Proceedings of the 8th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2026.36865

Generative Artificial Intelligence Driven Product Form Innovation Design Method

Chenlu Zhou

Tianjin Chengjian University, Tianjin, China
zhouchenlu02569@outlook.com

Abstract. Generative artificial intelligence provides rapid visual exploration for product design, but prompt-driven image generation alone often produces attractive concepts with weak semantic control, limited geometric feasibility, or excessive similarity to existing products. This study proposes a generative artificial intelligence driven product form innovation design method that couples semantic requirement encoding, geometry-constrained diffusion generation, novelty-aware candidate selection, and designer-guided refinement. Product requirements are translated into weighted form semantics and structural constraints. Stable Diffusion XL, ControlNet, and IP-Adapter are then combined to generate candidates while preserving functional silhouette and reference-level style information. A multi-objective form innovation score integrates semantic consistency, geometric novelty, manufacturing feasibility, and intra-set diversity. The method was validated through a desktop air-purifier form design experiment involving 36 designers and 12 expert evaluators. The proposed method achieved a semantic alignment score of 0.812 ± 0.031, a geometric novelty score of 0.384 ± 0.026, and an expert innovation rating of 6.12 ± 0.41, while reducing concept-development time to 18.6 ± 3.4 min. The results indicate that controlled generative exploration can improve product-form innovation without sacrificing structural coherence or design feasibility

Keywords: generative artificial intelligence, product form design, design innovation, diffusion model, human-AI co-design

1. Introduction

The application of artificial intelligence in design has advanced, but it has shifted from analytics-oriented automation to generation and exploration. Recent mapping of design research oriented towards artificial intelligence suggests that artificial intelligence has expanded from a single stage of design process to the whole design process including problem framing, ideation, representation, evaluation, and decision support [1]. This is even more true for generative AI because its value lies in extending the range of the artifacts which could be generated based on an incomplete or qualitative requirements. Generativity, from the perspective of design science, means ability to create previously unspecified solutions with meaningful design qualities [2].

It has significant importance for product form innovation. Early stages of industrial design require designers to explore the ratio, contour, balance, semantics, affordance, and manufacturing

DOI: 10.54254/2755-2721/2026.36865

aspects of the form at once. Artificial intelligence could enable combinational, exploratory, and transformational creativity through fast translation of textual and visual inputs into different design representations [3]. Moreover, the impact of generative AI goes beyond industrial design, as multimodal generation allows to reduce time of concept exploration and increase the number of potential solutions before engineering starts [4]. Therefore, large language models and multimodal artificial intelligence are currently investigated for product design and manufacturing applications, despite the fact that domain-independent models struggle with engineering constraints and physical reasoning [5].

Text-to-image-based workflows currently available are not sufficient for systematic product form innovation. An image generated from a prompt that is too novel can be non-functional, whereas strong conditional input can result in very conservative design forms. Generation of an image solely by prompt does not allow sufficient control over translation of semantics into form properties. For this reason, a coupled semantic-geometric workflow for design exploration is proposed in this study. This approach considers output of generative AI models as constrained design candidates instead of images. This method includes steps of requirement encoding, controlled diffusion, novelty estimation, manufacturing-oriented filtration, and designer selection.

2. Literature review

In recent years, product form studies have shifted from free image generation towards GenAI-assisted generation workflow. Generative AI has been paired with semantic image analysis, reference form extraction, Bézier curve editing, and CAD-based design refinement to assist the development of form for cars [6]. Additionally, diffusion models have been used in combination with Kansei engineering to link the generated appearance options with measurable emotional requirements, proving that generative models can be a part of the controlled product form design process rather than mere visualization tool [7]. The other research direction deals with the trade-off between the computation and qualitative goals. Controlled experiments with generative design tools have shown that the computational exploration can increase the size of objective space and aesthetic diversity; however, qualitative objectives still strongly rely on human evaluation [8]. Comparisons of AIGC with traditional ideation show that the computational design assistance does not automatically lead to greater creativity. Image quality, task specifics, and use of generated stimuli play an important role in the concepts generated by AIGC. It is crucial since product innovation involves both novelty and feasibility [9]. The same applies to recent investigations of AI-assisted design techniques – different problem-solving strategies lead to different levels of novelty, utility, and feasibility of design concepts [10]. Thus, the current literature indicates that the main research problem is not how well GenAI can generate product images anymore. The proposed methodology addresses this issue by providing a coupled generation mechanism.

3. Method

3.1. Semantic requirement and product form encoding

The first step of the proposed procedure is the separation of the design brief into semantic intentions and structural requirements that cannot be negotiated. While semantic intentions represent perceptual qualities of the product, structural requirements constrain geometry necessary for use and fabrication. In the validation case, user comments such as "clean," "quiet," "lightweight," "soft," and "technological" were mapped into a weighted semantic vector. The constrained language-model

DOI: 10.54254/2755-2721/2026.36865

parser yielded only adjective-object pairs with associated importance weights, so that the uncontrolled prompts could not directly affect the image generator.

The semantic layer has been mapped into the set of form variables that included global aspect ratio, contour curvature, corner softness, visual center, surface segmentation, opening rhythm, symmetry, and perceived mass. Structural variables comprised the allowed footprint size, air-inlet zone, outlet area, control interface region, base contact surface, minimal edge radius, and maximal extension from the primary envelope. Those variables formed the product-form control code. While the code did not specify any particular geometry, it limited the range within which the generative search can take place. This point is critical because the overly restrictive geometric conditioning limits the diversity, while unrestricted text conditioning typically generates visually plausible but functional nonsense.

A set of 2,460 normalized product images has been assembled into a reference archive from publicly available product catalogs and design award datasets. All backgrounds have been stripped, and all images have been rotated to the standard three-quarters view. The perceptual semantics was represented by CLIP ViT-L/14 embeddings, and geometric appearance – by DINOv2 embeddings. Additionally, canny contours and binary silhouettes have been generated.

3.2. Semantic-geometric coupled generative strategy

Candidate generation used Stable Diffusion XL 1.0 as the visual backbone. ControlNet-Canny constrained the principal silhouette, while IP-Adapter Plus supplied low-strength visual conditioning from selected reference forms. The two conditioning channels were deliberately assigned different roles. ControlNet preserved basic product structure, whereas IP-Adapter transferred only distributed form characteristics rather than reproducing a complete reference. Text conditioning contained the weighted semantic descriptors together with material-neutral industrial-design terms. The classifier-free guidance scale was fixed at 6.5, 40 denoising steps were used, and the ControlNet conditioning strength was set to 0.72. IP-Adapter strength was limited to 0.35 to prevent reference imitation.

Each generation cycle produced 24 candidates from six random latent seeds. After the first cycle, candidates that violated the predefined silhouette envelope or obscured functional air paths were removed. The surviving designs were embedded and distributed across the semantic-geometric space. Four representative candidates were then selected from different regions and used as weak visual references for the second generation cycle. This operation produced controlled divergence: semantic consistency was retained, but the generation trajectory moved away from a single seed or reference family.

The core candidate-ranking function was defined as shown in Equation (1):

$$S(x \mid t, R, S_k) = \alpha \frac{\mathbf{e}_t^{\mathrm{T}} \mathbf{e}_x}{\|\mathbf{e}_t\|_2 \|\mathbf{e}_x\|_2 + \varepsilon} + \beta \left[1 - \max_{r \in \mathbb{R}} \frac{\mathbf{z}_x^{\mathrm{T}} \mathbf{z}_r}{\|\mathbf{z}_x\|_2 \|\mathbf{z}_r\|_2 + \varepsilon}\right] + \tag{1}$$

$$\gamma exp\left[-\sum_{m=1}^{M} \left(\frac{\max(0, g_m(x))}{\sigma_m}\right)^2\right] + \delta\left[1 - \max_{y \in S_k} \frac{\mathbf{z}_x^{\mathrm{T}} \mathbf{z}_y}{\|\mathbf{z}_x\|_2 \|\mathbf{z}_y\|_2 + \varepsilon}\right] - \eta P_c(x)$$

Here, x denotes a generated candidate, t is the semantic target, $\mathscr{R}$ is the reference archive, and $\mathscr{S}_k$ is the set of previously selected candidates. The vectors $\mathbf{e}$ and $\mathbf{z}$ denote CLIP and DINOv2 representations. The first term measures semantic correspondence. The second rewards distance from existing products. The exponential term represents geometric feasibility through normalized constraint violations $g_m(x)$. The fourth term prevents selected alternatives from collapsing into the same form family. $P_c(x)$ penalizes contour discontinuities, blocked openings, disconnected components, and abnormal projections. The weights were set to 0.30, 0.25, 0.25, 0.15, and 0.05 after pilot calibration. As shown in Figure 1, the proposed method couples semantic requirements and structural constraints with controlled generative modeling, multi-objective innovation evaluation, and designer-guided local regeneration to support iterative product form exploration and refinement.

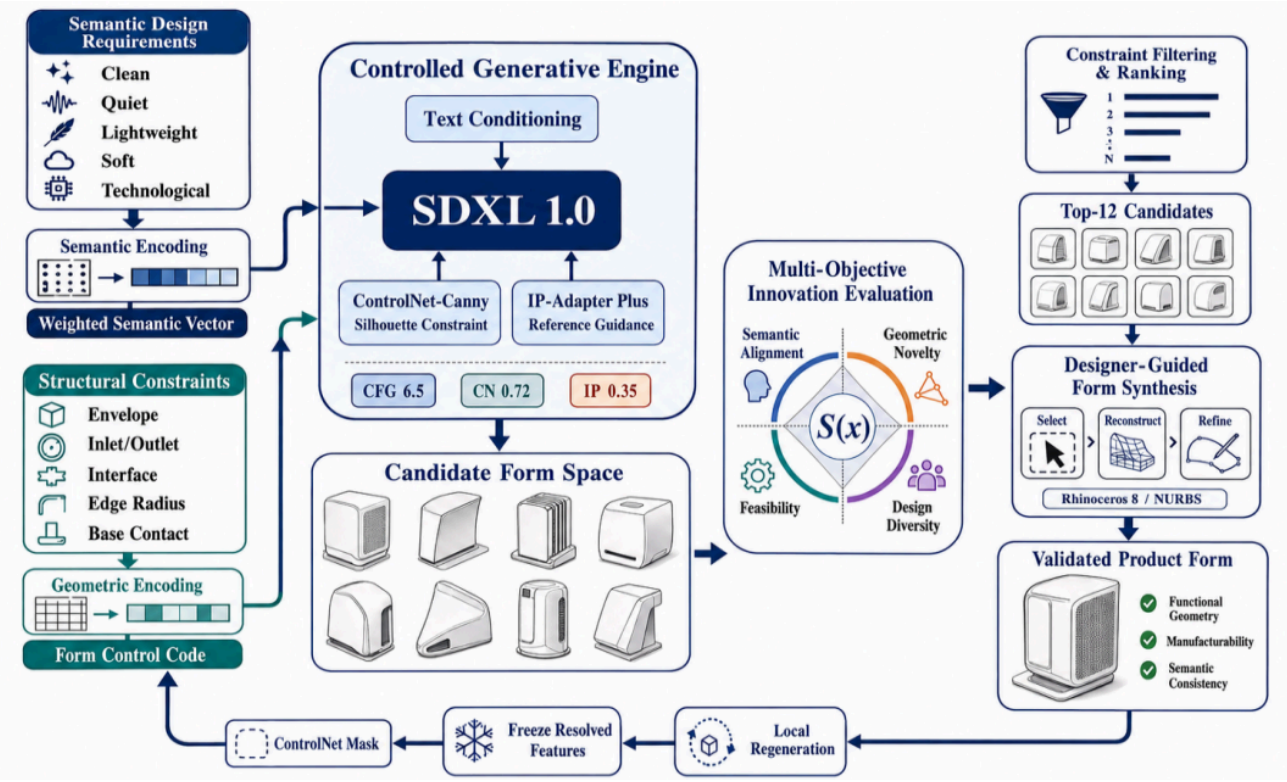

Figure 1. Semantic-geometric coupled generative AI framework for product form innovation design

3.3. Designer-guided form synthesis and refinement

The top 12 candidates were displayed on a semantic-novelty map rather than as a linear ranking. Designers selected three alternatives and reconstructed their main contour logic in Rhinoceros 8, transferring only proportion, curvature, segmentation, and interface placement into NURBS geometry while discarding AI-generated decoration.

A refinement loop allowed selected form features to remain fixed while weak regions were regenerated using ControlNet masks. This supported controlled form synthesis without altering resolved decisions. Final geometry was checked for minimum edge radius, stable base contact, unobstructed inlet and outlet zones, and manufacturable shell separation.

4. Experimental validation

4.1. Product design task and participants

A desktop air purifier was selected because its external form combines strong perceptual requirements with clear functional constraints. Three semantic targets were prepared: calm-clean, dynamic-technical, and soft-friendly. Thirty-six participants with at least two years of industrial-design or product-design training completed three controlled sessions. A Latin-square arrangement

DOI: 10.54254/2755-2721/2026.36865

counterbalanced the semantic target and design method to reduce order effects. Twelve industrial-design experts with a mean professional experience of 9.3 years independently evaluated the final concepts.

4.2. Design application procedure

Three design conditions were compared. The conventional condition used sketching and Rhinoceros without generative assistance. The prompt-only condition used the same SDXL model but accepted only textual prompts and had no geometry or novelty control. The proposed condition used the complete semantic-geometric workflow.

Each session began with a five-minute brief interpretation stage. Designers then generated or developed alternatives, selected one direction, and produced a standardized three-quarter-view concept. The maximum available time was 45 min. In the proposed condition, 48 first-cycle candidates were reduced automatically by geometric screening, after which 24 second-cycle candidates were produced. Twelve ranked alternatives were exposed to the designer. The selected concept was reconstructed as simplified NURBS geometry before final rendering, ensuring that evaluations were based on realizable form structure rather than uncontrolled image details.

4.3. Evaluation and statistical procedure

Automated evaluation included CLIP semantic alignment, DINOv2 archive distance for novelty, LPIPS-based within-set diversity, geometric-validity ratio, and task completion time. Experts rated form innovation, form coherence, semantic fit, and manufacturability on seven-point scales. Inter-rater consistency was evaluated with ICC(2,k). Repeated-measures ANOVA tested differences among the three methods. Greenhouse-Geisser correction was used when sphericity was violated, and Holm-adjusted paired comparisons were applied afterward. Partial eta squared was reported for effect magnitude, with significance set at 0.05.

5. Results

5.1. Product form quality

Expert agreement was high, with an ICC(2,k) of 0.861. The proposed method produced the highest innovation, semantic-fit, and form-coherence scores. Semantic alignment reached 0.812 ± 0.031, compared with 0.748 ± 0.042 for prompt-only generation and 0.691 ± 0.038 for conventional development. The method effect on semantic alignment was significant, with $F(2,70) = 72.84$, $p < 0.001$, and partial eta squared = 0.675. Novelty showed a similarly strong method effect, $F(2,70) = 58.17$, $p < 0.001$, partial eta squared = 0.624 (see Table 1).

Table 1. Comparative performance of the three product form design methods

Evaluation Metric	Conventional Design	Prompt-Only GenAI	Proposed Method	F	Partial η^2
Semantic alignment	0.691 ± 0.038	0.748 ± 0.042	0.812 ± 0.031**	72.84	0.675
Geometric novelty	0.278 ± 0.024	0.311 ± 0.031	0.384 ± 0.026**	58.17	0.624

DOI: 10.54254/2755-2721/2026.36865

Table 1. (continued)

Within-set diversity	0.302 ± 0.035	0.356 ± 0.032	0.423 ± 0.028**	49.63	0.586
Geometric validity	0.914 ± 0.026	0.612 ± 0.066	0.873 ± 0.041**	96.28	0.733
Expert innovation	4.89 ± 0.47	5.37 ± 0.52	6.12 ± 0.41**	61.45	0.637
Form coherence	5.72 ± 0.43	5.11 ± 0.58	6.08 ± 0.39**	38.91	0.526
Manufacturability	5.71 ± 0.46	4.76 ± 0.63	5.92 ± 0.44*	34.76	0.498
Completion time (min)	42.7 ± 6.5	14.2 ± 2.8	18.6 ± 3.4**	183.42	0.84

5.2. Efficiency and controlled exploration

However, the geometric validity of prompt-only generation dropped down to 0.612 ± 0.066. On the other hand, the new procedure took 18.6 ± 3.4 minutes to develop a concept with the geometric validity of 0.873 ± 0.041. Thus, compared to the traditional procedure, this was quite an effective way to reduce time spent on concepts' development without losing their quality as in the case of unrestricted generation. In addition, the greatest within-set diversity was provided by the proposed approach.

5.3. Ablation results

Removing semantic coupling reduced CLIP alignment from 0.812 ± 0.031 to 0.746 ± 0.036. Removing the novelty term reduced archive distance from 0.384 ± 0.026 to 0.319 ± 0.029, while removing geometric feasibility screening decreased expert manufacturability from 5.92 ± 0.44 to 4.81 ± 0.57. Eliminating the diversity term produced visually concentrated alternatives and lowered LPIPS diversity from 0.423 ± 0.028 to 0.351 ± 0.030. These results confirm that the improvement was produced by the coupled mechanism rather than diffusion generation alone.

6. Conclusion

This study develops a GenAI-driven product form innovation method that connects semantic design intention with controlled visual generation and geometry-oriented evaluation. Its central contribution is a coupled mechanism in which semantic correspondence, distance from existing products, structural feasibility, and candidate diversity jointly determine which AI-generated forms enter designer refinement. Experimental application to desktop air-purifier design shows that the method can produce more semantically consistent and innovative forms while maintaining manufacturability and substantially reducing concept-development time. The ablation results further demonstrate that unrestricted generative imagery is insufficient for reliable product innovation. GenAI becomes more useful when generation is treated as constrained design-space exploration and when designers retain control over selection, synthesis, and final geometric refinement.

References

[1] Figoli, F. A., Bruggeman, R., Rampino, L., & Ciuccarelli, P. (2025). AI-against-design map: A systematic review of 20 years of AI-focused studies in design research. *Design Studies, 96*, 101279.

DOI: 10.54254/2755-2721/2026.36865

[2] Bordas, A., Le Masson, P., Thomas, M., & Weil, B. (2024). What is generative in generative artificial intelligence? A design-based perspective. *Research in Engineering Design, 35*, 427–443.

[3] Furtado, L. S., Soares, J. B., & Furtado, V. (2024). A task-oriented framework for generative AI in design. *Journal of Creativity, 34*(2), 100086.

[4] Mariani, M., & Dwivedi, Y. K. (2024). Generative artificial intelligence in innovation management: A preview of future research developments. *Journal of Business Research, 175*, 114542.

[5] Mustapha, K. B. (2025). A survey of emerging applications of large language models for problems in mechanics, product design, and manufacturing. *Advanced Engineering Informatics, 64*, 103066.

[6] Lu, P., Hsiao, S. W., Tang, J., & Wu, F. (2024). A generative-AI-based design methodology for car frontal forms design. *Advanced Engineering Informatics, 62*, 102835.

[7] Yang, C., Liu, F., & Ye, J. (2023). A product form design method integrating Kansei engineering and diffusion model. *Advanced Engineering Informatics, 57*, 102058.

[8] Saadi, J. I., Chong, L., & Yang, M. C. (2024). The effect of targeting both quantitative and qualitative objectives in generative design tools on the design outcomes. *Research in Engineering Design, 35*, 409–425.

[9] Lin, H., Jiang, X., Deng, X., Bian, Z., Fang, C., & Zhu, Y. (2024). Comparing AIGC and traditional idea generation methods: Evaluating their impact on creativity in the product design ideation phase. *Thinking Skills and Creativity, 54*, 101649.

[10] Chulvi, V. (2025). Effectiveness of the AI using different typologies of design methods. *Research in Engineering Design, 36*, Article 7.