Journal of
Theoretical and Applied Electronic Commerce Research

MDPI

Article

Digital Tourism Commerce: Do Multisensory Mental Imagery and Imagined Presence Explain Consumers' Purchase Intentions?

Mihail Cristian Dițoiu [1], Cristina Nicolau [2,*], Otilia-Elena Platon [3] and Viorica-Mirela Ștefan-Duicu [3]

[1] The Faculty of Business and Administration, University of Bucharest, B-dul Regina Elisabeta No. 4–12, Etaj 1, Sector 3, 030018 Bucharest, Romania; cristian.ditoiu@faa.unibuc.ro
[2] Faculty of Economics Sciences and Business Administration, Transylvania University of Brașov, Colina Universității No. 1, Corp A, Etaj III, Sala AIII5, 500068 Brașov, Romania
[3] Faculty of Economics and Business Administration, Nicolae Titulescu University of Bucharest, Calea Văcărești, No. 185, Sector 4, 040051 Bucharest, Romania; otiliaplaton@univnt.ro (O.-E.P.); viorica.stefanduicu@univnt.ro (V.-M.Ş.-D.)
* Correspondence: cristina.nicolau@unitbv.ro

Abstract

Digital tourism increasingly relies on technology-enhanced visuals to shape consumer decision-making, yet most digital destination communication remains predominantly visual, limiting the multisensory richness of pre-consumption experiences. This study investigates, theoretically and empirically, how multisensory mental imagery (MI), imagined presence (IP), and purchase intention (PI) are related in digital tourism commerce, and whether visually embedded sensory metaphors can activate MI, strengthen IP, and be associated with higher PI in online tourism contexts. In a within-subject online experiment (n = 903), participants viewed, in a fixed order, a conventional destination visual followed by an AI-generated image augmented with cues suggestive of smell, taste, and touch, for two destination types (an exotic leisure destination and an urban city-break destination). For the exotic destination, sensory-enriched visuals were associated with small but reliable gains in perceived vividness, IP, and PI (d_z = 0.11–0.13, Holm-corrected); for the city-break destination, only PI increased (d_z = 0.09). Sense-specific imagery ratings rose markedly in both contexts, with gustatory and olfactory imagery emerging as the most responsive dimensions. Mediation analyses showed that IP is the strongest pathway linking MI to behavioral motivation, consistent with a sequential experiential mechanism that connects sensory design to consumer intention formation. Importantly, even subtle visual elements were sufficient to evoke sensory anticipation without requiring immersive technologies such as AR or VR. These findings advance research on digital tourism commerce, sensory marketing, and technology-mediated consumer behavior by showing that multisensory cues can be systematically embedded into static visuals to enrich experiential meaning. In practical terms, the results offer guidance for designing more engaging digital touchpoints that foster sensory imagination and support consumer decision-making in competitive online environments, with the caveat that the study measured PI rather than actual purchases.

Keywords: digital tourism commerce; technology-enhanced experiences; online consumer behavior; consumer decision-making; experiential digital marketing; purchase intention; digital experience design; sensory marketing

Check for updates

Academic Editor: Bassam Al-Romeedy

Received: 10 September 2026
Revised: 16 October 2026
Accepted: 19 October 2026
Published: 20 October 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The core of the travel process is the tourist experience, which encompasses the personal and subjective feelings, thoughts, and memories individuals construct through their interaction with a destination [1]. Moreover, the destination's physical and social environmental stimuli shape tourists' perceptions, emotions, and cognitive processes, ultimately producing behavioral responses [2]. In this light, this paper analyzes whether a sensory-driven sequence (Stimulus) in which the destination evokes cognitive-affective evaluation and experiential immersion (Organism) results in behavioral decisions (Response) [2], drawing on the Stimulus–Organism–Response (S-O-R) framework [3], which offers a robust theoretical foundation for explaining how tourists' behavioral responses emerge [4].

Destinations, perceived as physical places and experiences, also form their images around four pillars: attributes, functional and psychological characteristics, and holistic impressions (imagery) [5]. The 1988 definition of destination image (DI) has since evolved from a stereotype, the publicly held common mental picture [6], into a concept combining tangible (functional) and intangible (psychological) aspects [5] and has become multidimensional, mediating the relationship between e-reputation, DI, social media marketing efforts, and purchase intention [7]. In this context, DI requires multisensory integration to create a richer, deeper experience and strengthen the tourist-destination bond [8]. Accordingly, the tourist experience extends far beyond visual stimuli, encompassing a wide spectrum of sensory inputs such as sounds, smells, tactile impressions, and even tastes [9]. Building on this multisensory nature of experience, recent text-to-image models now enable destination marketers to digitally reproduce such cues with unprecedented precision: they can generate photorealistic imagery at negligible marginal cost and, crucially for sensory design, control individual pictorial elements so that specific sensory triggers such as visible steam, food textures, or warm lighting can be inserted into an otherwise unchanged scene. Large-scale evaluations show that these AI-generated visuals can match or even surpass human-made marketing images in perceived quality and consumer appeal [10], thereby offering a powerful tool for stimulating tourists' digital PI. At the same time, research on synthetic advertising indicates that consumer responses to AI-produced content hinge on disclosure, perceived authenticity, and trust in manipulated imagery [11], and early tourism-specific work shows that generative models already shape destination portrayals, including by reproducing destination stereotypes [12]. What remains untested is whether this new production capability creates experiential value in digital tourism commerce: whether sensory cues deliberately embedded through generative AI change what consumers imagine, how present they feel in a depicted destination, and, ultimately, what they intend to buy.

Against this background, the objective of the present research is to investigate, theoretically and empirically, how multisensory mental imagery (MI), imagined presence (IP), and purchase intention (PI) are related in digital tourism commerce, and through which experiential mechanism digitally conveyed sensory cues are associated with consumers' intention to purchase tourism products online. The study is therefore not confined to comparing two kinds of images. Theoretically, it integrates DI theory, sensory marketing, and MI research within the S-O-R logic to conceptualize a sequential experiential pathway in which sensory-enriched destination visuals (stimulus) activate MI and IP (organism), which are in turn associated with PI (response); four hypotheses derived from this framework are developed in Section 2. Empirically, it tests this pathway in a within-subject online experiment (n = 903) in which AI-generated, sensory-enriched visuals serve as the instrument for varying sensory input across two destination types, and it examines the MI $\rightarrow$ IP $\rightarrow$ PI sequence through correlational, between-person, and within-participant mediation analyses. The generative-AI image pair is thus the empirical vehicle of the inves-

tigation; its object is the experiential process that links sensory anticipation to consumer decision-making in online tourism environments.

The remainder of the paper is structured as follows. Section 2 reviews the literature on DI, MI, IP, and PI, and develops the sequential experiential pathway examined here. Section 3 presents the materials and methods: the research design and stimuli, the sample and data collection procedure, the measurement instrument, and the statistical strategy, including the handling of the fixed presentation order and multiple comparisons. Section 4 reports the results, covering condition effects for each destination, sense-specific analyses, correlation structures, and the between-person and within-participant mediation models. Section 5 discusses the findings, their theoretical and entrepreneurial implications, and the study's limitations. Section 6 concludes.

2. Literature Review

Multiple senses shape tourism experiences, combining visual impressions with sounds, smells, textures, and tastes into a single, coherent experience [13]. Most digital destination communication still privileges sight [14], reducing complex environments to static imagery and often creating a perceptual gap between online representations and embodied travel encounters [13,15,16]. In an increasingly crowded, digital marketplace, such representations matter because DI remains one of the strongest predictors of pre-consumption evaluation and travel choice [17,18]. As a result, digital touchpoints that evoke richer sensory expectations become a decisive factor in shaping consumers' willingness to commit to a destination and to translate interest into purchase intentions, particularly as digital tourism commerce increasingly relies on such multisensory cues to guide evaluative judgments and motivate transactional engagement.

DI is commonly understood as a composite of cognitive beliefs, affective responses, and behavioral tendencies toward a place [17,19]. Recent work extends this framework by introducing MI [16] as a complementary dimension that captures anticipatory impressions of smell, taste, and touch that emerge before physical arrival [15]. Sensory expectations influence not only how destinations are imagined but also how strongly they are desired and perceived. Together, these anticipatory sensory impressions play a central role in shaping consumers' motivational states, strengthening desire and increasing the likelihood that positive evaluations explain concrete purchase intentions within digital tourism commerce. Consistent with this, engaging multiple sensory channels or dimensions has been shown to deepen emotional connection and strengthen experiential meaning [13,20]. Brands are perceived more vividly when they stimulate more than one sense, influencing memory, attachment, and behavioral response [21].

Such multisensory stimulation not only enhances brand vividness but also primes consumers to generate richer internal simulations when they encounter destination cues online, deepening the cognitive–affective processes that shape their travel-related decision-making. When prospective tourism consumers perceive destination visuals online, they actively construct internal representations that may extend beyond vision to include imagined smells, tastes, or tactile sensations [22]. MI can foster IP, the subjective feeling of "being there" [23,24], which strengthens emotional anticipation and supports behavioral intention [25,26] and increases consumers' propensity to convert interest into actual purchasing decisions within digital tourism commerce environments. Despite growing interest in the sensory dimensions of tourism, most studies focus on on-site experiences or characterize sensory attributes in largely descriptive terms, leaving open questions about how digitally encoded sensory metaphors influence pre-travel evaluation [13,15]. These gaps highlight the need for more robust empirical approaches that systematically examine how digitally encoded sensory cues shape pre-travel evaluation and consumer purchase

decision-making. Moreover, quantitative frameworks that capture this digitally mediated sensory anticipation also remain scarce [27], reinforcing the need for research that advances new measurement strategies and offers novel empirical insight into how multisensory digital cues influence consumer responses in tourism contexts. Against this backdrop, the present article makes a timely contribution by examining whether visually embedded sensory metaphors can activate MI and IP in digital tourism environments, grounded in DI theory, sensory marketing, and MI research, which together provide the conceptual foundation for investigating these digitally mediated sensory effects.

Thus, this article examines whether visually embedded sensory metaphors can activate MI and IP in digital tourism contexts. We integrated olfactory, gustatory, and tactile cues into destination visuals and compared them with conventional imagery to assess their influence on MI, IP, and PI. The research hypotheses are as follows:

H1. *Sensory-enriched visuals increase MI compared with conventional destination imagery.*

H2. *MI positively affects IP in digital tourism contexts.*

H3. *IP positively affects PI in digital tourism contexts.*

H4. *In digital tourism contexts, IP mediates the relationship between MI and PI.*

By examining such mechanisms across two destination types, the research seeks to clarify how simple visual strategies can narrow the sensory gap between online encounters and anticipated travel experiences, offering both theoretical refinement and practical guidance for e-commerce strategies in tourism businesses and destination marketing alike, particularly in designing persuasive sensory-based tactics that strengthen consumer motivation and increase the likelihood of purchase. Sensory cues can shape judgments even without direct physical stimulation by triggering imagined sensory experiences [20,28], which underlines the importance of empirically examining how visually encoded sensory metaphors embedded in digital destination imagery affect multisensory MI, IP, and PI across different destination types, and of identifying the experiential processes through which sensory-enriched visuals influence pre-travel evaluation in the e-commerce of tourism products.

3. Materials and Methods

3.1. Research Design

The study employed an online paired A/B experimental design to assess how visually embedded sensory metaphors influence multisensory MI, IP, and PI in digital tourism contexts. This approach draws on tourism research showing that destination perception emerges through experiential and multisensory processes rather than visual evaluation alone [13,29]. Participants evaluated two categories of stimuli: conventional destination visuals (control condition) and sensory-enriched images embedding metaphorical cues intended to evoke olfactory, gustatory, and tactile associations (experimental condition). To account for contextual variability, we included two destination types: an exotic leisure destination and an urban city-break destination, reflecting established differences in DI formation across tourism typologies [17,18].

Visual metaphors served as indirect carriers of non-visual sensory information [30], drawing on cross-modal correspondences in which visual attributes activate associated sensory representations [31]. We produced all visual stimuli using generative artificial intelligence tools (ChatGPT 5.2; OpenAI, San Francisco, CA, USA) and then curated them to ensure consistent framing, resolution, and thematic coherence across conditions. Control images resembled standard tourism promotional content emphasizing scenery or architec-

tural landmarks (see Figures 1 and 2). Sensory-enriched visuals incorporated elements such as visible steam, surface condensation, food textures, warm lighting, atmospheric diffusion, tactile material cues, gustatory primes, and thermal contrast. We selected these features to stimulate multisensory MI, following sensory marketing evidence that visual inputs can evoke olfactory, gustatory, and tactile associations through cross-modal processing [20,31]. Generative artificial intelligence was used solely to create the stimuli, of which there were four in total: one conventional and one sensory-enriched image for each destination. All participants viewed the same four images, so destination type was a within-subject replication factor with a single exemplar per cell rather than a manipulated variable sampled across multiple images. We constructed each sensory-enriched image as an augmented version of its conventional counterpart, preserving the base scene, composition, framing, subject matter, and overall color palette. Images were matched in resolution and presented without labels, and participants were not informed of the conditions. Generative AI offered two advantages for stimulus construction: it could add sensory cues to an otherwise unchanged scene, which is difficult to achieve with photography, and it could hold visual style constant across conditions. At the same time, AI generation introduces potential boundary conditions, including perceived authenticity, trust in synthetic imagery, and the correspondence between the digital representation and the actual destination [32]. Stimulus order was held constant rather than randomized or counterbalanced: within each destination block, all participants viewed the conventional visual first and the sensory-enriched visual second. This was a deliberate design decision.



(**a**) (**b**)

Figure 1. Exotic leisure destination. (**a**) Control. (**b**) AI-enhanced.



(**a**) (**b**)

Figure 2. Urban city-break destination. (**a**) Control. (**b**) AI-enhanced.

First, a fixed baseline-then-treatment sequence follows the logic of paired A/B evaluation, in which the conventional visual serves as the reference point against which the incremental value of the manipulation is judged; within-subject comparison against a common referent increases statistical power and removes between-participant heterogeneity

from the treatment contrast [33–35]. Second, keeping the procedure identical for every respondent ensures that presentation order cannot covary with participant characteristics, so order operates as a constant rather than a confound; in unsupervised online experiments, procedural standardization is recommended precisely because it removes a source of uncontrolled variance [36]. Third, the sequence reproduces the typical decision context of digital tourism commerce, where consumers encounter conventional promotional imagery as the default and evaluate enhanced content relative to that prevailing standard [37]; within-subject presentation supplies the shared frame of reference that such relative judgments require [38]. Because we collected outcome measures immediately after each stimulus using parallel item sets, we estimate condition effects as within-person differences, which is the appropriate estimand for this research. We acknowledge that the lack of randomization confounds the manipulation with serial position, so the design alone cannot separate a treatment effect from a systematic second-stimulus effect. For this reason, the analysis draws on two internal checks that follow the logic of the nonequivalent dependent variables design [39,40]: outcomes that the manipulation should not move (e.g., DI and IP for cues targeting smell and taste) serve as comparison outcomes against which a generic position effect would be detected, and the sense-specific imagery items allow us to test whether the increase is confined to the modalities actually cued.

For each image used as a stimulus, respondents rated perceived visual vividness and attractiveness (DI), ease of mentally projecting oneself into the destination (IP) [41], and PI. After viewing paired images within each destination category, participants completed an additional block assessing multisensory MI, indicating how strongly the visuals suggested smells, tastes, and tactile sensations. The items used to measure each variable are presented in Table 1.

Table 1. Measurement scales for the latent variables.

Variable	Item
DI	This image is vivid and appealing.
IP	I can easily imagine myself being present at this destination.
PI	This image makes me want to purchase a tourism product (e.g., a trip or a stay) at the destination.
MI	This image evokes pleasant smells (e.g., fresh air, sea breeze, local food aromas); This image evokes taste-related experiences (e.g., local food or drinks); This image evokes tactile sensations (e.g., warmth, textures, etc.).

Open-ended responses provided qualitative insight into spontaneous sensory associations. This design integrates controlled visual manipulation with perceptual and experiential measures, enabling a systematic examination of how sensory metaphors embedded in static digital imagery shape pre-consumption tourism evaluation. The authors were responsible for the experimental design, measurement, data collection, analysis, and interpretation.

3.2. Population, Sampling, and Data Collection

The target population for this study consisted of Romanian adults who engage with digital tourism content and may consume online destination imagery. The study used non-probability convenience sampling, consistent with prior research using online perceptual instruments in tourism and digital marketing contexts. We collected data from 27 January to 2 February 2026, using an online questionnaire designed to evaluate visual perception, multisensory MI, IP, and PI in response to DI. We distributed the instrument online via social media channels. We invited all adults to participate and informed respondents that we

would keep their answers confidential. Informed consent was implied through voluntary survey completion; participants could discontinue at any time without consequences. The procedure aligns with minimal-risk ethical standards for anonymous online surveys and mirrors the methodological approach used in comparable digital tourism perception studies. We conducted all analyses on anonymized data.

We retained 903 valid responses for analysis after excluding incomplete submissions (with missing answers), straight-line patterns (we computed the within-respondent standard deviation across all Likert-scale perceptual items; cases with SD < 0.20 were classified as straight-liners and removed, as such low variability indicates non-differentiated responding), and completion-time outliers (completion times below 90 s or above 15 min were classified as outliers, as such durations are incompatible with attentive processing of the stimuli and questionnaire). We retained 93.58% of the responses received; the remaining 6.42% were incomplete (3.94%), straight-liners (1.14%), or completion-time outliers (1.34%). As a robustness check, re-running the main paired comparisons under alternative screening thresholds (straight-line criterion SD < 0.30; completion-time bounds of 60 s and 20 min) left all substantive conclusions unchanged.

The resulting sample composition provided experiential diversity while preserving cultural homogeneity, an important consideration in DI research [18]. Participants reported varied income levels, residential backgrounds, and educational attainment; however, as Table 2 shows, women aged 45–55 years constituted the largest group.

Table 2. Sample demographic profile (n = 903).

Variable	Category	%
Gender	Female (dominant)	94.91
	Male/no answer	4.76/0.33
Age	45–55 (dominant)	50.39
	35–44	29.01
	Other	20.60
Residence	Urban/rural	74.42/25.58
Education	Master's	54.71
	Bachelor's	38.43
	Other	6.86
Income	>€1000/month	58.69
	<€1000/month	18.83
	No answer	22.48

3.3. Measurement Instrument and Data Analysis

The measurement instrument used in this study consists of a structured questionnaire organized into two parts: six descriptive questions capturing respondents' age, gender, residence, education, income, and recent travel behavior, and twenty questions designed to assess visual and multisensory perception in digital tourism marketing contexts: 90% of items were measured on a seven-point Likert scale from 1 = strongly disagree to 7 = strongly agree, and 10% were open-ended questions.

A first set of items focused on the core visual-perceptual evaluation of each stimulus. The questions comprised two image-based evaluation blocks, one for the exotic destination and one for the city-break destination, each with parallel sets of items to ensure comparability across stimuli. For each image, respondents evaluated: (1) perceived vividness

and attractiveness (DI), (2) ease of mental immersion in the destination (IP), and (3) desire to purchase a tourism product at the destination (PI). These items capture the cognitive–affective appraisal of digital imagery and reflect established evidence that visual vividness and mental simulation precede digital engagement and subsequent PI.

A second set of items assessed multisensory metaphors, capturing the sensory cues that respondents perceived in the two AI-enhanced images. These items captured olfactory associations (e.g., sea breeze, fresh air, culinary aromas), gustatory associations (e.g., local food or drinks), and tactile sensations (e.g., warmth, textures, physical comfort). This multisensory block extends traditional visual evaluation scales by incorporating embodied cognition and sensory marketing principles, allowing the instrument to capture how digital imagery evokes cross-modal experiential cues. Administering multisensory items for the control stimuli would have been methodologically consistent with a paired A/B design, but it would have risked prompting respondents to infer sensory meanings that were never embedded in the stimuli. In our view, restricting multisensory assessment to the enriched images preserves the internal validity of the A/B design by ensuring that sensory responses are measured only where the experimental manipulation is present.

Finally, each image block concludes with an open-ended item inviting respondents to describe, in 1–2 sentences, the feelings or MI elicited by the visual stimuli. This qualitative component enriches the perceptual data by capturing spontaneous associations and metaphorical interpretations that Likert-type items may not fully reflect.

We developed the research instrument in Romanian and structured it according to validated perceptual and sensory evaluation scales. Its design ensures content validity through clear construct operationalization, parallel item structure across stimuli, and alignment with theoretical frameworks on digital imagery, sensory metaphors, and tourism experience formation. We assessed DI, IP, and PI using single-item indicators [42] adapted for digital destination contexts, a common approach in tourism perception research [13,17]. We assessed perceived vividness and attractiveness (DI) with one item that evaluates respondents' visual appraisal of the displayed images. Because the item combines vividness and attractiveness, the two attributes cannot be separated in these data. Single items were chosen for three reasons: each construct is concrete and singular from the respondent's perspective; respondents rated four stimuli in succession in an unsupervised online setting where repeated multi-item batteries would have inflated fatigue and dropout; and single-item measures show predictive validity comparable to multi-item scales for concrete constructs of this kind [42,43]. The trade-off is that we cannot estimate internal-consistency reliability for these variables, which we acknowledge as a limitation. Two consequences follow. First, because each construct is captured by one item, internal-consistency reliability (Cronbach's α) cannot be computed for DI, IP, or PI, and random measurement error in these items cannot be separated from true-score variance. Second, IP and PI were rated in the same format, by the same respondent, immediately after the same stimulus; the very high IP–PI correlation observed in the sensory-enriched condition ($r = 0.82–0.86$, versus 0.70–0.76 in the control condition; see Section 4) should therefore be read with this measurement strategy in mind, since adjacent single items of closely related constructs are prone to common method variance [44] and cannot establish discriminant validity. Mediation estimates involving IP and PI are accordingly interpreted as descriptions of the observed association structure rather than as evidence of two separable constructs.

In contrast, we operationalized multisensory MI using three items capturing perceived olfactory, gustatory, and tactile activation. Because smell, taste, and touch are conceptually distinct modalities, sense-specific results are treated as primary throughout, and the three items are combined into a composite MI only where internal consistency supports aggregation; internal consistency for the sensory-enriched condition is reported in Section 4. The

structure reflects prior work conceptualizing DI as a multisensory construct shaped by anticipatory sensory impressions [18,29].

Qualitative inputs complemented quantitative measures by capturing subjective sensory associations. Quantitative analyses included descriptive statistics for all study variables. Because each participant evaluated both the conventional (control) and the sensory-enriched (experimental) visual within each destination context, paired-samples *t*-tests compared mean ratings across conditions for perceived DI, multisensory MI (overall and sense-specific olfactory, gustatory, tactile), IP, and PI. Effect sizes (Cohen's d_z, the standardized mean difference for paired data) were calculated to assess practical relevance [45]. We examined associations among multisensory MI, IP, and PI through correlation analyses, reflecting experiential models linking sensory anticipation to behavioral motivation [13]. We reviewed qualitative responses thematically to identify recurring sensory metaphors and experiential patterns that support interpretation of the statistical outcomes.

Statistical analyses proceeded in three steps. First, condition effects were tested with paired-samples *t*-tests organized in two pre-specified families: Family A comprised nine tests on the core outcomes (DI/vividness, IP, PI) for the exotic destination, the city-break destination, and the pooled destination average, and Family B comprised twelve tests on olfactory, gustatory, and tactile imagery and their composite for the same three sets, giving k = 21 confirmatory paired comparisons in total. Type I error was controlled with the Holm–Bonferroni step-down procedure within each family ($\alpha = 0.05$, two-sided) [46]; Benjamini–Hochberg false-discovery-rate adjusted values [47] and a single-step Bonferroni correction over all 21 tests are additionally provided in the Supplementary Materials, and no decision changes under any of the three procedures. Effect sizes are reported as Cohen's d_z, the mean of the paired differences divided by the standard deviation of those differences, which is the standardized effect size that corresponds to the paired-samples *t*-test ($t = d_z\sqrt{n}$) [45]; 95% confidence intervals for d_z are based on the noncentral t distribution. Because d_z depends on the correlation between the two ratings of the same respondent, it is not directly comparable with the between-subjects Cohen's d; the Supplementary Materials therefore also report d_{av}, the mean difference divided by the average of the two condition standard deviations, which is comparable across designs [44]. In the main text, d_z is the only effect-size metric reported, and the subscript is used throughout to avoid confusion with the between-subjects d. Moreover, we computed Wilcoxon signed-rank tests and matched-pairs rank-biserial correlations as distribution-free checks, which corroborated every parametric decision. Second, we computed Pearson correlation matrices among all measured variables separately for each destination × condition combination (21 unique pairs per matrix), with Holm-adjusted *p*-values and Fisher-z 95% confidence intervals; we present Spearman coefficients and the full matrices in the Supplementary Materials.

Third, the hypothesized pathway DI → MI → IP → PI was examined with two complementary mediation models, both estimated with ordinary least squares (OLS) regression and 5000 percentile-bootstrap resamples for indirect effects. (a) A between-person serial mediation model was estimated separately within each destination × condition combination, with HC3 heteroscedasticity-robust standard errors; because each participant contributes exactly one observation to each of these models, observations within a model are independent, which is what justifies the use of OLS despite the repeated-measures structure across conditions. (b) The effect of the manipulation itself on the pathway was tested within the within-participant mediation framework [48], in the path-analytic form [49] for two-condition repeated-measures designs, in which difference scores (ΔIP and ΔPI, computed as sensory-enriched minus control) are regressed on ΔMI with mean-centered condition sums included as covariates; this approach accounts explicitly for the dependence between the two ratings supplied by the same respondent, which is why no additional

multilevel specification is required. Indirect effects were considered significant when the 95% percentile bootstrap confidence interval excluded zero. Only model (b) follows this within-participant framework; model (a) is a conventional between-person serial mediation estimated separately within each condition and does not test the manipulation. For both models, every indirect effect is reported in Section 4 as a point estimate with its 95% percentile bootstrap confidence interval (5000 resamples). Because stimulus order was fixed (Section 3.1), these indirect effects are interpreted as associations consistent with the hypothesized sequence rather than as estimates of causal effects. All analyses used the 903 complete cases; there were no missing values among retained respondents.

4. Results

Most participants reported recent leisure travel experience and urban residence. Although the present sample limits direct population-level generalization, it provides relevant insight into multisensory processing within an educated consumer segment highly engaged with digital visual environments.

Paired-samples *t*-tests compared evaluations of conventional versus sensory-enriched visuals within each destination block (n = 903, df = 902 throughout); unless marked as raw, the p-values below are Holm-adjusted within their family. See Tables 3 and 4. For the exotic destination, sensory-enriched imagery was associated with higher perceived visual vividness (M = 5.79, SD = 1.50, vs. M = 5.63, SD = 1.37; t = 3.33, p = 0.005, d_z = 0.11, 95% CI [0.05, 0.18]), higher IP (M = 5.86, SD = 1.46, vs. M = 5.71, SD = 1.45; t = 3.52, p = 0.003, d_z = 0.12 [0.05, 0.18]), and higher PI (M = 5.89, SD = 1.47, vs. M = 5.73, SD = 1.54; t = 4.03, p = 0.001, d_z = 0.13 [0.07, 0.20]). For the city-break destination, neither vividness (M = 5.82 vs. 5.77; t = 0.91, raw p = 0.364, d_z = 0.03 [−0.04, 0.10]) nor IP (M = 5.79 vs. 5.78; t = 0.29, raw p = 0.774, d_z = 0.01 [−0.06, 0.07]) differed between conditions, whereas PI was higher in the sensory-enriched condition (M = 5.76 vs. 5.65; t = 2.59, p = 0.048, d_z = 0.09 [0.02, 0.15]). Pooled across destinations, DI (d_z = 0.09 [0.02, 0.15], p = 0.048) and PI (d_z = 0.13 [0.07, 0.20], p = 0.001) were higher under sensory enrichment, while the pooled imagined-presence difference did not survive correction (d_z = 0.08 [0.01, 0.14]; raw p = 0.024, Holm-adjusted p = 0.072). Wilcoxon signed-rank tests replicated every decision. The Supplementary Materials report complete test statistics, adjusted p-values under all three correction procedures, and additional effect-size metrics.

Table 3. Descriptive statistics and paired differences (n = 903).

Variable		Conventional M (SD)	Sensory-Enriched M (SD)	Mean Difference 95% CI
Exotic	DI	5.632 (1.366)	5.788 (1.505)	0.156 [0.064, 0.248]
	IP	5.710 (1.447)	5.860 (1.458)	0.151 [0.067, 0.235]
	PI	5.729 (1.535)	5.894 (1.470)	0.165 [0.085, 0.245]
City break	DI	5.774 (1.196)	5.815 (1.464)	0.041 [−0.048, 0.130]
	IP	5.782 (1.261)	5.794 (1.455)	0.012 [−0.071, 0.095]
	PI	5.646 (1.373)	5.756 (1.500)	0.111 [0.027, 0.195]
Pooled	DI	5.703 (1.048)	5.802 (1.305)	0.099 [0.023, 0.174]
	IP	5.746 (1.121)	5.827 (1.292)	0.081 [0.011, 0.152]
	PI	5.687 (1.184)	5.825 (1.302)	0.138 [0.070, 0.206]

Table 4. Paired tests with adjusted p values and effect size intervals (n = 903).

Variable		t (902)	Raw p	Holm p	Cohen's d_z 95% CI
Exotic	DI	3.3336	0.000892	0.005353	0.111 [0.045, 0.176]
	IP	3.5172	0.000458	0.003206	0.117 [0.052, 0.182]
	PI	4.0345	0.000059	0.000534	0.134 [0.069, 0.200]
City break	DI	0.9078	0.364244	0.728487	0.030 [−0.035, 0.095]
	IP	0.2870	0.774140	0.774140	0.010 [−0.056, 0.075]
	PI	2.5936	0.009651	0.048255	0.086 [0.021, 0.152]
Pooled	DI	2.5574	0.010707	0.048255	0.085 [0.020, 0.150]
	IP	2.2608	0.024009	0.072027	0.075 [0.010, 0.141]
	PI	3.9910	0.000071	0.000569	0.133 [0.067, 0.198]

For sense-specific imagery items, ratings of olfactory, gustatory, and tactile imagery were high for the sensory-enriched visuals (means ranging from 5.06 to 5.84). The three-item composite showed high internal consistency in the sensory-enriched condition (Cronbach's α = 0.90 for the exotic and α = 0.93 for the city-break block [50]). Figure 1 (exotic leisure destination) and Figure 2 (urban city-break destination) present representative examples of conventional and sensory-enriched visuals for both destination contexts. Each figure contrasts baseline imagery with its sensory-augmented counterpart, illustrating how visual metaphors (e.g., food, steam, texture, atmospheric elements) were embedded to cue non-visual sensory associations. Sense-specific analyses showed a consistent ordering across the two destinations. Gustatory imagery was the most responsive dimension in both blocks (M sensory = 5.59 exotic and 5.84 city break). Olfactory and tactile imagery increased to a similar, though somewhat smaller, extent in the exotic block (olfactory: 5.06, tactile: 5.24), while in the city-break block olfactory imagery (5.80) clearly exceeded tactile imagery (5.36). The pattern indicates that the visually encoded metaphors primarily activated gustatory and olfactory representations, consistent with the food-, steam-, and aroma-related content of the embedded cues, while tactile imagery, for which fewer cues were embedded, was comparatively less responsive, particularly in the urban block.

Within the sensory-enriched condition, multisensory imagery correlated positively with IP (exotic r = 0.60; city break r = 0.70) and with PI (r = 0.60 and r = 0.67, respectively), and IP correlated strongly with PI (r = 0.82 exotic; r = 0.86 city break); all Fisher-z 95% confidence intervals exclude zero and all Holm-adjusted $p < 0.001$. The full Pearson and Spearman matrices for every destination × condition combination, with raw and adjusted p-values and confidence intervals, are provided in the Supplementary Materials. The very high presence–intention correlations (r = 0.82–0.86) exceed what is typical for empirically distinct constructs and may partly reflect the conceptual proximity of the two single items and common method variance arising from same-source, same-format measurement [44] (see the measurement note in Section 3.3); discriminant validity between IP and PI therefore cannot be established in these data, and mediation estimates involving these variables should be read with caution.

Comparative analyses across destination types suggest that sensory enrichment operates differently depending on contextual framing. Exotic imagery consistently improved DI, IP, and PI, whereas city-break imagery showed limited sensitivity to sensory augmentation, with statistically meaningful change only for PI. This pattern suggests that exotic destinations may support more holistic experiential amplification through sensory cues, while urban contexts appear to rely more on cognitively familiar schemas and localized

stimuli. These contextual differences suggest that sensory-evocative design may strengthen pre-travel evaluation through distinct experiential pathways, ultimately shaping how strongly consumers translate perceptual responses into concrete intentions to purchase tourism products.

Open-ended responses supported the quantitative findings while highlighting distinct experiential profiles across destination types. Exotic visuals elicited descriptions centered on relaxation, calmness, sea breeze, sunlight, sand, and beach atmosphere, reflecting a predominantly environmental and embodied sensory orientation. By contrast, city-break imagery generated narratives anchored in urban rituals such as coffee, croissants, breakfast scenes, and, in particular, the smell and taste of coffee, emphasizing olfactory and gustatory dominance in urban contexts. Across both destinations, participants frequently reported mentally projecting themselves into the scene, providing qualitative confirmation that IP is central to pre-consumption destination evaluation. Together, these qualitative patterns illustrate how sensory-evocative imagery strengthens pre-travel mental simulation, ultimately reinforcing consumers' intention to purchase tourism products by making the destinations feel more vivid, desirable, and experientially accessible.

The between-person serial mediation models (DI → MI → IP → PI), estimated within each destination × condition combination, supported the hypothesized pathway in the sensory-enriched condition. For the exotic destination, the total association between DI and PI was B = 0.69, 95% CI [0.62, 0.76]; the serial indirect path through MI and IP was B = 0.09 [0.06, 0.13], the path through MI alone B = 0.07 [0.04, 0.10], the path through IP alone B = 0.31 [0.24, 0.39], and the direct path B = 0.22 [0.13, 0.32] (R^2 = 0.71 for the final equation). The city-break model showed the same structure (total B = 0.82 [0.76, 0.87]; serial indirect B = 0.11 [0.07, 0.16]; via IP B = 0.41 [0.30, 0.53]; direct B = 0.25 [0.12, 0.37]; R^2 = 0.77). Because MI was not measured in the control condition, that condition cannot display any MI-related pathways; only the DI → IP → PI route was present (indirect B = 0.36 [0.29, 0.44] exotic; B = 0.45 [0.35, 0.55] city break), reflecting the baseline cognitive route through which conventional imagery influences imagined presence and subsequent intention. For these results, see Table 5. Complete coefficient tables (unstandardized B, HC3 robust SE, 95% CI, standardized β, R^2) for all twelve equations, and all bootstrap indirect-effect estimates, are given in the Supplementary Materials.

Table 5. Indirect associations and contrasts from the replacement analyses.

Variable		a	b	Indirect	95% Bootstrap CI
Between-person MI → IP → PI	Exotic	0.5533	0.7219	0.3994	[0.3319, 0.4697]
	City break	0.6674	0.7904	0.5275	[0.4499, 0.6074]
Within-participant → IP → PI	Exotic	0.1506	0.6618	0.0997	[0.0437, 0.1556]
	City break	0.0122	0.8081	0.0098	[−0.0591, 0.0771]

The within-participant analysis [48,49], which tests the manipulation itself, refined this overall pattern. The condition effect on multisensory imagery was very large in both blocks (a-path: mean ΔMI = 3.53, SE = 0.05, exotic; 3.96, SE = 0.05, city break). The change in IP strongly predicted the change in PI in both blocks (b-path: B = 0.64, SE = 0.04, exotic; B = 0.80, SE = 0.03, city break; both $p < 0.001$). However, the magnitude of a respondent's imagery change predicted the magnitude of the presence change only for the city-break destination (B = 0.26, SE = 0.05, $p < 0.001$; serial indirect effect of condition through ΔMI and ΔIP = 0.82, 95% CI [0.50, 1.17]; percentile bootstrap, 5000 resamples), not for the exotic destination (B = 0.06, SE = 0.06, $p = 0.318$; serial indirect 0.13 [−0.11, 0.40]). The total within-person condition effect on PI was small in both blocks (mean ΔPI = 0.17 [0.08, 0.25]

exotic; 0.11 [0.03, 0.19] city break). Taken together, the models support a sequential pathway in which IP links imagery to intention: between persons in the sensory-enriched condition and within persons for the city-break destination; however, for the exotic destination, the magnitude of the imagery change was not systematically related to the magnitude of the presence change. Because stimulus order was fixed (Section 3.1) and IP and PI rest on single items whose discriminant validity could not be established (Section 3.3), these estimates, including the bootstrapped indirect effects, describe associations consistent with the proposed sequence rather than confirmed causal effects.

In sum, the empirical findings provide varying degrees of support for the four research hypotheses. The results indicate that sensory-enriched destination imagery can stimulate multisensory mental imagery and enhance consumers' purchase intentions, although the strength and nature of these effects vary across destination contexts and analytical models. Table 6 summarizes the findings and hypothesis validation.

Table 6. Summary of hypothesis testing.

	Hypothesis	Validation	Empirical Evidence
H1	Sensory-enriched visuals increase MI compared with conventional destination imagery.	Partially supported	High sensory imagery ratings were observed for both destinations, particularly for gustatory and olfactory imagery. However, MI was not measured in the conventional condition, which precludes a direct statistical comparison.
H2	MI positively affects IP in digital tourism contexts.	Supported, with contextual limitations	MI was positively correlated with IP for both the exotic destination (r = 0.60) and the city-break destination (r = 0.70), both $p < 0.001$. Within-participant changes were significant only for the city-break destination.
H3	IP positively affects PI in digital tourism contexts.	Supported	IP was strongly positively correlated with PI for the exotic destination (r = 0.82) and the city-break destination (r = 0.86), both $p < 0.001$. Changes in IP also significantly predicted changes in PI in both contexts ($p < 0.001$).
H4	IP mediates the relationship between MI and PI.	Partially supported	Between-person mediation analyses supported the indirect relationship in both contexts. Within-participant sequential mediation was significant for the city-break destination but not for the exotic destination.

H1 proposed that sensory-enriched visuals increase mental imagery (MI) compared with conventional destination imagery. The findings support this hypothesis, as the sensory-enriched images elicited substantial olfactory, gustatory, and tactile imagery in both destination contexts. Gustatory and olfactory imagery were particularly responsive to the sensory cues embedded in the images. Nevertheless, because the study administered multisensory imagery measures only for the sensory-enriched images, it does not provide a direct paired statistical comparison of overall MI between the conventional and sensory-enriched conditions. Therefore, the observed sensory responses support H1, but this evidence should be interpreted in light of the measurement design.

H2 proposed a positive relationship between MI and imagined presence (IP). The results support this hypothesis at the associative level. MI was positively correlated with IP in both destination contexts under the sensory-enriched condition: r = 0.60 for the exotic destination and r = 0.70 for the city-break destination; both $p < 0.001$. However, the within-participant mediation analysis showed that changes in MI were significantly associated with changes in IP only for the city-break destination, not for the exotic destination. Thus, the evidence supports H2 overall, while indicating that the relationship may depend on destination context and the analytical approach.

H3 proposed that IP positively affects PI. The findings support this hypothesis through strong positive associations between IP and PI in both destination contexts: r = 0.82 for the exotic destination and r = 0.86 for the city-break destination; both $p < 0.001$. In addition, the within-participant mediation models showed that changes in IP significantly predicted changes in PI for both destinations (B = 0.64 and B = 0.80, respectively; both $p < 0.001$). These results indicate that consumers who experience a stronger sense of imagined presence tend to report higher purchase intentions. Nevertheless, given the observational nature of the relationships between these constructs and the study's fixed stimulus order, the findings should be interpreted as evidence of positive associations rather than definitive causal effects.

H4 proposed that IP mediates the relationship between MI and PI. The results provide partial support for this hypothesis. The between-person mediation analyses identified significant indirect associations through IP in both destination contexts, with the corresponding confidence intervals excluding zero. However, the within-participant analyses revealed that the sequential indirect effect of sensory enrichment through changes in MI and IP was significant for the city-break destination but not for the exotic destination. Consequently, H4 is partially supported: the proposed mediation mechanism is evident in the between-person models and in the within-participant analysis for the city-break destination, but it is not consistently confirmed across both destination contexts.

Taken together, the findings highlight the role of multisensory mental imagery and imagined presence in explaining consumers' purchase intentions in digital tourism environments. While the results generally support the proposed experiential framework, the context-dependent findings for H2 and H4 suggest that the relationships between sensory imagery, imagined presence, and purchase intention are not uniform across destination types. Furthermore, because the study measured purchase intention rather than actual purchasing behavior, the findings concern consumers' stated intentions and should not be interpreted as evidence of realized purchases.

5. Discussion

This study aimed to investigate, theoretically and empirically, how multisensory MI, IP, and PI relate in digital tourism commerce, using visually encoded sensory metaphors in AI-generated destination imagery, contrasted with conventional destination visuals, to vary sensory input. Against this objective, the results address three questions: whether sensory-enriched digital content activates MI, whether MI links to IP and PI, and through which experiential pathway this occurs. Using a large Romanian sample ($n = 903$) and a paired experimental design, the findings provide convergent quantitative and qualitative evidence that sensory-enriched visuals facilitate richer experiential anticipation in pre-consumption tourism contexts. Sensory-enhanced images were associated with substantially higher multisensory imagery in both destination blocks and small but reliable gains in DI, IP, and PI for the exotic destination, whereas for the city-break destination only PI differed; by far the largest condition differences were observed in the sense-specific imagery items, particularly gustatory and olfactory activation. These results reinforce earlier conceptualizations of tourism as an inherently multisensory experience and extend DI theory by empirically showing that sensory impressions can be activated before physical travel through visual design alone [13,15,16,18].

A central contribution of the study is the identification of a sequential experiential pathway in which visually induced multisensory imagery is linked to stronger IP, which in turn is associated with higher PI. This sequential pattern points to IP as an important link between sensory imagery and behavioral motivation. It also aligns with MI and embodied simulation perspectives in consumer research [51], which assume that evocative

stimuli enable individuals to mentally rehearse future experiences [52], thereby increasing affective engagement and decision confidence [26]. In digital tourism environments, IP appears to bridge abstract online representations and anticipated physical experiences, both cognitively and emotionally. This mechanism also aligns with place attachment, defined as the emotional bond individuals develop toward destinations through sensory interaction, mental associations, and memorable experiences [53]. Digitally induced multisensory imagery may represent an early stage in place attachment formation, occurring before physical visitation. In S-O-R terms, the findings locate the organism stage in the experiential variables MI and IP rather than in the visual stimulus itself: it is the extent to which consumers imagine sensory qualities and feel present in the destination, rather than mere exposure to a richer image, that is associated with their behavioral response. This is the central answer to the research objective: in digital tourism commerce, PI appears to be formed through an experiential sequence in which sensory anticipation feeds the mental simulation of being there, with the image acting as the trigger of that sequence rather than as its explanation.

From a sensory marketing perspective, the study's strong gustatory and olfactory effects are particularly noteworthy. Visual cues associated with food, steam, texture, and atmospheric warmth were accompanied by large increases in taste- and smell-related imagery, with tactile imagery, for which fewer cues were embedded, responding least, particularly in the urban block. This finding supports cross-modal correspondence research suggesting that vision can reliably activate non-visual sensory representations through learned associations and perceptual inference [31]. Importantly, the present results indicate that multisensory engagement in digital tourism does not require immersive hardware technologies such as virtual reality or tactile interfaces [54]. Instead, relatively subtle visual elements embedded in static imagery can elicit meaningful sensory simulations [28], offering a potentially scalable approach for experiential destination communication, within the boundary conditions discussed herein.

Sensory enrichment produced consistent perceptual and intentional effects for the exotic context, whereas city-break imagery showed more limited responsiveness, with statistically meaningful change emerging primarily for PI. Qualitative responses help clarify this divergence. Exotic visuals elicited diffuse, atmospheric descriptions related to relaxation, breeze, sunlight, and bodily comfort, reflecting a holistic form of sensory immersion. In contrast, city-break imagery triggered more localized, concrete associations centered on coffee, breakfast rituals, and the smells and tastes of urban environments.

Together, the patterns suggest two distinct modes of digitally mediated multisensory processing: an atmospheric mode characteristic of exotic destinations, marked by embodied environmental immersion, and an object-based mode in urban contexts, anchored in specific sensory cues such as coffee or food. While urban sensory stimuli appear easier to verbalize because of their concreteness, exotic imagery seems to activate broader experiential states that are better captured by quantitative measures of multisensory imagery and IP. The distinction offers a novel conceptual lens for understanding how different destination typologies engage the senses online.

The present research advances the tourism and digital tourism commerce literature in several ways. First, it empirically operationalizes multisensory MI within a digital destination context, integrating olfactory, gustatory, and tactile dimensions into pre-consumption evaluation models. Although prior studies have acknowledged the multisensory nature of tourism experiences, quantitative examinations of digitally mediated sensory anticipation remain scarce [13,15]. This expanded operationalization not only advances theoretical understanding but also equips tourism businesses with actionable insights for designing

more persuasive digital content, enabling them to strategically leverage multisensory cues to strengthen consumer engagement and potentially raise PI for tourism products.

Second, the study shows how metaphorical design can systematically embed sensory information in visual content, providing a practical framework for activating non-visual senses in static digital media. This extends sensory marketing research by showing that cross-modal cues can be deliberately engineered at the level of visual composition, enabling tourism businesses to strategically enhance the experiential richness of their digital touchpoints and strengthen the persuasive power of their promotional materials in ways that may support customer engagement and, ultimately, PI; whether these effects lead to conversion requires behavioral validation.

Third, the empirically supported pathway linking multisensory MI, IP, and PI offers correlational evidence consistent with experiential consumption models that emphasize mental simulation as a driver of tourist decision-making [26,55]. By positioning IP as a mediating construct, the study clarifies how sensory perception translates into behavioral intention in online tourism environments. Together, these findings underscore the need to advance empirical research on digitally mediated sensory processing, as understanding this experiential pathway is essential to explaining how online destination cues shape tourist decision-making in increasingly visual and competitive digital environments.

The findings also carry managerial and entrepreneurial implications. For tourism marketers and destination managers, the findings suggest that visual aesthetics alone cannot convey experiential richness. Digital touchpoints, such as destination websites or apps, social media content, and booking platforms, can be enhanced by incorporating sensory metaphors that implicitly cue smell, taste, and touch alongside visual appeal. Practical applications include emphasizing texture in food imagery, conveying temperature through lighting and color, suggesting aromas via contextual elements (e.g., steam or proximity shots), and visually communicating comfort or materiality in accommodation visuals. For entrepreneurs, the central implication is one of access: because generative AI produces photorealistic, sensory-enriched imagery at negligible marginal cost [56], experiential visual communication is no longer the preserve of destination management organizations with professional production budgets. Small and resource-constrained ventures (guesthouses, family-run restaurants, independent tour operators, rural and emerging destinations) can embed sensory cues into their own promotional content without photographers, studios, or immersive hardware, narrowing a differentiation gap that previously favored incumbents. The findings also point to opportunities for new ventures, such as specialized services that design, test, and localize sensory-enriched content for the hospitality sector, and travel-tech startups that integrate sensory-cue generation into booking or listing platforms. The low cost of producing stimulus variants additionally makes an experimentation-driven approach feasible even for microenterprises: founders can iterate sensory designs and measure response directly rather than committing to a single campaign. Such strategies can strengthen emotional engagement and stated purchase intention without additional technological infrastructure; because the study measured intention rather than actual bookings or conversions, claims about downstream purchase behavior would be premature, and these design recommendations should be validated in field settings, for example through A/B tests on live booking platforms, a validation step that the same low production costs place within the reach of small firms as well as large ones. Sensory-enriched visuals may be particularly valuable in highly competitive destination markets, where differentiation increasingly depends on experiential rather than informational attributes. Moreover, the identified differences between exotic and urban contexts suggest that sensory design strategies should adapt to destination typology, prioritizing atmospheric immersion for leisure destinations and object-centered cues for

city environments, a contingency that entrepreneurs targeting a specific niche can exploit more quickly than broad-portfolio incumbents.

Before turning to the limitations, we must distinguish these findings from a systematic second-stimulus effect, since the sensory-enriched visual always followed the conventional one. A generic position effect, whether produced by contrast with the preceding image, by response drift, or by respondents inferring that the second image is the one of interest [10,11,57], would leave a clear observable signature: it would elevate ratings of the second image on all outcomes and in all blocks, regardless of what the image depicts. The data do not show this signature. First, within the city-break block, IP and perceived vividness did not differ between the two images, and only PI changed; a position effect operating on rating behavior would not spare two of the three outcomes measured with the same items immediately after the same image. Second, the size of the effect tracks the content of the manipulation: the sense-specific imagery items, which the embedded cues directly target, moved by an order of magnitude more (d_z = 1.69–2.44) than the experiential outcomes ($d_z \leq 0.13$), and within the imagery items, the gustatory and olfactory dimensions (matching the food-, steam-, and aroma-related cues) responded most, while tactile imagery, for which fewer cues were embedded, responded least in the urban block.

A uniform second-stimulus advantage would not produce this graded, content-matched profile; specificity and coherence of this kind are recognized grounds for preferring a treatment explanation over a design artifact when randomization is unavailable [12,39,40]. Third, open-ended responses to enriched visuals spontaneously referred to smells, tastes, and temperatures tied to the specific cues shown, rather than to the images being generically better or more appealing. Finally, the biases most often associated with fixed sequences pull in opposite directions: habituation and fatigue depress responses to later stimuli and would therefore bias the paired comparison against the hypothesis, whereas contrast and inferred-demand effects would inflate it. We also note that perceived vividness itself increased in the exotic block, so the manipulation should be understood as enriching the stimulus as a whole, its visual surface together with its cross-modal associations, rather than as adding purely non-visual information; the selectivity arguments above address position effects, not the internal decomposition of the enrichment. Because the pattern across outcomes, senses, and blocks is selective rather than uniform, we consider a systematic second-stimulus effect an unlikely account of the results, while recognizing that only randomized or counterbalanced presentation can rule it out conclusively.

Several limitations should be acknowledged. First, generalizability is limited by the sample composition: respondents were recruited through non-probability convenience sampling on social media, were culturally homogeneous, and were predominantly women aged 45–55 (94.9% female), so the findings characterize an educated, digitally engaged consumer segment rather than tourism consumers at large, and replication in gender-balanced and cross-cultural samples is needed. Second, all stimuli were AI-generated; perceived authenticity, trust in synthetic imagery, and the correspondence between the generated scene and the actual destination are potential boundary conditions [32] that were not measured and may moderate the effects, particularly as consumers become more aware of, and more skeptical toward, AI-produced marketing content. Third, only one image pair represented each destination. This was a deliberate trade-off. The paired A/B logic required each sensory-enriched image to be an augmentation of its own conventional baseline, with scene, composition, framing, and palette held constant (Section 3.1), and every additional matched pair would have lengthened an unsupervised online session in which respondents already rated four stimuli in succession, increasing fatigue and dropout; destination type was therefore treated as a within-subject replication context rather than as a factor sampled across stimuli. The cost is that, with a single exemplar

per cell, destination type is fully confounded with the idiosyncratic properties of the two images used, such as the attractiveness of the base scene and the number, salience, and kind of sensory cues that could plausibly be embedded in it. The exotic–urban differences reported above, including the atmospheric versus object-based modes proposed earlier in this section, therefore cannot be separated from image-specific effects and should be read as hypotheses generated by this image pair rather than as established destination-type effects. Likewise, the inferential statistics warrant generalization to new respondents viewing these images, not to new images of the same destination type [58]; establishing the latter requires stimulus sampling across several exemplars per destination and condition, as noted among the future research directions below. Fourth, DI, IP, and PI were measured with single items (item wording in Section 3.3); although single items perform acceptably for concrete constructs [42,43], internal-consistency reliability cannot be estimated, construct nuances are not captured, and the very high correlation between IP and PI (r = 0.82–0.86) may partly reflect this measurement strategy together with common method variance [44], so discriminant validity between these constructs remains to be demonstrated with multi-item instruments. Fifth, the sensory-imagery items were strongly floored in the control condition, which limits the interpretability of standardized effect sizes for these items and precludes a reliable control-condition composite. Sixth, the study relied on self-reported measures and static visual stimuli, and it examined PI rather than actual purchase behavior.

Future research could incorporate physiological indicators, eye-tracking, behavioral outcomes such as click-through or booking data, or immersive media formats to further elucidate sensory engagement mechanisms. Longitudinal designs would also help determine whether digitally induced multisensory imagery translates into actual travel behavior. Acknowledging these limitations highlights the need for continued empirical investigation into digitally mediated multisensory processing, as further methodological refinement is essential to advancing theoretical understanding and strengthening the practical relevance of sensory-based strategies in digital tourism communication.

A further limitation concerns the fixed order of stimulus presentation. Because the sequence was constant, with the conventional visual always shown before the sensory-enriched visual within each destination block, the design cannot statistically separate the effect of the sensory manipulation from possible order, contrast, or priming effects, which are a recognized concern in within-subject research [33,39,59]. Randomization was not implemented because the study was designed as a baseline-referenced paired evaluation in which every respondent judged the enhanced visual against the same conventional standard (Section 3.1); however, counterbalancing the order between respondents would have preserved this logic while allowing the position effect to be estimated directly, and we regard its absence as a genuine weakness. The internal checks presented herein indicate that the observed effects are selective across outcomes, sensory modalities, and destination blocks rather than uniform, and that the biases most likely to arise under a fixed sequence, namely habituation and fatigue, would bias the comparison against the hypothesis, so that the small effects observed (Cohen's $d_z \leq 0.13$ for the core outcomes) may be conservative. These observations mitigate but do not remove the concern. Accordingly, the differences reported here should be interpreted as within-subject associations under a fixed baseline-then-treatment sequence, rather than as isolated causal effects of the sensory cues, and the mediation results should be read as evidence that the data are consistent with the proposed sequential pathway rather than as confirmation of it. Future studies should implement counterbalanced or randomized sequences, ideally complemented by a between-subjects arm, to quantify residual position effects and strengthen causal inference.

Finally, future work may explore individual difference variables, such as sensory sensitivity, travel motivation, or prior destination familiarity, as potential moderators of

responses to sensory-enriched digital content. Extending the framework to additional destination types and cultural settings would further clarify how multisensory imagery shapes tourism decision-making in digital environments. Four avenues appear especially promising: (1) counterbalanced or randomized presentation with a between-subjects arm, to quantify position effects and test first-exposure generalization; (2) multi-item measurement of IP and PI, with procedural and statistical remedies for common method variance, to establish discriminant validity; (3) stimulus sampling across multiple image exemplars per destination and per condition, ideally with pre-tested perceptual matching, to separate destination-type effects from image-specific characteristics; and (4) moderation analyses of perceived authenticity and trust in AI-generated imagery, which are likely to shape the persuasive value of synthetic sensory cues as disclosure of AI use becomes standard practice.

6. Conclusions

This article set out to investigate, theoretically and empirically, how multisensory MI, IP, and PI are related in digital tourism commerce and how digitally conveyed sensory cues are associated with consumers' intention to purchase tourism products online; the paired-image experiment was the empirical means to this end, not an end in itself. With respect to this objective, the study provides empirical evidence that visually encoded sensory metaphors can meaningfully enrich digital tourism experiences by activating multisensory MI and, for the exotic destination in particular, by being associated with stronger IP and higher PI. Compared with conventional destination visuals, sensory-enriched imagery elicited stronger olfactory, gustatory, and tactile associations, with gustatory and olfactory imagery emerging as the most responsive dimensions. These effects were accompanied by higher experiential vividness and greater behavioral motivation, primarily in exotic destination contexts. Beyond documenting main effects, the findings clarify an experiential pathway through which sensory design operates in digital environments. Visually induced multisensory imagery facilitated IP, which in turn was the strongest correlate of PI. The mediation analysis indicates that IP is not merely an accompanying experiential state but a statistically significant mediator linking MI to behavioral intention; given the fixed-order within-subject design, this pattern should be interpreted as consistent with, rather than demonstrating, a causal sequence. This sequential process underscores the central role of mental simulation in pre-consumption decision-making and demonstrates how carefully designed visual cues can partially compensate for the absence of physical sensory input in online tourism communication.

The results also highlight important contextual nuances. Exotic destinations primarily evoked diffuse, atmospheric forms of sensory immersion linked to bodily relaxation and environmental qualities, whereas city-break imagery activated more localized, object-centered sensory associations, particularly around food and coffee-related cues. This distinction suggests that different destination typologies engage distinct modes of multisensory processing, with implications for tailoring sensory metaphors across tourism contexts. Theoretically, this research extends DI and sensory marketing frameworks by empirically integrating olfactory, gustatory, and tactile imagery into digital pre-consumption evaluation models. In doing so, it specifies, within the S-O-R logic, an experiential sequence (MI $\rightarrow$ IP $\rightarrow$ PI) through which online destination cues are associated with consumer decision-making in digital tourism commerce; the AI-generated image pair served as the vehicle for varying sensory input, whereas the contribution lies in the pathway it allowed us to test. By showing that non-visual sensory cues can be systematically embedded in static visuals, the study points to a potentially scalable complement to immersive technologies, pending validation in real purchase environments, and advances understanding of how experiential meaning is constructed before physical arrival. Practically, the findings provide actionable

guidance for tourism marketers, destination managers, and entrepreneurs. Rather than relying solely on visual aesthetics or hardware-based solutions, marketers can enhance digital touchpoints with strategic sensory metaphors that suggest aroma, texture, warmth, or taste. Such design choices can transform conventional online content into richer experiential interfaces, strengthening emotional connection and destination appeal in competitive digital marketplaces.

Although the sample was skewed in terms of gender (it was predominantly women), the perceptual mechanisms observed here point to broader relevance for digitally mediated tourism experiences. Future research should examine these processes across more diverse populations, media formats, and destination types, and explore whether digitally evoked MI explains actual travel behavior over time. In conclusion, the study underscores the value of multisensory thinking in digital tourism design and contributes to a deeper understanding of how sensory imagination and IP shape tourist decision-making in the pre-consumption stage, before the journey even begins. This integrative perspective highlights the importance of designing digital tourism experiences that deliberately activate multisensory imagination, enabling businesses to create more compelling, emotionally resonant digital environments that support consumer decision-making well in advance of actual travel. Because the evidence concerns stated intentions rather than realized purchases, future field research should validate these design recommendations against behavioral outcomes.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/jtaer21100378/s1, descriptive statistics; paired-test results with adjusted *p*-values under the Holm, Benjamini–Hochberg, and Bonferroni procedures, effect-size confidence intervals, and nonparametric checks; Pearson and Spearman correlation matrices for each destination × condition combination; full regression coefficient tables for the serial mediation models; bootstrap indirect-effect estimates; and the within-participant mediation results.

Author Contributions: Conceptualization, M.C.D.; methodology, M.C.D. and C.N.; software, M.C.D.; validation, O.-E.P. and V.-M.Ş.-D.; formal analysis, M.C.D. and C.N.; investigation, O.-E.P. and V.-M.Ş.-D.; resources, M.C.D.; data curation, M.C.D.; writing—original draft preparation, M.C.D.; writing—review and editing, C.N., O.-E.P. and V.-M.Ş.-D.; visualization, M.C.D. and C.N.; supervision, M.C.D. and C.N.; project administration, M.C.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study because no sensitive information was collected and the personal data of all participants included in the study were fully anonymized.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tanriverdi, H.; Ciki, K.D.; Uygun, N.; Ciki, S. Narratives of medical tourists travelling for hair transplant: A case of Turkiye. *Tour. Recreat. Res.* **2025**, *51*, 374–386. [CrossRef]
2. Tanriverdi, H.; Ciki, K.D.; Ogretmenoglu, M.; Uygun, N. Unveiling the key dimensions shaping visitor experiences in cave tourism through a content analysis approach: The case of Antalya's show caves. *Curr. Issues Tour.* **2025**, *29*, 4051–4066. [CrossRef]
3. Mehrabian, A.; Russell, J.A. *An Approach to Environmental Psychology*; MIT Press: Cambridge, MA, USA, 1974.

4. Yang, W.; Chen, Q.; Huang, X.; Xie, M.; Guo, Q. How do aesthetics and tourist involvement influence cultural identity in heritage tourism? The mediating role of mental experience. *Front. Psychol.* **2022**, *13*, 990030. [CrossRef] [PubMed]
5. Echtner, C.M.; Ritchie, J.R.B. The meaning and measurement of destination image. *J. Tour. Stud.* **2003**, *14*, 37–48.
6. Pearce, P.L. *The Ulysses Factor*; Springer: New York, NY, USA, 1988.
7. Baber, R.; Baber, P. Influence of social media marketing efforts, e-reputation and destination image on intention to visit among tourists: Application of S-O-R model. *J. Hosp. Tour. Insights* **2023**, *6*, 2298–2316. [CrossRef]
8. Wang, M.; Xiao, X.; Li, P.; Zhang, H.; Zhang, J. The magic of audio-visual integration in simulated crowding images of culture-oriented tourism destinations: An eye-tracking analysis on tourists' visual attention and behavioural intentions. *Curr. Issues Tour.* **2025**, *29*, 3812–3831. [CrossRef]
9. Çıkı, S. The Application of SOR Theory in Tourism Research: Mapping WOS Data-Based Research. *Pr. I Stud. Geogr.* **2025**, *70*, 21–36. [CrossRef]
10. Strack, F. "Order effects" in survey research: Activation and information functions of preceding questions. In *Context Effects in Social and Psychological Research*; Schwarz, N., Sudman, S., Eds.; Springer: Berlin/Heidelberg, Germany, 1992; pp. 23–34. [CrossRef]
11. Zizzo, D.J. Experimenter demand effects in economic experiments. *Exp. Econ.* **2010**, *13*, 75–98. [CrossRef]
12. Zhu, J.; Zhan, L.; Tan, J.; Cheng, M. Tourism destination stereotypes and generative artificial intelligence (GenAI) generated images. *Curr. Issues Tour.* **2025**, *28*, 2721–2725. [CrossRef]
13. Agapito, D.; Valle, P.; Mendes, J. The sensory dimension of tourist experiences: Capturing meaningful sensory-informed themes in Southwest Portugal. *Tour. Manag.* **2014**, *42*, 224–237. [CrossRef]
14. Buhalis, D.; Law, R. Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tour. Manag.* **2008**, *29*, 609–623. [CrossRef]
15. Xiong, J.; Hashim, N.H.; Murphy, J. Multisensory image as a component of destination image. *Tour. Manag. Perspect.* **2015**, *14*, 34–41. [CrossRef]
16. Trang, N.T.; Yoo, J.J.-E.; Joo, D.; Lee, G. Incorporating senses into destination image. *J. Destin. Mark. Manag.* **2023**, *27*, 100760. [CrossRef]
17. Pike, S. Destination image analysis—A review of 142 papers from 1973 to 2000. *Tour. Manag.* **2002**, *23*, 541–549. [CrossRef]
18. Tasci, A.D.A.; Gartner, W.C. Destination image and its functional relationships. *J. Travel Res.* **2007**, *45*, 413–425. [CrossRef]
19. Baloglu, S.; McCleary, K.W. A model of destination image formation. *Ann. Tour. Res.* **1999**, *26*, 868–897. [CrossRef]
20. Krishna, A. An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *J. Consum. Psychol.* **2012**, *22*, 332–351. [CrossRef]
21. Lindstrom, M. *Brand Sense: Build Powerful Brands Through Touch, Taste, Smell, Sight, and Sound*; Free Press: New York, NY, USA, 2005.
22. MacInnis, D.J.; Price, L.L. The role of imagery in information processing: Review and extensions. *J. Consum. Res.* **1987**, *13*, 473–491. [CrossRef] [PubMed]
23. Steuer, J. Defining Virtual Reality: Dimensions Determining Telepresence. *J. Commun.* **1992**, *42*, 73–93. [CrossRef]
24. Lombard, M.; Ditton, T. At the Heart of It All: The Concept of Presence. *J. Comput.-Mediat. Commun.* **1997**, *3*, JCMC321. [CrossRef]
25. Tussyadiah, I.P.; Wang, D.; Jung, T.; Tom Dieck, M.C. Virtual Reality, Presence, and Attitude Change: Empirical Evidence from Tourism. *Tour. Manag.* **2018**, *66*, 140–154. [CrossRef]
26. Kim, M.; Kim, J.-H.; Park, M.; Yoo, J. The roles of sensory perceptions and mental imagery in consumer decision-making. *J. Retail. Consum. Serv.* **2021**, *61*, 102517. [CrossRef]
27. Agapito, D. The senses in tourism design: A bibliometric review. *Ann. Tour. Res.* **2020**, *83*, 102934. [CrossRef]
28. Elder, R.S.; Krishna, A. The visual depiction effect in advertising: Facilitating embodied mental simulation through product orientation. *J. Consum. Res.* **2012**, *38*, 988–1003. [CrossRef]
29. Dițoiu, M.-C.; Căruntu, A.L. Sensory experiences regarding five-dimensional brand destination. *Procedia Soc. Behav. Sci.* **2014**, *109*, 301–306. [CrossRef]
30. Lakoff, G.; Johnson, M. *Metaphors We Live by*; University of Chicago Press: Chicago, IL, USA, 1980.
31. Spence, C. Managing sensory expectations concerning products and brands: Capitalizing on the potential of sound and shape symbolism. *J. Consum. Psychol.* **2012**, *22*, 37–54. [CrossRef]
32. Campbell, C.; Plangger, K.; Sands, S.; Kietzmann, J. Preparing for an era of deepfakes and AI-generated ads: A framework for understanding responses to manipulated advertising. *J. Advert.* **2022**, *51*, 22–38. [CrossRef]
33. Greenwald, A.G. Within-subjects designs: To use or not to use? *Psychol. Bull.* **1976**, *83*, 314–320. [CrossRef]
34. Charness, G.; Gneezy, U.; Kuhn, M.A. Experimental methods: Between-subject and within-subject design. *J. Econ. Behav. Organ.* **2012**, *81*, 1–8. [CrossRef]
35. Keren, G. Between- or within-subjects design: A methodological dilemma. In *A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues*; Keren, G., Lewis, C., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 1993; pp. 257–272.
36. Reips, U.-D. Standards for Internet-based experimenting. *Exp. Psychol.* **2002**, *49*, 243–256. [CrossRef]

37. Morales, A.C.; Amir, O.; Lee, L. Keeping it real in experimental research—Understanding when, where, and how to enhance realism and measure consumer behavior. *J. Consum. Res.* **2017**, *44*, 465–476. [CrossRef]
38. Birnbaum, M.H. How to show that 9 > 221: Collect judgments in a between-subjects design. *Psychol. Methods* **1999**, *4*, 243–249. [CrossRef]
39. Shadish, W.R.; Cook, T.D.; Campbell, D.T. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*; Houghton Mifflin: Boston, MA, USA, 2002.
40. Cook, T.D.; Campbell, D.T. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*; Rand McNally: Rosemont, IL, USA, 1979.
41. Witmer, B.G.; Singer, M.J. Measuring presence in virtual environments: A presence questionnaire. *Presence Teleoperators Virtual Environ.* **1998**, *7*, 225–240. [CrossRef]
42. Bergkvist, L.; Rossiter, J.R. The Predictive Validity of Multiple-Item versus Single-Item Measures of the Same Constructs. *J. Mark. Res.* **2007**, *44*, 175–184. [CrossRef]
43. Fuchs, C.; Diamantopoulos, A. Using single-item measures for construct measurement in management research: Conceptual issues and application guidelines. *Die Betriebswirtschaft* **2009**, *69*, 195–210.
44. Podsakoff, P.M.; MacKenzie, S.B.; Lee, J.-Y.; Podsakoff, N.P. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *J. Appl. Psychol.* **2003**, *88*, 879–903. [CrossRef] [PubMed]
45. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Front. Psychol.* **2013**, *4*, 863. [CrossRef] [PubMed]
46. Holm, S. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* **1979**, *6*, 65–70.
47. Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **1995**, *57*, 289–300. [CrossRef]
48. Judd, C.M.; Kenny, D.A.; McClelland, G.H. Estimating and testing mediation and moderation in within-subject designs. *Psychol. Methods* **2001**, *6*, 115–134. [CrossRef] [PubMed]
49. Montoya, A.K.; Hayes, A.F. Two-condition within-participant statistical mediation analysis: A path-analytic framework. *Psychol. Methods* **2017**, *22*, 6–27. [CrossRef] [PubMed]
50. Cronbach, L.J. Coefficient alpha and the internal structure of tests. *Psychometrika* **1951**, *16*, 297–334. [CrossRef]
51. Barsalou, L.W. Grounded cognition. *Annu. Rev. Psychol.* **2008**, *59*, 617–645. [CrossRef] [PubMed]
52. Kosslyn, S.M.; Ganis, G.; Thompson, W.L. Neural foundations of imagery. *Nat. Rev. Neurosci.* **2001**, *2*, 635–642. [CrossRef] [PubMed]
53. Stăncioiu, A.-F.; Dițoiu, M.-C. Place Attachment. In *Encyclopedia of Tourism*, 2nd ed.; Jafari, J., Xiao, H., Eds.; Springer Nature: Berlin/Heidelberg, Germany, 2024; pp. 796–797. [CrossRef]
54. Slater, M.; Wilbur, S. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence Teleoperators Virtual Environ.* **1997**, *6*, 603–616. [CrossRef]
55. Slusariuc, G.C.; Simina, A.E. Evolutions of Rural Tourism in Romania and Europe. *Ovidius Univ. Ann. Econ. Sci. Ser.* **2018**, *18*, 182–185.
56. Hartmann, J.; Exner, Y.; Domdey, S. The power of generative marketing: Can generative AI create superhuman visual marketing content? *Int. J. Res. Mark.* **2025**, *42*, 13–31. [CrossRef]
57. Tourangeau, R.; Rips, L.J.; Rasinski, K. *The Psychology of Survey Response*; Cambridge University Press: Cambridge, UK, 2000. [CrossRef]
58. Wells, G.L.; Windschitl, P.D. Stimulus sampling and social psychological experimentation. *Personal. Soc. Psychol. Bull.* **1999**, *25*, 1115–1125. [CrossRef]
59. Poulton, E.C. Unwanted range effects from using within-subject experimental designs. *Psychol. Bull.* **1973**, *80*, 113–121. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.