Color Navigation by Qualitative Attributes for Fashion Recommendation

Yeongnam Chae, Jiu Xu, Björn Stenger and Soh Masuko
Rakuten Institute of Technology, Rakuten, Inc.,
1-14-1 Tamagawa, Setagaya-ku, Tokyo, Japan
{yeongnam.chae, jiu.xu, bjorn.stenger, so.masuko}@rakuten.com

***Abstract*—This paper proposes a novel method to navigate a color palette using attributes recognized from speech input. Our target application is a fashion recommender system for mobile e-commerce. Starting with a selected color, a user can request to show items of a different color by qualitative attributes (*e.g.* 'a little cuter'). These attributes are mapped to a query vector within the Lab color space in order to select the next color. The system distinguishes 85 attributes, each with three different possible magnitudes. This color navigation by speech was demonstrated in a mobile fashion recommender system. The proposed model is validated in a user study with 196 subjects.**

I. Introduction

The use of mobile devices for e-commerce has seen a strong increase in volume in recent years, as more and more products can be conveniently placed at the consumers' fingertips. More recently, e-commerce chatbots have been launched as a new shopping platform to enhance the shopping experience. However, challenges remains between real-world shopping experiences and digital shopping experiences. For example, one challenge is allowing the user to efficiently navigate the product space via emotional attributes. We consider the scenario in which users are presented with fashion items of a particular color and wish to explore items of related colors with a particular quality. For example, when a user is presented with a dress in burgundy red, she may wish to also look at other dresses in a 'more cheerful' color. In order to support this, we introduce attribute-based color navigation. An attribute, input via speech, is converted to a query vector in the Lab-color space by interpolation in polar coordinates. As a use case of the proposed method, we developed *fashion concierge* mobile application and implemented the attribute-based color navigation as shown in Fig. 1.

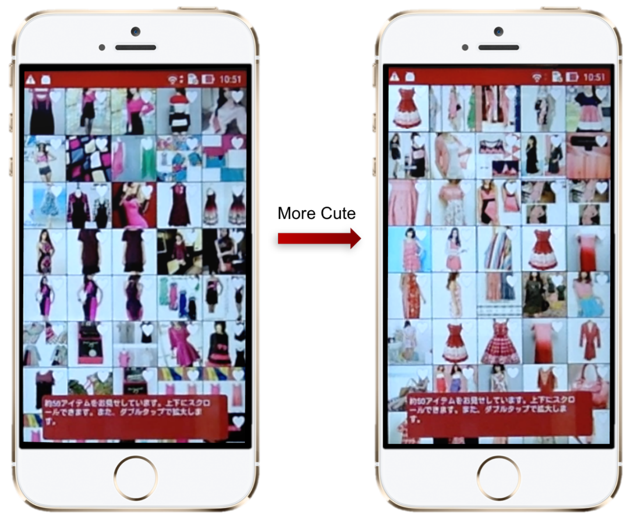

Fig. 1. **Mobile Fashion Concierge.** *The proposed fashion recommender system using speech input for color navigation. The displayed product selection is updated based on the input of qualitative attributes, which is mapped to a vector in the color space.*

II. Color Navigation

In this section, we describe the details of the color navigation by using qualitative attributes. We use the Lab color space, which is designed to approximate the human visual perception of colors [1]. The Euclidean distance in Lab color space approximately corresponds to the perceptual color differences [2]. It makes that we can use linear regression to model the perceptual color differences in the human vision.

In order to efficiently search data sets with millions of items, we use a discrete color palette (JIS Z8102 [3], Japanese Industrial Standard) containing 269 colors. This standard links color values with color names and attributes. In the JIS color description system, the hue dimension is discretized to 20 bins and colors within each bin are arranged in 2D on the brightness and saturation plane. Associated with each color is a qualitative attribute, such as "brighter" or "subdued".

A. Query Vector Generation

To interpret the attribute-based query, we map it to a query vector in color space. In order to establish the relation between color values and attributes, we make use of a public dataset relating example color palettes and attributes [4]. Some relationships such as "vivid" are picked up from the JIS color system.

To navigate along the 'attribute axis' we define a path from the current source color to the target color. The search query is required to lie on this path, the distance from the source color is dependent on the magnitude value in the search query ("a little", "more", "much more").

There are many ways to define a path between a source and target color pair. We consider two different options: (1) linear interpolation in the 3-dimensional Lab space, and (2) linear interpolation in the polar coordinate representation of

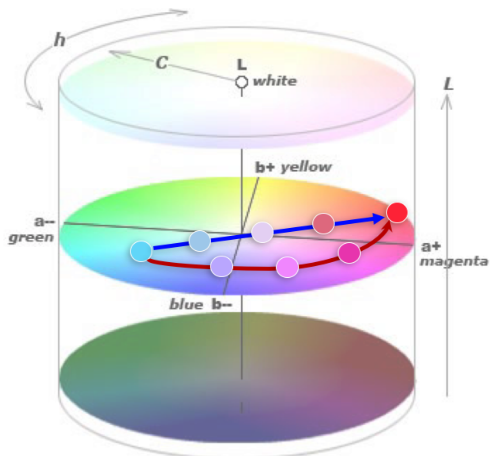

Fig. 2. **Interpolation paths.** *Interpolation path in a Euclidean and polar coordinate system on Lab space.*

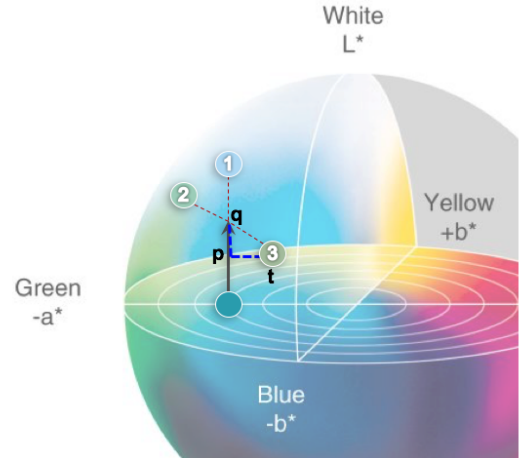

Fig. 3. **Distance measure.** *Vector projection based distance measure that includes directional information.*

the chromaticity components.

The polar coordinate interpolation is defined as:

$$\begin{bmatrix} L \\ r \\ \theta \end{bmatrix} = \begin{bmatrix} L_s \\ r_s \\ \theta_s \end{bmatrix} + \rho \begin{bmatrix} L_t - L_s \\ r_t - r_s \\ \theta_t - \theta_s \end{bmatrix} \tag{1}$$

where $\boldsymbol{s}$, $\boldsymbol{d}$ are source and target color vectors, L is the luminance value in Lab space, and r_r and θ_i are the magnitude and angle in the chromaticity plane, shown in Fig. 2. ρ is the magnitude parameter which determines the degree of change with adverbs such as "more", "much more" and "a little". It is decided by splitting the path equivalently.

B. Navigating the Color Palette

To navigate the color palette, we find the closest palette color from the query vector. Using the Euclidean distance, however, For example, in Fig. 3, the Euclidean distances from query vector $\boldsymbol{q}$ for "brighter" to candidate target vector 1, 2, 3 is the same, while perceptually 1 is most similar, given that it has the same chromaticity value as $\boldsymbol{q}$. We therefore add separate luminance and chromaticity components in the distance computation. For a more general formulation, we project the target color vector of the palette color onto the query vector and define the distance measure as the sum of Euclidean distances on the path from the query to the target via the projected vector, shown in Fig.3. The distance is calculated as follows:

$$d(\boldsymbol{q}, \boldsymbol{t}) = ||\boldsymbol{q} - \boldsymbol{p}|| + ||\boldsymbol{p} - \boldsymbol{t}||, \tag{2}$$

where $\boldsymbol{q}$ is the query vector, $\boldsymbol{p}$ is the projected vector and $\boldsymbol{t}$ is the target vector in color palette.

For each attribute together with its magnitude we calculate a graph to navigate all 269 colors in the JIS palette. An example relation for "more vivid" and "more mellow" colors are presented in Fig.4. Each vertex in the relation graph corresponds to a color in the JIS palette and each directed edge represents the relationship defined by the attribute between two colors. Colors of convergence, i.e. end points in the navigation by this attribute, and sample navigations are found by traversing paths in this graph.

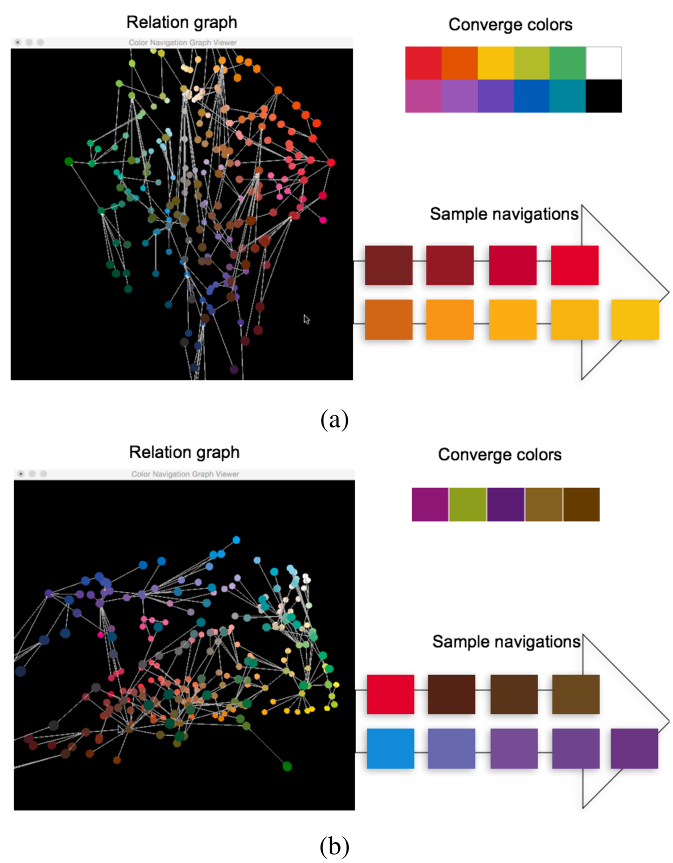

Fig. 4. **Sample relation.** *(a) "more vivid" relation graph and sample navigations (b) "more mellow" relation graph and sample navigations.*

III. Demonstration in a Fashion Recommender

We implemented a fashion recommender system, *Fashion Concierge*, with approximately 400,000 items extracted from a large Internet marketplace. This system displays fashion items of the selected color as determined by attribute-based query. In the pre-processing stage, each item is assigned to a JIS color based on the most dominant color in the images. We used the Google speech API [5] for speech recognition.

A. Item Image Segmentation

To determine the dominant color of each fashion image, we extract the foreground region of an image based on visual saliency. First, we select 300 potential object regions using selective search [6] and accumulate the weights for each window, including overlapping areas so as to generate a heat map for saliency detection. The heat map is binarized using Otsu's method for minimizing the intra-class variance [7], see Fig. 5.

B. Item Color Recognition

After the image segmentation step, dominant colors from the segmented image are acquired. The color values of all pixels are clustered using K-Means with $K = 10$ clusters. For each cluster center we search for the most similar color in Lab space and merge duplicated colors if there is more than one center belonging to the same color in the the palette. Thus, for an input image we obtain a table with the color palette index and its corresponding percentage in the image.

IV. Results

To validate the proposed navigation model, we carried out a survey with 194 users comparing the two different methods to interpolate in the color space based on the same attribute-based input. Each user is presented with 30 samples where each contains an attribute and magnitude of change, e.g. "a little more cheerful". Below each attribute query we display (in randomized order) the current (source) color and the results of the proposed query color using (a) linear interpolation in 3D space and (b) interpolation in polar coordinate space. Users select the color that they consider to more accurately describe the change defined by the attribute, also being given the option of 'neither'. The overall result is shown in Fig. 6. The colors suggested by polar coordinate interpolation show consistently better agreement with the descriptions by attribute. The performance increases for larger magnitudes of change.

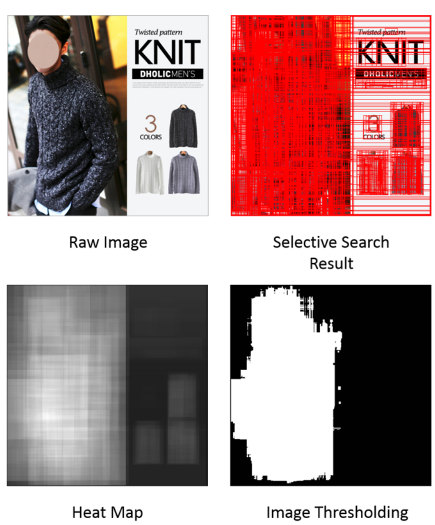

Fig. 5. **Dominant color extraction:** *To obtain the dominant color within an input image (left) we find potential object regions using selective search, accumulate their weights, and perform thresholding (right) to obtain an approximate foreground mask.*

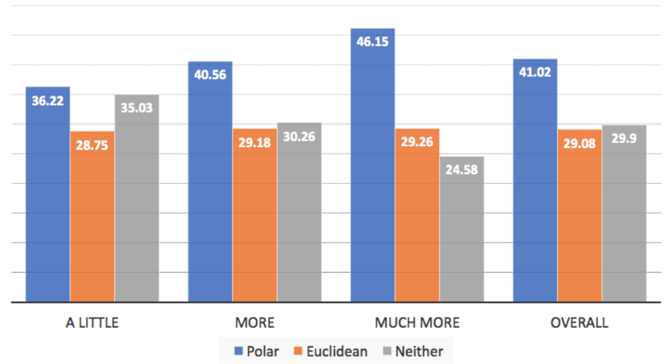

Fig. 6. **User study result.** *The query interpolated on polar coordination is closer to humans feeling than the query interpolated on euclidean coordination.*

V. Conclusion

This paper proposed a novel method to navigate a color palette via an attribute-based query. By interpolating the path between input color and target color in polar coordinates, the attribute-based query is converted to a query vector in the Lab color space. By measuring the distance between candidate color and query vector with the proposed distance measure, we navigate on JIS color palette. We validated the proposed model in a user study. In the future we consider expanding the navigation from a single to multiple target colors simultaneously.

References

[1] Lab color space. (2017, June 1). In Wikipedia, The Free Encyclopedia. Retrieved 04:08, June 13, 2017, from https://en.wikipedia.org/w/index.phptitle=Lab_color_space&oldid=783299531

[2] Anil K. Jain. "Fundamentals of Digital Image Processing." (1989) pp. 68, 71, 73. ISBN 0-13-336165-9.

[3] Japanese Industrial Standards, Names of non-luminous object colours, JISZ8102:2001.

[4] Imagination of Color, http://usui-tosou.net/iroimage.html#chic

[5] Google Cloud Speech API, https://cloud.google.com/speech/

[6] J.Uijlings, K.vandeSande, T.Gevers, A.Smeulders "Selective search for object recognition." International Journal of Computer Vision 104.2 (2013) 154–171

[7] N. Otsu "A threshold selection method from gray- level histograms." IEEE Transactions on Systems, Man, and Cybernetics 9.1 (1979): 62-66