Use of Personal Color and Purchasing Patterns for Distinguishing Fashion Sensitivity

Takanobu Nakahara[✉]

School of Commerce, Senshu University,
2-1-1 Higashimita, Tama-ku, Kawasaki-shi, Kanagawa, Japan
nakapara@isc.senshu-u.ac.jp

Abstract. In this research, we focus on customers with high or low fashion sensitivity to determine how they differ regarding their color choices and the garments that they purchase. Customers with high fashion sensitivity tend to purchase items that are more expensive than those purchased by customers with low fashion sensitivity, and therefore the high-sensitivity customers contribute more to sales. Furthermore, the purchasing characteristics of customers with high fashion sensitivity represent important information for product development. We ascertain these features by comparing two customer groups based on purchasing data from e-commerce apparel sites, including customer ID and garment color information, and also questionnaire data. Specifically, we enumerate emerging patterns to make a classification model of high or low fashion sensitivity by adding information about personal color preferences based on color psychology using the concept of four seasons.

Keywords: Emerging patterns · Hierarchical structure
Sensory marketing · Classification model

1 Introduction

Our objective herein is to ascertain the predilection for purchasing garments on e-commerce apparel sites by using color information of the garments from the perspective of sensory marketing. Sensory marketing is an emerging paradigm for both businesses and consumers. It is intended to aid or influence a person's thinking both consciously and unconsciously through the five senses of sight, sound, smell, touch, and taste. Of these, we focus on sight, which is considered to be one of the strongest human senses [1] with regard to sensory stimuli and perception, to determine the features of fashion sensitivity.

Sensory perception is known to change according to experience. For example, bitterness is a signal of things that are inherently poisonous, so preferences for coffee, beer, and other bitter food and beverages arise via a transition from discomfort to pleasure through repeated consumption. Also, by listening repeatedly to music that is initially of no interest, it can become favorable. As in such

phenomena, it is known that sensitivity in general changes with the number of contacts, which is known as the "mere-exposure effect" [2].

In this research, we focus on customers with either high or low fashion sensitivity to assess how they differ in color preferences and the garments that they purchase. Customers with high fashion sensitivity tend to purchase items that are more expensive than those purchased by customers with low fashion sensitivity, and therefore the former customers contribute more to sales. Furthermore, the purchasing characteristics of customers with high fashion sensitivity represent important information for product development. We ascertain these features by comparing two customer groups based on purchasing data from e-commerce apparel sites, including customer ID and garment color information, and also questionnaire data. The purchasing data covers approximately 100,000 customers whose purchasing behavior was collected each day for 12 months. The questionnaire data were collected from 3,000 customers and comprise their answers to roughly 100 questions. These data are courtesy of a data analysis competition held in fiscal 2016 and sponsored by the Joint Association Study Group of Management Science (JASMAC).

In computational experiments, we construct a classification model that distinguishes between customers with high or low fashion sensitivity by using emerging patterns (EP) [3]. These distinguish between high and low fashion sensitivity by adding information about personal color preferences based on color psychology using the concept of four seasons [4]. Ultimately, the goal is to make customers with low fashion sensitivity more interested in fashion and thereby increase the sales of e-commerce sites by repeatedly applying external stimuli based on their preferences.

In Sect. 2, we present the method for enumerating EPs and we construct a classification model using them. In Sect. 3, we present and analyze the results obtained by applying our proposed method to actual data. Finally, in Sect. 4, we draw conclusions and suggest future work.

2 Methodology

In this section, we describe the personal color used as color information for the first time and describe the pattern mining method and model construction using the color information.

2.1 Personal Color

The color information registered in the data does not completely represent the color of the clothes. For example, even if clothes has multiple colors, only the most representative color among them is registered by text. However, it is difficult to express the color of clothes with only monochrome text information, and even if analyzed using only that information, it is impossible to obtain correct results. Therefore, to handle uncertain color information, we use the personal color to classify colors into groups, and analysis is performed based on those groups.

Figure 1 shows the personal color based on the four seasons method. In personal color, we handle two basic colors: a warm base such as yellow that represents a warm feeling, and a cool base such as blue that expresses a cold feeling. All colors belong to those two bases, except for white, gray, and black. Furthermore, when a certain color belongs to the warm base, it is divided into seasons of spring or autumn, and if it belongs to the cool base it is divided into summer or winter. Therefore, a certain color can be represented by the hierarchical structure of the bases of warm and cool and the four seasons. Note that one color may belong to multiple seasons; for example, burgundy belongs to both spring and autumn.

In fact, each color belongs to the Practical Color Coordinate System (PCCS), which is a discrete color space indexed by hue and tone as developed by the Japan Color Research Institute [5]. Figure 2 shows an image of the PCCS, which categorizes colors by tone. It divides individual hues into 12 tones (vivid, soft, pale, etc.) based on the impressions that they impart in terms of vividness. A vivid tone is a grouping that is close to a pure color. Raising the lightness produces a pale, light tone and lowering the lightness produces a deep, dark tone described as dark grayish [6].
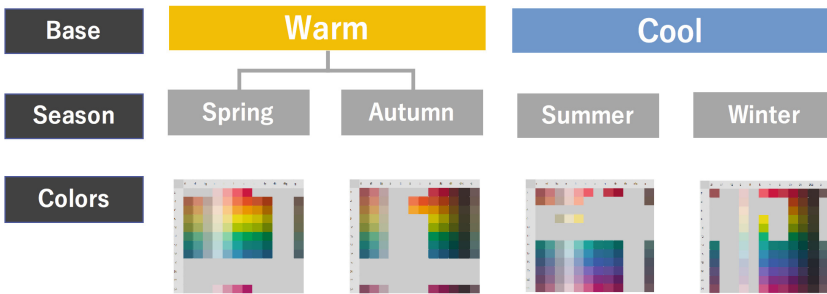

Fig. 1. Personal color. One color is represented by the hierarchical structure of the bases of warm and cool and the four seasons. (Color figure online)

For the color information of the data to be used, the color names of the clothes are entered as text such as "Burgundy," and it is necessary to match those names with the PCCS color information handled by the four seasons method. In this research, matching was done using a "color search dictionary"[1], which is a color search site. We inputted the color name and acquired the PCCS information from the website.

Table 1 lists the personal characteristics of the colors belonging to each season from the study of color psychology.

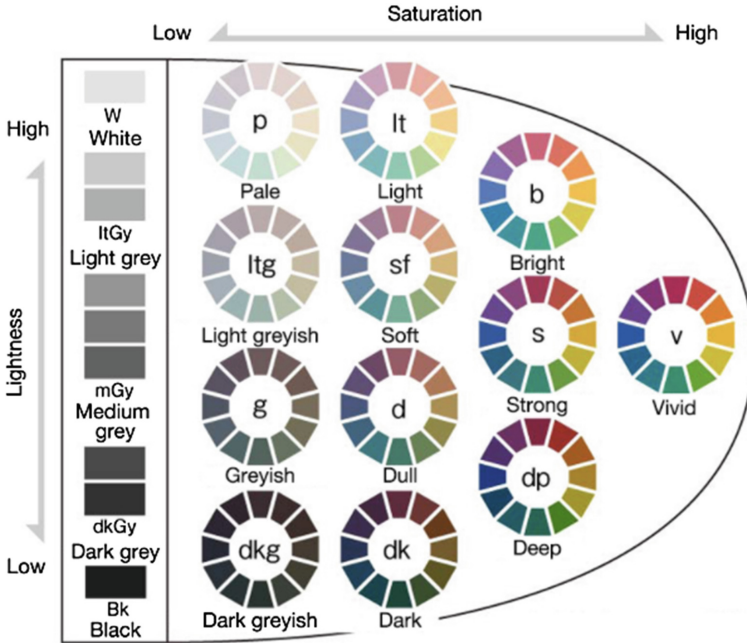**Spring.** People of spring type have cute, fun, and lovely clothes with youthful bright colors.

---

[1] http://www.colordic.org/search/.

Fig. 2. Image of the Practical Color Coordinate System (PCCS), which categorizes colors by tone (adapted from [6]).

Table 1. Characteristics of each season based on color psychology

	Personality impressions
Spring	bright, youthful, active, lovable, pop
Autumn	calm, trendy, mature, gorgeous, sophisticated
Summer	quiet, refreshing, gentle, trendy, precise
Winter	sharp, vivid, clear, luxury, modern

Autumn. People of autumn type have an adult atmosphere with a natural feeling and wear natural clothes that look nice and calm.

Summer. People of summer type wear clothes that give a soft, elegant atmosphere, cool and sweet impression.

Winter. People of winter type generally have a distinctive and unique atmosphere and their clothes colors look sharp and full of contrast.

Using these characteristics of the seasons, we can estimate a person's personality from the colors that they choose.

2.2   Classification Model

In constructing the classification model, this study defines those customers with high fashion sensitivity as a positive class and those with low fashion sensitivity as a negative class and uses these classes as dependent variables. Fashion sensitivity is determined by the scores of the four questionnaire items relating to fashion consciousness. Specifically, the four items are "Clothing is one way of showing my personality," "Fashion is part of my lifestyle," "I often look at what other people are wearing," and "One's value can be increased or decreased by clothing." A person who answered "yes" to more than three of these four questions was defined as having high sensitivity (678 people), and someone who answered "no" to all four questions was defined as having low sensitivity (471 people).

The explanatory variables are created by EPs, namely itemsets whose supports change considerably from one class to another, capturing discriminating features that sharply contrast instances between the classes. Specifically, when extracting an EP as an explanatory variable, the item used is the purchased item of clothing and its color pair. For example, if a customer purchased a burgundy T-shirt, the item is expressed as "T-shirts_burgundy." To enhance the reliability of the classification model, we exploit the hierarchical structure in personal colors. For instance, the color "Burgundy" is contained in the season category "winter," and these two categories are contained in the base category "Cool." Therefore, if a customer purchases a burgundy T-shirt, then "T-shirts_winter" and "T-shirts_cool" are also created as items in addition to "T-shirts_burgundy."

To find characteristic purchasing patterns for high and low fashion sensitivity, we use the support and growth rate to enumerate EPs whose values are greater than or equal to the minimum support and minimum growth rate, respectively, which are the thresholds for the user. With the positive transaction set expressed as D_p and the negative set as D_n, the support of the positive class regarding two items a and b is defined as follows:

$$support_{D_p}(a,b) = \frac{|Occ_p(a,b)|}{|D_p|},\tag{1}$$

In the above formula, $Occ_p(a,b)$ represents the transaction set D_p in which items a and b co-occur.

The growth rate of the positive class relative to the negative class for two items a and b is defined as follows:

$$GR_{D_n \to D_p}(a,b) = \frac{support_{D_p}(a,b)}{support_{D_n}(a,b)}.\tag{2}$$

This formula represents the ratio of the co-occurrence probability of the positive class to that of the negative class. If the ratio is greater than 1.0, we can say that items a and b have a co-occurrence pattern that is distinctive of the positive class (an EP). The same method is applied to the negative class, and EPs that are characteristic of the positive and negative classes are enumerated.

Explanatory variables other than the pattern for constructing the model utilize the fashion questionnaire and demographic attributes such as gender and age. Table 2 lists the explanatory variables used.

Table 2. List of explanatory variables

Explanatory variables	Description
Questionnaire (97 questions)	About participation in events, fashion issues, fashion view, fashion change, a feeling of happiness, life values, fashion consciousness, purchase timing, etc.
Percentage purchased in each season	Total purchase percentage for each customer will be 100%
Emerging patterns	Minimum growth rate = 1.1, minimum support = 0.1
Gender	Male, Female
Age	Under 20, early 20s, late 20s, early 30s, late 30s, 40 and over

The classification model is a logistic regression model. With the dependent variable in the classification model as $y \in \{0, 1\}$ (0: negative, 1: positive) and with the p explanatory variable vectors as $\mathbf{x} = (x_1, x_2, \cdots, x_p)$, the logistic regression model is expressed as Eq. (3):

$$\Pr(y = 1|\mathbf{x}) = f\left(\boldsymbol{\beta}^\top \mathbf{x} + \beta_0\right), \tag{3}$$

where $f(\cdot)$ is a logistic function defined as $f(a) = 1/(1 + \exp(-a))$; $\boldsymbol{\beta} \in \mathbb{R}^p$ is a regression coefficient vector and $\beta_0 \in \mathbb{R}$ is a constant term, both of which are estimated from the training samples. Moreover, we use variable selection that involves selecting a set of relevant explanatory variables from many candidates and using them to construct a statistical model. This procedure facilitates interpretation of the subsequent analysis of the statistical model and enhances the model's predictive performance by preventing overfitting [7]. We apply to a stepwise method that begins with no explanatory variables, whereupon the variable that leads to the largest decrease in Akaike's information criterion is added or eliminated iteratively.

3   Calculation Experiment

A model is constructed by the method described in the previous section for 678 customers with high fashion sensitivity and 471 with low fashion sensitivity. In the model construction, the training and test data were randomly sampled from the overall data set at a ratio of 9:1. The prediction accuracy is 78.1%[2], indicating a relatively high correct rate.

Table 3 gives the results for the explanatory variables selected by the stepwise method. Positive coefficients are the choice factors of high-sensitivity customers,

---

[2] The correct-answer rate when randomly selecting two customer sets was 59%.

and those with negative coefficients are the choice factors of low-sensitivity customers. As a remarkable explanatory variable for customers with high sensitivity, fashion perspective is an important factor. For example, high awareness of fashion appeared from statements such as "fashion expresses my identity," "I care how my fashion is viewed by other people," "I try new fashion as much as possible," and "Buying clothes makes me feel refreshed and relieves my stress." These customers incorporate information about new fashion, they enjoy fashion, and they find that it relieves stress. In short, fashion is part of their lives. Meanwhile, customers with low sensitivity are not conscious about fashion because they cannot judge what to wear by themselves and instead choose clothes by listening to the opinions of friends and shop assistants. Also, the life values of those customers are different. Customers with high sensitivity value wealth (Q3_17), evaluation (Q3_16), and personality (Q3_6), whereas customers with low sensitivity value history (Q3_14), latest technology (Q3_7), and stability (Q3_1). These factors are important features representing the differences between the two customer groups, but it is difficult to apply them directly to sale promotion because it is difficult to change personality and fashion consciousness. Therefore, by interpreting the EP that expresses characteristic purchasing behavior, we obtain a hypothesis that drives customers with low sensitivity to purchase.

Figure 3 shows the EPs resulting from the selected explanatory variables of Table 3. Each row represents a pattern and each column represents an item that could be included in the pattern, such as cloth and its color. Patterns 1–4 are characteristics of customers with high sensitivity and patterns 5 and 6 are characteristics of customers with low sensitivity.

	T-shirt/Cut and sewn		Knit/Sweater	Shirt/Blouse	Pants			Sneaker	Parker
	Winter	White	Gray	Summer	Gray	Black	Beige	Warm	Cool
Pattern 4				●		●			
Pattern 2								●	●
Pattern 3	●			●					
Pattern 1							●		
Pattern 5		●							
Pattern 6			●						

Fig. 3. Emerging patterns (EPs) in the model. Each row is a pattern and each column is an item that could be included in the pattern, such as cloth and its color.

"Pattern 5" involves a cut-and-sewn white T-shirt, which is an EP for low-sensitivity customers. "Pattern 6" involves a gray knitted sweater as an EP for low-sensitivity customers. Both involve the purchase of only a single item of clothing, and its chosen color is white or gray, which are common colors. It is understood that customers with low fashion sensitivity select safe colors that are easy to match with any clothes.

Table 3. Explanatory variables used in the model

	Explanatory variable	Coefficient	
Q8_8	**Fashion view: Fashion is an expression of my identity.**	2.6851	*
Q8_14	**Fashion view: I care how my fashion is viewed by other people.**	2.1953	*
Q8_5	Fashion view: I am sensitive to fashion trends	1.7848	*
Q8_11	**Fashion view: I try new fashion as much as possible.**	1.5204	*
Q8_16	**Fashion view: Buying clothes makes me feel refreshed and relieves my stress.**	1.3391	*
Q8_6	Fashion view: I have a different style from the people around me	1.2825	*
Q8_12	Fashion view: I am careful to choose clothes that are suitable for the time, place, and occasion	1.1888	*
Pattern 1	High sensitivity: Pants_beige	1.0774	*
Q8_20	Fashion view: If I like the quality and function of the clothes, their brands do not matter	0.9669	*
Q8_29	Fashion view: I have specific favorite brands	0.8722	*
Pattern 2	High sensitivity: Sneakers_Warm & Parka_Cool	0.7384	
Pattern 3	High sensitivity: T-shirt/Cut-and-sewn_Winter & Pants_Gray	0.7084	
Q8_19	Fashion view: I use clothes carefully even when it becomes old	0.7074	*
Q8_3	Fashion view: When deciding what to wear, I am conscious of the opposite sex	0.6686	
Q8_27	Fashion view: For some items I like, I have the same item in different colors	0.6594	*
Pattern 4	High sensitivity: Shirt/blouse_Summer & Pants_Summer & Pants_Blue	0.6342	
Q2_11	Conscious: Conscious of brands when shopping	0.6038	*
Q3_6	Life values: Personality (uniqueness, heterogeneity, special, etc.)	0.591	*
Q8_1	Fashion view: I have imitated celebrity fashion	0.59	*
Q8_17	Fashion view: I do not care about dress casually at home	0.5352	*
Q1_3	Event: I have participated in barbecues and outdoor events	0.5318	*
Q3_16	Life values: Emphasize evaluation from people close to me	0.4553	*
Q3_17	Life values: Wealth (money, assets, savings, money making, etc.)	0.3132	
Q6_7	Purchase time of clothes: before travel/event	0.2607	
Q9_1	Fashion change: Interest in fashion	0.2126	
Q3_14	Life values: History (old, established, traditional, ruins, culture, etc.)	−0.218	
Q3_7	Life values: State of the art (science, technology, IT, pioneers, etc.)	−0.2982	
Q2_12	Conscious: I try to find cheap things even if it takes time	−0.3433	
Pattern 5	Low sensitivity: T-shirt/Cut-and-sewn_White	−0.4705	
Gender	Male	−0.5912	
Q3_1	Life values: Stable (secure, conservative, etc.)	−0.6723	*
Q1_4	Event: I have participated in a music festival	−0.6783	*
Q8_2	**Fashion view: When choosing clothes, I decide after listening to the opinions of friends and shop assistants.**	−0.708	*
Q7_5	**Fashion problem: I do not know what to wear depending on the situation.**	−0.7081	*
Q1_8	Event: I have participated in a Halloween party	−0.7413	
Pattern 6	Low sensitivity: Knit/Sweater_Gray	−0.7883	*
Q7_2	Fashion problem: I do not know the differences among brands	−1.1195	
Ratio	Purchase ratio in Summer	−6.4492	*

*: p-value < 0.05

By contrast, the common feature of the outstanding patterns for fashion-sensitive customers is selecting multiple clothing categories and various colors, and different levels of item appear in these patterns. This is due to the effect of hierarchical classification. For example, "Pattern 4" shows that high-sensitivity customers purchases shirts/blouses of summer colors and blue pants. Also, "Pattern 2" shows that high-sensitivity customers purchase sneakers of the warm category and parkers of the cool category. In this way, customers with high fashion sensitivity purchase clothes that show consciousness of total coordination considering personal color and do not purchase white or gray clothing alone.

It is important to change the consciousness of customers with low fashion sensitivity by referring to the purchasing behavior of customers with high fashion sensitivity and repeatedly presenting coordination to which customers with low fashion sensitivity can refer.

4   Conclusion

In this research, we have used purchasing-history data from e-commerce apparel sites and proposed a method of using EPs by personal color as a method of handling color information. Handling of colors is a problem with apparel-based data, and it is difficult to obtain appropriate results from analyzing monochrome text information as-is. We showed that we can construct a highly accurate classification model that can interpret meaning by using hierarchical classification using personal color as proposed in this research. Customers with high fashion sensitivity are highly conscious about trends and coordination, and it became clear that they are selecting colors while being conscious of multiple fashion items and personal colors. Meanwhile, customers with low fashion sensitivity find that they cannot choose clothes according to the situation and instead buy clothes without referring to the opinions of friends or shop assistants. In addition, the item to be purchased belongs to a single category and has common colors such as white and gray are selected. To improve sales, it is necessary to provide fashion information considering total coordination and personal color and to develop sensibility about fashion by making contact repeatedly.

References

1. Hultén, B.: Sensory Marketing: Theoretical and Empirical Grounds (Routledge Interpretive Marketing Research), p. 175. Taylor and Francis (2015)
2. Zajonc, R.B.: Attitudinal effects of mere exposure. J. Pers. Soc. Psychol. **9**(2, Pt. 2), 1–27 (1968). https://doi.org/10.1037/h0025848
3. Dong, G., Li, J.: Efficient mining of emerging patterns: discovering trends and differences. In: Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 43–52 (1999)

4. Kamiyama, Y., Nagasaka, N.: The Textbook of Personal Color. Shinkigensha (2004). (in Japanese)
5. https://en.wikipedia.org/wiki/Practical_Color_Coordinate_System
6. http://art-design-glossary.musabi.ac.jp/tone/
7. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**, 1157–1182 (2003)