This is the accepted version of the publication Sun, H., Xu, H., Sun, S., Li, H., & Wang, S. (2025). Beyond Visual Appeal: The Impact of Multisensory Experience of Hotel Marketing and Review Images on Sales. Journal of Travel Research, 0(0). Copyright © 2025 The Author(s). DOI: 10.1177/00472875241310818.

Beyond Visual Appeal: The Impact of Multisensory Experience of Hotel Marketing and Review Images on Sales

Haoqiang Sun[a,b], Haozhe Xu[a], Shaolong Sun[a,b,*], Hengyun Li[c], Shouyang Wang[d,e]

[a]School of Management, Xi'an Jiaotong University, Xi'an 710049, China
[b]System Behavior and Management Laboratory, Xi'an Jiaotong University, Xi'an 710049, China
[c]School of Hotel and Tourism Management, The Hong Kong Polytechnic University, Hong Kong SAR, China
[d]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[e]School of Entrepreneurship and Management, ShanghaiTech University, 393 Middle Huaxia Road, Pudong, Shanghai, 201210
*Corresponding author. School of Management, Xi'an Jiaotong University, Xi'an, 710049, China. E-mail address: sunshaolong@xjtu.edu.cn (S. L. Sun)

Acknowledgements

None.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received the support of research funds from the National Key R&D Program for Young Scientists (Project No. 2022YFF0903000), the National Natural Science Foundation of China (Project No. 72101197) and the Fundamental Research Funds for Xi' an Jiaotong University (Humanities and Social Sciences) Free Exploration Student Project (Project No. SK2025033).

Beyond Visual Appeal: The Impact of Multisensory Experience of Hotel Marketing and Review Images on Sales

ABSTRACT

Images, with their direct content experience, have become essential in tourists' narratives of accommodation sharing. Although visual content plays an important role in influencing user decisions, research on hotel images often neglects the significance of sensory elements. Sensory elements in visual content help shape tourists' perceptions and evaluations of their accommodation experiences. This study explores how sensory elements in marketing and review images affect hotel sales from a sensory perspective. The research employs a combination of hybrid machine vision and econometric methods to examine how visual, tactile, auditory, and smell elements influence hotel sales. These results reveal the relationship among sensory elements in marketing and review images, their discrepancies, and hotel sales. The primary theoretical significance of this study is in analyzing how sensory elements in image content influence hotel sales, providing a unique sensory perspective for tourist decision-making and hotel marketing strategies.

Keywords: Marketing images; Review images; Multisensory experience; Machine vision; Hotel sales

1. INTRODUCTION

Visual content is paramount in today's digital landscape, particularly in the hotel industry. The growth of social media platforms has significantly amplified the role of visual elements in driving hotel sales (An et al., 2020). Visual content not only shapes customer perceptions but also influences booking decisions by providing both latent and explicit information (Li et al., 2023a). Given the uncertainty surrounding hotel quality as an experiential product, images are crucial for offering insights into a hotel's offerings and ambiance (Li et al., 2022; Ma et al., 2023).

Explicit images, such as clear photographs of hotel rooms, facilities, and views, provide immediate and straightforward information that enhances the effectiveness of online hotel reviews (An et al., 2020). In contrast, more subtle visual content—referred to here as latent image content, can influence tourist booking behaviors by offering deeper, more nuanced insights into the hotel's atmosphere and overall experience (Hou & Pan, 2023). For instance, Zhan et al. (2024) demonstrated that such latent content can effectively convey emotions, with tools like DeepSentiBank identifying sentiments like "lonely city," "lovely beach," and "excited crowd." Advances in machine vision have further enhanced the accuracy of these emotional interpretations, making visual content increasingly influential in shaping consumer perceptions. As McGrath (2017) noted, the human brain processes visual information faster than text due to our reliance on vision as a primary sensory input. Consequently, visuals not only serve as an essential source of information but also quickly evoke emotional and cognitive responses, making them crucial for capturing and retaining tourist attention (Sun et al., 2024).

Recent advances in social media and mobile devices have increased visual data in the hospitality industry, including marketing and review images (Xiao et al., 2022). Marketing images are designed by hotels to highlight appealing aspects and promote their brand and services. In contrast, review images are spontaneous visuals shared by customers, providing authentic feedback on their experiences (Bufquin et al., 2020). In digital marketing, hotels incorporate multisensory experiences to attract customers beyond visual appeal (Lv et al., 2024). The focus in hotel marketing has expanded from purely visual elements to include auditory, tactile, and smell dimensions, which are critical in shaping customer perception and decision-making (Liu et al., 2022a). The traditional emphasis on visual cues has shifted to multisensory marketing strategies (Lee et al., 2019). Visual elements are integrated with auditory, tactile, and smell elements, offering a comprehensive multisensory experience. Factors such as lobby music (Wen et al., 2020), signature dish flavors (Han & Hyun, 2017), and unique hotel scents (Errajaa et al., 2021) significantly enhance the overall customer experience. This shift from a unidimensional to a multidimensional approach has fundamentally altered hotel marketing dynamics.

Multisensory experiences, including visual, auditory, tactile, and smell elements, play a crucial role in hotel sales (Kim & Perdue, 2013). On online platforms, high-quality images and videos are essential for capturing customer attention and interest (Zhu et al., 2024). Research shows that visually appealing content significantly influences customers' perceptions and decision-making (Li et al., 2024). The portrayal of facilities, rooms, and services plays a key role in shaping initial customer expectations and preferences. Auditory experiences in hotels, often undervalued, profoundly impact customer emotions and satisfaction (Alpert et al., 2005). The strategic implementation of music and ambient sounds can foster a hospitable environment and augment the overall visitor experience. Studies (Wen et al., 2020) suggest that integrating auditory and visual elements elevates perceived quality and luxury, thereby affecting purchase decisions. Tactile experiences in hotels, like the comfort of soft beds and furnishings, impress customers and contribute to increased return rates (Liu et al., 2022a). In the hotel industry, smell experiences are key to customer satisfaction. The presentation and quality of food leave lasting impressions on customers, influencing repeat patronage and reviews (Okumus et al., 2018). Smell elements, like distinctive lobby or room fragrances, subtly but powerfully create memorable experiences (Gambetti & Han, 2022). Well-chosen scents can elicit positive emotions and associations, which are crucial for enhancing customer loyalty (Slåtten et al., 2009).

Given the importance of multisensory elements, numerous studies have begun exploring their roles in customer behavior and marketing (Shahid et al., 2022). According to Kirillova and Chan (2018), the appearance and interior visual design of hotels directly influence customers' first impressions and overall experiences. Hotels with high aesthetic value are more likely to be booked and provide more reliable service quality (Hou et al., 2023; Hou & Pan, 2023; Kirillova & Chan, 2018). Wen et al. (2020) investigated the significance of auditory stimuli, discovering that music has the capacity to elicit emotional responses from customers. They noted that suitable music fosters a pleasant atmosphere, positively affecting customers' emotional experiences. Studies have also further explored the roles of tactile and smell senses. Tactile perception plays a crucial role in customer purchasing decisions (Liu et al., 2022a). Furthermore, the utilization of fragrances that complement the brand image can amplify customers' positive reactions to the environment, thus elevating their satisfaction, propensity to return, and perceptions of the products and services (Errajaa et al., 2021).

However, research on the impact of multisensory experiences on hotel sales, particularly regarding image sensory experiences, remains lacking (Liu et al., 2022a). Image content provides a comprehensive multisensory experience, engaging not only the visual sense but also stimulating other senses, thereby facilitating full immersion (Liu et al., 2022a). Specifically, a distinct divergence exists between the purposes of marketing and review images, resulting in varying influences of sensory elements within hotel marketing contexts (Li et al., 2023b). Predominantly, research has concentrated on visual aspects, neglecting the combined effects of auditory, tactile, and smell experiences (Lv et al., 2024).

To address the above research gaps, this study aims to investigate the impact of multisensory experiences in review and marketing images (i.e., visual, auditory, tactile, smell) on sales using evidence from Xi'an hotels. Both the hotel marketing image sample and the tourist online review sample (images uploaded by customers) were analyzed. First, we examine the impact of multisensory experiences in marketing images on hotel sales. Second, we calculate the sensory elements associated with review images and explore their effects on hotel sales. Lastly, by considering the sensory disparities between marketing images and review images, we explore their effect on hotel sales. This study contributes to the literature in several ways: First, it represents an initial attempt to reveal the role of review image sensory experience by adopting a hybrid machine vision method and econometric modeling with big data. Second, this study enriches the research related to hotel marketing plans by analyzing the impact of marketing images on hotel sales. Third, this study expands the literature by revealing the effects of multisensory elements disparities in images, specifically in visual, auditory, tactile, and smell dimensions. Fourth, this study uses a machine learning model to construct a nonlinear relationship between marketing images, review images, and sales, and further explains this relationship using SHAP (SHapley Additive exPlanations) values.

2. LITERATURE REVIEW AND THEORETICAL BACKGROUND

2.1 *Theoretical foundation of Multisensory Experiences in Hotel Sales*

This study investigates how multisensory experiences of images influence hotel sales, using Sensory Marketing Theory (Kulkarni & Kolli, 2022). Currently, hotel marketing strategies recognize visual content as a key determinant for market positioning and brand cultivation (Li et al., 2022). This recognition stems from visual content's ability to offer customers an intuitive experience (Li et al., 2023c). Hotels enhance potential customers' sensory experiences by optimizing visual content (i.e., review images and marketing images), thereby increasing the hotel's appeal and sales outcomes (Yim & Khuntia, 2021). This marketing strategy leverages the power of visual elements to showcase the hotel's facilities and services, thereby stimulating the interest and booking intentions of the target customer group (Dai et al., 2022; Zhang & Luo, 2023).

However, hotel marketing models have shifted from single-element visual marketing to multisensory marketing due to the diversification of social media content. Existing research has confirmed the profound impact of sensory marketing in shaping customer preferences and decision-making processes, ultimately aiding in deepening the understanding of customer subconscious marketing strategies (Booth, 2014; Kulkarni & Kolli, 2022; Petit et al., 2019). Sensory Marketing Theory explains this phenomenon. Sensory Marketing Theory (Booth, 2014; Lindstrom, 2006; Malhotra, 1984; Petit et al., 2019; Shahid et al., 2022) emphasizes leveraging human senses to affect customer perception, emotion, preference, and behavior. It posited that engaging the senses substantially augments the appeal of products, brands, and retail environments, shaping purchasing decisions and brand allegiance. The theory spans visual, auditory, smell, gustatory, and tactile marketing domains (Lindstrom, 2006). Specifically, visual marketing employs imagery, colors, and design to create memorable brand experiences. Concurrently, auditory, smell, gustatory, and tactile marketing strategies aim to create an immersive sensory experience (Liu et al., 2022a). Understanding of the sensory allure of visual content will elucidate its distinct influences on hotel sales. However, hotel marketing faces limitations in exploring multisensory experiences derived from visual content, including auditory, smell, gustatory, and tactile experiences (Ali & Ahmed, 2019). The impact of image content on hotel sales is particularly challenging.

2.2 *The impact of visual content on hotel sales*

Visual content includes images, graphics, and videos used to convey information, express perspectives, or showcase products and services (Liu et al., 2023a; Zhan et al., 2024; Zhang et al., 2023). In the hotel industry, visual content is mainly categorized into user-generated and hotel marketing visual content (Li et al., 2023b; Phillips et al., 2017). User-generated visual content includes images and videos shared by regular customers on social media platforms. These may feature their experiences during hotel stays or their opinions on the hotel's services, facilities, or environment (Li et al., 2022). In contrast, visual content in hotel marketing is crafted by the hotel management team or marketing personnel to enhance the hotel's brand image, allure potential clientele, and boost reservation rates (Hou et al., 2023). Visual content significantly impacts sales because it plays a crucial role in both customer choices and hotel marketing efforts (Deng & Liu, 2021). For customers, visual content is crucial in selecting hotels. According to the Dual Coding Theory (Paivio, 1991), customers rely on visual stimuli, such as high-resolution images and videos, to evaluate a hotel's ambiance, facilities, and location (Qian et al., 2023). This strategy effectively mitigates the shortcomings of textual descriptions, providing customers with a more thorough and accurate comprehension of the hotel's offerings (Van Looy, 2021). Furthermore, visual content significantly boosts customers' trust in online reviews (An & Ozturk, 2022). Authentic images and videos embedded within reviews offer potential customers reliable feedback on hotel experiences. This stimulates their interest and can potentially sway booking decisions (Li et al., 2019; Zhu et al., 2024). For hotel managers, analyzing customer-generated visual content provides valuable insights into customer preferences and needs, allowing them to refine room layouts, facilities, and marketing strategies (Ryu et al., 2012). Additionally, visual content serves as a crucial tool for hotel brand building and image enhancement (Li et al., 2023a). High-quality images and videos effectively highlight a hotel's distinctive features, attracting potential customers and thereby influencing hotel sales (Hou et al., 2023; Zhu et al., 2024).

Given the significance of visual content in hotel sales, many researchers have studied its influence on customers (An et al., 2020; Deng & Liu, 2021; He et al., 2022; Li et al., 2022; Li et al., 2024; Marder et al., 2021), as shown in **Table 1**. These researchers analyze diverse indicators extracted from visual content to capture customer attention, including image color (Hou et al., 2023; Zhang & Luo, 2023), aesthetic composition features (Hou et al., 2023; Hou & Pan, 2023; Marder et al., 2021), semantic content (Deng & Liu, 2021; Yang et al., 2023; Zhan et al., 2024), and emotional experiences (Hou et al., 2023; Li et al.,2018; Li et al., 2022; Li et al., 2024). They study how these factors influence variables like helpfulness (Li et al., 2022; Ma et al., 2018), enjoyment (Li et al., 2022; Yim et al., 2021; Zhu et al., 2024), and intention (Back et al., 2020; Bufquin et al., 2020; Ryu et al., 2012). Hou and Pan (2023) extend the analysis of aesthetic composition attributes in tourism and hotel marketing photographs and their effects on customer engagement. Li et al. (2022) explore the emotional disparities between user-generated text and images and their effects on perceived usefulness and enjoyment. This study uses deep learning techniques (Arefieva et al., 2021) and econometric models to conduct the analysis. However, studies on the impact of disparities between marketing and review images are lacking. Such disparities might lead to varied interpretations and perceptions of the information presented to customers, thereby influencing their cognitive processes and attitudes towards the hotel (Xiao et al., 2022).

Marketing images are often carefully curated and edited to highlight the hotel's strengths, usually presenting an idealized view (Deng & Liu, 2021). Conversely, review images may provide a more genuine reflection of real hotel experiences, including both positive and negative aspects (Xiang et al., 2024). Consequently, disparities between marketing and review images can prompt changes in customer expectations and attitudes towards the hotel, affecting their decision-making processes (Xiao et al., 2022). Additionally, prior research has mainly focused on visual analysis, neglecting other sensory dimensions like auditory, tactile, and smell experiences. However, images provide a comprehensive sensory experience, engaging not only the visual sense but also stimulating other senses, facilitating complete immersion (Countryman & Jang, 2006; Errajaa et al., 2021; Gambetti & Han, 2022; Zhang et al., 2023).

Insert Table 1

2.3 *Multisensory elements of images and hotel sales*

Competition within the hotel industry is intensifying, with sensory characteristics playing an increasingly pivotal role in shaping customer experiences and driving sales (Ali et al., 2019; Li et al., 2018). This significance is attributed to the fact that sensory perceptions fundamentally mediate our interaction with the environment, guiding cognition and behavioral responses through sensory information and resultant subjective experiences (Kirillova & Chan, 2018).

Recent advancements in sensory marketing research underscore the potential of utilizing sensory elements, such as visual, auditory, tactile, and smell, to significantly bolster hotel sales (Liu et al., 2022a). Kim and Perdue (2013) observed that customer hotel selection is influenced by both cognitive factors, such as price and service quality, and emotional and sensory attributes, including comfort and ambiance. This finding highlights the strategic importance of integrating sensory attributes into marketing efforts. Additionally, Kim et al. (2021a) demonstrated that sensory characteristics could evoke customers' imagination about a hotel's ambiance. Sensory information processing ultimately leads to perception, thereby affecting emotional and cognitive reactions, which in turn shape customer attitudes and behaviors (Kim & Perdue, 2013).

However, a significant portion of the extant literature concerning sensory cues and hotel sales has focused predominantly on the collective effects of the five senses, frequently overlooking the unique impacts attributed to individual sensory cues (Liu et al., 2022a). Such a generalized methodology might obscure the nuanced influences that each sense exerts on customer emotions and cognitive processes. For instance, visual elements hold paramount importance in crafting an ambient atmosphere and projecting a brand's image, significantly influencing emotional experiences through aspects such as color, layout, and décor (Kirillova & Chan, 2018). Conversely, auditory elements, through music and sound effects, serve to effectively modulate emotional states and cognitive perceptions (Wen et al., 2020).

2.3.1 Visual Perception of Images

Visual elements, particularly images, hold paramount importance in captivating attention due to their rich and engaging content (Sun et al., 2024). Our visual system's sensitivity to attributes such as color (Elliot & Maier, 2007), shape (Zhan et al., 2024), and texture (Zhan et al., 2024) makes images a potent asset in hotel marketing strategies. Strategically designed visuals attract customer interest and enhance the appeal of promotional offers (Hou et al., 2023).

Color psychology explains how colors significantly influence emotions, cognition, and behavior (Elliot & Maier, 2007). For example, blue is associated with tranquility and relaxation, making it suitable for bedrooms and relaxation areas (Chung & Saini, 2022). Warm colors like red and yellow stimulate energy and appetite, creating a lively atmosphere (Zhang et al., 2022). Metallic colors such as gold and silver are associated with luxury and wealth, enhancing perceptions of a hotel's high-end status (Kim et al., 2024).

Furthermore, research by Zhan et al. (2024) indicates that the intrinsic semantic content of images conveys specific information and serves as heuristic cues, enhancing viewers' desire to travel and visit. In the hotel industry, review images showing friendly staff and satisfied customers effectively communicate the hotel's professional service and comfortable experience, stimulating potential customers' booking desires (An & Ozturk, 2022). Similarly, promotional materials with appealing food images trigger taste sensations, evoke anticipation for culinary experiences, and increase enjoyment and satisfaction (Booth, 2014).

2.3.2 Auditory Perception of Images

The relationship between visual and auditory perception is well established. The "Kiki-Köhler's 1929 "Kiki-Bouba" experiment demonstrates a strong cross-modal association between shapes and sounds (Barany et al., 2023). This experiment shows how the brain naturally connects visual shapes with specific sounds, such as sharp shapes with sharp sounds like "Kiki" and round shapes with round sounds like "Bouba". This illustrates the interconnected nature of sensory experiences.

According to Multisensory Integration Effects (Calvert et al., 2004; Stein & Stanford, 2008), sensory channels interact closely, enabling visual stimuli to elicit responses across other sensory modalities. Visual information influences not only visual perception but also overall sensory experience by triggering associations with auditory and tactile sensations (Barany et al., 2023; Theeuwes et al., 2007). For instance, viewing an image of a crowded street outside a hotel can evoke auditory associations like the sounds of footsteps, conversations, and car horns (Heung & Gu, 2012). This multisensory representation enriches the overall perception of the scene, enhancing customer experience through a more immersive interaction (Barany et al., 2023; Marks, 1975).

2.3.3 Tactile Perception of Images

The significance of tactility in the hospitality sector, particularly in hotels, is immense (Schmidt et al., 2014). Strategic selection of materials and textures profoundly impacts guest satisfaction by enriching their sensory experience, bolstering positive perceptions of service quality, and enhancing sales performance (Hu et al., 2021; Liu et al., 2022a).

Moreover, the human brain's ability to cross-modally integrate visual and tactile information further amplifies this effect (Martino & Marks, 2000). This phenomenon, known as cross-modal texture transfer, demonstrates how visual stimuli evoke tactile responses (O'Callaghan et al., 2018; Sann & Streri, 2007). When guests view textured images, their brains activate areas responsible for processing touch sensations, creating an illusion of physical contact. Essentially, the visual cortex's interpretation of texture details triggers a tactile response, simulating the sensation of touch through sight alone (Ali & Ahmed, 2019).

Multisensory Integration Effects (Calvert et al., 2004; Stein & Stanford, 2008) elucidates this process, highlighting the importance of image semantics in evoking tactile perception. When people see pictures showing places with lots of touchy-feely stuff, like super soft beds and cozy blankets, they often start imagining how it would feel to touch those things (Liu et al., 2022a). This occurs because our brains connect visual stimuli with past tactile experiences, making us feel the warmth and softness (Martino & Marks, 2000). Our brain's integration of different senses allows us to anticipate how nice and comfortable it will be (Schmidt et al., 2014). Visitors planning a trip may start imagining the positive sensations awaiting them. This mental preparation enhances the actual experience, making it even more enjoyable upon arrival. This deepens and enhances their overall enjoyment (Schmidt et al., 2014).

2.3.4 Smell Perception of Images

Similar to tactile perception, the semantic content of food images is crucial in evoking taste sensations (Errajaa et al., 2021). When people see images of familiar foods, such as a juicy steak or a steaming bowl of soup, they can mentally simulate the taste of visual stimuli related to food can trigger taste memories and cravings (Gambetti & Han, 2022). Food images, especially when combined with smell cues, can enhance the taste experience by triggering associated memories (Han & Hyun, 2017). This suggests that in the hotel industry, using high-quality food images in marketing can create a more immersive and appealing dining experience for potential customers (Lindstrom, 2006).

We have summarized representative studies on sensory characteristics over recent years, encompassing visual, auditory, tactile, and smell aspects, as depicted in **Table 2**. However, most existing research focuses on textual data, often overlooking the rich sensory information in images. Since images convey customer emotions and shape initial hotel impressions (Li et al., 2023b), analyzing sensory characteristics in images is essential. This study aims to explore the impact and emotional influence of sensory characteristics in image data.

Insert Table 2

Based on the above literature review and theoretical analysis, our study establishes the following framework for extracting sensory features from images. As shown in **Figure 1**, the framework identifies analytical processes for visual, auditory, tactile, and smell aspects. This study, based on six theories and experiments, analyzes the relationship between image data attributes and sensory cognition. Additionally, the study empirically examines the relationship between multisensory experience and hotel sales.

Insert Figure 1

3. RESEARCH METHODOLOGY

3.1 *Online review and marketing data*

Ctrip, a widely used online review platform in China, was the primary data source for this study. Hotels, as experience-centric services, prompt customers to rely heavily on online reviews for accommodation choices. Our analysis included 405 hotels with diverse star ratings in Xi'an, spanning various city districts and attractions. To address heterogeneity in hotel services, we employed stratified sampling. Initially, hotels were categorized by location and star rating. Then, the sample size for each star category was determined based on the overall profitability of each region. The final sample selection prioritized hotels with both marketing and review images to enhance data relevance, as shown in **Figure 2**. Therefore, this study obtained 405 hotels with marketing images and collected all review data for these hotels. Ultimately, we retained online reviews with both text and images, comprising 87,800 review images and 4,224 marketing images.

The Ctrip dataset includes: (1) hotel-level data, including names, IDs, locations, marketing images, and other pertinent attributes; (2) review data, featuring ratings, texts, and images for each hotel. We analyzed hotels ranging from 2 to 5 stars in Xi'an, between August 19, 2017, and August 20, 2021, focusing on establishments with both marketing and review images and textual reviews. Sales volume, indicated by total review counts (Ye et al., 2009; Öğüt & Onur Taş, 2012), was the dependent variable in our model. These reviews are verified reviews from real customers on Ctrip. Control variables included each hotel's star rating, *kf_num*, *location*, establishment *age*, *RatGood*, *RatImage*, *RatOverall*, *ReadPre*, *RatPos*, and *RatGood*RatImage*. *kf_num* indicates the number of guest rooms in a hotel. *Location* was measured by proximity to the airport, with a value of 1 assigned to hotels within 2 km and 0 to those beyond 2 km. *RatGood* is the proportion of customers with positive ratings, *RatImage* refers to the proportion of reviews containing images for a hotel, *RatOverall* refers to the hotel's rating on Ctrip, *ReadPre* refers to the readability of the hotel's description text, and *RatPos* refers to the proportion of positive text reviews. We also considered the interaction effect between *RatGood* and *RatImage*, given that extremely satisfied or dissatisfied customers are more likely to upload images and leave reviews, while other customers may not take any action.

3.2 *Methods for sensory information mining*

Our variable measurement framework is shown in **Figure 2**. The extraction process is divided into two main categories: extracting multisensory characteristics from marketing images and review images. The extraction dimensions are categorized into four aspects: visual, auditory, tactile, and smell. The specific extraction process is detailed in the following subsections.

Insert Figure 2

3.2.1 Visual content extraction

Color and semantic content in images significantly influence customers' visual perception of hotels (Chung & Saini, 2022; Sun et al., 2024; Zhan et al., 2024). From the perspective of visual sensory perception, we employed color features to encapsulate one dimension of sensory vision, using *HSV* (Hue, Saturation, Value) values to delineate color vision in hotel image data (Sun et al., 2024). *HSV* represents color, where H stands for hue, S for saturation, and V for value. Hue defines the type of color, saturation measures the intensity or purity, and value determines the brightness or darkness (Hou et al., 2023). The *HSV* color model, being more intuitive than the *RGB* (Red, Green, Blue) space, closely corresponds with human visual perception (Hou et al., 2023). We extracted the color vision attributes of marketing and review images for each hotel, designating them as variables *Visual_H*, *Visual_S*, and *Visual_V*.

The objects depicted in images represent their semantic content, and we extracted these objects to capture the semantic dimension of visual perception. We employed the yolov5 algorithm (Liu et al., 2022b; Zhang et al., 2019) to identify objects within images **(Appendix A)**. Examples of test results are shown in **Table 3**. The model's adaptable features make it suitable for analyzing diverse hotel amenities, enhancing its ability to extract data from images. This capability aligns with our investigation requirements. Post-training, the model detected all requisite object types for our study, encompassing prevalent hotel amenities such as beds, tables, and toilets, identifying 36 object classes, as shown in **Table 4**. These 36 classes were segregated into distinct themes, with objects in guest rooms grouped under 'Room and Suite' and those in bathrooms under 'Bathroom' (**Appendix B**). We determined the proportions of six thematic categories within the marketing and review images for each hotel, designated as variables *Visual_Bathroom*, *Visual_Facilities*, *Visual_Guest-room*, *Visual_Bed*, *Visual_gourmet*, *Visual_Customer*. A descriptive analysis was conducted on the statistical outcomes of the visual sensory feature indicators, as delineated in **Table 5**.

Insert Table 3

Insert Table 4

Insert Table 5

3.2.2 Auditory information extraction

Our research mainly illustrates auditory perception using the shapes and semantic content of images. As demonstrated by Köhler's experiment, shapes are associated with sounds. Simon Lacey and his team expanded Köhler's work by transforming symbols into sounds, enhancing the understanding of audiovisual associations (Lacey et al., 2020). They paired auditory pseudowords with visual shapes and used representational similarity analysis (RSA) to link acoustic and visual properties. This method connected 90 different shapes, varying in protrusion and contour, with sounds rated from 1 (most rounded) to 7 (most sharp), effectively bridging the gap between sight and sound.

Building on Simon Lacey's research (Lacey et al., 2020), we focused on entire images and used the Canny algorithm to extract their contours. The Canny algorithm, an edge detection method, uses multi-stage processing to deliver clear and precise edge identification. It effectively suppresses noise and offers adjustable edge detection results. Using the OpenCV package, we compared the similarity of image contours with the 90 shapes from Simon Lacey's research. The roundness/sharpness level of the shape with the highest similarity was used as an auditory attribute sensory feature of the image, as shown in **Figure 3.**

Insert Figure 3

Finally, we calculated the roundness/sharpness levels for the shapes of all images for each hotel and averaged them to obtain the roundness/sharpness level for each hotel, defined as the variable *Auditory_KB*. This variable represents customers' auditory perception of the hotel.

Human perception is not limited to just the apparent features of images; the semantic features of images can also greatly influence perception (Kim et al., 2021; Zhan et al., 2024). To correlate the semantic features of images with auditory sensory characteristics effectively, we used the DeepSentiBank model to extract adjective-noun pairs from the images (He et al., 2022; Hou et al., 2023; Sun et al., 2024). This model is a concept classifier trained on over a million geotagged images and can generate 2,089 adjective-noun pairs (ANPs) composed of 231 adjectives and 424 nouns, transforming image information into text.

In the adjective-noun construct, nouns play a crucial role in conveying multisensory experiences. Therefore, we used nouns to encapsulate the sensory dimension inherent in the entire adjective-noun pairing (He et al., 2022). To transform nouns into sensory features, we used the methodology proposed by Lv et al. (2024), which employs direct sensory stimuli as the criterion. For instance, the noun "Birds" is associated with auditory responses and thus classified as an auditory adjective-noun pair, whereas "Food" is related to smell responses and is categorized as a smell adjective-noun pair. Five doctoral students with relevant research experience independently conducted this encoding process. They were instructed to record the first sensory experience that came to mind when encountering the nouns. Before formal encoding, each person practiced by encoding 30 nouns and cross-checked the results to ensure consistency. After verification, they discussed and agreed upon a consistent set of encoding results, then proceeded with the formal encoding. Next, we input the nouns into GPT-4.0, asking it to encode the data with different sensory features (**Appendix C**). We compared our encoding results with those from GPT-4.0 and found over 95% similarity, proving the reliability and representativeness of our results. We identified the ten adjective-noun pairs with the highest confidence levels for each image to outline its adjective-noun pair characteristics. Subsequently, we quantified the prevalence of adjective-noun pairs across all images for each hotel and independently calculated the proportions of adjective-noun pairs associated with auditory, tactile, and smell senses, designating these as variables *Auditory_ANPs*, *Tactile_ANPs*, and *Smell_ANPs*, respectively. Partial results of auditory adjective-noun pairs are presented in **Table 6**.

Insert Table 6

As shown in **Table 7**, we conducted a descriptive analysis of the statistical results for the auditory sensory feature indicators.

Insert Table 7

3.2.3 Tactile and Smell perception extraction

To extract tactile perception information from hotel images, we used the LBP (Local Binary Patterns) algorithm to capture texture features **(Appendix D)**. LBP is invariant to grayscale and rotation, making it ideal for hotel images. It identifies local variations in pixel values, indicating surface details and tactile sensations. The LBP algorithm compares each neighboring pixel to the central pixel, assigning a value of 1 if the neighbor is greater and 0 otherwise. This process creates LBP images. We then normalized the histograms of these images and calculated their mean and variance (Ahonen et al., 2006).

Variance reflects texture disparities: higher variance indicates rougher textures, while lower variance suggests smoother textures. We calculated the texture variance for each hotel image and then computed the average texture variance for each hotel, defining this as the variable *Tactile_LBP*, representing customers' tactile perception of the hotel.

Tactile perceptual information is derived from the encoding results based on ANPs. Similarly, smell perceptual information is obtained through the encoding results based on ANPs. This approach is consistent with the method used for extracting auditory perceptual information. Partial results of tactile ANP and smell ANP are shown in **Table 8**.

Insert Table 8

As shown in **Table 9**, we conducted a descriptive analysis of the statistical results for the tactile and smell sensory feature indicators.

Insert Table 9

3.2.4 Disparities in sensory features between marketing images and review images

Marketing images are promotional strategies used by managers to maximize profits. However, these strategies often lack authenticity. In contrast, review images, genuinely uploaded by customers, effectively reflect the true state of the hotel (Sun et al., 2024; Zhan et al., 2024). When there is a significant difference between marketing and review images, customers often find marketing images unreliable. To predict the impact of disparities between marketing and review images on hotel sales, we define the difference using Formula 1 (Li et al., 2022).

$$disparity = \left| PRO - REV \right| \tag{1}$$

Where *PRO* represents the characteristics of marketing images, and *REV* represents the characteristics of review images, with the results shown in **Table 10**.

Insert Table 10

4 MODEL SPECIFICATION

4.1. *Regression Analysis*

We used Poisson regression, a method in generalized linear regression, to analyze count-dependent variables (Zhu et al., 2024). Logarithmic transformation was applied to all skewed variables (*Visual_H*, *Visual_S*, *Visual_V*, *kf_nums*). First, we tested the primary impact of sensory features in marketing images on hotel sales using all marketing image samples. The specific econometric model is shown in **Equation (2)**:

$$\begin{aligned}&Hotel\_sales_{ijk} = \beta_0 + aPRO\_Visual_{ijk} + bPRO\_Auditory_{ijk} + cPRO\_Tactile_{ijk}\\&+dPRO\_Smell_{ijk} + \pi Controls_{ijk} + \varepsilon_{ijk}\end{aligned} \tag{2}$$

where i, j, and k respectively denote the review, reviewer, and hotel. *PRO* represents marketing image data.

To analyze the perceptual characteristics of review images, we used all review image samples to test the primary impact of sensory features on hotel sales. The specific econometric model is shown in **Equation (3)**:

$$\begin{aligned}&Hotel\_sales_{ijk} = \beta_0 + aREV\_Visual_{ijk} + bREV\_Auditory_{ijk} + cREV\_Tactile_{ijk}\\&+dREV\_Smell_{ijk} + \pi Controls_{ijk} + \varepsilon_{ijk}\end{aligned} \tag{3}$$

where i, j, and k respectively denote the review, reviewer, and hotel. *REV* represents review image data.

To investigate the impact of sensory disparities on hotel sales, we compared marketing images to review images. Our analysis tested the primary effect of perceptual variability using samples from both categories of images. The specific econometric model for this is shown in **Equation (4)**:

$$
\begin{aligned}
&Hotel\_sales_{ijk} = \beta_0 + aDIS\_Visual_{ijk} + bDIS\_Auditory_{ijk} + cDIS\_Tactile_{ijk} \\
&+dDIS\_Smell_{ijk} + \pi Controls_{ijk} + \varepsilon_{ijk}
\end{aligned} \tag{4}
$$

where i, j, and k respectively denote the review, reviewer, and hotel. *DIS* represents the difference between review images and marketing images.

4.2 *Feature Importance Analysis*

To examine the effect of different perceptual types on hotel sales, we proposed an analytical framework. This framework includes two methods: category-level feature importance and individual-level feature importance (Zhou et al., 2021). First, we used a standard feature permutation strategy ("permutation feature importance") to assess the importance of features at the sensory category level, as shown in **Equation 5**. Specifically, we defined feature importance based on the change in *Log-Likelihood*. This change is calculated by subtracting the *Log-Likelihood* value obtained when a certain category of sensory features is removed from the value when it is not removed. The *Log-Likelihood* represents the natural logarithm of the *likelihood* function, which evaluates the probability of observing the given data under specific model parameters. It serves as an effective measure of error in generalized linear regression models. Finally, to explore the importance of individual dimension-level sensory features further, we used a correlation analysis algorithm to calculate the relationship between individual-level sensory variables and hotel sales.

$$
IMP_i = \frac{LogLike_{chan_i} - LogLike_{orig}}{LogLike_{orig}} \tag{5}
$$

Where $LogLike_{chan_i}$ refers to the log-likelihood value of the generalized linear regression model after removing the i_{th} sensory category, and $LogLike_{orig}$ is the log-likelihood value of the generalized linear regression model when no sensory features are removed.

5. EMPIRICAL RESULTS

This paper analyzes the visual, auditory, tactile, and smell sensory elements in marketing and review images and how these elements shape consumers' hotel experiences. By thoroughly exploring the sensory information in these images, this study aims to quantitatively assess the impact of these multisensory experiences on hotel sales. Furthermore, it investigates the disparities between marketing and review images in conveying multisensory experiences and their roles in hotel sales performance.

5.1 *The effect of sensory features of marketing images on hotel sales*

We initially examined how multisensory experiences in marketing images influence hotel sales. Using data from 405 hotels, we utilized Equation (2) to evaluate hotel sales and presented the findings in **Table 11**. All models meet the overdispersion criteria, confirming the suitability of our generalized linear regression models for the primary analysis.

In **Table 11**, Model 1 includes control variables. Model 2 incorporates the variables of interest (*PRO_Visual_H*, *PRO_Visual_S*, *PRO_Visual_V*, *PRO_Visual_Bathroom*, *PRO_Visual_Facilities*, *PRO_Visual_Guest_room*, *PRO_Visual_Bed*, *PRO_Visual_gourmet*, *PRO_Visual_Customer*, *PRO_Auditory_KB*, *PRO_Auditory_ANPs*, *PRO_Tactile_LBP*, *PRO_Tactile_ANPs*, and *PRO_Smell_ANPs*). The estimated results indicate that sensory variables (visual, auditory, tactile, smell) in marketing images impact hotel sales. Certain visual elements in hotel marketing images positively impact hotel sales. For example, elements like *PRO_Visual_Bathroom* (coefficient=1.5604, $p<0.01$), *PRO_Visual_Facilities* (coefficient=1.6473, $p<0.01$), and *PRO_Visual_gourmet* (coefficient=0.6128, $p<0.01$) in hotel marketing images. This effect is due to customers' initial exposure to these images, leading to heightened attention to depicted comfort and quality. Consequently, this enhances perceived value, thereby increasing sales.

Additionally, color visual perception plays a regulatory role in hotel sales. According to color psychology (Elliot & Maier, 2007), color information in visual perception directly influences customers' first impressions. However, hotels often use exaggerated marketing images. Often, only the promotional images of luxury hotels are consistent with the actual situation. Therefore, this phenomenon often has a counterproductive effect, as customers overlook color comfort, such as *Log PRO_Visual_H* (coefficient=-0.2930, $p<0.01$), *Log PRO_Visual_S* (coefficient=-0.0191, $p<0.01$), *Log PRO_Visual_V* (coefficient=-0.5249, $p<0.01$). This indicates that the current trend of exaggerated marketing photography has led to an evident lack of trust in visual perception among tourists.

The auditory sensory aspect, such as *PRO_Auditory_ANPs* (coefficient=0.4912, $p<0.01$), positively impacted hotel sales. This indicates that a higher proportion of auditory features correlates with higher hotel sales. *PRO_Auditory_KB* (coefficient=-0.0551, $p<0.01$) represents the roundness/sharpness of sounds. Its negative coefficient suggests that sharper sounds are associated with lower hotel sales. Hotels are meant for relaxation and sleep. After a busy day, customers usually prefer a calm atmosphere, and loud noises can be bothersome. Similarly, both tactile and smell sensory aspects positively influenced hotel sales. When hotels comprehensively showcase tactile and smell characteristics, it allows customers to gain a more holistic understanding of the hotel. For example, displaying an image of gourmet food on the homepage can instantly evoke a smell response in customers, thereby increasing sales.

Insert Table 11

5.2 *The effect of sensory features of review images on hotel sales*

Table 12 presents the estimated effects of review imagery on hotel sales. Model 3 includes relevant control variables, namely *Stars*, *Location*, *Age*, *Log kf_num*, *RatGood*, *RatImage, RatOverall, ReadPre, RatPos,* and *RatGood*RatImage*. Model 4 examines the influence of multisensory experiences (visual, auditory, tactile, and smell) as independent variables. Results from Model 4 indicate that visual elements in review images, reflecting customers' true perceptions, greatly impact hotel sales. Specifically, the significant negative coefficient of *REV_Visual_V* (coefficient=-1.1136, p<0.01) indicates that authentic visuals in reviews enable customers to make more informed decisions, negative affecting hotel revenue. Conversely, realistic color representation, as shown by *REV_Visual_H* (coefficient=0.3621, p<0.01) and *REV_Visual_S* (coefficient=0.6908, p<0.01), enhances hotel sales. In the visual perception, aspects like *REV_Visual_Facilities* (coefficient=0.4886, p<0.01), *REV_Visual_Bed* (coefficient=0.7891, p<0.01), *REV_Visual_gourmet* (coefficient=0.6937, p<0.01), *REV_Visual_Customer* (coefficient=1.3526, p<0.01) are beneficial for sales. However, the coefficient for *REV_Visual_Guest_room* (coefficient=-0.2568, p<0.01) suggests a negative effect on hotel sales.

Undoubtedly, auditory sensory elements, as exemplified by *REV_Auditory_KB* (coefficient=0.2364, p<0.01) and *REV_Auditory_ANPs* (coefficient=0.3070, p<0.01), exert a positive influence on hotel sales. Similarly, tactile sensory components in review images, particularly *REV_Tactile_LBP* (coefficient=-54.4237, p<0.01) and *REV_Tactile_ANPs* (coefficient=0.8931, p<0.01), significantly affect hotel sales. *REV_Tactile_LBP* is significantly negatively correlated with sales, whereas *REV_Tactile_ANPs* show a significant positive correlation. Additionally, smell aspects, as shown by *REV_Smell_ANPs* (coefficient=1.8191, p<0.01), positively affect hotel sales. Illustrative images, especially those showing delicious food, can activate the sense of smell and positively influence customer decisions.

Insert Table 12

5.3 *The effect of sensory features disparities on hotel sales*

To explore the impact of sensory disparities on hotel sales, we selected the absolute values of sensory features differences between marketing and review images as the independent variables. This includes 14 sensory variables across four dimensions (visual, auditory, tactile, smell). The estimation results in **Table 13** passed the overdispersion test.

Overall, most sensory disparities negatively affect sales, as shown by *DIS_Visual_S* (coefficient=-0.4339, p<0.01), *DIS_Visual_Bathroom* (coefficient=-0.5378, p<0.01), *DIS_Visual_Guest_room* (coefficient=-0.2997, p<0.01), *DIS_Auditory_ANPs* (coefficient=-1.0180, p<0.01), *DIS_Tactile_LBP* (coefficient=-188.0747, p<0.01), *DIS_Tactile_ANPs* (coefficient=-0.6718, p<0.01), and *DIS_Smell_ANPs* (coefficient=-0.3921, p<0.01). A significant discrepancy between marketing and review images can create distrust among consumers. They may believe the hotel is deliberately using misleading images to boost sales, while they trust tourist-uploaded review images more (Uthaisar et al., 2024).

Our research further reveals a phenomenon: specific sensory disparities can positively impact hotel sales, such as *DIS_Visual_H* (coefficient=0.1845, p<0.01), *DIS_Visual_Facilities* (coefficient=1.5792, p<0.01), *DIS_Visual_Bed* (coefficient=0.2937, p<0.01), *DIS_Visual_gourmet* (coefficient=0.1752, p<0.01), *DIS_Visual_Customer* (coefficient=1.1636, p<0.01), and *DIS_Auditory_KB* (coefficient=0.0340, p<0.01). This may be due to the hotel's marketing strategies—carefully selecting and optimizing promotional images to compensate for shortcomings. Simultaneously, customer-uploaded review images display the hotel's actual conditions. According to Expectation-Confirmation Theory (Oliver, 1980), consumer satisfaction and purchasing decisions largely depend on comparing expectations and actual experiences. When promotional images raise customer expectations by highlighting unique selling points and review images show the real situation, if the actual experience exceeds expectations, this positive confirmation can significantly enhance customer satisfaction. Increased satisfaction enhances the hotel's market attractiveness and competitiveness, directly promoting sales growth.

Insert Table 13

5.4 *Ranking of Sensory Feature Importance Based on Regression Models*

5.4.1 Category-level feature

In **Table 14**, using the feature permutation strategy, we determined the relative importance of sensory categories. It is clear that visual sensory elements are paramount for both marketing and review images, while smell sensory elements have the least impact. This suggests that customers predominantly rely on visual cues to form their perceptions of a hotel brand, as images cannot effectively transmit smell experiences. This insight has significant implications for hotel visual marketing strategies and design aesthetics. In marketing images, tactile sensory elements matter more than auditory ones, highlighting the effectiveness of emphasizing textures and touch to attract customers. Conversely, auditory sensory elements are prioritized over tactile ones in review images. This indicates that sonic characteristics, such as the definition and acoustics of shapes, along with auditory ANPs, are crucial in attracting attention and fostering positive associations.

Insert Table 14

5.4.2 *Individual-level feature*

The relationship between the importance of single-dimension category-level sensory features and hotel sales, calculated using a correlation analysis algorithm, are shown in **Figure 4**.

Insert Figure 4

Figure 4(a) shows that *PRO_Visual_V* (0.17) and *PRO_Visual_S* (0.14) have the strongest correlations with sales among marketing images. This highlights the significant effect of color in the visual sensory domain on sales figures. Incorporating visually compelling colors and elements can significantly enhance customer engagement and satisfaction with hotel images. Following these, *PRO_Auditory_KB* (0.10) ranks third, indicating customers' notable emphasis on auditory characteristics of marketing images. Therefore, hotels should incorporate enjoyable auditory elements, such as music or ambient sounds, into their marketing materials. This enhances the appeal and memorability of the content, aligning with customer preferences for auditory experiences. Other sensory features have weaker correlations with hotel sales (below 0.1), suggesting their relatively minor influence. Specifically, *PRO_Smell_ANPs* (0.01) shows a negligible impact of smell perception on hotel sales, aligning with previous findings. Therefore, when developing marketing strategies, prioritizing innovation and design in both visual and auditory aspects is advisable. This ensures alignment with customer perceptions, maximizing the effectiveness and impact of these initiatives.

Figure 4(b) shows that in review images, *REV_Visual_S* (0.25) and *REV_Visual_V* (0.23) have strong correlations with sales, reaffirming the significance of color perception. Notably, *REV_Smell_ANPs* (0.23) and *REV_Tactile_ANPs* (0.22) are the third and fourth most influential factors, highlighting customers' substantial interest in smell and tactile experiences from previous visitors. This finding contradicts the minor significance of smell perception in overall feature importance. However, it emphasizes the significant impact of specific and personal scent experiences mentioned in reviews, which can greatly influence perceptions. Unlike controlled marketing images, authentic, customer-shared smell narratives in reviews are more likely to engender trust and empathy, amplifying their sales impact.

6. NON-LINEAR HOTEL SALES ANALYSIS

6.1 *Predicting Hotel Sales with Multisensory variable*

In the second phase of our analysis, we used a machine learning classifier to explore the nonlinear relationship between image sensory characteristics and hotel sales. We employed the non-parametric Gradient Boosting Machine method (**Appendix E**) (Parsa et al., 2020; Zhou et al., 2021). This method improves prediction accuracy by incrementally building decision tree models, each trained on the residuals of the previous model. Each new model corrects the errors of the previous one, ultimately forming a powerful ensemble to enhance predictive performance. We allocated 80% of the data as the training set and 20% as the test set, using a grid search to select the optimal parameters for the Gradient Boosting Machine.

Next, to explain the nonlinear relationship between the sensory variables and sales, we used SHAP to interpret the model. SHAP values, based on cooperative game theory, provide a sound method for distributing the output of a coalition among its members (Chen et al., 2021; Zhou et al., 2021). In our context, the coalition consists of interpretable model input features (e.g., image sensory features), and the output is the prediction value (e.g., hotel sales) given those features. SHAP values measure a feature's importance by the change in the model's output when the feature is known versus unknown (Liu et al., 2023b). This allocation adheres to local accuracy and consistency, making it highly reliable. Additionally, the SHAP method effectively explains complex model outputs, providing a more intuitive and credible analysis of feature importance, making it suitable for our data (Zhou et al., 2021).

To clarify the results, we used density scatter plots to summarize the impact of all features across the entire dataset. The plot combines feature importance with feature effects.

Each point represents a SHAP value for a feature and an instance. The y-axis position is determined by the feature, and the x-axis position by the SHAP value. The color represents the feature value from low to high. Overlapping points are jittered in the y-axis direction, showing the distribution of the SHAP values per feature. Features are ordered by importance. The greater the difference between the maximum and minimum SHAP values for a variable, the more significant its impact on the model (Zhou et al., 2021). Red SHAP values indicate a positive influence on sales, suggesting that an increase in the feature value tends to increase the model's sales prediction. Conversely, blue SHAP values signify a negative impact on sales, implying that an increase in the feature value leads to a decrease in predicted sales.

6.2 *Explanation of SHAP Value Results*

Figure 5(a) shows that *RatPos*, *RatGood*, and *kf_num* positively influence the model output, with higher values generally corresponding to higher SHAP values. This indicates that increases in these factors substantially enhance the predicted value. The *Stars* variable also has a positive influence, with higher star ratings generally associated with higher predicted outcomes, though SHAP values vary. Conversely, *RatImage* negatively impacts the model; higher image ratios tend to decrease the predicted value, while lower ratios enhance it. *PRO_Visual_Facilities* consistently exerts a positive effect on predictions, while *PRO_Visual_Guest_room* shows a mixed impact, indicating uncertainty or complexity in its relationship with the model output. *PRO_Auditory_KB* negatively influences the model, with sharper sounds leading to lower predicted values. *PRO_Visual_gourmet* has a dispersed impact, with some high-value samples positively influencing predictions, though the overall effect remains minimal. Other variables have minimal impact, with SHAP values concentrated near the zero axis.

Figure 5(b) confirms that *RatPos*, *RatGood*, and *kf_num* significantly positively impact the model, while RatImage has a negative effect. High values of *REV_Visual_S* (color saturation) positively influence the model output, suggesting that more vibrant colors enhance predictions, likely by increasing visual appeal. Similarly, higher values of *REV_Smell_ANPs* positively affect predictions, indicating that scent-related features significantly contribute to user satisfaction. *REV_Visual_Customer* also positively impacts the model, indicating that visual aspects of customer interactions significantly enhance satisfaction or effectiveness. Other variables exhibit minimal impact, with SHAP values concentrated near the zero axis, indicating their limited contribution to the model's predictions. We discuss the results in more detail in **Appendix F**.

Insert Figure 5

7. DISCUSSION AND IMPLICATIONS

7.1 *Conclusion and discussion*

The visual experiences presented by marketer-generated and reviewer-generated images have been recognized (Deng & Liu, 2021; He et al., 2023; Li et al., 2023b; Zhan et al., 2024), but the other multisensory experiences they evoke have yet to be fully explored. Current research primarily focuses on visual content in a vast amount of image data (Deng & Liu, 2021; Zhan et al., 2024), but few studies investigate multisensory experiences from images. This study explores the influence of multisensory experiences from hotel marketing and review images on sales. We use hybrid machine vision models to extract sensory experiences from large datasets. Then, we use econometric models to examine their impact on sales. Our research elucidates the impact of multisensory experiences on hotel sales and their differences, closing existing knowledge gaps. Our conclusions are summarized as follows:

Firstly, we found that multisensory experiences in marketing images significant impact hotel sales. This result suggests that multisensory aspects of images can enhance marketing strategies. Li et al. (2023c) discovered that visual content in online reviews can enhance customer engagement, which is crucial for businesses. Our research further indicates that the sensory elements of visual, auditory, tactile, and smell present in marketing images can serve as a rich source of sensory cues to affect hotel sales. Additionally, the exaggerated visual colors in marketing images are one of the key factors contributing to customers' distrust, which in turn affects hotel sales.

Secondly, the review images accurately reflect customers' true feelings, especially the impact of authentic color visuals on hotel sales. The authenticity in review images makes sensory experiences have a larger impact on hotels, particularly in visual, auditory, and tactile perceptions. As An et al. (2020) mentioned, customers tend to express their real-life accommodation experiences in images. Our research further demonstrates that sensory experiences in user-generated review images significantly impact hotel sales.

Thirdly, Li et al. (2022) demonstrated the role of review image sentiment and text-image sentiment disparity. Different from prior studies, we introduce the concept of sensory disparities between marketing and review images and analyze their effect on hotel sales. The study results show that significant disparities in multisensory experiences between marketing and review images often negatively impact hotel sales. This negative impact is related to sensory perceptions like visual, auditory, tactile, and smell disparities, leading to a lack of trust among tourists. However, we found some factors that produced the opposite result (e.g., *DIS_Visual_H*, *DIS_Auditory_KB*). The greater the disparities in these sensory elements, the more beneficial they are to hotel sales. This phenomenon supports the Expectation-Confirmation Theory from a sensory perspective (Oliver, 1980). Enhancing customer expectations through marketing images and fulfilling or exceeding them with authentic review images can boost hotels' cognitive appeal.

Finally, our study reveals that the sensory characteristics in marketing and review images have different significance for hotel sales. Visual sensory characteristics are the most significant in both marketing and review images. Smell sensory aspects are least impactful in marketing images but gain importance in review images due to enhanced authenticity and trustworthiness. In marketing images, tactile sensory features are more crucial than auditory features, whereas in review images, auditory features surpass tactile ones. Visual and auditory variables are among the top three most impactful on sales in marketing images, with smell variables ranking lowest. In review images, visual variables have the highest impact on sales, and the significance of smell variables increases, highlighting their role in enhancing sensory experience authenticity. Overall, visual sensory characteristics dominate in both marketing and review contexts, while smell sensory aspects become more pronounced in review images, emphasizing authenticity's role in influencing customer trust and perceptions.

7.2 *Theoretical Contributions*

This study makes several significant theoretical contributions. First, it addresses a gap in the existing literature on multisensory experiences and emotional analysis in online reviews, which has mainly focused on textual content (e.g., Deng & Liu, 2021; Ma et al., 2023; Xiao et al., 2022). While the impact of visual content in images has gained attention in recent years for corporate marketing (e.g., An et al. 2020; Li et al. 2022), most studies have focused on the sentiment in online review images. They have overlooked the multisensory experiences associated with image features, particularly the effect of marketing and review image characteristics on customers' sensory experiences (Zhan et al., 2024). Our research offers a methodology for extracting sensory features from images, enriching the theoretical understanding of multisensory experiences related to images. Specifically, we extracted visual, auditory, tactile, and smell features from images and validated their impact on hotel sales. The results indicate that multisensory experiences play a crucial role in customers' decision-making processes, providing a theoretical basis for hotels to formulate targeted marketing strategies.

Secondly, we investigated the effect of disparities in sensory features between marketing and review images on hotel sales, exploring the relationship between these disparities and hotel sales. Our findings suggest that the disparity between marketing and review images may reduce their impact on sales. Therefore, our insights into the sensory disparities between customer generated images and marketer generated images expand the existing body of research on hotel marketing strategies on social media (An et al. 2020; Li et al. 2022). Our study also extends the application of visual marketing in hotel management, contributing insights to the literature on sensory research, hotel marketing, and user-generated content.

Thirdly, this study enhances the understanding of factors affecting hotel sales and expands knowledge about the sensory value of image content in online reviews (An et al. 2020; Hou & Pan, 2023). The sensory experiential value of images plays a crucial role in attracting customers and shaping their attitudes and behaviors (Park & Nicolau, 2015). Despite its importance, there has been limited research examining the impact factors or moderating effects of sensory experiential value on hotel sales. This study examines different sensory experiential values and their impact on hotel sales. It also prioritizes these values in both review and marketing images, enriching the existing literature.

Last but not least, this study enriches the literature on customer and marketer web information search by investigating the impact of sensory experiences from marketer generated images and customer generated images on hotel sales in online reviews. Visual information, including review images and marketing images, plays an increasingly important role in customer decision-making (An et al. 2020; Li et al. 2022; Li et al. 2023a). Existing research primarily focuses on the presence and quantity of images in online reviews (Li et al. 2022; Li et al. 2023c). This study examines the sensory effects of image information, providing a unique perspective on how multisensory experiences of images influence customer decision-making and marketing strategies.

7.3 *Practical implications*

This research also offers practical contributions. People increasingly rely on marketing content and user-generated content for decision-making, especially in the hospitality industry (Li et al., 2023b; Liu et al., 2022a; Xiao et al., 2022; Zhan et al., 2024). Images, as carriers of customer emotions, are core to customer decision-making (Li et al., 2023a; Liu et al., 2023a). Our research provides the following practical contributions:

Firstly, our analysis introduces new perspectives for hotel managers, encouraging them to explore sensory features in images to better understand customers' perceptions of hotels. This perspective helps managers to understand customers' emotions and expectations more accurately, which can lead to the provision of more tailored services in the hotel industry. Additionally, this method can reveal the specific sensory elements that customers value most in review images.

Secondly, our research cautions hotel managers that excessively enhancing the appeal of marketing images may result in substantial discrepancies with actual user reviews, which potentially breeds customer distrust. When crafting marketing strategies, ensuring consistency between marketing images and real customer experiences is crucial. This helps prevent misinformation and strengthens customers' trust in the hotel's authenticity.

Thirdly, we found that online review platforms could integrate image visual algorithms to detect sensory features in user-generated content. This technology allows platforms to analyze user-generated content more accurately, presenting more positive and informative content to customers. This suggestion aims to improve user-generated content quality and optimize the customer experience, better meeting sensory needs and encouraging positive participation and interaction.

Lastly, our study highlights that hotel managers need to pay attention to images containing negative sensory features, as these may reflect potential flaws in the hotel. By analyzing negative feedbacks in-depth, managers can gain a comprehensive understanding of customer experience issues. This enables them to make targeted improvements, respond promptly to customer needs, enhance service quality, and strengthen the brand image.

7.4 *Limitations and future research*

Several limitations apply to this research and should be addressed in future studies. Firstly, the data of our study only covers hotels in Xi'an, and the time span of the data includes the period of the Covid Pandemic, which may limit the generalizability of our findings. To increase the external validity, future research could expand the sample to include hotels from different geographical locations and diverse cultural backgrounds, and to test if our research findings are applicable during the post-pandemic period. Secondly, the perspective proposed in this study provides a sensory experience analysis framework for image data research. However, different sensory dimensions have various characteristics, and customers from different regions have distinct preferences. Future research could explore variations in sensory characteristics among customers with different cultural backgrounds and preferences. Thirdly, this study focused only on certain sensory features such as visual, auditory, tactile, and smell. In reality, human perception extends beyond these aspects. Future research could incorporate a more comprehensive set of sensory features to thoroughly investigate the impact of sensory experiences on customer perceptions.

REFERENCES

Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. *IEEE transactions on pattern analysis and machine intelligence*, 28(12), 2037-2041. https://doi.org/10.1109/TPAMI.2006.244

Ali, E. H. M., & Ahmed, M. O. (2019). Sensory marketing and its effect on hotel market-share: Perception of hotel customers. *Journal of Tourism and Hospitality Management*, *7*(1), 116-126. https://doi.org/10.15640/jthm.v7n1a12

Alpert, M. I., Alpert, J. I., & Maltz, E. N. (2005). Purchase occasion influence on the role of music in advertising. *Journal of Business Research*, *58*(3), 369-376. https://doi.org/10.1016/S0148-2963(03)00101-2

An, Q., & Ozturk, A. B. (2022). Assessing the effects of user-generated images on hotel guests' price, service quality, overall image perceptions and booking intention. *Journal of Hospitality and Tourism Technology*, *13*(4), 608-625. https://doi.org/10.1108/JHTT-05-2021-0146

An, Q., Ma, Y., Du, Q., Xiang, Z., & Fan, W. (2020). Role of user-generated images in online hotel reviews: An analytical approach. *Journal of Hospitality and Tourism Management*, *45*, 633-640. https://doi.org/10.1016/j.jhtm.2020.11.002

Arefieva, V., Egger, R., & Yu, J. (2021). A machine learning approach to cluster destination image on Instagram. *Tourism Management*, *85*, 104318. https://doi.org/10.1016/j.tourman.2021.104318

Back, R. M., Park, J. Y., Bufquin, D., Nutta, M. W., & Lee, S. J. (2020). Effects of hotel website photograph size and human images on perceived transportation and behavioral intentions. *International Journal of Hospitality Management*, *89*, 102545. https://doi.org/10.1016/j.ijhm.2020.102545

Barany, D. A., Lacey, S., Matthews, K. L., Nygaard, L. C., & Sathian, K. (2023). Neural basis of sound-symbolic pseudoword-shape correspondences. *Neuropsychologia*, *188*, 108657. https://doi.org/10.1016/j.neuropsychologia.2023.108657

Booth, D. A. (2014). Measuring sensory and marketing influences on customers' choices among food and beverage product brands. *Trends in food science & technology*, *35*(2), 129-137. https://doi.org/10.1016/j.tifs.2013.11.002

Bufquin, D., Park, J. Y., Back, R. M., Nutta, M. W., & Zhang, T. (2020). Effects of hotel website photographs and length of textual descriptions on viewers' emotions and behavioral intentions. *International Journal of Hospitality Management*, *87*, 102378. https://doi.org/10.1016/j.ijhm.2019.102378

Calvert, G., Spence, C., & Stein, B. E. (Eds.). (2004). *The handbook of multisensory processes*. MIT press.

Chen, S. X., Wang, X. K., Zhang, H. Y., Wang, J. Q., & Peng, J. J. (2021). Customer purchase forecasting for online tourism: A data-driven method with multiplex behavior data. *Tourism Management*, *87*, 104357. https://doi.org/10.1016/j.tourman.2021.104357

Chung, M., & Saini, R. (2022). Color darkness and hierarchy perceptions: How consumers associate darker colors with higher hierarchy. *Psychology & Marketing*, *39*(4), 820-837. https://doi.org/10.1002/mar.21623

Countryman, C. C., & Jang, S. (2006). The effects of atmospheric elements on customer impression: The case of hotel lobbies. *International Journal of Contemporary Hospitality Management*, *18*(7), 534-545. https://doi.org/10.1108/09596110610702968

Dai, F., Wang, D., & Kirillova, K. (2022). Travel inspiration in tourist decision making. *Tourism Management*, *90*, 104484. https://doi.org/10.1016/j.tourman.2021.104484

Deng, N., & Liu, J. (2021). Where did you take those images? Customers' preference clustering based on facial and background recognition. *Journal of Destination Marketing & Management*, *21*, 100632. https://doi.org/10.1016/j.jdmm.2021.100632

Elliot, A. J., & Maier, M. A. (2007). Color and psychological functioning. *Current directions in psychological science*, *16*(5), 250-254. https://doi.org/10.1111/j.1467-8721.2007.00514.x

Errajaa, K., Legohérel, P., Daucé, B., & Bilgihan, A. (2021). Scent marketing: Linking the scent congruence with brand image. *International Journal of Contemporary Hospitality Management*, *33*(2), 402-427. https://doi.org/10.1108/IJCHM-06-2020-0637

Gambetti, A., & Han, Q. (2022). Camera eats first: exploring food aesthetics portrayed on social media using deep learning. *International Journal of Contemporary Hospitality Management*, *34*(9), 3300-3331. https://doi.org/10.1108/IJCHM-09-2021-1206

Han, H., & Hyun, S. S. (2017). Impact of hotel-restaurant image and quality of physical-environment, service, and food on satisfaction and intention. *International Journal of Hospitality Management*, *63*, 82-92. https://doi.org/10.1016/j.ijhm.2017.03.006

He, Z., Deng, N., Li, X., & Gu, H. (2022). How to "read" a destination from images?

machine learning and network methods for DMOs' image projection and photo evaluation. *Journal of Travel Research*, *61*(3), 597-619. https://doi.org/10.1177/0047287521995134

Heung, V. C., & Gu, T. (2012). Influence of restaurant atmospherics on patron satisfaction and behavioral intentions. *International Journal of Hospitality Management*, *31*(4), 1167-1177. https://doi.org/10.1016/j.ijhm.2012.02.004

Hou, J. R., Zhang, J., & Zhang, K. (2023). Pictures that are worth a thousand donations: How emotions in project images drive the success of online charity fundraising campaigns? An image design perspective. *Management Information Systems Quarterly*, *47*(2), 535-584. https://doi.org/10.25300/MISQ/2022/17164

Hou, L., & Pan, X. (2023). Aesthetics of hotel images and its impact on customer engagement: A computer vision approach. *Tourism Management*, *94*, 104653. https://doi.org/10.1016/j.tourman.2022.104653

Hu, M., Lu, Y., Zhuang, M., Zhang, X., Zhang, H., Zhang, Y., ... & Liu, P. (2021). Development of tranquility perception scale: From customers' perspective. *Journal of Hospitality and Tourism Management*, *49*, 418-430. https://doi.org/10.1016/j.jhtm.2021.10.008

Kim, D., & Perdue, R. R. (2013). The effects of cognitive, affective, and sensory attributes on hotel choice. *International Journal of Hospitality Management*, *35*, 246-257. https://doi.org/10.1016/j.ijhm.2013.05.012

Kim, H., Hwang, Y., Gim, J., & Cheng, Y. (2024). When are vivid hotel photos effective? The moderating effects of resource scarcity and brand level. *International Journal of*

Hospitality Management, *116*, 103617. https://doi.org/10.1016/j.ijhm.2023.103617

Kim, M., Lee, S. M., Choi, S., & Kim, S. Y. (2021). Impact of visual information on online customer review behavior: Evidence from a hotel booking website. *Journal of Retailing and Customer Services*, *60*, 102494. https://doi.org/10.1016/j.jretconser.2021.102494

Kirillova, K., & Chan, J. (2018). "What is beautiful we book": Hotel visual appeal and expected service quality. *International Journal of Contemporary Hospitality Management*, *30*(3), 1788-1807. https://doi.org/10.1108/IJCHM-07-2017-0408

Kulkarni, P., & Kolli, H. (2022). Sensory Marketing Theory: How Sensorial Stimuli Influence Customer Behavior and Subconscious Decision-Making. *Journal of Student Research*, *11*(3). https://doi.org/10.13140/RG.2.2.12606.36165

Lacey, S., Jamal, Y., List, S. M., McCormick, K., Sathian, K., & Nygaard, L. C. (2020). Stimulus parameters underlying sound - symbolic mapping of auditory pseudowords to visual shapes. *Cognitive Science*, *44*(9), e12883. https://doi.org/10.1111/cogs.12883

Lee, M., Lee, S., & Koh, Y. (2019). Multisensory experience for enhancing hotel guest experience: Empirical evidence from big data analytics. *International Journal of Contemporary Hospitality Management*, *31*(11), 4313-4337. https://doi.org/10.1108/IJCHM-03-2018-0263

Li, C., Kwok, L., Xie, K. L., Liu, J., & Ye, Q. (2023). Let photos speak: The effect of user-generated visual content on hotel review helpfulness. *Journal of Hospitality & Tourism Research*, *47*(4), 665-690. https://doi.org/10.1177/10963480211019113

Li, H., Ji, H., Liu, H., Cai, D., & Gao, H. (2022). Is a picture worth a thousand words?

Understanding the role of review photo sentiment and text-photo sentiment disparity using deep learning algorithms. *Tourism Management*, *92*, 104559. https://doi.org/10.1016/j.tourman.2022.104559

Li, H., Liu, H., Shin, H. H., & Ji, H. (2024). Impacts of user-generated images in online reviews on customer engagement: A panel data analysis. *Tourism Management*, *101*, 104855. https://doi.org/10.1016/j.tourman.2023.104855

Li, H., Zhang, L., & Hsu, C. H. (2023). Research on user-generated images in tourism and hospitality: A systematic review and way forward. *Tourism Management*, *96*, 104714. https://doi.org/10.1016/j.tourman.2022.104714

Li, H., Zhang, L., Guo, R., Ji, H., & Yu, B. X. (2023). Information enhancement or hindrance? Unveiling the impacts of user-generated images in online reviews. *International Journal of Contemporary Hospitality Management*, *35*(7), 2322-2351. https://doi.org/10.1108/IJCHM-03-2022-0291

Li, X., Shi, M., & Wang, X. S. (2019). Video mining: Measuring visual information using automatic methods. *International Journal of Research in Marketing*, *36*(2), 216-231. https://doi.org/10.1016/j.ijresmar.2019.02.004

Li, Z., Fan, Y., Liu, W., & Wang, F. (2018). Image sentiment prediction based on textual descriptions with adjective noun pairs. *Multimedia Tools and Applications*, *77*, 1115-1132. https://doi.org/10.1007/s11042-016-4310-5

Lindstrom, M. (2006). Brand sense: How to build powerful brands through touch, taste, smell, sight and sound. *Strategic Direction*, *22*(2). https://doi.org/10.1108/sd.2006.05622bae.001

Liu, C. R., Wang, Y. C., Kuo, T. M., Chen, H., & Tsui, C. H. (2022). Memorable dining experiences with five senses: Conceptualization and scale development. *Journal of Hospitality and Tourism Management*, *53*, 198-207. https://doi.org/10.1016/j.jhtm.2022.11.001

Liu, W., Quijano, K., & Crawford, M. M. (2022). YOLOv5-Tassel: Detecting tassels in RGB UAV imagery with improved YOLOv5 based on transfer learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 8085-8094. https://doi.org/10.1109/JSTARS.2022.3206399

Liu, X., Nicolau, J. L., Law, R., & Li, C. (2023). Applying image recognition techniques to visual information mining in hospitality and tourism. *International Journal of Contemporary Hospitality Management*, *35*(6), 2005-2016. https://doi.org/10.1108/IJCHM-03-2022-0362

Liu, Z., Jiang, P., Wang, J., Du, Z., Niu, X., & Zhang, L. (2023). Hospitality order cancellation prediction from a profit-driven perspective. *International Journal of Contemporary Hospitality Management*, *35*(6), 2084-2112. https://doi.org/10.1108/IJCHM-06-2022-0737

Lv, X., Zhang, C., & Li, C. (2024). Beyond image attributes: A new approach to destination positioning based on sensory preference. *Tourism Management*, *100*, 104819. https://doi.org/10.1016/j.tourman.2023.104819

Ma, S., Li, H., Hu, M., Yang, H., & Gan, R. (2023). Tourism demand forecasting based on user-generated images on OTA platforms. *Current Issues in Tourism*, 1-20. https://doi.org/10.1080/13683500.2023.2216882

Ma, Y., Xiang, Z., Du, Q., & Fan, W. (2018). Effects of user-provided images on hotel review helpfulness: An analytical approach with deep leaning. *International Journal of Hospitality Management*, *71*, 120-131. https://doi.org/10.1016/j.ijhm.2017.12.008

Malhotra, N. K. (1984). Information and sensory overload. Information and sensory overload in psychology and marketing. *Psychology & Marketing*, *1*(3‐4), 9-21. https://doi.org/10.1002/mar.4220010304

Marder, B., Erz, A., Angell, R., & Plangger, K. (2021). The role of photograph aesthetics on online review sites: Effects of management-versus traveler-generated images on customers' decision making. *Journal of Travel Research*, *60*(1), 31-46. https://doi.org/10.1177/0047287519895125

Marks, L. E. (1975). On colored-hearing synesthesia: cross-modal translations of sensory dimensions. *Psychological bulletin*, *82*(3), 303. https://doi.org/10.1037/0033-2909.82.3.303

Martino, G., & Marks, L. E. (2000). Cross-modal interaction between vision and touch: the role of synesthetic correspondence. *Perception*, *29*(6), 745-754. https://doi.org/10.1068/p2984

McGrath, T. (2017). Social listening meets image sharing: A picture is worth 1,000 words. *Marketing Science Institute*.

O'Callaghan, G., O'Dowd, A., Simões-Franklin, C., Stapleton, J., & Newell, F. N. (2018). Tactile-to-visual cross-modal transfer of texture categorisation following training: an fMRI study. *Frontiers in Integrative Neuroscience*, *12*, 24. https://doi.org/10.3389/fnint.2018.00024

Öğüt, H., & Onur Taş, B. K. (2012). The influence of internet customer reviews on the online sales and prices in hotel industry. *The Service Industries Journal*, *32*(2), 197-214. https://doi.org/10.1080/02642069.2010.529436

Okumus, B., Koseoglu, M. A., & Ma, F. (2018). Food and gastronomy research in tourism and hospitality: A bibliometric analysis. *International Journal of Hospitality Management*, *73*, 64-74. https://doi.org/10.1016/j.ijhm.2018.01.020

Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, *17*(4), 460-469. https://doi.org/10.1177/002224378001700405

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *45*(3), 255. https://doi.org/10.1037/h0084295

Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online customer reviews. *Annals of Tourism Research*, *50*, 67–83. https://doi.org/10.1016/j.annals.2014.10.007

Parsa, A. B., Movahedi, A., Taghipour, H., Derrible, S., & Mohammadian, A. K. (2020). Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accident Analysis & Prevention*, *136*, 105405. https://doi.org/10.1016/j.aap.2019.105405

Petit, O., Velasco, C., & Spence, C. (2019). Digital sensory marketing: Integrating new technologies into multisensory online experience. *Journal of Interactive Marketing*, *45*(1), 42-61. https://doi.org/10.1016/j.intmar.2018.07.004

Phillips, P., Barnes, S., Zigan, K., & Schegg, R. (2017). Understanding the impact of online reviews on hotel performance: An empirical analysis. *Journal of Travel Research*, *56*(2), 235-249. https://doi.org/10.1177/0047287516636481

Qian, L., Guo, J., Qiu, H., Zheng, C., & Ren, L. (2023). Exploring destination image of dark tourism via analyzing user generated images: A deep learning approach. *Tourism Management Perspectives*, *48*, 101147. https://doi.org/10.1016/j.tmp.2023.101147

Ryu, K., Lee, H. R., & Kim, W. G. (2012). The influence of the quality of the physical environment, food, and service on restaurant image, customer perceived value, customer satisfaction, and behavioral intentions. *International journal of contemporary hospitality management*, *24*(2), 200-223. https://doi.org/10.1108/09596111211206141

Sann, C., & Streri, A. (2007). Perception of object shape and texture in human newborns: evidence from cross‐modal transfer tasks. *Developmental science*, *10*(3), 399-410. https://doi.org/10.1111/j.1467-7687.2007.00593.x

Schmidt, T. T., Ostwald, D., & Blankenburg, F. (2014). Imaging tactile imagery: changes in brain connectivity support perceptual grounding of mental images in primary sensory cortices. *Neuroimage*, *98*, 216-224. https://doi.org/10.1016/j.neuroimage.2014.05.014

Shahid, S., Paul, J., Gilal, F. G., & Ansari, S. (2022). The role of sensory marketing and brand experience in building emotional attachment and brand loyalty in luxury retail stores. *Psychology & Marketing*, *39*(7), 1398-1412. https://doi.org/10.1002/mar.21661

Slåtten, T., Mehmetoglu, M., Svensson, G., & Sværi, S. (2009). Atmospheric experiences that emotionally touch customers: a case study from a winter park. *Managing Service Quality: An International Journal*, *19*(6), 721-746.

https://doi.org/10.1108/09604520911005099

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature reviews neuroscience*, *9*(4), 255-266. https://doi.org/10.1038/nrn2331

Sun, H., Xu, H., Wu, J., Sun, S., & Wang, S. (2024). Let pictures speak: hotel selection-recommendation method with cognitive image attribute-enhanced knowledge graphs. *International Journal of Contemporary Hospitality Management*, *36*(12), 4296-4318. https://doi.org/10.1108/IJCHM-12-2023-1849

Theeuwes, J., van der Burg, E., Olivers, C. N., & Bronkhorst, A. (2007). Cross-modal interactions between sensory modalities: Implications for the design of multisensory displays. *Attention: From theory to practice*, 196-205.

Uthaisar, S., Eves, A., & Wang, X. L. (2024). Tourists' online information search behavior: Combined user-generated and marketer-generated content in restaurant decision making. *Journal of Travel Research*, *63*(6), 1549-1573. https://doi.org/10.1177/00472875231195314

Van Looy, A. (2021). A quantitative and qualitative study of the link between business process management and digital innovation. *Information & Management*, *58*(2), 103413. https://doi.org/10.1016/j.im.2020.103413

Wen, H., Leung, X., & Pongtornphurt, Y. (2020). Exploring the impact of background music on customers' perceptions of ethnic restaurants: The moderating role of dining companions. *Journal of Hospitality and Tourism Management*, *43*, 71-79. https://doi.org/10.1016/j.jhtm.2020.02.007

Xiang, K., Tong, Y., Wu, M. Y., & Gao, F. (2024). Performing multi-subject interactions in hotels during crises: Evidence from textual and visual data. *Tourism Management*, *101*, 104811. https://doi.org/10.1016/j.tourman.2023.104811

Xiao, X., Fang, C., Lin, H., & Chen, J. (2022). A framework for quantitative analysis and differentiated marketing of tourism destination image based on visual content of images. *Tourism Management*, *93*, 104585. https://doi.org/10.1016/j.tourman.2022.104585

Yang, Y., Wang, Y., & Zhao, J. (2023). Effect of user-generated image on review helpfulness: Perspectives from object detection. *Electronic Commerce Research and Applications*, *57*, 101232. https://doi.org/10.1016/j.elerap.2022.101232

Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management*, *28*(1), 180-182. https://doi.org/10.1016/j.ijhm.2008.06.011

Yim, D., Malefyt, T., & Khuntia, J. (2021). Is a picture worth a thousand views? Measuring the effects of travel images on user engagement using deep learning algorithms. *Electronic Markets*, 1-19. https://doi.org/10.1007/s12525-021-00472-5

Zhan, L., Cheng, M., & Zhu, J. (2024). Progress on image analytics: Implications for tourism and hospitality research. *Tourism Management*, *100*, 104798. https://doi.org/10.1016/j.tourman.2023.104798

Zhang, K., Chen, Y., & Li, C. (2019). Discovering the tourists' behaviors and perceptions in a tourism destination by analyzing photos' visual content with a computer deep learning model: The case of Beijing. *Tourism Management*, *75*, 595-608.

https://doi.org/10.1016/j.tourman.2019.07.002

Zhang, M., & Luo, L. (2023). Can customer-posted images serve as a leading indicator of restaurant survival? Evidence from Yelp. *Management Science*, *69*(1), 25-50. https://doi.org/10.1287/mnsc.2022.4359

Zhang, S., Lee, D., Singh, P.V. and Srinivasan, K. (2022), "What makes a good image? Airbnb demand analytics leveraging interpretable image features", *Management Science*, *68*(8), 5644-5666. https://doi.org/10.1287/mnsc.2021.4175

Zhou, M., Chen, G. H., Ferreira, P., & Smith, M. D. (2021). Customer behavior in the online classroom: Using video analytics and machine learning to understand the consumption of video courseware. *Journal of Marketing Research*, *58*(6), 1079-1100. https://doi.org/10.1177/00222437211042013

Zhu, J., & Cheng, M. (2024). Automatic videos analytics in tourism: A methodological review. *Annals of Tourism Research*, *108*, 103800. https://doi.org/10.1016/j.annals.2024.103800

Zhu, J., Cheng, M., & Wang, Y. (2024). Viewer In-Consumption Engagement in Pro-Environmental Tourism Videos: A Video Analytics Approach. *Journal of Travel Research*, 00472875231219634. https://doi.org/10.1177/00472875231219634

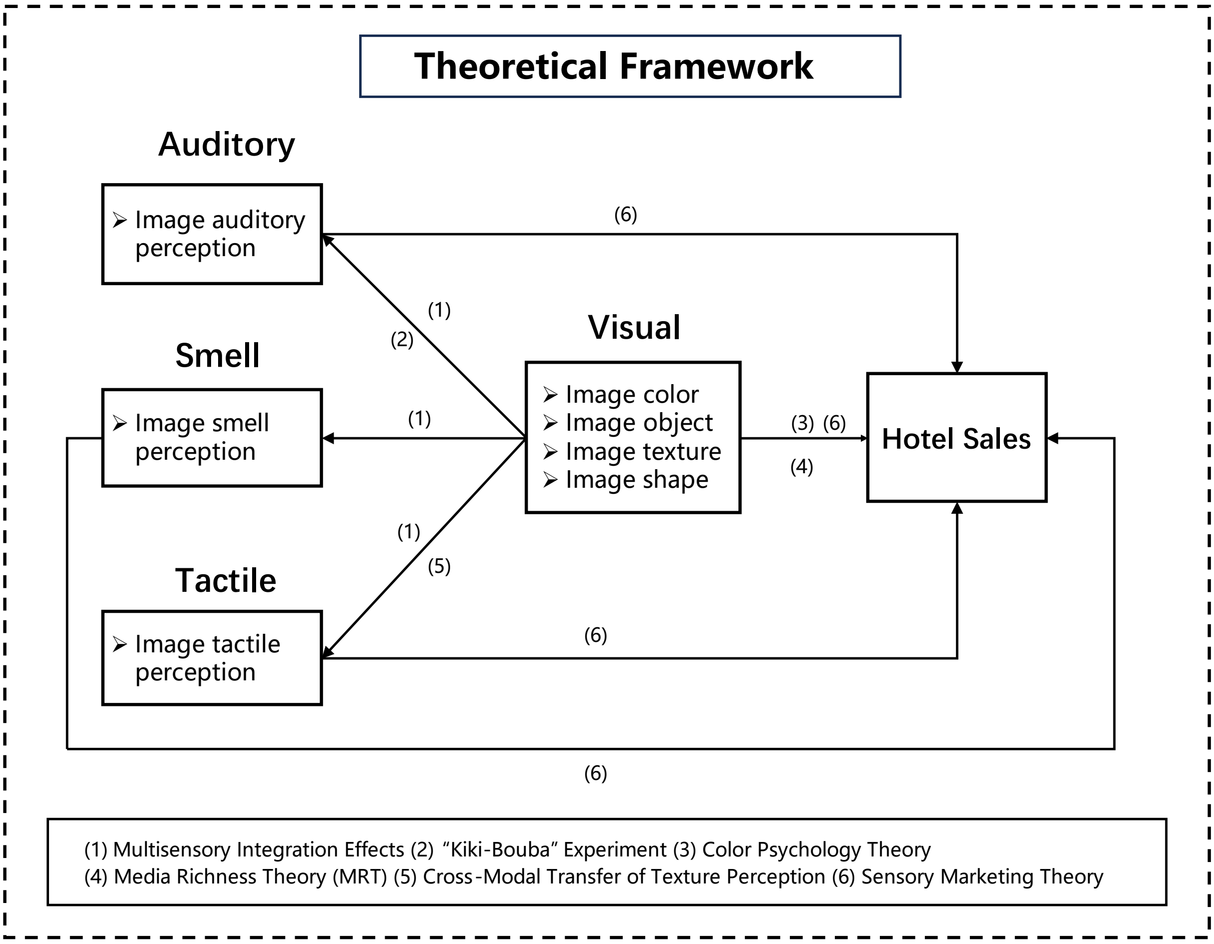


Figure 1 Theoretical framework of this research. (*Source*. Authors.)

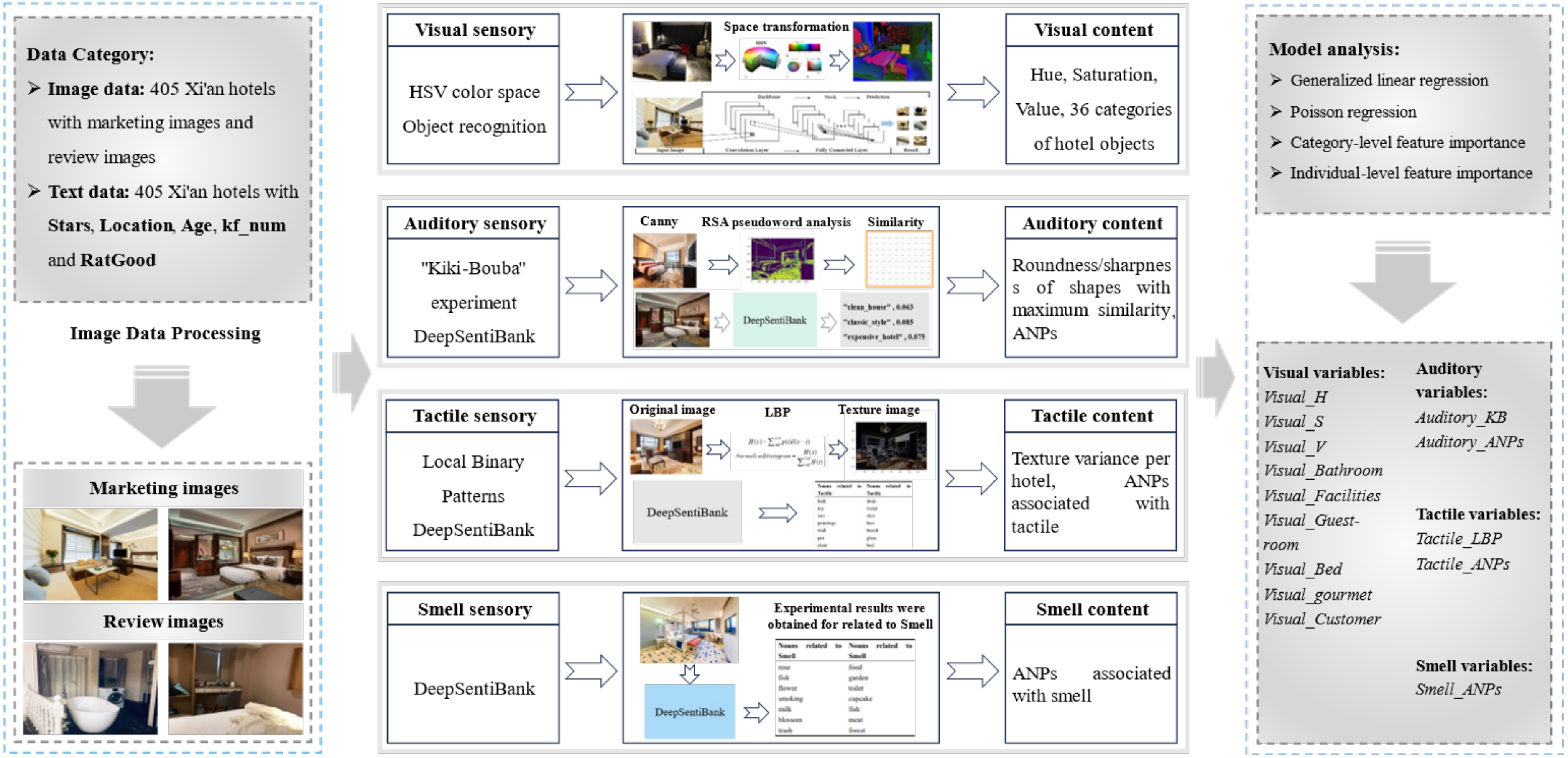


Figure 2 Methodological framework for extraction of sensory variables. (*Source*. Authors.)

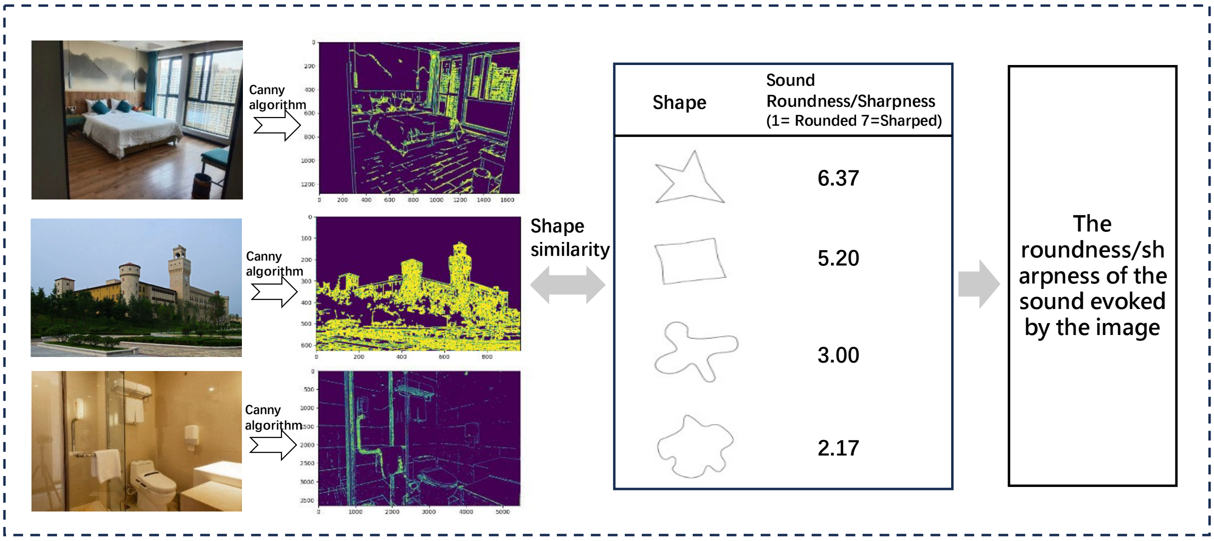
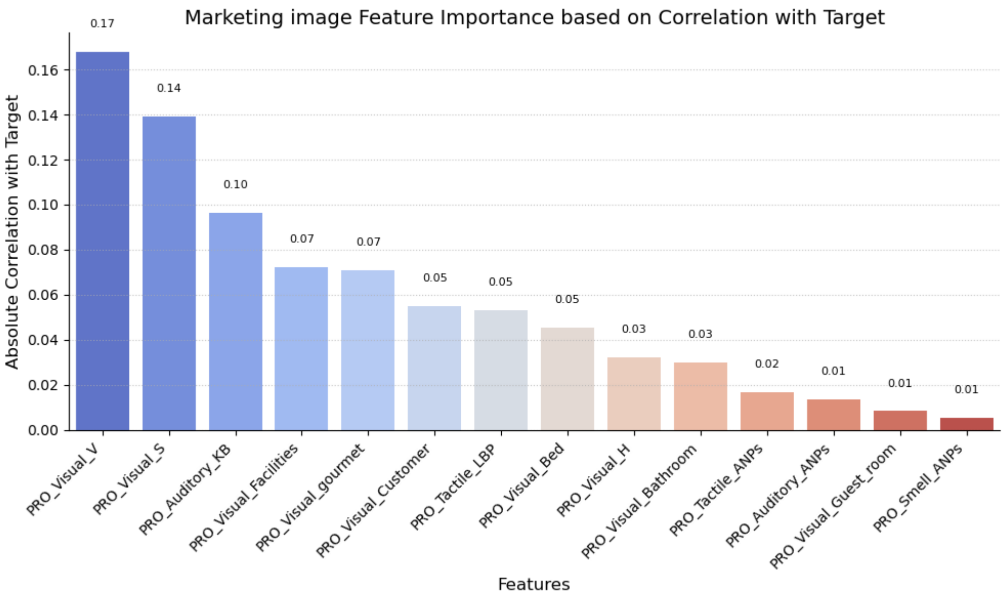
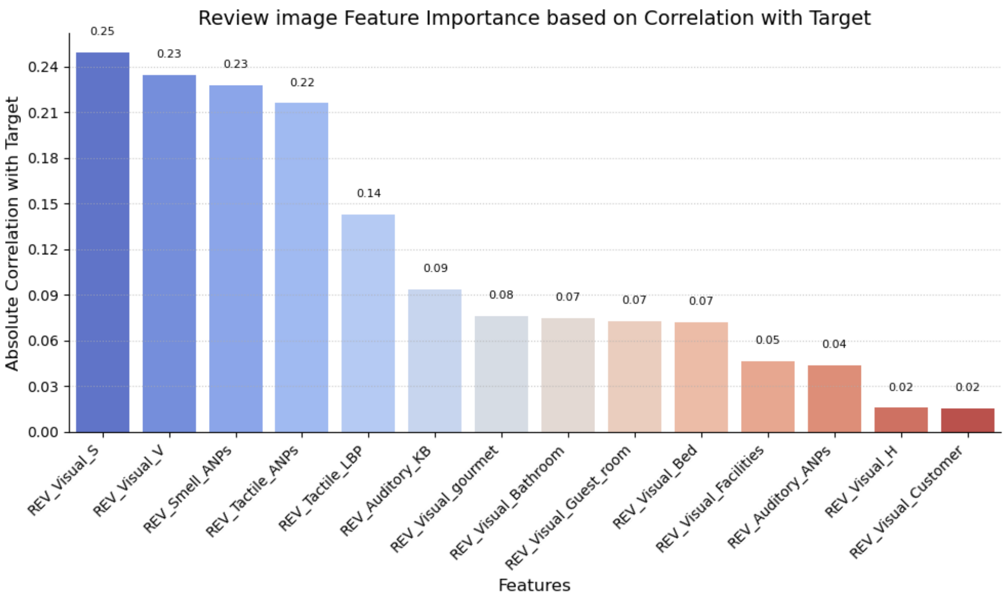


Figure 3 Methodological framework for extracting roundness and sharpness in hotel images. (*Source*. Authors.)



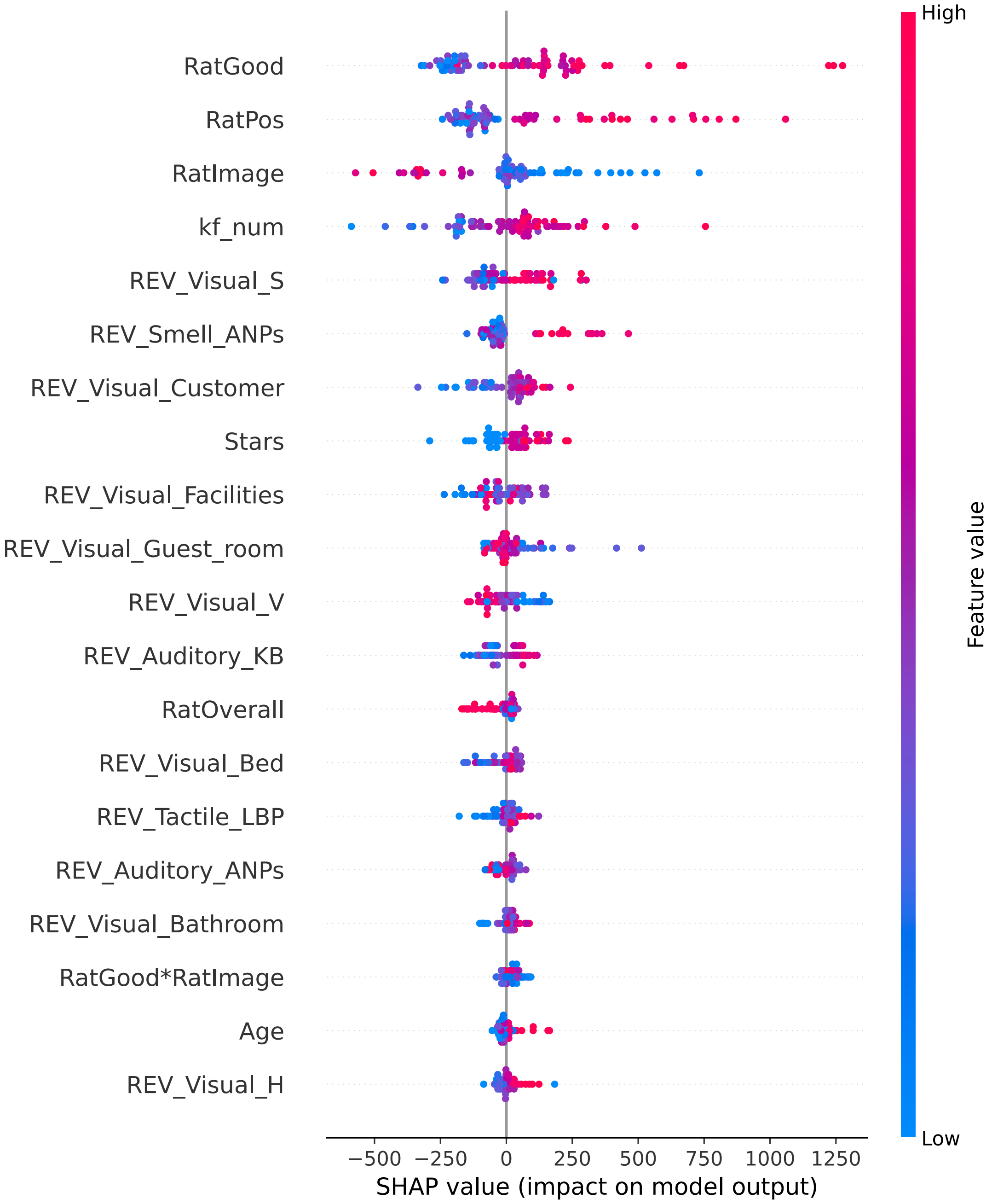
(a)Marketing images



(b)Review images

Figure 4 Results of the importance of unidimensional category-level sensory features based on the correlation analysis algorithm. (*Source*. Authors.)

(a) Marketing images



(b) Review images

Figure 5 Analysis of non-linear relationships based on shape values. (*Source*. Authors.)

Table 1 Recent research in visual data: types, methods, and contributions.

Authors	Title	Method	Data	Key empirical contribution
Hou and Pan (2023)	Aesthetics of hotel photos and its impact on consumer engagement: A computer vision approach.	Deep convolutional neural network model	Thumbnail images and managerial photos of hotels from TripAdvisor	The study extends the analysis of tourism and hospitality marketing photos by focusing on perceptive features, bridging aesthetic research from offline to online contexts.
(Kim et al., 2021)	Impact of visual information on online consumer review behavior: Evidence from a hotel booking website.	Empirical analysis; hypothesis testing	Visual and textual information from Agoda.com	The study examines the impact of visual and textual information on consumers' online review behaviors, using empirical data from the hotel booking website Agoda.com to analyze these effects.
(Li et al., 2023a)	Let photos speak: The effect of user-generated visual content on hotel review helpfulness.	Econometric analyses; image-processing techniques	Image and text data based on 12138 hotels from Qunar.com	The study addresses the need to better understand the impact of visual content on online reviews and demonstrates the analysis of publicly accessible internet data through an interdisciplinary approach that integrates econometric models with deep-learning techniques.
(Li et al., 2022)	Is a picture worth a thousand words? Understanding the role of review photo sentiment and text-photo sentiment disparity using deep learning algorithms.	Deep learning; Econometrics	Yelp's Las Vegas restaurant data (2005–2021) includes 401,269 reviews and 336,092 photos, with 105,237 reviews featuring photos.	Images have become integral to consumers' sharing of consumption experiences due to their abilities of carrying rich and vivid information. This study investigates the impacts of restaurant review photo sentiment on customers' perceived review usefulness and enjoyment using deep learning and econometric model analysis. This study contributes to the electronic word-of-mouth literature as well as to the application of machine learning technologies in computer vision to tourism and hospitality research.
(Zhan et al., 2024)	Progress on image analytics: Implications for tourism and hospitality research.	Overview and analysis	27 articles in tourism and hospitality using images as the data	The study advances the definitions, features, and theories related to images, presenting a methodological framework for image-related studies that complements the dominant textual analysis in tourism and hospitality research.

			sources	
Deng and Liu (2021)	Where did you take those photos? Tourists’ preference clustering based on facial and background recognition.	DeepSenti-Bank; deep learning	The dataset consists of 14,886 photos from Beijing in Instagram	The study introduces a novel approach to analyzing tourists' travel patterns and preferences using facial and photo content recognition techniques, filtering photos with tourists' faces and classifying different groups by age and gender.
Zhu and Cheng (2024)	Automatic videos analytics in tourism: A methodological review.	Review Analysis; Theoretical Methods	Videos data	This study provides a critical review of the progress of automatic video analytics in tourism and hospitality and a guiding framework by detailing theoretical and methodological issues with this new form of knowledge production. The research offers a blueprint for future tourism research endeavors tapping into the potential of videos as a data source.

Table 2 Summary of image sensory research: visual, auditory, tactile and smell aspects.

Feature Type	Feature	Author	Research method	Main finding
Visual content	Visual design	Kirillova and Chan (2018)	Survey and empirical research	The appearance and interior design of a hotel are directly linked to tourists' first impressions and overall experiences. Hotels with high aesthetic value are more likely to be booked and offer more reliable and assured service quality.
	Colors	Countryman and Jang (2006)	Structural equation modeling	Effective color matching can guide tourists' emotional experiences and is essential for shaping the hotel's overall atmosphere and personality, thereby boosting sales.
	Object	(He et al., 2022)	Machine learning algorithm	Objects in images can capture tourists' attention. By quantitatively analyzing the frequency of object appearances in online reviews, it is possible to identify which objects are likely to positively impact sales.
Auditory information	Music	(Wen et al., 2020)	PLS-SEM	Music can evoke emotional responses in tourists, and appropriate music creates a pleasant atmosphere, positively influencing customer emotions.
	Ambience	Heung and Gu (2012)	Factor analysis and multiple regression analysis	The atmosphere of a hotel directly impacts customer satisfaction. A quiet, pleasant environment not only enhances the customer experience but also increases their likelihood of returning.
Tactile perception	Texture	(Liu et al., 2022a)	—	Texture plays a crucial role in tactile perception and significantly influences consumer purchasing decisions.
Smell perception	Smell	(Errajaa et al., 2021)	A factorial design	Using fragrances that align with the brand image can enhance guests' positive responses to the environment, thereby improving satisfaction, increasing the willingness to revisit, and enhancing the perception of products and services.
	Food	Han and Hyun (2017)	—	The quality of food is directly tied to customer satisfaction. Delicious flavors, fresh ingredients, and a variety of dishes can

significantly boost satisfaction. Additionally, the presentation and plating of food have a notable impact on consumers' sensory experiences.

Table 3 Hotel object type detection results based on YOLOv5.

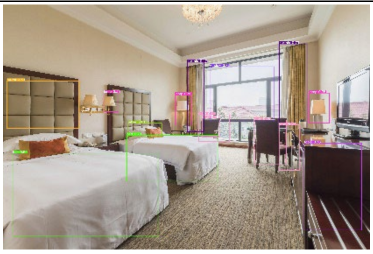
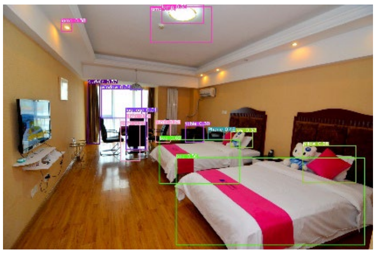
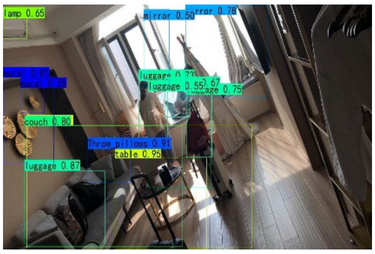
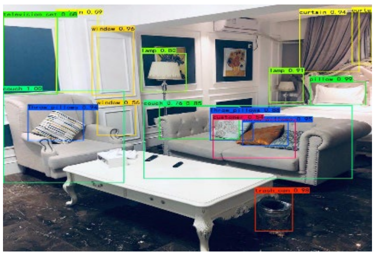
<table>
<tr><th colspan="3">Marketing Image</th></tr>
<tr>
<td>
Type of test:
Mural; Bed; Lamp; Pillow; Curtain; Window; Chair; Table</td>
<td>
Type of test:
Lamp; Curtain; Luggage; Chair; Table; Bed; Phone; Pillow</td>
<td>
Type of test:
Curtain; Window; Table; Chair; Bed</td>
</tr>
<tr><th colspan="3">Review Image</th></tr>
<tr>
<td>
Type of test:
Lamp; Mirror; Luggage; Couch; Throw_pillows; Table</td>
<td>
Type of test:
Couch; Throw_pillows; Window; Lamp; Television-set</td>
<td>
Type of test:
Toilet-paper; Towel; Water-cup, eta</td>
</tr>
</table>

Table 4 Object theme classification results based on visual sensory features.

Theme	Types included
Bathroom	Washstand, Towel, Bathtowel, Bathroom, Toilet-paper, Bathtub, Toiletries, Toilet, Bathrobes, Trash.
Facilities	Television-set, Phone, Washer, Air-conditioning, Treadmill.
Room and Suite	Trashcan, Window, Curtain, Table, Chair, Lamp, Couch, Book, Mirror, Mural, Shoe.
Bed	Bed, Pillow, Throw-pillows.
Food and Water	Food, Water, Tea-set, Kettle, Water-cup.
Tourists	Customer, Luggage.

Table 5 Description of visual sensory factors based on color features and object features.

Variable	Definition	Mean	Min	Max
All visual sensory features (with review images)				
Visual_H	Average hue of review images per hotel.	57.9831	36.8292	92.2824
Visual_S	Average saturation of review images per hotel.	78.8285	46.6822	109.9921
Visual_V	Average value of review images per hotel.	126.3075	98.5742	151.2335
Visual_Bathroom	Percentage of objects included in the bathroom theme of review images in each hotel.	0.0871	0.0000	0.5000
Visual_Facilities	Percentage of objects included in the facilities theme of review images in each hotel.	0.0279	0.0000	0.1183
Visual_Guest-room	Percentage of objects included in the guest-room theme of review images in each hotel.	0.5995	0.3412	1.0002
Visual_Bed	Percentage of objects included in the bed theme of review images in each hotel.	0.1492	0.0000	0.4667
Visual_gourmet	Percentage of objects included in the gourmet theme of review images in each hotel.	0.0701	0.0000	0.5003
Visual_Customer	Percentage of objects included in the customer theme of review images in each hotel.	0.0659	0.0000	0.3750
All visual sensory features (with marketing images)				
Visual_H	Average hue of marketing images in each hotel.	72.5853	25.3254	123.7091
Visual_S	Average saturation of marketing images in each hotel.	87.5686	43.6392	151.4201
Visual_V	Average value of marketing images in each hotel.	152.7487	102.4123	213.3485
Visual_Bathroom	Percentage of objects included in the bathroom theme of marketing images in each hotel.	0.00352	0.0000	0.1544
Visual_Facilities	Percentage of objects included in the facilities theme of marketing images in each hotel.	0.0177	0.0000	0.1433
Visual_Guest-room	Percentage of objects included in the guest-room theme of marketing images in each hotel.	0.7514	0.3333	1.0006
Visual_Bed	Percentage of objects included in	0.1892	0.0000	0.5714

	the bed theme of marketing images in each hotel.			
Visual_gourmet	Percentage of objects included in the gourmet theme of marketing images in each hotel.	0.0334	0.0000	0.2784
Visual_Customer	Percentage of objects included in the customer theme of marketing images in each hotel.	0.0055	0.0000	0.1003

Table 6 Auditory perception classification based on ANPs.

Nouns related to Auditory	Nouns related to Auditory
Birds	Alarm
Guitar	Cars
Piano	Street
Ears	Laughter
Noise	Accident
Concert	Cinema
Interview	Combat
Scream	

Table 7 Description of Auditory sensory factors based on Kiki-Bouba experiment and ANPs.

Variable	**Definition**	**Mean**	**Min**	**Max**
All auditory sensory features (with review images)				
Auditory_KB	Average roundness/sharpness level of the shape of review images in each hotel.	3.6853	2.5888	5.0410
Auditory_ANPs	Percentage of auditory ANPs of review images in each hotel.	0.2652	0.1611	0.4234
All auditory sensory features (with marketing images)				
Auditory_KB	Average roundness/sharpness level of the shape of marketing images in each hotel.	3.4116	2.3242	5.4476
Auditory_ANPs	Percentage of auditory ANPs of marketing images in each hotel.	0.3561	0.1156	0.5901

Table 8 Tactile and smell perception classification based on ANPs.

Nouns related to Tactile	**Nouns related to Tactile**	**Nouns related to Smell**	**Nouns related to Smell**
Bath	Desk	Rose	Food
Toy	Water	Fish	Garden
Sun	Skin	Flower	Toilet
Paintings	Hair	Smoking	Cupcake
Wall	Beach	Milk	Fish
Pet	Glass	Blossom	Meat
Chair	Bed	Trash	Forest

Table 9 Description of tactile and smell sensory factors based on LBP and ANPs.

Variable	Definition	Mean	Min	Max
All tactile and smell sensory features (with review images)				
Tactile_LBP	Average tactile LBP score of review images in each hotel.	0.0006	0.0001	0.0015
Tactile_ANPs	Percentage of tactile ANPs score of review images in each hotel.	0.4485	0.1692	0.7031
Smell_ANPs	Percentage of smell ANPs score of review images in each hotel.	0.1958	0.0000	0.4848
All tactile and smell sensory features (with marketing images)				
Tactile_LBP	Average tactile LBP score of marketing images in each hotel.	0.0008	0.0003	0.0016
Tactile_ANPs	Percentage of tactile ANPs score of marketing images in each hotel.	0.5122	0.3324	0.7021
Smell_ANPs	Percentage of smell ANPs score of marketing images in each hotel.	0.2232	0.0638	0.4185

Table 10 Description of differences in sensory features between marketing images and review images.

Variable	Definition	Mean	Min	Max
DIS_Visual_H	the absolute value of the difference between *PRO_Visual_H* and *REV_Visual_H*.	0.2703	0.0039	0.9438
DIS_Visual_S	the absolute value of the difference between *PRO_Visual_S* and *REV_Visual_S*.	0.1994	4.5861	0.7965
DIS_Visual_V	the absolute value of the difference between *PRO_Visual_V* and *REV_Visual_V*.	0.1932	0.0036	0.5184
DIS_Visual_Bathroom	the absolute value of the difference between *PRO_Visual_Bathroom* and *REV_Visual_Bathroom*.	0.0849	0.0000	0.5000
DIS_Visual_Facilities	the absolute value of the difference between *PRO_Visual_Facilities* and *REV_Visual_Facilities*.	0.0283	0.0000	0.1010
DIS_Visual_Guest-room	the absolute value of the difference between *PRO_Visual_Facilities* and *REV_Visual_Facilities*.	0.1652	0.0000	0.5167
DIS_Visual_Bed	the absolute value of the difference between *PRO_Visual_Guest-room* and *REV_Visual_Guest-room*.	0.0900	0.0018	0.4180
DIS_Visual_gourmet	the absolute value of the difference between *PRO_Visual_gourmet* and *REV_Visual_gourmet*.	0.0585	0.0000	0.5000
DIS_Visual_Customer	the absolute value of the difference between *PRO_Visual_Customer* and *REV_Visual_Customer*.	0.0620	0.0000	0.3750
DIS_Auditory_KB	the absolute value of the difference between *PRO_Auditory_KB* and *REV_Auditory_KB*.	0.5822	0.0068	2.0339
DIS_Auditory_ANPs	the absolute value of the difference between *PRO_Auditory_ANPs* and *REV_Auditory_ANPs*.	0.1029	0.0005	0.3183
DIS_Tactile_LBP	the absolute value of the difference between *PRO_Tactile_LBP* and *REV_Tactile_LBP*.	0.0002	3.4805	0.0012
DIS_Tactile_ANPs	the absolute value of the difference between *PRO_Tactile_ANPs* and *REV_Tactile_ANPs*.	0.0923	0.0008	0.3564
DIS_Smell_ANPs	the absolute value of the difference between *PRO_Smell_ANPs* and *REV_Smell_ANPs*.	0.0739	2.2822	0.2869

Table 11 The effect of sensory features of marketing images on hotel sales.

	Hotel_sales	
	Model 1	**Model 2**
Constant	0.9851***	3.1828***
	(0.061)	(0.073)
Stars	-0.0485***	-0.0325***
	(0.002)	(0.002)
Location	0.0119***	0.0345***
	(0.003)	(0.003)
RatOverall	-0.3602***	-0.2781***
	(0.008)	(0.009)
ReadPre	-0.0210***	-0.0215***
	(0.002)	(0.002)
RatPos	4.6470***	4.7176***
	(0.034)	(0.035)
Age	0.0054***	0.0059***
	(0.000)	(0.000)
Log kf_num	0.3553***	0.3775***
	(0.003)	(0.003)
RatGood	3.6232***	3.9602***
	(0.051)	(0.052)
RatImage	-3.6973***	-1.2164***
	(0.375)	(0.381)
RatGood*RatImage	1.3990***	-1.5218***
	(0.410)	(0.418)
Log PRO_Visual_H		-0.2930***
		(0.006)
Log PRO_Visual_S		-0.0191***
		(0.007)
Log PRO_Visual_V		-0.5249***
		(0.013)
PRO_Visual_Bathroom		1.5604***
		(0.085)
PRO_Visual_Facilities		1.6473***
		(0.050)
PRO_Visual_Guest_room		-0.3404***
		(0.024)
PRO_Visual_Bed		0.1219***
		(0.027)
PRO_Visual_gourmet		0.6128***
		(0.036)
PRO_Visual_Customer		-0.4193***
		(0.077)
PRO_Auditory_KB		-0.0551***

		(0.003)
PRO_Auditory_ANPs		0.4912***
		(0.028)
PRO_Tactile_LBP		303.2766***
		(7.181)
PRO_Tactile_ANPs		1.3620***
		(0.027)
PRO_Smell_ANPs		1.3296***
		(0.026)
Log-Likelihood	-92075	-85918

Notes: *, **, *** indicates 10%, 5%, and 1% significance level.

Table 12 The effect of sensory features of review images on hotel sales.

	Hotel_sales	
	Model 3	**Model 4**
Constant	0.9851***	3.0192***
	(0.061)	(0.104)
Stars	-0.0485***	-0.0475***
	(0.002)	(0.002)
Location	0.0119***	0.0190***
	(0.003)	(0.003)
RatOverall	-0.3602***	-0.2503***
	(0.008)	(0.009)
ReadPre	-0.0210***	-0.0162***
	(0.002)	(0.002)
RatPos	4.6470***	2.8288***
	(0.034)	(0.027)
Age	0.0054***	0.0050***
	(0.000)	(0.000)
Log kf_num	0.3553***	0.3608***
	(0.003)	(0.003)
RatGood	3.6232***	1.7945***
	(0.051)	(0.047)
RatImage	-3.6973***	-3.2086***
	(0.375)	(0.385)
RatGood*RatImage	1.3990***	1.5829***
	(0.410)	(0.421)
Log REV_Visual_H		0.3621***
		(0.011)
Log REV_Visual_S		0.6908***
		(0.012)
Log REV_Visual_V		-1.1136***
		(0.023)
REV_Visual_Bathroom		-0.0479
		(0.034)
REV_Visual_Facilities		0.4886***
		(0.068)
REV_Visual_Guest_room		-0.2568***
		(0.023)
REV_Visual_Bed		0.7891***
		(0.027)
REV_Visual_gourmet		0.6937***
		(0.028)
REV_Visual_Customer		1.3526***
		(0.039)
REV_Auditory_KB		0.2364***

		(0.006)
REV_Auditory_ANPs		0.3070***
		(0.042)
REV_Tactile_LBP		-54.4237***
		(12.656)
REV_Tactile_ANPs		0.8931***
		(0.044)
REV_Smell_ANPs		1.8191***
		(0.044)
Log-Likelihood	-92075	-87695

Note: *, **, *** indicates 10%, 5%, and 1% significance level.

Table 13 The effect of sensory differences in marketing and review images on hotel sales.

	Hotel_sales	
	Model 5	**Model 6**
Constant	0.9851***	4.2972***
	(0.061)	(0.048)
Stars	-0.0485***	-0.0045*
	(0.002)	(0.002)
Location	0.0119***	0.0207***
	(0.003)	(0.003)
RatOverall	-0.3602***	-0.2673***
	(0.008)	(0.009)
ReadPre	-0.0210***	-0.0189***
	(0.002)	(0.002)
RatPos	4.6470***	2.9741***
	(0.034)	(0.027)
Age	0.0054***	0.0056***
	(0.000)	(0.000)
Log kf_num	0.3553***	0.3275***
	(0.003)	(0.003)
RatGood	3.6232***	1.7355***
	(0.051)	(0.047)
RatImage	-3.6973***	-1.8123***
	(0.375)	(0.382)
RatGood*RatImage	1.3990***	-0.4077
	(0.410)	(0.419)
DIS_Visual_H		0.1845***
		(0.007)
DIS_Visual_S		-0.4339***
		(0.010)
DIS_Visual_V		0.0283*
		(0.014)
DIS_Visual_Bathroom		-0.5378***
		(0.029)
DIS_Visual_Facilities		1.5792***
		(0.068)
DIS_Visual_Guest_room		-0.2997***
		(0.014)
DIS_Visual_Bed		0.2937***
		(0.020)
DIS_Visual_gourmet		0.1752***
		(0.025)
DIS_Visual_Customer		1.1636***
		(0.034)
DIS_Auditory_KB		0.0340***

		(0.004)
DIS_Auditory_ANPs		-1.0180***
		(0.022)
DIS_Tactile_LBP		-188.0747***
		(7.930)
DIS_Tactile_ANPs		-0.6718***
		(0.022)
DIS_Smell_ANPs		-0.3921***
		(0.025)
Log-Likelihood	-92075	-88836

Note: *, **, *** indicates 10%, 5%, and 1% significance level.

Table 14 Sensory category-level feature importance results based on feature ranking strategies.

Marketing images					
Type	PRO_Visual_HSV	PRO_Visual_YOLO	PRO_Auditory	PRO_Tactile	PRO_Smell
Importance	0.01916656	0.01734726	0.00396311	0.01437375	5.91523e-16
Rank	Rank1	Rank2	Rank4	Rank3	Rank5
Review images					
Type	REV_Visual_HSV	REV_Visual_YOLO	REV_Auditory	REV_Tactile	REV_Smell
Importance	0.05524578	0.01754223	0.00835983	0.00043638	4.58423e-16
Rank	Rank1	Rank2	Rank3	Rank4	Rank5

Supplemental Material

Appendix A

1. YOLOv5 algorithm principle

YOLOv5 (You Only Look Once version 5) is a deep learning-based object detection algorithm, representing one of the latest iterations in the YOLO series. Known for its rapid detection speed and high accuracy, YOLOv5 is extensively used in real-time object detection tasks.

To explain the working principle of the YOLOv5 model, this paper will address the following two aspects:

（1）YOLOv5 Model Architecture

Backbone: The primary network responsible for extracting image features. YOLOv5 typically employs CSPDarknet53 as its backbone, a Convolutional Neural Network (CNN) structure designed to efficiently extract image features.

Neck: This section utilizes the Feature Pyramid Network (FPN) and Path Aggregation Network (PAN) to fuse feature maps from various levels, thereby improving the detection of small objects.

Head: The final section of the network responsible for predictions. It comprises several convolutional layers that generate bounding boxes, class probabilities, and object confidence scores.

（2）YOLOv5 target detection process

Pre-processing: The input image is resized to a fixed dimension and normalized.

Feature extraction: The input image passes through the Backbone network, extracting feature maps at multiple scales. These maps contain varying levels of information, aiding in the detection of targets of different sizes.

Feature fusion: Feature maps at different levels are fused through the FPN and PAN structures in the Neck to enhance the model's ability to detect targets at various scales.

Prediction: The fused feature maps pass through the Head section to generate bounding boxes, category probabilities, and object confidence. Each bounding box is associated with a predicted category and corresponding probability.

Non-Maximum Suppression (NMS): To handle multiple bounding boxes predicted for the same target, NMS filters the final detection results by removing redundant boxes with significant overlap, retaining only the best one.

2. Training and detection process based on the data of this study

In deep learning, the benefits of transfer learning have made it a preferred method for object detection. However, in specific scenarios, pre-trained YOLOv5 models may pose challenges to the reliability of analysis. To ensure reliable detection results, this study employed a custom dataset for training. Furthermore, to assess visual element standardization, we meticulously designed the labeling method, as illustrated in Table 4.

To standardize and analyze hotel visual elements, we defined specific visual element types and enlisted Ph.D. students with relevant expertise to refine the classifications. This process resulted in six themes: Bathroom, Facilities, Room and Suite, Bed, Food and Water, and Tourists. These themes served as primary categories with corresponding object counts: 10, 5, 11, 3, 5, and 2, respectively.

Consequently, we used these six themes and 36 object types as dataset labels. To ensure dataset reliability, we utilized LabelMe to annotate both Marketing and Review Images. Ultimately, a combined dataset of Marketing and Review Images was constructed for training, as shown in Figure B1. Figure B2 illustrates the detection performance of the custom-trained YOLOv5 model on our dataset. The trained YOLOv5 model demonstrates excellent accuracy and effectiveness, yielding reliable detection results for extracting sensory features.

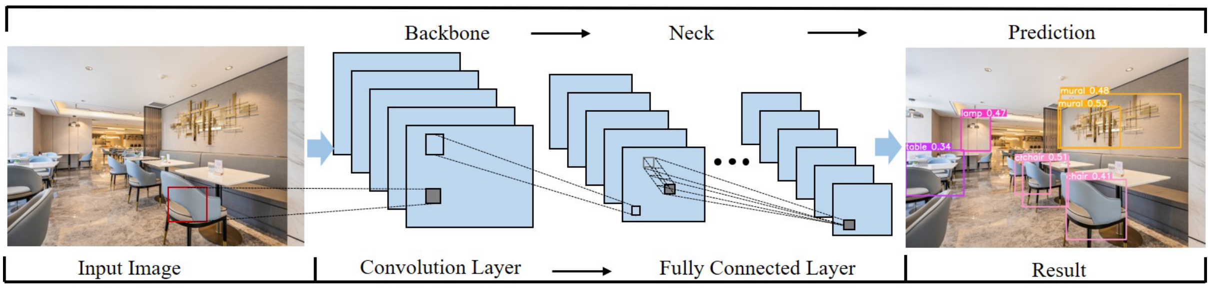


Figure A1. Yolov5 detection framework.

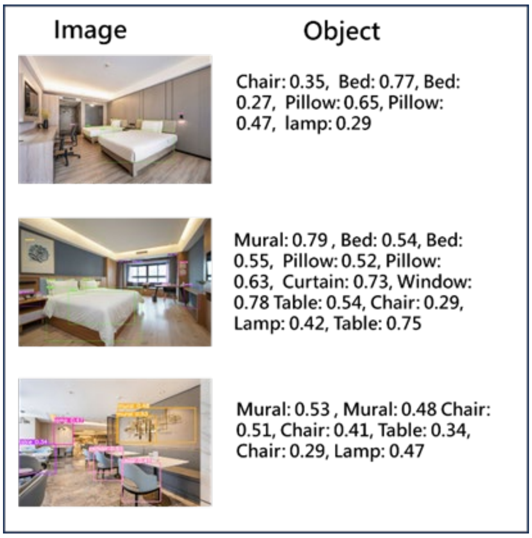


Figure A2. Yolov5 test results.

Appendix B

To ensure precise categorization, we enlisted 10 experts in travel and hotel studies to classify 36 object classes. The experts worked independently, without discussion, leading to varied individual categorizations. By comparing these categorizations, we identified the most recurrent themes.

After the experts completed their classifications, we organized a focused discussion session. Each expert presented their results and explained the rationale behind their categorizations. This step enabled the experts to understand each other's perspectives and the basis for their classifications.

Following the presentations, we conducted a comparative analysis to identify the most frequently occurring themes. This analysis revealed the commonalities and divergences in the experts' classifications. The identified themes were then collectively discussed, allowing for a comprehensive review of the results.

We employed the following methods to assess the reliability of the classification results:

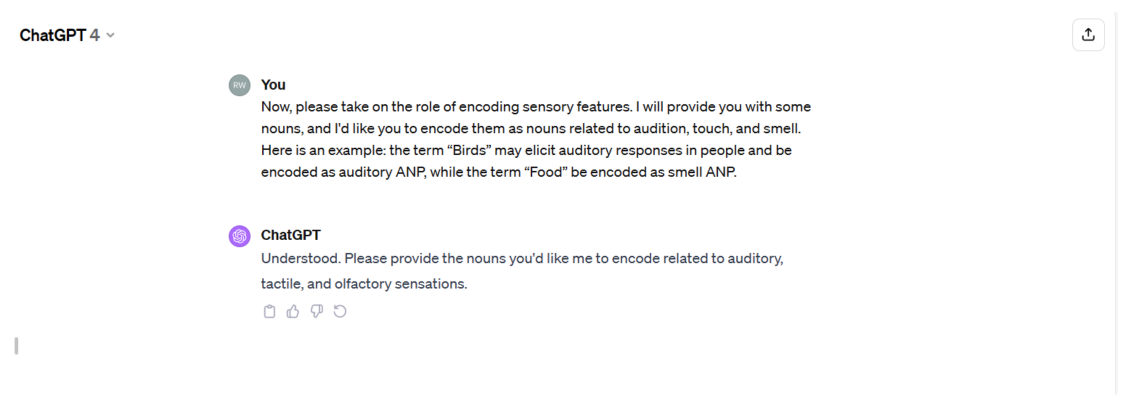
(1) Degree of Consensus Calculation: We calculated the frequency of each object's appearance in various themes and, following the majority rule, assigned the object to the theme where it appeared most frequently. This method quantified the level of agreement among the experts. A higher consensus indicated a stronger agreement on classification standards, thereby enhancing the reliability of the results.

(2) Discussion and Comparative Analysis: Through focused discussions and detailed comparative analysis, we addressed and eliminated individual discrepancies in the classification results. This process involved an in-depth discussion of the differences in the experts' classification results, identifying the causes of these discrepancies, and reaching a consensus through collective decision-making. This step was crucial for optimizing the classification standards and ensuring the accuracy of the classification results.

(3) Final Review and Agreement: After a thorough review of the statistical results, all participants jointly evaluated the classification results. The experts discussed any remaining issues and ultimately reached a unanimous agreement, confirming the reliability of the classification results. This process ensured that the classification results underwent multiple rounds of validation and adjustments, achieving a high level of accuracy and consistency.

Appendix C

Table C1 Chatgpt prompt.

Chatgpt prompt
Prompt: Now, please take on the role of encoding sensory features. I will provide you with some nouns, and I'd like you to encode them as nouns related to audition, touch, and smell. Here is an example: the term “Birds” may elicit auditory responses in people and be encoded as auditory ANP, while the term “Food” be encoded as smell ANP.


Appendix D

Using the LBP algorithm, we compared each neighboring pixel to the central pixel. If a neighboring pixel's value was greater than that of the central pixel, it was set to 1; otherwise, it was set to 0. This process produced LBP images from the hotel data. Subsequently, we analyzed the images generated by the LBP algorithm. First, we normalized the histograms of the images as shown in Formula 1. Then, we calculated the mean and variance of these histograms, as detailed in Formulas 2 and 3.

$$\left.\begin{aligned} H(x) &= \sum_{i=0}^{L-1} p(i)\delta(x-i) \\ NormalizedHistogram &= \frac{H(x)}{\sum_{i=0}^{L-1} H(i)} \end{aligned}\right\} \tag{1}$$

Where $H(x)$ is the value of the histogram at grayscale level x, $p(i)$ represents the number of pixels at grayscale level i in the image, L is the total number of grayscale levels, and $\delta(x-i)$ is a delta function, indicating non-zero values at grayscale level i.

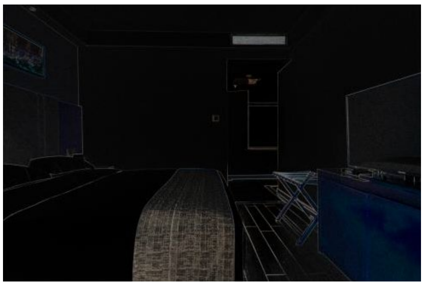
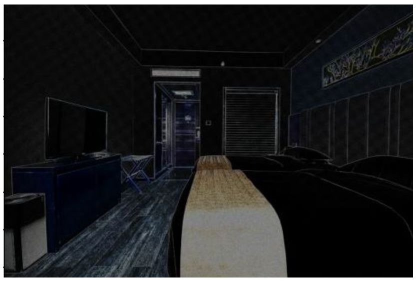
$$\upsilon = \frac{1}{N}\sum_{i=0}^{L-1} i * p(i) \tag{2}$$

$$\delta^2 = \frac{1}{N}\sum_{i=0}^{L-1} (i-\upsilon)^2 * p(i) \tag{3}$$

μ represents the mean, N is the total number of data points, and $p(i)$ is the number of pixels at grayscale level i in the histogram.

Based on calculations, we found that the variance of LBP images could reflect differences in texture features. A higher variance indicates significant texture differences, suggesting a rough and uneven surface, which contributes to a coarse visual impression. Conversely, a smaller variance suggests minimal texture differences, resulting in a smoother appearance. To validate the correlation between variance values and human tactile perception, we conducted a simple test involving three doctoral students. Participants were shown two sets of images without prior knowledge of the study's purpose. The first set contained images with high texture variance, while the second set featured images with low texture variance. They were then asked to identify which set of images appeared coarser. The test results **(Table D1)** showed that images with higher variance were perceived as coarser, consistent with our theory.

Table D1 Variance values and human tactile perception test results.

Variance values and human tactile perception test results		
	Image with high texture variance	**Image with low texture variance**
LBP pattern image		
The proportion of individuals perceiving a coarser texture	**100%**	**0%**

Appendix E

Gradient Boosting Machine (GBM) is a machine learning algorithm based on ensemble learning principles, widely used in regression and classification tasks. The basic principle is to build a strong predictive model by integrating multiple weak learners (usually decision trees), thereby improving overall predictive performance. GBM adopts an incremental optimization method, where each round of training attempts to correct the errors from the previous round, ultimately forming a powerful ensemble model.

During the training process of GBM, each new learner is trained to fit the residuals from the previous round, i.e., the difference between the model's predicted values and the actual values. This process can be seen as continually correcting model errors, thereby gradually enhancing the model's predictive capability. Specifically, GBM uses the gradient descent algorithm to guide the training direction of new learners. In each step, it calculates the gradient of the loss function relative to the current model's predicted values and then uses this gradient to train the new learner, effectively reducing the residuals.

Appendix F

The results in **Figure 5(a)** indicate that *RatPos* has a significant positive impact on the model output. High-value samples correspond to higher SHAP values, suggesting that increased positive ratings substantially enhance the predicted value. Similarly, *RatGood* also positively influences the model output, with higher ratings directly increasing the predicted value. The effect of *kf_num* on the model is relatively dispersed; however, overall, a greater number of guest rooms is associated with higher SHAP values, indicating that an increase in guest rooms typically enhances the prediction outcome. Stars also influence the model output. Although SHAP values are distributed on both sides of the axis, high star ratings (in red) predominantly correspond to positive SHAP values, suggesting that higher star ratings generally lead to higher predicted values. Therefore, it can be inferred that star ratings positively influence the predicted values. In contrast, *RatImage* does not entirely have a positive impact on the model output. Low image ratio samples (in blue) mostly correspond to positive SHAP values, while high image ratio samples (in red) are more concentrated on the negative side of the SHAP value axis. This indicates that within the model, a higher image ratio may reduce the predicted value, while a lower image ratio tends to enhance it.

For *PRO_Visual_Facilities*, a higher proportion of visual facilities corresponds mostly to positive SHAP values, indicating a clear positive impact on prediction outcomes. For *PRO_Visual_Guest_room*, red and blue dots are intermixed along the SHAP value axis and are distributed on both positive and negative sides, suggesting a high degree of uncertainty or complexity in this variable's impact on model predictions. This often indicates that there is no singular, clear directional relationship between the variable and the model's predicted outcome.

Regarding *PRO_Auditory_KB*, higher sound sharpness is often associated with negative SHAP values, while lower sharpness corresponds to positive SHAP values, indicating that its influence on the model is negative. The SHAP values for *PRO_Visual_gourmet* are more dispersed, with some high-value samples corresponding to positive SHAP values, suggesting that gourmet elements positively influence certain predictions. However, the overall impact is minimal. For the remaining variables, SHAP values are concentrated near the zero axis, with colors blending together, indicating that their impact on model output is minimal and contributes little to the prediction outcomes.

The results in **Figure 5(b)** indicate that *RatPos*, *RatGood*, and *kf_num* have a significant positive impact on the model output, while RatImage has a significant negative impact, consistent with the previous analysis. High values of the *REV_Visual_S* features (indicated by red dots) are mainly concentrated on the positive axis, showing that higher values of this feature positively impact the model output. This suggests that increased color saturation may lead the model to predict more favorable outcomes. The likely reason is that highly saturated visual effects generally attract more attention from users or consumers, enhancing visual impact and thereby improving overall user experience and satisfaction. High values of *REV_Smell_ANPs* (red dots) show a positive impact on SHAP values, indicating that higher values of *REV_Smell_ANPs* positively affect the model output. This suggests that scent-related features positively contribute to this model. The reason could be that scent is crucial to the user experience in certain products or services, and a pleasant scent experience can increase consumer satisfaction, thereby improving the model's positive predictions. Higher values of *REV_Visual_Customer* (red dots) are mainly concentrated on the positive axis, indicating that higher values of this feature positively impact the model output. This indicates that visual features related to customers (such as the visual presentation of customer service or the design of customer interfaces) significantly enhance customer satisfaction or interaction effectiveness. The SHAP values of the remaining variables are concentrated near the zero axis, with colors blending together, indicating minimal impact on the model output and little contribution to the prediction outcomes.