PRISM: A System for Weighted Multi-Color Browsing of Fashion Products

Donggeun Yoo[1], Kyunghyun Paeng[1], Sunggyun Park[1], Jungin Lee[1 2], Seungwook Paek[1 2], Sung-Eui Yoon[1], In So Kweon[1]

[1]KAIST
Daejeon, 305-701, Korea

[2]Cldi Inc.
Daejeon, 302-847, Korea

ABSTRACT

Multiple color search technology helps users find fashion products in a more intuitive manner. Although fashion product images can be represented not only by a set of dominant colors but also by the relative ratio of colors, current online fashion shopping malls often provide rather simple color filters. In this demo, we present PRISM (Perceptual Representation of Image SiMilarity), a weighted multi-color browsing system for fashion products retrieval. Our system combines widely accepted backend web service stacks and various computer vision techniques including a product area parsing and a compact yet effective multi-color description. Finally, we demonstrate the benefits of PRISM system via web service in which users freely browse fashion products.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Storage and Retrieval; I.2.6 [**Artificial Intelligence**]: Learning

Keywords

Web fashion product images, Multi-color search

1. INTRODUCTION

With the rapid growth of web technology, we are now facing massive amounts of online products. Among these, fashion items occupy the biggest portion of online products [7]. Latest web-based fashion shopping services provide basic color-based browsing features[1 2 3] since the most distinctive property of fashion products is visual appearance which plays the most important role in buying decision [1]. However, the degree of freedom in the color-based approaches is highly restricted to choosing a set of *binary* color filters, i.e. users are allowed to control only the existence of particular color in the search results.

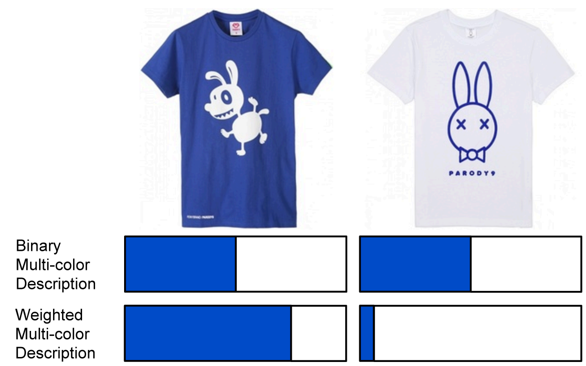

Figure 1: A motivating example. In conventional binary multi-color search engine, the system cannot distinguish between the two t-shirt images.

Fig. 1 shows an example illustrating the limitation. The two images of t-shirts contain identical dominant colors, blue and white, with different ratios. In a conventional color-based search system using binary color filters, these two products will be represented very similarly to each other. For more refined results, users need to spend more time to search manually. Thus, a service that provides *weighted* multi-color search is required to provide a much faster way to find the desired products.

Multiple color search engine for fashion products is differentiated from conventional systems in two major aspects. First, weighted multi-color search requires a high dimensional spatial database (HDSD) while binary color filtering can be easily implemented using relational database management systems (RDBMS). Deploying HDSD in commercial systems with a large number of products and users requires query time to be comparable to that of RDBMS while maintaining visual accuracy. Second, parsing the product area in a fashion product image is essential to improve the quality of search results. The visual information outside the actual product area should not be considered during the visual description process. This task becomes more challenging when product images contain noisy information such as outdoor background, retailer logos and fashion models wearing the products.

This demo presents PRISM (Perceptual Representation of Image SiMilarity), a system for fast weighted multiple color search of fashion products on the web. PRISM combines

[1]`http://www.ebay.com`
[2]`http://www.shopstyle.com`
[3]`http://www.polyvore.com`

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2577012.

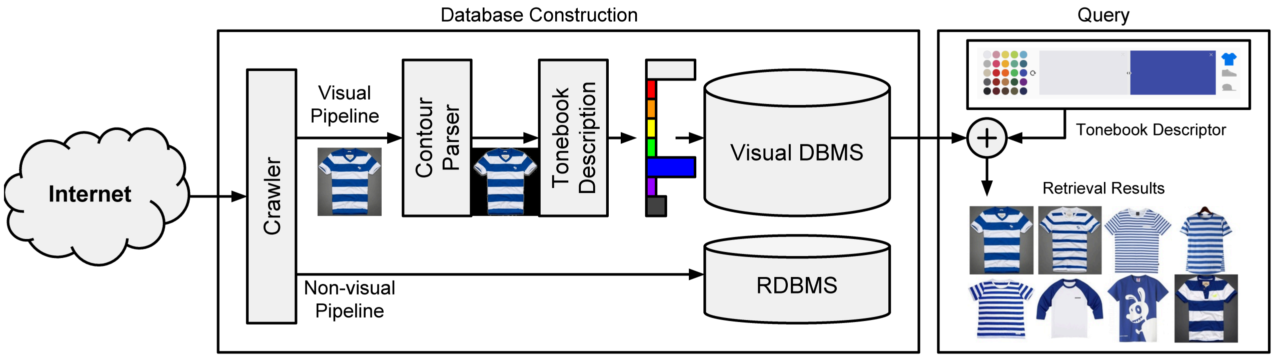

Figure 2: The overview of PRISM system.

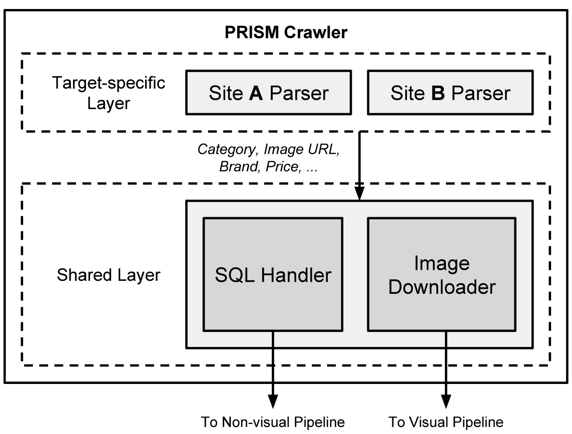

Figure 3: The two-layer structure of our crawler.

various computer vision techniques together with widely accepted web backend stacks. In particular, we focus on two factors in designing multiple color search systems: (1) an efficient product parsing from fashion product images, (2) a fast in-memory HDSD with perceptual color clustering. First, we describe a product area parsing process specially designed for fashion items. From the parsed area, the color distribution is encoded into a small vector, also known as a descriptor. In order to increase search speed, a compact color descriptor called *tonebook* which reflects color similarity measure in human visual perception is developed. This encoding technique reduces the memory requirement and the query time simultaneously while retaining the search quality. We store the tonebook descriptors from product images in our in-memory HDSD achieving less than 100 ms of query time.

2. FRAMEWORK OF PRISM SYSTEM

Fig. 2 shows the overall block diagram of our system. The PRISM backend system is composed of three modules - crawler, contour parser and weighted multi-color description engine.

2.1 Crawler

Our crawler collects a large amount of fashion product information from online shopping malls. To support diverse data structures of shopping malls with minimal development cost, we used a two-layer architecture consisting of target-specific and shared layers, as depicted in Fig. 3. Only the target-specific layer is modified when new shopping malls are added. The shared layer receives structured data of products from the target-specific layer and initiates two-way data processing pipelines. The non-visual pipeline directly stores the textual meta-data of the product into RDBMS such as MySQL. The visual pipeline converts the downloaded product image into a compact descriptor vector for weighted multi-color search. These pipelines are synchronized to the primary key of the product information in RDBMS.

2.2 Contour Parser

Our fashion product parsing algorithm consists of three stages. The first stage is a contour extraction stage which extracts outermost contour from a given image. As a pre-processing, the input image is converted into an edge image by Canny edge detector [3]. Then we applied the contour searching algorithm implemented in OpenCV[4] to find every contour from the edge image. The contour with the maximum area is chosen for the next stage.

In the second stage, a category-aware contour classifier decides whether the given product contour candidate is correct or not. We used 'Bag of Hash Bins' approach[4] to obtain a contour descriptor. Using a set of contour descriptors from our training dataset, support vector machine (SVM) classifiers are trained for every product categories. Once a contour from a query is classified positive, the inside of the contour is directly used as a mask for multi-color description.

If the product category is for an upper body (jacket, shirt, dress, etc.), the contour descriptor with negative confidence is reexamined using face [9] and skin detector [5]. We mark the detected area as a strong background and refine the contour using the GrabCut segmentation algorithm [6]. The resulting contour is classified again by the same SVM classifier and passed on to the next stage when the classification result is positive.

2.3 Weighted Multi-color Description

The objective of designing weighted multi-color description is to encode color distributions where the human's perceptual color similarity is reflected, and to make the descriptors compact for fast search. We encode the color distribu-

[4] `http://opencv.org/`

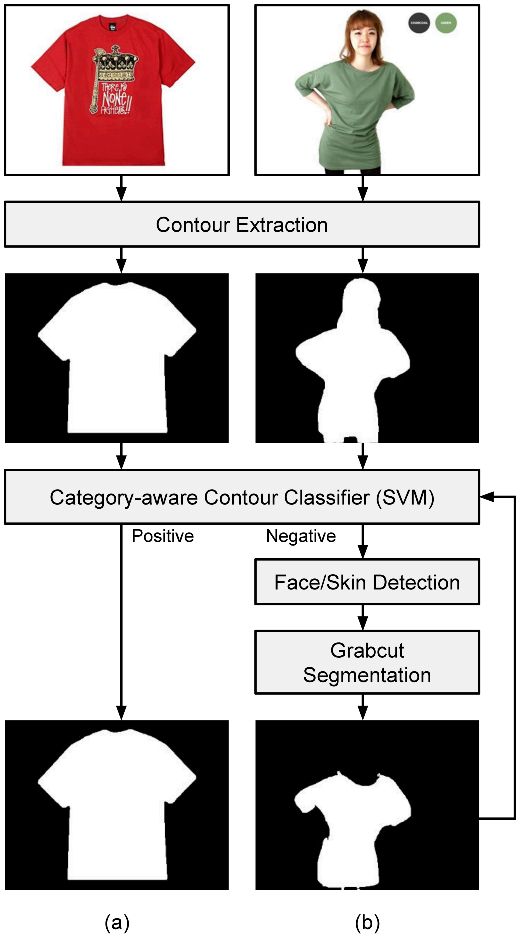

Figure 4: Contour parsing process.

tions of the product area in each image based on 'Bag of Color' (BOC) model [10], which describes an image as a histogram of color-word appearences. However, if we use the BOC model-based framework naïvely, one significant problem arises: BOC utilizes k-means clustering algorithm that may separate two colors which are not distinguishable with the human eye, or group those which are distinguishiable on the other hand since k-means algorithm often uses a distance metric that does not take into consideration human's perceptual color model.

To solve this problem, we propose a human-assisted method for generating a color codebook, named tonebook, to reflect the human's perceptual color model. We collected 5,000 training images of fashion items and over-segmented them into super-pixels [2]. Each super-pixel is given a label which is one of {Red, Orange, Yellow, Green, Blue, Navy, Purple, Black, Gray, White}. With these labeled super-pixels, we trained a transformation matrix, which maps the original color space into a new perceptual color space maximizing the margin between each colors by using metric learning framework [8]. In this space, we trained a color codebook containing color vocabularies by k-means. After obtaining the color codebook, we nonlinearly map the BOC space to a new perceptual BOC space again. This stage is where human-assisted grouping takes place. We grouped color-

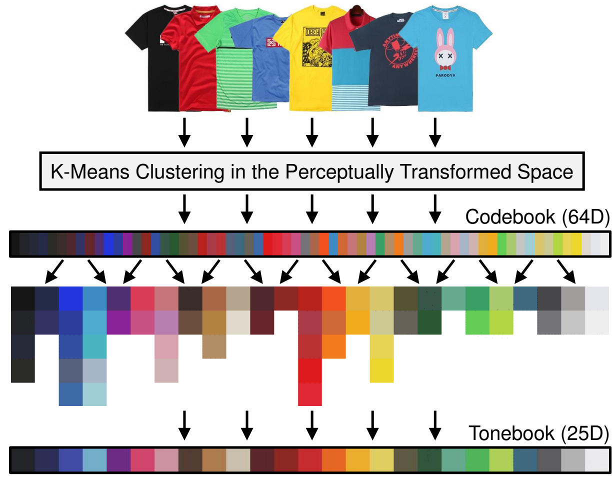

Figure 5: Two-stage tonebook generation process.

vocabularies which are often regarded as the same colors in fashion items. This tonebook generation process is described in Fig. 5. The dimensionality reduction gives two adventages: (1) better invariance to illumination and shade, (2) compact image representation, which enables scalable and fast search since a database image is represented by a vector with a much smaller dimension.

Given a tonebook and a linear transformation matrix, which are learned before, we describe all images as following. For each image, color pixels within a product area are mapped to the perceptual color space by the transformation matrix. The transformed vectors are assigned to the nearest color vocabularies, and the color vocabularies are grouped in the same way as explained above. The grouped vocabularies are L_1 normalized, resulting in a final tonebook descriptor. When a query input is given, we search vectors by computing all L_1-distances between the query.

We used CIELAB as the initial color space before transformation. The sizes of the color codebook and the tonebook are 64 and 25, respectively. To evaluate the two perceptual transformation steps (one for the color codebook and the other for the tonebook), we compare our results with [10] using a color codebook of size 25. As shown in Fig. 6, the retrieval results obtained by the proposed method contain more products with desired color ratio.

3. DEMONSTRATION

We demonstrate the benefits of our PRISM system via a web service named LookPickr[5]. LookPickr provides two use cases; weighted multi-color search and visual similarity search. These two scenarios use the same backend system but with different query interfaces.

With the multi-color search interface, users can select multiple colors from tonebook cluster centers and manipulate their relative ratio to generate a more fine-grained query input (Fig. 7). In addition, a target product category also can be set by users. Given a query input consisting of multi-color information and target product category, a master of PRISM system distributes the query to HDSDs containing tonebook descriptors of selected categories. The master then merges

[5]http://www.lookpickr.com

Figure 6: Comparisons of top-12 retrieval results.

the retrieval results from the selected HDSDs and reranks the final product list according to the similarity scores. Using this weighted multi-color search, users can find the desired products in a faster and more controlled manner.

The second use case is a product similarity search. If a user wants to find other products similar to a specific product in our service, he or she can directly use the product as a query input. The tonebook descriptor in HDSD associated to the query product is used, and a target product category is set to the same one as the query product. Users can repeatedly use similarity-based search to jump over the various products. This feature provides enjoyable shopping experience by increasing chances to discover products.

4. CONCLUSIONS

In this demo, we present PRISM, a weighted multi-color browsing system for fashion products. PRISM integrates various computer vision techniques: product area parsing and a compact yet effective multi-color description. We demonstrate the benefits of PRISM system via LookPickr service in which users freely browse fashion products in a very intuitive manner. In the future, we plan to introduce other visual features such as pattern and silhouette into PRISM system.

5. ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2010-0028680).

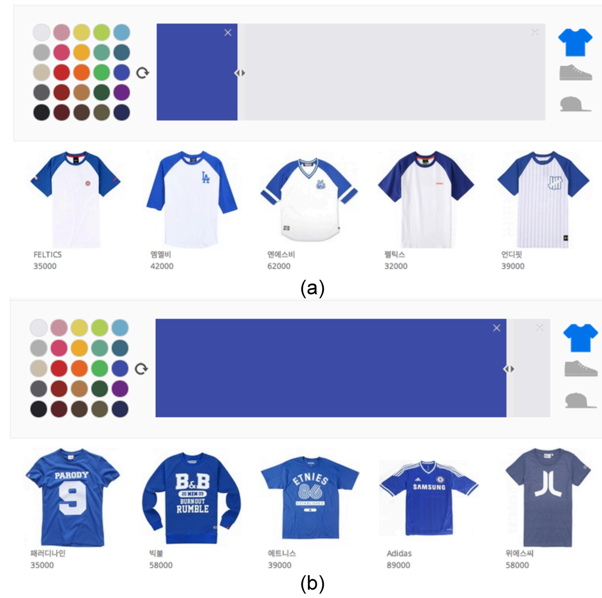

Figure 7: The web-based user interface and top-5 results of blue/white queries where each major color is (a) white and (b) blue.

6. REFERENCES

[1] How do colors affect purchases? URL `http://blog.kissmetrics.com/color-psychology`.

[2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2282, 2012.

[3] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[4] J. He, J. Feng, X. Liu, T. Cheng, T.-H. Lin, H. Chung, and S.-F. Chang. Mobile product search with bag of hash bits and boundary reranking. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2012.

[5] M. Jones and J. Rehg. Statistical color models with application to skin detection. *International Journal of Computer Vision*, 46(1):81–96, 2002.

[6] C. Rother, V. Kolmogorov, and A. Blake. Grabcut: Interactive foreground extraction using iterated graph cuts. In *ACM SIGGRAPH*, 2004.

[7] K. Saleh. Us online retail sales - statistics and trends [infographic], July 2013. URL `http://www.invesp.com/blog/ecommerce/us-online-retail-sales.html`.

[8] M. Schultz and T. Joachims. Learning a distance metric from relative comparisons. In *Advances in Neural Information Processing Systems*, 2003.

[9] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2001.

[10] C. Wengert, M. Douze, and H. Jégou. Bag-of-colors for improved image search. In *Proceedings of the 19th ACM International Conference on Multimedia*, 2011.