Optimizing Product Placement for Virtual Stores

Wei Liang[1,4] * Luhui Wang[1] Xinzhe Yu[1] Changyang Li[2] Rawan Alghofaili[2] Yining Lang[3] Lap-Fai Yu[2]

[1]Beijing Institute of Technology [2]George Mason University [3]Alibaba Group
[4]Yangtze Delta Region Academy of Beijing Institute of Technology, Jiaxing, China

Figure 1: Taking as input a 3D virtual store scene and some virtual products as shown on the left, our approach optimizes the product placement considering both product exposure and spatial constraints. A user navigates the virtual store with the optimized product placement in virtual reality as shown on the right.

Abstract

The recent popularity of consumer-grade virtual reality devices has enabled users to experience immersive shopping in virtual environments. As in a real-world store, the placement of products in a virtual store should appeal to shoppers, which could be time-consuming, tedious, and non-trivial to create manually. Thus, this work introduces a novel approach for automatically optimizing product placement in virtual stores. Our approach considers product exposure and spatial constraints, applying an optimizer to search for optimal product placement solutions. We conducted qualitative scene rationality and quantitative product exposure experiments to validate our approach with users. The results show that the proposed approach can synthesize reasonable product placements and increase product exposures for different virtual stores.

Index Terms: Human-centered computing—Human computer interaction (HCI)—;——Computing methodologies—Virtual reality

1 Introduction

Virtual reality (VR) retail is an emerging shopping experience, offering many advantages over conventional shopping. It can be regarded as an extension of online shopping, offering benefits such as reducing the overall operating costs of retailers and enabling customers to shop without having to leave their comfortable couches at home. The COVID-19 pandemic may accelerate the shift towards a more digital shopping world [17]. The concept of Metaverse may also spur the progress of VR retail [1], encouraging the community to explore how emerging VR technologies may shape future shopping experiences.

In contrast to traditional online shopping, VR retail enables customers to shop in innovative ways [42]. VR retail offers immersive shopping experiences and enables customers to virtually try on items akin to shopping in real stores. The physical analogies help users act out with concepts already understood in real stores, reducing the cost and time of learning something new and unfamiliar. In addition, a virtual store could be tailored through computational design approaches to specific user's preferences and interests, potentially enhancing shoppers' satisfaction, improving shopper's engagement with brands, and bringing more revenues to retailers.

Designing the product placement is vital for a VR store creation task. Like in a traditional brick-and-mortar store, a well-designed placement contributes a lot to a positive shopping atmosphere, leading to higher consumer satisfaction, better consumer relationship, and more importantly, boosting product sales [20]. The task is usually accomplished manually by designers according to their professional knowledge of the store's targets, considering the spatial constraints and product exposure in the 3D virtual scene. Retailers need to change their product placement to display new products regularly and keep the stores fresh. However, designing and maintaining a VR store incurs recurrent efforts and costs.

To this end, we propose to automate the product placement for VR stores, facilitating the intuitive and quick adjustment of the product placement for retailers. Although the placement design considerations may vary from designer to designer, there are still some common and key rules to follow in computational design. One key factor to consider is enhancing the exposure time of products to gain more customers' attention. Some market studies have shown that the attention of customers in a store is highly relevant to their purchase behaviors [23,24]. Because of the profit-oriented feature of stores, we consider generating the product placement with respect to the overall product exposure. Another consideration is introducing spatial constraints to improve the appeal of the product placement. For example, products in a virtual store are placed at spacious locations to motivate shoppers' buying behaviors [31]. Thus we design a computational framework that encodes the above two considerations to optimize a product placement automatically.

More concretely, given a virtual store scene and products to be displayed, we formulate the product placement synthesis as an opti-

*e-mail: liangwei@bit.edu.cn(Corresponding author)

mization problem. Against a total cost function that encodes spatial constraints and product exposure constraints, the optimizer searches for a placement iteratively through the Markov chain Monte Carlo algorithm. To experiment with our approach, we conducted a qualitative scene rationality evaluation experiment and a quantitative exposure evaluation experiment to validate the virtual store designs synthesized by our approach. Moreover, we show our approach's capability for generating product placements for different types of virtual stores. The results show that our approach enhances product exposure and optimizes product placement more effectively than alternative approaches. The main contributions of this paper are summarized as follows:

- We propose a new problem statement for optimizing product placements in virtual stores.
- We devise a computational design approach for synthesizing product placements in VR stores driven by product exposure prediction and spatial constraints.
- We validate the effectiveness of the proposed approach by conducting user studies on the synthesized stores to investigate spatial rationality and product exposure.

2 Related Work

We discuss the background and prior VR shop applications. Additionally, we review the previous works on visual attention and scene synthesis, which inspire the constraints employed in our framework to guide product placement optimization, product exposure, and spatial constraints.

2.1 Virtual Reality Shopping

Virtual Reality shopping applications allow customers to shop in a VR environment, replicating in-store experiences. Prior researches have explored and facilitated virtual store design in different ways, for example, through design strategy [26], interaction mode [36], and application [40]. Xi and Hamari [42] reviewed the latest research and development in VR shopping; the potential VR shopping themes; and the research gaps that need to be filled to realize VR shopping.

VR systems have been applied for marketing by companies such as eBay [11]. There are also research efforts in this direction. Several works explored certain factors for predicting shopping behaviors based on 3D simulation technology. Using 3D simulation technology capable of offering a realistic virtual experience, they tested their hypotheses. Jiang et al. [22] used 3D computer graphics as experimental stimuli to investigate the effect of visual complexity in a fashion store on affective and behavioral responses. Alawadhi et al. [2] ran a virtual experiment for customers to investigate the essential effects of the allocation of products on perceived crowding, which indirectly but significantly affects the shopping behaviors of customers. Creating virtual stores manually is costly. Hence the design process could benefit from automation. Compared with previous works, we take account of both the product exposure and the spatial constraints to optimize product placement automatically.

2.2 Visual Attention

Recent works inspired our exposure formulation in visual attention studies. Visual attention refers to the ability of the human visual system to rapidly select the most relevant information in the visual field. One of the most popular visual saliency models was proposed by Itti et al. [21], namely, the data-driven attention model. It computes the multi-scale feature contrasts of input images by using a difference of Gaussian and linearly combining the feature conspicuous maps to produce a master saliency map. After nearly 20 years of development, researchers have proposed many computational visual attention models [41], including Bayesian surprise models, task-driven models, etc.

Eye-tracking technology with visual attention models finds applications in areas such as skill training, pictorial database query, and advertising design. Lang et al. [27] used eye-tracking headset to improve driving habits. Li et al. [28] collected visual attention to evaluate the earthquake safety drill effect. Tabbaa et al. [38] captured eye-tracking behavior in 360-VEs and presented a novel affective dataset. Meyer et al. [33] combined head- and eye movement features to recognize activities. Chang et al. [7] detected visual attention to record visual content of personal interest. An obvious application of eye-tracking and visual attention models is to enhance advertising design, e.g., predicting users' attention when browsing complex web pages [10]. Alghofaili et al. [3] utilized visual attention models to find the best panel placement by maximizing the likelihood that the panels would be seen in a 3D room. Their focus is on placing 2D elements in virtual environments such as museums. In our approach, we devise an optimization-based model for product placement to satisfy some design constraints specific to placing products (e.g., the visual balance). Furthermore, we added a perceptual study as a means of validating our approach.

We applied visual attention models to the scene synthesis problem, optimizing 3D visual elements in a 3D virtual environment. We obtained the product exposure data of 3D virtual products through experiments via a FOVE eye-tracking virtual reality headset. Subsequently, the data were utilized for training a regressor for predicting product exposure.

2.3 3D Scene Synthesis

Automatic 3D scene synthesis is an active research topic. Researchers have developed approaches for synthesizing 3D scenes, layouts, and virtual environments, such as workspace [29], furniture layouts [45], and partial scenes [14]. Zhang et al. [46] provides a recent review on automatic 3D indoor scene synthesis.

Distinct from coarser-level layout synthesis, we focus on optimizing fine-level item placement. Majerowicz et al. [30] presented a data-driven method specifically designed for artifact arrangement, which automatically populated empty surfaces with different believable arrangements of artifacts in a given style. Fender et al. [13] presented a system that empirically analyzed user behaviors in a space and automatically suggested positions and sizes.

Product placement is one application of scene synthesis. Compared to typical indoor synthesis problems, store synthesis should consider not only the basic scene layout but also how to place each product reasonably and attract attention. To overcome the challenge of designing a virtual store, we encode the rational consideration of a general scene, human knowledge about product arrangement, and the estimation of product exposure into our optimization-based approach.

3 Product Placement Optimizer

Given a virtual store scene and some virtual products as input, our approach employs an optimization algorithm to find a desired product placement solution for the virtual store. The framework of our approach is illustrated in Fig. 2. We use a clothing store as an illustrative example to facilitate our technical discussion. Note that our framework can be generalized to synthesize other types of stores, as demonstrated in the experiments.

This section discusses the optimization process to generate a product placement with high product exposure and a rational product layout in a virtual store.

The optimizer consists of a total cost function and an optimizing algorithm. The total cost function is devised to evaluate each generated product placement, which is informed by some store design articles and an interview with experts [12, 19, 39]. It consists of an exposure cost and a spatial cost. The product exposure term, which is based on a Random Forest regressor trained to predict shoppers' gaze duration on products, guides the optimizer to improve product exposure in the virtual store. The spatial term encodes the design

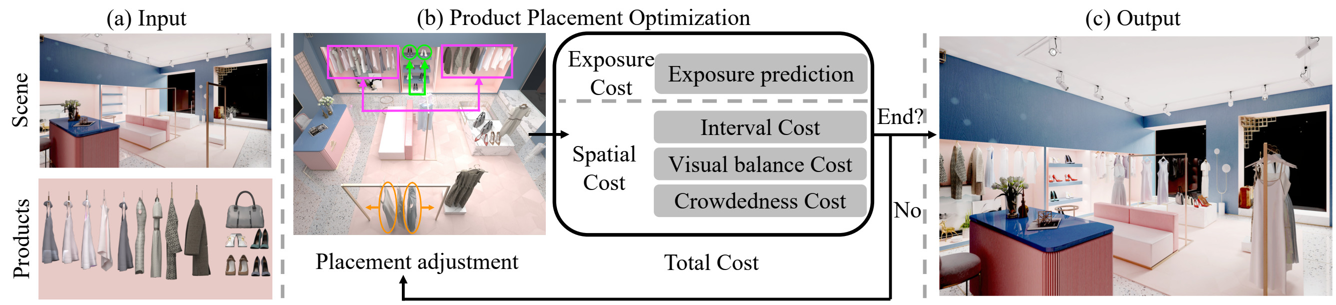


Figure 2: The overview of our approach. (a) The input is an empty virtual store scene and 3D virtual products. (b) Our approach optimizes the product placement against exposure and spatial cost terms iteratively. The exposure term is defined on the basis of exposure prediction. The spatial cost encodes three constraints, i.e. interval, visual balance, and crowdedness. (c) The output is an optimized product placement.

priors for arranging products reasonably to improve shopping experiences by avoiding crowdedness, ensuring even distribution and keeping visual balance.

Formally, a product placement is defined as $P = \{p_i = (l_i, \theta_i), i \in \{1, 2, \cdots, N\}\}$, where N is the number of products. For the i-th product, $l_i = (x_i, y_i, z_i)$ is the coordinates of its center, representing the product's location; θ_i is the angle of the front surface relative to the nearest wall plane, representing the product's orientation. The front surface means the surface the retailer wants to show to customers. The wall plane is defined as the back of the product showing area vertically relative to the access area.

The total cost is written as:

$$C_{\text{Total}}(P) = \omega_{\text{Exposure}} C_{\text{Exposure}}(P) + \omega_{\text{Spatial}} C_{\text{Spatial}}(P), \quad (1)$$

where C_{Exposure} and C_{Spatial} are exposure cost and spatial cost, respectively; ω_{Exposure} and ω_{Spatial} are their weights and are set as 0.5 empirically, considering the importance of the two costs is equivalent. The details of the cost terms are in the following.

Against the cost function, the optimizer searches for an optimal product placement iteratively through an MCMC (Markov Chain Monte Carlo) algorithm [4].

3.1 Exposure Cost

Considering that products should be placed to attract more exposure from the customers [19], we design the exposure cost term to evaluate the received exposure under the current placement. We train a regressor to estimate the product exposure of a virtual product placed at a certain location. On the basis of the regressor, we define the exposure cost over all products to guide the optimizer to select a solution with a higher predicted overall exposure:

$$C_{\text{Exposure}}(P) = 1 - \frac{1}{\sum_{i=1}^{N} w_i} \sum_{i=1}^{N} w_i e_i, \quad (2)$$

where e_i is the predicted product exposure of product i at its current location, in proportion to the product exposure duration. $w_i \in [0, 1]$ is an exposure weight for the i-th product and is set as $\frac{1}{N}$ by default. This weight is designed to enable adjustment, i.e. increasing the exposure consideration for a specific product. A lower cost value of $C_{\text{Exposure}}(P)$ means a higher total product exposure duration.

Product Exposure Regressor. To predict the product exposure for each virtual product, we apply a data-driven approach, Random Forest regressor [37] to learn the relationship between product exposure and a feature vector encoding product attributes and its location attributes. The training process and prediction process are implemented by the scikit-learn library. Fig. 3 visualizes predicted product exposures in a store.

Features. Per a consultation interview with store staff and interior designers, as well as product placement references [34], we extract four relatively general features to train the regressor. The features cover the products' intrinsic attributes and relative position in a scene, which can also be extracted from unseen products easily.

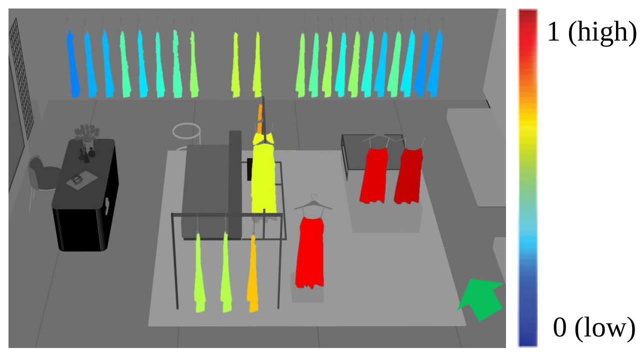


Figure 3: Visualization of predicted product exposures for the items. A redder color refers to a higher exposure value.

(1) *Size*. Previous works suggested that the size of a product affects the user's attention during the shopping process [47]. In order to extract the size, we first calculate a tight bounding box with respect to the product's dimensions. Then, we take the length, width, and height of the bounding box as the size feature.

(2) *Texture.* Inspired by previous findings that the product's color affects customers' purchase decisions [18], we use texture intensity as a texture feature. It is defined as how much light a product reflects. The higher the value is, "the brighter" the product appears to be. We define r, g, and b as the average red, green, and blue values of all the pixels in the product texture image. The intensity is obtained as $I = (r + g + b)/3$ akin to [16]. These give us four feature values to measure the product intensity.

(3) *Height from the Floor.* Previous works found that the height of a product from the floor could affect the attention it receives [6]. This feature could reflect how well the product can be viewed in the scene. It is defined as the relative height of a product from the floor.

(4) *Distance from Landmarks.* Literature [5] suggested that brands in the horizontal center receive more visual attention. Hence we use the shortest distance between the product and any entrance, as well as the distance between the product and the center of the store, as the distance from landmarks feature. Note, the designer could also include distances from other places of interest in addition to the entrances and the store center.

Training Data. We obtained training data through a data collection process. The placement was generated randomly at the beginning. Participants wore a FOVE eye-tracking VR headset for navigation. A capsule collider with a virtual camera on the top is used as the participant's avatar. The participant moved in the virtual scene by a gamepad to control the moving direction of the capsule collider. While the view of the observation, i.e., the virtual camera's orienta-

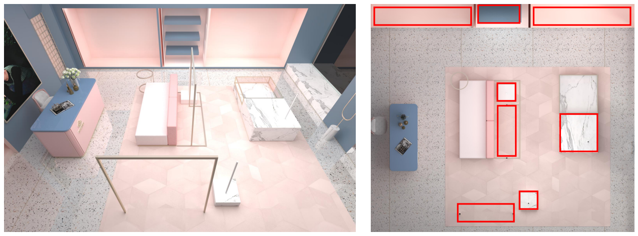


Figure 4: Left: An input 3D scene. Right: The placement regions. Each red rectangle refers to a distinct placement region.

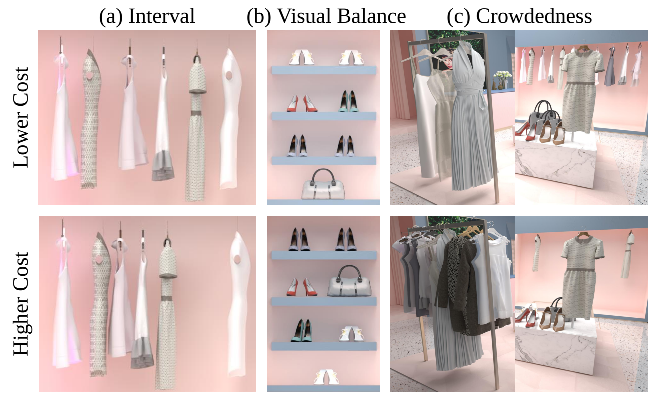


Figure 5: Examples placements with a different cost value of (a) interval cost, (b) visual balance cost, and (c) crowdedness cost.

tion, is synchronized with the direction of the headset. Participants always start walking from a predefined entrance position. We requested the participants to walk and look around in the store as if they were shopping in an actual store. Then the participant's movement and gaze data during the navigation were recorded. FOVE provides a visual frame rate of 70 fps with less than one-degree tracking error. We calibrated the headset before each test.

23 participants were recruited to gather the product exposure data, most of whom were students and college staff aged 19 to 30. Participants reported normal or corrected-to-normal vision. We created three virtual scenes, namely, clothing Store, grocery Store, and toy Store, referring to real-world shop layouts. Each participant tests three types of stores in a random order. Participants were given 5 minutes in each store to view products, counting as one session. After one session, we summarized a virtual product's viewing periods as its product exposure.

In the training process, each data sample input refers to a feature vector of a product with normalized exposure time. To alleviate the bias of individual observations and habits, we use normalized exposure time rather than absolute exposure time as input.

The original exposure time is normalized to $[0,1]$ based on the maximum and minimum exposure time. We use the exposure time of the product with the largest amount of gaze duration received from the participants in every single data collection session as the maximum exposure time, and use 0 as the minimum exposure time for normalization. For example, for a participant navigating a virtual store in a session, if the exposure time of a T-shirt was 10.15 seconds, and the exposure time of the product (e.g., a dress) that received the maximum exposure was 20.30 seconds, then the normalized exposure data of the T-shirt is 0.5.

3.2 Spatial Cost

We include a spatial cost, which evaluates how reasonable the products are placed, to encourage the optimizer to favor solutions with a rational product layout. The spatial cost incorporates the following three spatial constraints: (1) placing products with equally-spaced intervals on shelves; (2) keeping visual balance by placing large products at lower locations and small products at higher locations; and (3) avoiding too many products on one shelf. These constraints were devised by consulting 5 experts, including store staff and interior designers, and with reference to store design books [12, 39].

To encode these considerations, we define the spatial cost as:

$$C_{\text{Spatial}}(P) = \lambda_{\text{i}} C_{\text{i}}(P) + \lambda_{\text{v}} C_{\text{v}}(P) + \lambda_{\text{c}} C_{\text{c}}(P), \tag{3}$$

where C_{i}, C_{v}, and C_{c} are the interval cost, visual balance cost; and crowdedness cost, respectively; λ_{i}, λ_{v}, and λ_{c} are their corresponding weights set as 0.3, 0.3, and 0.4, by default. We slightly favor the crowdedness cost because it affects the number of goods within one areas, and then further affects the users' overall experiences.

Interval. A layout with an even product placement helps relieve shoppers' negative feelings due to clutter [39]. Thus, we design an interval cost to encourage the optimizer to choose a solution with even intervals among products. The store scene is divided into regions $R = \{r_m\}$. Each region r_m refers to a surface capable of supporting products as shown by the red rectangles in Fig. 4. For example, products can be placed on a table or hung on the shelves. This step can be done manually by the user or processed by some furniture affordance prediction algorithms [15].

The interval cost is defined as:

$$C_{\text{i}}(P) = \frac{1}{|R|} \sum_{m=1}^{|R|} \frac{1}{N_{\text{m}}} \sigma_m^2, \tag{4}$$

where $|R|$ is the number of regions. Assume $D(l_i, l_{i+1})$ is the space intervals of adjacent products calculated by the Euclidean distance between them. σ_m^2 is the variance of intervals $D(l_i, l_{i+1})$ in region r_m. N_{m} is a normalization constant and is set as the square of the longest diagonal of the r_m region.

Visual Balance. A common assumption in placing products is to keep visual balance, i.e. placing larger products at lower places. Considering larger products carry stronger visual weights [32], we use the front face area of a product's bounding box to define the visual balance cost as:

$$C_{\text{v}}(P) = \frac{1}{N} \sum_{i} \frac{1}{N_{\text{v}}} a_i h_i, \tag{5}$$

where h_i is the product's height (from the product's bottom to the ground); a_i is the area of the front face of the bounding box of product i; N_{v} is a normalization constant defined as the product of the tallest shelf's height and the largest front face's area.

Since $a_i h_i$ increases when large products are placed high, placing large products at lower shelves implies a lower cost. Fig. 5 (b) shows an example placement with a smaller visual balance cost (top) and a placement with a larger visual balance cost (bottom).

Crowdedness. Previous research found that products placed at spacious locations motivate shoppers' buying behaviors [31]. This implies that all regions should have similar product placement crowdedness so that the products would not pile up in one region, which may affect the shopping experience.

We devise the crowdedness cost to model the difference in occupancy rates across all regions. A solution with less difference will be preferred. The cost is defined as:

$$C_{\text{c}}(P) = \frac{1}{N_{\text{c}}} \sum_{m=1}^{|R|-1} \sum_{n=m+1}^{|R|} \left| \frac{\sum_i s_i}{S_n} - \frac{\sum_j s_j}{S_m} \right|, \tag{6}$$

where $|R|$ is the number of regions. $N_{\text{c}} = \frac{|R|(|R|-1)}{2}$ is the normalization parameter. S_n and S_m are the areas of the n-th region and m-th region, respectively. s_i and s_j are the areas of the i-th and j-th objects in the n-th region and m-th region.

Fig. 5 (c) shows an example. Compared to the bottom placement, the top placement puts more products in a larger shelf.

3.3 Optimization

Based on the defined total cost, we use a MCMC optimization framework to search for reasonable product placement. To speed up the optimization, we assume that one type of products is placed in one region. So we group products according to their categories at the beginning. Then, we initialize the virtual store scene randomly, that is, the 3D products are randomly placed in the placement regions. Each placement region is only filled with products of the same type.

At the beginning of the optimization, each virtual product is initialized with a random position and orientation. At each iteration of the optimization, a new product placement P' is proposed. To explore the solution space effectively, we define three types of moves to propose the new placement:

Product Translation. Select and move a product slightly based on its current location within its current placement region. This move adjusts the configuration of the placement within a local region.

Product Swapping. Swap the location and orientation of a randomly-selected product pair. This move enables more rapid exploration of the solution space and prevents the optimizer from being stuck at local minima.

Group Swapping. Swap two randomly selected groups. Two groups of products are selected at random and their placement regions are interchanged. This move induces a rearrangement of the placement.

Each proposed move is accepted with the following probability based on the Metropolis criterion [8]:

$$\alpha\left(P'|P\right) = \min\left(1, \frac{f\left(P'\right)}{f\left(P\right)}\right), \tag{7}$$

where $f\left(P\right) = e^{-\frac{1}{t}C_{\text{Total}}(P)}$ and t is the temperature parameter for simulated annealing [25]. The optimization proceeds iteratively. We set $t = 300$ empirically at the beginning of the optimization, allowing the optimizer to explore the solution space more aggressively with a high temperature. The value of t drops by 0.5 every iteration of the optimization until it reaches 1, allowing the optimizer to refine the solution near the end of the optimization. We terminate the optimization if the absolute change in the total cost value is less than 5% over the past 50 iterations. Obtaining a solution in our experiments required up to 500 iterations.

4 Perceptual Study

We conducted perceptual studies to evaluate the effectiveness of our approach and investigated whether our synthesized placement provides immersive shopping experiences and induces product exposure. We conducted three user studies: (1) a general evaluation for validating scene rationality; (2) an exposure evaluation for validating the exposure enhancement effectiveness; (3) a spatial constraints evaluation for validating the effectiveness of spatial constraints.

Participants. We recruited 15 participants, including 8 males and 7 females aged from 18 to 50. All the subjects reported normal or corrected-to-normal vision with no color blindness. Note that none of the participants took part in the previous training data collection process.

Procedure. Each participant was asked to navigate different scenes and observe the products for a given period of time via a FOVE VR headset with a gamepad, akin to the training data collection process. Before the evaluation, each participant went through a 5-minute warm-up session to get familiar with the navigation. A VR warm-up application allowed participants to learn the basics of operating a FOVE VR headset.

Participants experienced three types of scenes one by one, i.e. clothing store, toy store, and grocery store. For each type, the results of all four compared approaches were given to the participants randomly to avoid any carryover effects. In each experiment, participants spent max 5 minutes navigating each scene. Instructions for each experiment were provided via a window in the VR environment. The participants read the instructions. When they were ready to proceed to the next window, they could use the controller to enter the experiment scenes. During the experiment, the participants were not explicitly told by which approach the current store was synthesized. Participants were told to report any sickness or discomfort with the apparatus at any point during the experiment and that they could terminate their session at any time. When the participants completed the task in one scene, they were asked to rate the results by answering some questions.

Table 1: Statistics of different scenes.

Scenes	No. of products	No. of product types	No. of shelves
Clothing Store	38	4	6
Toy Store	102	10	8
Grocery Store	290	40	13

4.1 Synthesized Product Placement Results

We tested our approach in three different virtual store scenes as shown in Fig. 6, consisting of a Clothing Store, a Grocery Store, and a Toy Store. The scenes are with different numbers of shelves and products. Table 1 shows some statistics of the three scenes.

For the Clothing Store, the products include clothes, shoes, and bags. The placement generated by our approach is shown in Fig. 6(a). The products are moderately spaced. We see that our optimizer placed the more appealing products, such as the evening dress with a bright color, near the center of the room.

For the Toy Store, the products include dolls, toy cars, and robots. The placement generated by our approach is shown in Fig. 6(b). There is no occlusion. The distance between the toys is moderate. Most of the large toys are placed below the small ones. Besides, the appealing products, such as the big bear toys, are put by the optimizer near the center of the shelf, to gain more exposure.

For the Grocery Store, the products include beverages, wine, candies, etc. The placement generated by our approach is shown in Fig. 6(c). Due to the large number of products used, the products are placed more densely yet they are spaced regularly from each other. Products with a big scale (e.g., the blue bags) are put at the bottom of the shelf. Besides, the products in bright colors (e.g., the bright red cans and yellow candies) are placed in the middle of the shelves. They are expected to receive more attention.

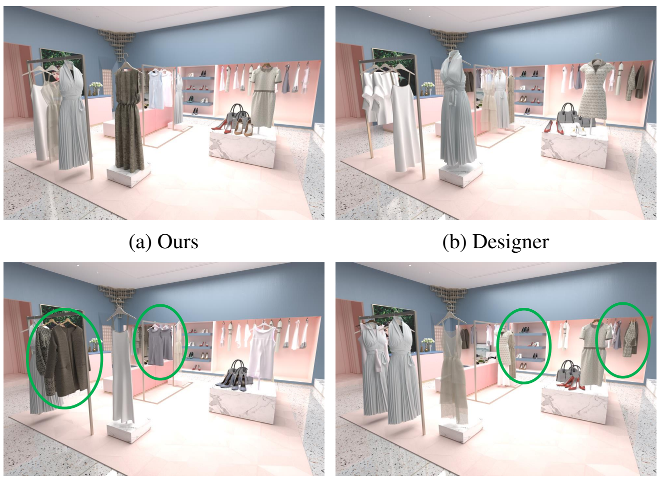
4.2 General Scene Rationality Evaluation

In this study, we evaluated the efficacy of our approach and its overall performance. We compared our approach with three other baseline approaches to synthesize store product placement, using one same random scene as the initialization. The comparison approaches include: (a) Designer approach: placement created by 3 designers with 2 to 5 years of retail store design experience, each designer completing one scene's placement design; (b) W/O EC: placement optimized by our approach without considering the exposure term; (c) W/O SC: placement optimized by our approach without considering the spatial term. The placements generated from the compared four approaches are shown in Fig. 7.

Metrics. We design a rating questionnaire to investigate the participants' feedback, measuring three aspects of the experiences in the virtual scenes. The questionnaire includes three questions: (1) *"I found the product placement realistic"*, evaluating whether our approach arranged the store akin to the actual stores; (2) *"I found the product placement convenient"*, evaluating whether the placement meets the essential function of a store, i.e. providing convenient observation for users; (3) *"I found the overall in-store experience satisfying"*, evaluating the overall experiences. We try to examine the efficacy of the whole designed pipeline. Thus the questionnaire mainly focused on the users' overall experiences rather than one

(a) Clothing Store (b) Toy Store (c) Grocery Store

Figure 6: Product placement results. Selected views of the virtual stores with products placed by the optimizer.



(a) Ours (b) Designer (c) W/O EC (d) W/O SC

Figure 7: Product placements optimized by different approaches starting with a randomized placement. Our synthesized result (a) is similar to the result (b) created by the designer. Compared approaches synthesized unrealistic placements that are undesirable. For example, the result (c) synthesized without the exposure cost has the less eye-catching products, such as the dark clothing items, placed at the center of the room. For the W/O EC approach the spatial constraint cost will be the least. In result (d) synthesized without the spatial cost, the clothing items are colliding. For the W/O SC approach, the exposure cost will be the least. See supplementary material for more results.

specific aspect. We also designed the corresponding ablation study to investigate the efficiency of the designed cost terms further.

We opt to use a 5-point Likert scale, with 1 meaning "strongly disagree" and five meaning "strongly agree". Fig. 8 shows the box plots of the ratings for the stores created by the four approaches. We carried out Friedman tests and found that for all metrics the difference was statistically significant among four compared approaches over three scenes. Thus, we further conducted post-hoc tests, and the overall results are shown in Table 2. We will discuss the details in the subsequent sections. Please see supplementary for the detailed testing results on each scene.

Placement Realism. The Friedman tests results revealed a significant difference among the four compared approaches ($\chi^2 = 97.80, p < 0.05, df = 3$) at $\alpha = 0.05$ significance level. A post-hoc test using Wilcoxon Signed-Ranks Test with Bonferroni correction (at the correlated significance level of $\alpha = 0.0125$) revealed a significant difference that the mean rating of our approach (M = 4.31) was statistically higher than that of the W/O SC approach (M = 2.09) ($W = 0, p < 0.013, r = 0.62$). However, the post-hoc test did not find any significant difference between the mean rating of our approach (M = 4.31) and that of the Designer approach (M = 4.44) ($W = 114.00, p = 0.157, r = 0.64$).

Interestingly, we did not find a significant difference between the ratings of participants using our approach (M = 4.31) and the ratings of participants using W/O EC approach (M = 4.04) ($W = 37.50, p = 0.027, r = 0.23$). We believe that this is due to the efficacy of the spatial cost term. Dropping the exposure cost term did not result in an unrealistic synthesized placement, since the placement realism was mainly affected by the spatial cost term.

The results indicated that, considering placement realism, human-designed product placement is not a clear winner over our synthesized ones. The results of the statistical tests also supported that our approach synthesized more realistic results than the W/O SC approach. We validated the effectiveness of the spatial cost term in our optimization as the W/O SC approach got the lowest score significantly in this task. Furthermore, we investigate each spatial constraint in detail in Sec. 4.4.

Placement Convenience. The Friedman test results revealed a significant difference among the four approaches ($\chi^2 = 74.69, p < 0.05, df = 3$) at $\alpha = 0.05$ significance level. We conducted a post-hoc Wilcoxon rank-test with Boneferroni correction. The results showed that the placement of our approach (M = 4.02) was reported to be more convenient than that of the W/O SC approach (M = 2.20) ($W = 0, p < 0.013, r = 0.60$) and that of the W/O EC approach (M = 3.07) ($W = 74.50, p < 0.013, r = 0.45$). The post-hoc test showed no significant difference between the convenience score of the placement synthesized by our approach (M = 4.02) and that of the placement synthesized by the Designer approach (M = 4.16) ($W = 205.00, p = 0.423, r = 0.08$).

The results showed that the placement convenience effectiveness of our approach is similar to that of the Designer approach. They also indicated that the spatial cost term increased placement convenience over the compared scenes. Based on more realistic space allocation, the placement synthesized by our approach was more convenient for customers to get what they wanted compared to that synthesized by the W/O SC approach.

In-store Experience Satisfaction. The Friedman tests results revealed a significant difference among the four approaches ($\chi^2 = 80.42, p < 0.05, df = 3$) at the $\alpha = 0.05$ significance level. The results of post-hoc tests using Wilcoxon Signed-Ranks Test with Bonferroni correction (at the correlated significance level of $\alpha = 0.013$) did not find a significant difference in the ratings of participants using our approach (M = 4.27) and the Designer approach (M = 4.18) ($W = 91.50, p = 0.373, r = 0.09$). The results also indicated that the mean score rating of our approach (M = 4.27) was significantly higher than that of the W/O SC approach (M = 2.33) ($W = 7.00, p < 0.013, r = 0.58$) and that of the W/O EC approach (M = 3.33) ($W = 60.50, p < 0.013, r = 0.45$) in the overall scenes.

The results showed that the participants generally reported similar satisfactory ratings on in-store experience after navigating the store scene synthesized by our approach compared to that synthesized by the Designer approach. Through the statistical analysis, it also indicated that the participants generally had better experiences on the store scene synthesized by our approach, compared to that syn-

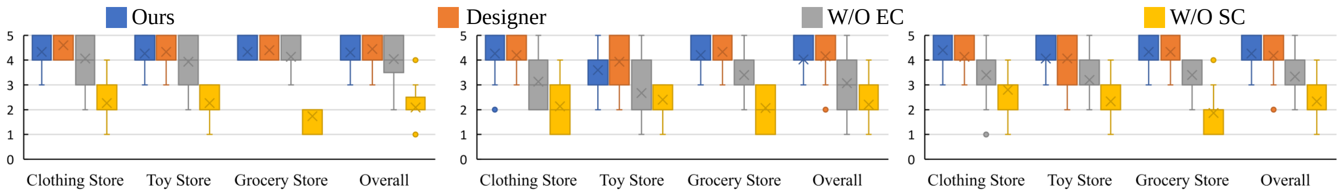


Figure 8: The participants' ratings on product placements synthesized by our approach, the design approach (Designer), our approach without the exposure cost term (W/O EC), and our approach without the spatial cost term (W/O SC). We opted to use a 5-point Likert scale, meaning that the higher the number is, the better the user's experiences are.

Table 2: The results of post-hoc tests in our scene rationality evaluation. We conducted a Wilcoxon signed-rank test with Bonferroni correction to compare the approaches. The p-values smaller than $\alpha = 0.0125(0.05/4)$, which reject the null hypothesis H0, are in bold.

Approach	Realism	Convenience	User Experience
Ours & Designer	$W = 114.0, p = 0.157, r = 0.15$	$W = 205.0, p = 0.423, r = 0.08$	$W = 91.5, p = 0.373, r = 0.09$
Ours & W/O EC	$W = 37.5, p = 0.027, r = 0.23$	$\boldsymbol{W = 74.5, p < 0.001, r = 0.45}$	$\boldsymbol{W = 60.5, p < 0.001, r = 0.45}$
Ours & W/O SC	$\boldsymbol{W = 0, p < 0.001, r = 0.62}$	$\boldsymbol{W = 0, p < 0.001, r = 0.60}$	$\boldsymbol{W = 7.0, p < 0.001, r = 0.58}$

thesized by the W/O SC approach and the W/O EC approach. Please refer to the supplementary material for more details.

4.3 Exposure Evaluation

In this study, we evaluated the exposure improvement effectiveness of our approach by investigating total product exposure time under different conditions. We took the virtual Clothing Store as an example. We collected the participants' eye-tracking data and calculated the amount of the total exposure time of all products.

There was no significant difference between our approach and the Designer approach in terms of exposure time in the placements. However, our approach synthesized placements that received significantly longer exposure time compared to the placements synthesized by the W/O EC approach and by the W/O SC approach.

We conducted Friedman tests on the total exposure time of all products. The results revealed a significant difference among the four approaches ($\chi^2 = 11.267, p < 0.05, df = 3$) at the $\alpha =$ 0.05 significance level. The amount of total product exposure (M = 117.54) using the Designer approach was not significantly different from the exposure time (M = 113.29) of our approach ($W = 63.00, p = 0.510, r = 0.12$). The Wilcoxon Signed-rank test with Bonferroni correction showed a significant difference ($W = 108.00, p < 0.013, r = 0.50$) in the exposure time of the product placement optimized by our approach (M = 113.29) and the exposure time of that of the W/O EC approach (M = 89.25). We also found a significant difference ($W = 119.00, p < 0.013, r = 0.61$) between the exposure time of the product placement optimized by our approach (M = 113.29) and the exposure time of product placement optimized by the W/O SC approach (M = 72.04).

The results showed that the total exposure of scenes synthesized by the W/O EC approach is significantly lower than that synthesized by our approach. As the placement of the W/O EC approach was optimized without exposure consideration, the attractive locations were likely occupied by less eye-catching products. Thus the exposure enhancement effectiveness of attractive places could be reduced. Conversely, our approach tended to fill attractive locations with eye-catching products.

Interestingly, we found that the total exposure time of the placement synthesized by the W/O SC approach is even lower than that of the W/O EC approach. Compared to the result of the W/O EC approach, it is clear that the result of the W/O SC approach is even less reasonable, and participants spent less time. In this regard, our spatial cost term imposes the effect of increasing exposure.

To further explore the effectiveness of the exposure cost, we compared our approach and the W/O EC approach by visualizing

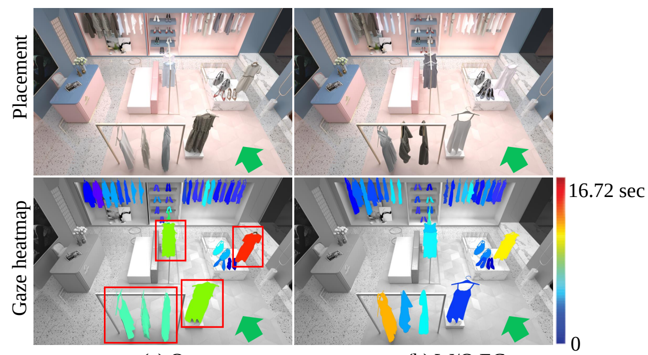


Figure 9: Gaze heatmaps on products placed by our approach (a) with and (b) without the exposure cost term. The top row shows the product placements synthesized by the respective approaches. The bottom row shows the gaze heatmaps of participants navigating in the virtual stores synthesized. Blue and red refer to low and high average exposure time captured from the participants. Our approach puts the more attractive products, such as the long dresses, at positions expected to receive more attention (shown in red boxes). Overall, the exposure cost leads to an increase in the product exposure particularly at some popular locations such as those near the entrance (depicted by the green arrow).

the amounts of average exposure time received from the participants in Fig. 9. Overall, the heatmap of our approach is redder than that of the W/O EC approach, suggesting our product placement received more exposure. We find that our approach generally puts more eye-catching products, such as the long dresses, at positions expected to receive more attention (e.g., positions shown in red boxes), which is conducive to improving the overall product exposure time.

While our approach increased product exposure significantly in the scene overall and at some popular positions, it did not reduce the product exposure at the other positions. We tested this through statistical analysis. To further compare the product exposure differences of different positions, akin to [35], we grouped the products by their placement region and compared the total exposure of each placement region under the two conditions respectively. The results showed that our approach increased the product exposure at some popular positions significantly, but it did not significantly cut down the product exposure at other positions.

Through the statistical analysis, we find that our synthesized placement does not show a significant difference in product exposure

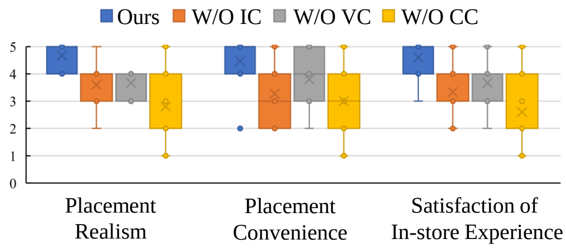


Figure 10: The participants' ratings on product placements synthesized by our approach, W/O IC, W/O VC, and W/O CC approach.

compared to the placement created by designers. The statistical analysis also indicates that our approach outperforms the other two approaches in product exposure.

4.4 Spatial Constraints Evaluation

Since the spatial cost is important for specifying the spatial constraints for product placements, we further evaluated how each spatial cost term affected the results using an ablation study. We take the virtual Clothing Store as an example to carry out this study with the same metrics in Sec. 4.2 and the results are shown in Fig. 10. The compared approaches are Ours, W/O IC (without Interval Constraints), W/O VC (without Visual Balance Constraints), and W/O CC (without Crowdedness Constraints). The participants' ratings are shown in Fig. 10. Regarding the three evaluations, the completed model of our approach achieved the highest score over all metrics.

The Friedman tests with significance level of $\alpha = 0.05$ were conducted to examine the differences among the four compared approaches. The results showed that the differences were significant among the four compared approaches over all evaluation metrics, i.e., placement realism ($\chi^2 = 17.863, p < 0.05, df = 3$), placement convenience ($\chi^2 = 11.238, p < 0.05, df = 3$), and in-store experience satisfaction ($\chi^2 = 19.173, p < 0.05, df = 3$).

We also used Wilcoxon Signed-Ranks tests to conduct the post-hoc tests with a significance level of $\alpha = 0.0125$. Please refer to the supplementary material for more details. Overall, the results suggested that our approach with all three cost terms achieved better performance ratings than the other three compared approaches with a significant difference. That is, the lack of any spatial constraints may cause significant performance degradation over all of the metrics.

Among those three spatial considerations, omitting the crowdedness term may affect participants' experiences most. For example, without crowdedness consideration, the rating of Satisfaction of In-store Experience decreased by 43.48%. We find that without the crowdedness consideration (W/O CC), most products were placed in one region, leaving other regions empty, due to the reason that in that region products receive higher exposure. However, it is not rational, so participants gave lower ratings on the realism aspect. Additionally, if most products are crowded in one region, some products may pile together. Participants felt that those occluded products might not be reachable even if they wanted to buy those products, resulting in lower ratings in convenience. Finally, the W/O CC achieved the lowest score in the in-store experience satisfaction.

Omitting the visual balance term has less influence on the ratings of convenience. The reason may be that the visual balance term does not affect the function of the store. Unlike the case without the interval constraints, products were distributed irregularly, which may cause overlapping among products, resulting in a lower score on the convenience rating.

5 Discussion

5.1 User Feedback

The participants gave us some additional feedback after the evaluation. In the scene rationality evaluation, many participants stated that they could identify some scenes without spatial considerations. Some participants claimed that the placement was quite awkward in these scenes so they did not want to stay too long. Overall, most participants were aware of the enhancement of placing products (e.g., the products were more regularly placed) and stated that such placements look more realistic.

Most participants commented that they had better immersive experiences during the navigation of virtual stores with a more pleasant product placement. They stated that a pleasant product placement improved shopping satisfaction and shopping behaviors, which might engage participants better and motivate them to spend more time shopping in the store. Some participants claimed that they had a good impression on more products in such scenes.

Some participants commented on other design aspects of the virtual stores. For example, for the virtual clothing store, some participants commented that the store's lighting design could be improved. While we focused on how to synthesize realistic product placement, we could further explore how to synthesize a virtual store considering some physical environmental aspects such as lighting.

A few participants were satisfied with the scene where the products were gathered at the entrance of the store even though the shelves in the store were very sparse. One of those participants believes that for small stores, placing things densely in conspicuous locations (such as near the entrance) will help him make quick choices and save him time for selection. Therefore, the degree of crowdedness is a factor of consideration by most but not all people.

5.2 Analysis of Product Exposure Prediction

To evaluate exposure prediction performance, we calculate the root mean square error (RMSE) between the ground truth i.e., the recorded gaze duration on one virtual product in data collection process and the predicted exposure. Our data collection assignments resulted in a total of $8,809$ data samples. Prior to training our regressors, we processed the raw data by normalizing the product exposure time as described. We randomly sampled a $2,202$ test set (about 25% of the entire dataset) before training our regressor. The remaining samples were used for training.

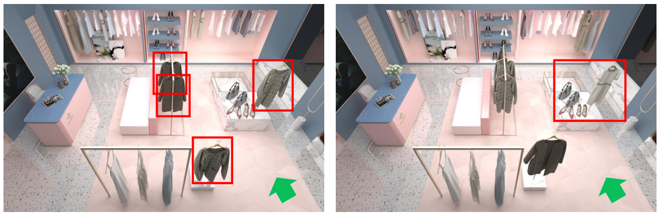
We experimented with training different types of regressors. For the Support Vector Machine regressor, we used an Epsilon-Support Vector Regressor with an ε of 0.01 and an error term penalty parameter C of $1,000$. For the Decision Tree regressor, we set the maximum depth to 5. For the Random Forests regressor, we set the maximum depth to 5 for all 5 trees in the forest. We used these hyper-parameters which yielded the highest accuracies using grid searches with 10-fold cross-validation.

In 10-fold cross-validation of the training set with $6,606$ samples, we obtained the RMSE of 28.02%, 27.57%, and 27.19% for the Support Vector Machine, Decision Tree, and Random Forests respectively. The Random Forests regressor attains the smallest root mean square error on our test set. We chose to use the Random Forests regressor for our optimization for its lowest prediction error.

5.3 Applications

Our optimization framework provides flexibility for the retailer to generate solutions that match different goals or constraints. Our approach enables retailers to participate in the product placement process and tailor their favorite placement. Here are two examples.

Product Promotion. In some cases, the retailer may want to promote certain products. For example, they want to enhance the exposure of some specific products. By increasing the corresponding exposure weights w_i of these priority products in Equation 2, the retailer can prompt our optimizer to prioritize placing these products at locations that are more noticeable to enhance the products' exposure. Fig. 11(a) shows an example. In this example, the exposure weights of the corresponding products (enclosed in red) are set as $w_i = 0.3$ while the exposure weights of the other products are set



(a) Product promotion (b) Product promotion

Figure 11: Applications. (a) The products desired to have more exposures (enclosed in red) are placed by the optimizer at the locations that will receive more visual attention. (b) The user fixed the locations of some products (enclosed in red). The optimizer changes other products' locations while keeping these products fixed.

as $w_i = 0.1$. Starting from a randomized placement, our approach synthesized a placement where the high-priority products are placed at locations expected to gain more exposure.

Location Preference. In some other cases, the retailer may want to fix some products' locations. For example, a retailer may want some products to be placed near the exit of a store regardless of where the other products are placed. Our approach can achieve this feature by using a hard constraint to fix the locations of certain specified products while the optimizer modifies the locations and orientations of other products. Fig. 11(b) shows an example. In this example, five clothing item locations (enclosed in red) are fixed by the user, and our optimizer is asked to modify the locations and orientations of the other products.

5.4 Future Work

In a real-world store, designers refer to the opinion that humans are torn between neophilia (the allure of anything new) and neophobia (the fear of anything new) [9]. Thus, most retailers opt to adjust the store partially to balance the fresh and familiar outlook of the store. A rational way is to change the product placement often and to sustain the scene layout over a long time. In line with such observations, it makes sense that we use a given synthesized scene as input, such as using a scene downloaded from online 3D scene repositories or generated by automatic scene synthesis algorithms [44, 45].

Besides using elaborately designed store scenes, one future direction is to apply our approach to 3D-reconstructed stores of the real world. Take the Metaverse as an application example. It needs a way to bridge the gap between reality and virtual world. With a 3D-reconstructed or a 3D-scanned store as input, retailers can build up a virtual twin store in Metaverse. Then one can apply our pipeline to synthesize product placements for an existing store conveniently for redesigning a store or creating a virtual store. We could also use other scene modeling techniques to create a store layout as input. For example, we can use the Sketch2Scene [43] approach to synthesize a 3D store scene from a sketch.

In our work, we considered exposure and spatial considerations in the optimization framework, which is sufficient to illustrate our core idea. Experiments have shown that considering these two items can obtain results comparable to those of designers. While the absence of either cost leads to worse results. Our framework also allows more considerations to meet retailers' personalized requirements.

Future work may consider commercial factors in the optimization. For example, grocery stores typically have popular products like milk placed in the farthest region to prompt customers to pass by more products as they walk to get the milk. Inspired by this, we may add estimated user paths as a consideration in our optimization. Also, we can introduce specific commercial purposes to guide the scene synthesis, such as increasing the store's revenue. To achieve this goal, we may use each product's profit to weigh the exposure cost and improve the exposure time of the product with a high profit.

We introduce spatial constraints in our pipeline akin to conventional design rules in the real world. The reason behind the consideration is that we want to reduce the learning cost. For most people, the shopping environment metaphors enable them to transfer the shopping experiences without extra effort. Of course, as virtual environments are fully controllable, we may present products in "magical" ways, e.g., showing products that float in the air, animating the displayed products, which may result in more imaginative and creative layouts.

For commercial applications, it seems promising to take advantage of customer behavior data to personalize virtual stores. In that case, state-of-the-art recommendation techniques can be integrated with our approach. For example, based on a customer's previous shopping data and his profile, the system may model his shopping preferences, predicting whether some products are more attractive for the specific user. With such a recommendation model, a company may synthesize a virtual store with product placements personalized with respect to the customer's background, habits, interesting products, etc. Doing so may further improve shopping experiences.

5.5 Limitations

For the purpose of placement, we divide the products into different categories. To devise general placement rules, we only consider the correspondence between products and shelves (e.g., clothes are on clothing shelves, shoes are on shoe shelves). There is no fine-grained categorization of products such as differentiating between dresses and skirts. Depending on the requirements of a specific store, retailers may employ more fine-grained product categorization and design of more specific placement rules, such as putting products of similar categories together.

We experimented with three common spatial considerations in our approach. Our optimization can be extended to consider other spatial constraints such as symmetry in the placement. The retailer may also add other aesthetic constraints to enhance the store's outlook and the shopping experience. In some cases, the designed cost terms may fail to guide the optimizer in searching for a reasonable solution. For example, a severely unbalanced number of products with different sizes within one local area may influence the cost of visual balance.

Our approach focuses on the layout of products in the store. We note that physical environment factors that our approach does not consider, such as the lighting conditions of the store, may have an impact on the appearance of products. For example, lights shone from a different angle or with a different color may change the visual appearance of the products, which may change the product exposure and shopping experience.

6 Conclusion

Our work introduces a novel computational design approach to optimize product placements for virtual stores. Given a store scene and some 3D virtual products, our approach synthesizes product placements considering spatial and product exposure factors. We conducted a user study and demonstrated that our approach provided pleasant visual experiences and increased exposure to virtual products. Our approach would be useful for VR store creation, especially for users who have little knowledge about retail strategies, visual displays, or virtual reality development.

There are some benefits of using our optimization approach. First, our approach provides an automatic way to generate product placement, which enables a virtual store's retailer to change the store layout conveniently without the help of professional designers. It may encourage more retailers to set up virtual stores. Second, our approach enhances the overall exposure of products in a store through optimization, which may increase profits for retailers. Third, the computational design approach is based on an optimization framework that can be extended by designers (or retailers) to incorporate other store design factors such as lighting and decoration styles.

References

[1] D. Adams et al. Virtual retail in the metaverse: Customer behavior analytics, extended reality technologies, and immersive visualization systems. *Linguistic and Philosophical Investigations*, (21):73–88, 2022.

[2] A. Alawadhi and S.-Y. Yoon. Shopping behavioral intentions contributed by store layout and perceived crowding: An exploratory study using computer walk-through simulation. *Journal of Interior Design*, 41(4):29–46, 2016.

[3] R. Alghofaili, M. S. Solah, H. Huang, Y. Sawahata, M. Pomplun, and L.-F. Yu. Optimizing visual element placement via visual attention analysis. In *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 464–473, 2019.

[4] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan. An introduction to mcmc for machine learning. *Machine learning*, 50(1):5–43, 2003.

[5] A. S. Atalay, H. O. Bodur, and D. Rasolofoarison. Shining in the center: Central gaze cascade effect on product choice. *Journal of Consumer Research*, 39(4):848–866, 2012.

[6] P. Chandon, J. W. Hutchinson, E. T. Bradlow, and S. H. Young. Does in-store marketing work? effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of marketing*, 73(6):1–17, 2009.

[7] Y. Chang, Y. Zhao, M. Dong, Y. Wang, Y. Lu, Q. Lv, R. P. Dick, T. Lu, N. Gu, and L. Shang. Memx: An attention-aware smart eyewear system for personalized moment auto-capture. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(2):1–23, 2021.

[8] S. Chib and E. Greenberg. Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335, 1995.

[9] P. Cowan. Neophobia and neophilia: new-object and new-place reactions of three rattus species. *Journal of Comparative and Physiological Psychology*, 91(1):63, 1977.

[10] E. Cutrell and Z. Guan. What are you looking for? an eye-tracking study of information usage in web search. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 407–416, 2007.

[11] ebay. World's first virtual reality department store, May 2016.

[12] C. Ebster. *Store design and visual merchandising: creating store space that encourages buying*. Business Expert Press, 2011.

[13] A. Fender, D. Lindlbauer, P. Herholz, M. Alexa, and J. Müller. Heatspace: Automatic placement of displays by empirical analysis of user behavior. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 611–621, 2017.

[14] M. Fisher, M. Savva, and P. Hanrahan. Characterizing structural relationships in scenes using graph kernels. In *ACM transactions on graphics (TOG)*, vol. 30, p. 34, 2011.

[15] T. Germer and M. Schwarz. Procedural arrangement of furniture for real-time walkthroughs. In *Computer Graphics Forum*, vol. 28, pp. 2068–2078, 2009.

[16] R. C. Gonzalez, R. E. Woods, and S. L. Eddins. *Digital image processing using MATLAB*. Pearson Education India, 2004.

[17] J. Grashuis, T. Skevas, and M. S. Segovia. Grocery shopping preferences during the covid-19 pandemic. *Sustainability*, 12(13):5369, 2020.

[18] E. Horská and J. Berčík. The influence of light on consumer behavior at the food market. *Journal of food products marketing*, 20(4):429–440, 2014.

[19] P. Huddleston, B. K. Behe, S. Minahan, and R. T. Fernandez. Seeking attention: an eye tracking study of in-store merchandise displays. *International Journal of Retail & Distribution Management*, 2015.

[20] H. Hwang, B. Choi, and M.-J. Lee. A model for shelf space allocation and inventory control considering location and inventory level effects on demand. *International Journal of Production Economics*, 97(2):185–195, 2005.

[21] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, 20(11):1254–1259, 1998.

[22] J. Y. Jang, E. Baek, S.-Y. Yoon, and H. J. Choo. Store design: Visual complexity and consumer responses. *International Journal of Design*, 12(2), 2018.

[23] O. Jolliet, A. S. Ernstoff, S. A. Csiszar, and P. Fantke. Defining product intake fraction to quantify and compare exposure to consumer products, 2015.

[24] H. Khachatryan, A. Rihn, B. Behe, C. Hall, B. Campbell, J. Dennis, and C. Yue. Visual attention, buying impulsiveness, and consumer behavior. *Marketing Letters*, 29(1):23–35, 2018.

[25] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.

[26] I. Krasonikolakis, A. Vrechopoulos, A. Pouloudi, and S. Dimitriadis. Store layout effects on consumer behavior in 3d online stores. *European Journal of Marketing*, 2018.

[27] Y. Lang, L. Wei, F. Xu, Y. Zhao, and L.-F. Yu. Synthesizing personalized training programs for improving driving habits via virtual reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 297–304. IEEE, 2018.

[28] C. Li, W. Liang, C. Quigley, Y. Zhao, and L.-F. Yu. Earthquake safety training through virtual drills. *IEEE transactions on visualization and computer graphics*, 23(4):1275–1284, 2017.

[29] W. Liang, J. Liu, Y. Lang, B. Ning, and L.-F. Yu. Functional workspace optimization via learning personal preferences from virtual experiences. *IEEE transactions on visualization and computer graphics*, 25(5):1836–1845, 2019.

[30] L. Majerowicz, A. Shamir, A. Sheffer, and H. H. Hoos. Filling your shelves: Synthesizing diverse style-preserving artifact arrangements. *IEEE transactions on visualization and computer graphics(TVCG)*, 20(11):1507–1518, 2013.

[31] L. M. Marx and G. Shaffer. Slotting allowances and scarce shelf space. *Journal of Economics & Management Strategy*, 19(3):575–603, 2010.

[32] P. Merrell, E. Schkufza, Z. Li, M. Agrawala, and V. Koltun. Interactive furniture layout using interior design guidelines. *ACM transactions on graphics (TOG)*, 30(4):1–10, 2011.

[33] J. Meyer, A. Frank, T. Schlebusch, and E. Kasneci. A cnn-based human activity recognition system combining a laser feedback interferometry eye movement sensor and an imu for context-aware smart glasses. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(4):1–24, 2021.

[34] J. L. Orquin and S. M. Loose. Attention and choice: A review on eye movements in decision making. *Acta psychologica*, 144(1):190–206, 2013.

[35] M. Siegrist, R. Leins-Hess, and C. Keller. Which front-of-pack nutrition label is the most efficient one? the results of an eye-tracker study. *Food Quality and Preference*, 39:183–190, 2015.

[36] M. Speicher. Shopping in virtual reality. In *IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1–2, 2018.

[37] L. M. Surhone, M. T. Tennoe, S. F. Henssonow, and L. Breiman. Random forest. *Machine Learning*, 45(1):5–32, 2010.

[38] L. Tabbaa, R. Searle, S. M. Bafti, M. M. Hossain, J. Intarasisrisawat, M. Glancy, and C. S. Ang. Vreed: Virtual reality emotion recognition dataset using eye tracking & physiological measures. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 5(4):1–20, 2021.

[39] L. W. Turley and R. E. Milliman. Atmospheric effects on shopping behavior: a review of the experimental evidence. *Journal of business research*, 49(2):193–211, 2000.

[40] A. Verhulst, J.-M. Normand, C. Lombart, and G. Moreau. A study on the use of an immersive virtual reality store to investigate consumer perceptions and purchase behavior toward non-standard fruits and vegetables. In *IEEE Virtual Reality (VR)*, pp. 55–63, 2017.

[41] W. Wang and J. Shen. Deep visual attention prediction. *IEEE Transactions on Image Processing*, 27(5):2368–2378, 2017.

[42] N. Xi and J. Hamari. Shopping in virtual reality: A literature review and future agenda. *Journal of Business Research*, 134:37–58, 2021.

[43] K. Xu, K. Chen, H. Fu, W.-L. Sun, and S.-M. Hu. Sketch2scene: sketch-based co-retrieval and co-placement of 3d models. *ACM Transactions on Graphics (TOG)*, 32(4):1–15, 2013.

[44] Y.-T. Yeh, L. Yang, M. Watson, N. D. Goodman, and P. Hanrahan. Synthesizing open worlds with constraints using locally annealed reversible jump mcmc. *ACM Transactions on Graphics (TOG)*, 31(4):1–11, 2012.

[45] L.-F. Yu, S. K. Yeung, C.-K. Tang, D. Terzopoulos, T. F. Chan, and

S. Osher. Make it home: automatic optimization of furniture arrangement. *ACM Transactions on Graphics (TOG)*, 30(4):86, 2011.

[46] S.-H. Zhang, S.-K. Zhang, Y. Liang, and P. Hall. A survey of 3d indoor scene synthesis. *Journal of Computer Science and Technology*, 34(3):594–608, 2019.

[47] N. Zuschke. The impact of task complexity and task motivation on in-store marketing effectiveness: An eye tracking analysis. *Journal of Business Research*, 116:337–350, 2020.