance (EMD) as a similarity metric to compare the query palette color histograms and the color palettes of the images in the collection. This first step is the same as a conventional color retrieval framework and provides an initial set of images whose colors match the query. We then find additional images that contain similar semantic content to the initial retrieved results; to do this, we leverage recent advances in deep learning for semantic image retrieval. Since these additional images are semantically similar to the initial query results, which share similar colors with the query, we postulate that these additional images stand a high chance of being good candidates for recoloring.

These additional images now need to be modified such that their colors better match the query colors. There are a number of methods available that can be used for image recoloring. Most notable are color transfer methods (e.g., [RAGS01, PKD05, XM09, PR11, NKB14, FSDH14]) that modify an image's colors to better match a target image's colors. These methods often work in a global manner and focus on aligning the color histograms between the source and target images. There are also methods that segment the image by color or simple region-based semantics to do local color manipulation (e.g., [TJT05, WDK*13]). Recently, palette-based methods have been proposed that have proved highly effective in manipulating an image's color appearance. Wang et al. [WYW*10] proposed a data-driven method to recolor an image to a particular color theme in order to enhance the image's appearance. Recent work by Chang et al. [CFL*15] proposed a method that can realistically recolor an image by modification of the palette colors. Our system builds upon this latter method, and leverages the earth mover's distance optimization to assist in determining how to map the query colors to the additional images' palettes and perform the image recoloring. Figure 1-(B) shows an example of the results obtained with our framework.

While this short paper represents work in progress, our work offers the following contributions: (1) we propose a framework for expanding the search results obtained by color query by recoloring images in the image collection to match the query; (2) we demonstrate a prototype of this framework based on semantic image search and palette-based image recoloring. The remainder of this paper details our prototype and discusses a number of open challenges for this problem.

2. Proposed Framework

Figure 2 provides an illustrative overview of our proposed framework. Each step in our framework is described in this section.

Step 1: Initial query with color-based search. We have adopted a palette-based approach to do both the image query and recoloring. To this end, each image is described by a small set of colors–the color palette that represents the most dominant colors in the image. We use a method similar to Chang et al. [CFL*15] to extract the palette colors, $\{c_k\}, k = 1, 2, \ldots, K$, and the corresponding weights, $\{w_k\}, k = 1, 2, \ldots, K$, that record the number of pixels in the image associated with each color. In our implementation, we allow K to range from 2 to 5 colors and use a dynamic k-means clustering approach to select the best K for a given image. This is applied as a preprocessing step to all images in the collection.

Given a query image, I_q, its color palette is determined and images with similar color palette are found in the collection using the earth mover's distance (EMD). The EMD [RTG00] between distributions P and Q is the minimum cost way to transform one distribution to the other by moving or flowing the weight in P to match the weight distribution in Q, where the moving cost is weight times distance moved. Let $P = \{(p_i, w_{pi})\}, i = 1, 2, \ldots, m$ and $Q = \{(q_j, w_{qj})\}, j = 1, 2, \ldots, n$ be the two distributions. The total cost to match P and Q with a flow $F = (f_{ij})$ is

$$\mathrm{cost}(F, P, Q) = \sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} d_{ij},$$

where d_{ij} is the L_2 distance between p_i and q_j and f_{ij} is the amount of weight moved from p_i to q_j. Then $\mathrm{EMD}(P, Q) = \min_F \mathrm{cost}(F, P, Q)$. In our approach, the color palette and associated weights are the distributions with $P = \{(c_k, w_k)\}, k = 1, 2, \ldots, K$ and $Q = \{(c'_k, w'_k)\}, k = 1, 2, \ldots, K'$. The terms c_k and c'_k represent the colors of the source and target color palette, respectively. Once the flow giving the minimum cost is found, we can rank all the images in the collection. We can either return all images under a certain cost threshold, or, as done in our prototype, return a fixed number of the topmost m images (in our case $M = 12$) as the initial query results $I = \{I_1, I_2 \ldots, I_M\}$.

Step 2: Query expansion with semantic search. Our approach also relies on semantic information in each image to expand the initial color query. To associate semantics with each image, we use a convolutional neural network model trained on the ADE20K dataset [ZZP*16] that provides 2693 semantic classes that describe both foreground objects and background scenes. After passing the image through the neural network, we obtain a semantic vector $Sem = (sem_s), s = 1, 2, \ldots, S$ where each element sem_s is a confidence value for each of the classes. Semantic vectors of all images in the dataset are precomputed and stored in the database.

This semantic information is used to retrieve more images from the initial results. For each image in the initial results I_i, we find the top 10 additional images with similar semantics. For our prototype system, we used a fixed number of additional images; however, this can be easily modified to be variable in size based on a threshold. This is done by computing the L_2 distance of the Sem vectors of all images in the collection. This collective set of additional images is denoted as $\{I'_i\}$.

Step 3: Palette-based recoloring of the expanded query images. To perform the palette-based recoloring, we need to establish a mapping between the source image's color palette and that target color image's palettes. Our approach leverages the optimization procedure of the EMD and uses the optimal flow matrix to determine the assignment between the source and target images. We let the distance in the EMD be the L_2 distance between two palette colors on the a and b channels of Lab color space: $d_{ij} = \|c_i - c'_j\|_2$. Using the EMD optimization procedure in [RTG00], the optimal flow matrix $F^* = \{f^*_{ij}\}$ is obtained. The term f^*_{ij} is the flow from c_i to c'_j. We assign the color c_i to the color c'_{j*} where $j^* = \mathbf{argmax}_j(f^*_{ij})$, which means that we recolor c_i in the expanded result image to the color in the initial (semantically similar) result image that is matched most to c_i in the EMD matching of the color palettes.

After obtaining the assignment between the query palette and target palettes, we use a modified version of Chang et al. [CFL*15]'s method to recolor the source image. The approach in [CFL*15] proposes a method that uses the Lab color space and modifies the target image's colors by modifying all three values. We found that for our application, using only the chromaticity component (i.e., the a and b channels) yields better results than also modifying the L channel. Thus, in contrast with [CFL*15] our recoloring does not affect the overall brightness of the target image.

3. Results and Evaluation

We have developed a fully functional interface that allows image search, expansion, and recoloring to be performed in real time. This search system is implemented on a machine with an Intel Xeon E3-1270 3.50GHz processor and 32GB RAM. The database we use is the MIRFLICKR-1M dataset [HTL10], which contains one million Flickr images. Using precomputed color palettes and semantic vectors, our approach can retrieve the initial 12 color results and additional 120 semantic results in roughly 300ms. The additional 120 images can be recolored in approximately 80ms per image.

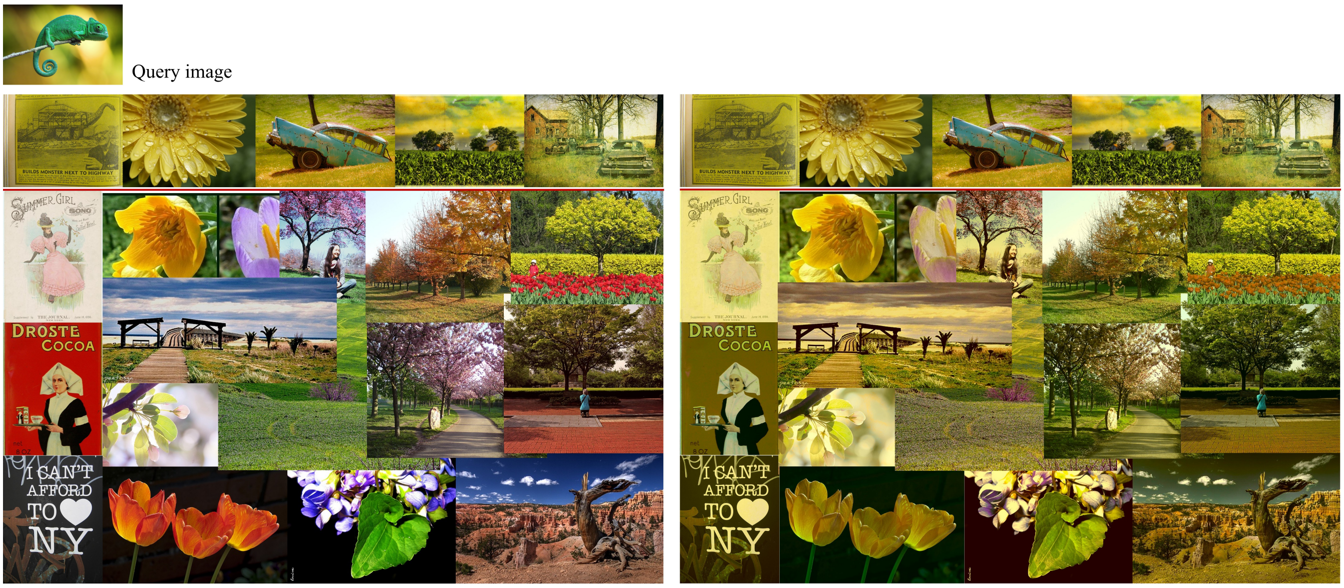
Figure 3 shows typical results of each step for a query image. Through the color transfer method, images searched by semantic vectors are recolored to share similar color to initial results. As seen in Figure 3-(B), no noticeable noise or block artifacts are introduced by the recoloring method.

Figure 3: *Results of our image search, expansion, and recoloring prototype. (A) Initial images searched by color palette. (B) Additional results by recolor expansions. (C) Expansions searched by semantics in terms of initial results.*

Figure 4 shows another example for a user-specified query image. In Figure 4-(A), the color-based query results are shown; below the red line show the additional images found via the semantic search. Figure 4-(B) shows the results that would be returned to the user. The additional images have been recolored using the method described in Section 2.

4. Summary and Discussion

This paper presents a preliminary prototype system for a novel color-based search system that expands color query results by recoloring images that otherwise would not be returned by the query. There are a number of open challenges needed in order to improve the current system.



(A) Results after search without recoloring candidate images.

(B) Results after search with expanded recolored images.

Figure 4: *Demo result of our image query expansion system. (A) Images above the red line are the initial results based on the color query. Images below the red line are the additional images searched based on semantics. (B) The bottom part contains additional images recolored to fit the query to produce the expanded query results.*

First, a formal metric for predicting if an image is suitable to be recolored by a given set of target colors is needed. This is related to being able to quantitatively measure the results of recoloring or color transfer. While there are metrics for comparing color transfer methods–for example, the color-mapping-round-trip [FPC*15]–there is still room for improvement. Also, it would be computationally advantageous to be able to determine the suitability for recoloring without having to recolor the images first.

There is also a need for recoloring methods that incorporate object-level semantics. Current methods can inadvertently recolor objects such as faces, foliage, and skies in an unnatural manner. This is a serious issue that will likely benefit from recent advances in deep learning. Lastly, many high-quality recoloring methods are computationally expensive and not suitable for real-time use on a large number of full-resolution images. These challenges require further investigation and would help to advance color transfer and image recoloring for use in the consumer application.

References

[CFL*15] CHANG H., FRIED O., LIU Y., DIVERDI S., FINKELSTEIN A.: Palette-based photo recoloringg. *ACM Transactions on Graphics 34*, 4 (2015), 139:1–139:11. 2, 3

[FPC*15] FARIDUL H. S., POULI T., CHAMARET C., STAUDER J., REINHARD E., KUZOVKIN D., TREMEAU A.: Colour mapping: A review of recent methods, extensions and applications. *Computer Graphics Forum 35*, 1 (2015), 59–88. 4

[FSDH14] FRIGO O., SABATER N., DEMOULIN V., HELLIER P.: Optimal transportation for example-guided color transfer. In *ACCV* (2014). 2

[HTL10] HUISKES M. J., THOMEE B., LEW M. S.: New trends and ideas in visual concept detection: The mir flickr retrieval evaluation initiative. In *ACM Multimedia Information Retrieval* (2010). 3

[NKB14] NGUYEN R., KIM S., BROWN M.: Illuminant aware gamut-based color transfer. *Computer Graphics Forum 33*, 7 (2014), 319–328. 2

[PKD05] PITIE F., KOKARAM A. C., DAHYOT R.: N-dimensional probability density function transfer and its application to color transfer. In *ICCV* (2005). 2

[PR11] POULI T., REINHARD E.: Progressive color transfer for images of arbitrary dynamic range. *Computers & Graphics 35*, 1 (2011), 67–80. 2

[RAGS01] REINHARD E., ASHIKHMIN M., GOOCH B., SHIRLEY P.: Color transfer between images. *Computer Graphics and Applications 21* (2001), 34–41. 2

[RTG00] RUBNER Y., TOMASI C., GUIBAS L. J.: The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision 40*, 2 (2000), 99–121. 2, 3

[TJT05] TAI Y.-W., JIA J., TANG C.-K.: Local color transfer via probabilistic segmentation by expectation-maximization. In *CVPR* (2005). 2

[WDK*13] WU F., DONG W., KONG Y., MEI X., PAUL J.-C., ZHANG X.: Content-based colour transfer. *Computer Graphics Forum 32*, 1 (2013), 190–203. 2

[WYW*10] WANG B., YU Y., WONG T.-T., CHEN C., XU Y.-Q.: Data-driven image color theme enhancement. *ACM Transactions on Graphics 29*, 6 (2010), 146:1–146:10. 2

[XM09] XIAO X., MA L.: Gradient-preserving color transfer. *Computer Graphics Forum 28*, 7 (2009), 1879–1886. 2

[ZZP*16] ZHOU B., ZHAO H., PUIG X., FIDLER S., BARRIUSO A., TORRALBA A.: Semantic understanding of scenes through the ADE20K dataset. *arXiv preprint - arXiv:1608.05442* (2016). 3