DesignGPT: Multi-Agent Collaboration in Design

Shiying Ding[1], Xinyi Chen[1], Yan Fang[1], Wenrui Liu[1], Yiwu Qiu[2], Chunlei Chai[1*]

1. State Key Lab of CAD&CG, Zhejiang University, Hangzhou, Zhejiang, China
2. ZAOWUYUN Inc., Hangzhou, Zhejiang, China

*Corresponding author

Email: {chatdsy,cadrenaline,22251360,liuwenrui}@zju.edu.cn,
qiuyiwu@zaowuyun.com, dishengchai@126.com

Abstract—Generative AI faces many challenges when entering the product design workflow, such as interface usability and interaction patterns. Therefore, based on design thinking and design process, we developed the DesignGPT multi-agent collaboration framework, which uses artificial intelligence agents to simulate the roles of different positions in the design company and allows human designers to collaborate with them in natural language. Experimental results show that compared with separate AI tools, DesignGPT improves the performance of designers, highlighting the potential of applying multi-agent systems that integrate design domain knowledge to product scheme design.

Index Terms—Creativity Support, Design Methods, LLM agent, Industrial Design, artificial intelligence

I. INTRODUCTION

In recent years, the development of artificial intelligence (AI) has provided new possibilities for various aspects of design, including supporting designers' innovative designs [1]. Design is a complex and creative process, and designers rely on their design thinking and design abilities to solve complex design problems. Due to individual cognitive abilities and experiences, designers often face challenges in the process of conceptualizing innovative solutions. Using a variety of generative AI tools is one of the solutions [2]. Image generation AI tools such as MidJourney [3] and Stable Diffusion [4] mainly generate rich images by inputting text prompts, which can provide inspiration for designers and improve the efficiency of product design. Text generative AI is developing rapidly, and large language model (LLM) such as ChatGPT [5] has the ability to understand and generate text, and have been used in design education [3] [6] and the generation of design solutions.

However, the existing generative AI embedded workflow often faces challenges in interface usability and interaction patterns, and there is a natural gap between design thinking and machine thinking. How to make AI better understand design thinking is one of the challenges for designers to interact with AI. Human-centered intelligent product design needs to bridge the gap between design thinking and machine thinking [7].

Design thinking mainly focuses on people, helping designers observe user behavior, explore solutions, and optimize design concepts through means such as brainstorming, social thinking, visual thinking, and prototype practice. These stages require support from different design abilities, including user analysis, sketch design, and CMF (color, material&Finishing) design. Existing AI design tools rarely consider the entire design process, and switching between different generative AI models can easily interrupt the creator's design thinking.

Design tools based on artificial intelligence continue to emerge in professional software to assist engineering and industrial designers in completing complex manufacturing and design tasks. These tools assume more agency roles than traditional Computer Aided Design tools and are often described as "co-creators" [8].

Agents usually refer to entities that exhibit intelligent behavior and have qualities such as autonomy, reactivity, initiative, and social ability [9]. Compared with the symbolic logic, response or reinforcement learning techniques commonly used by AI agents in the past, Large Language Model (LLM) has shown strong abilities in knowledge acquisition and natural language interaction [10], and has become an ideal choice for building intelligent agents. Agent [11] systems based on large language models have been proven to be feasible simulations of human behavior [12] [13]. By introducing standardized operating procedures (SOPs) [14], it can effectively reduce the problems of psychedelic, repetitive, incoherent, and invalid feedback in the output of large language models, and has good application prospects [15].

We have developed a system DesignGPT, which is designed by multiple AI agents in collaboration with designers. Specifically, we have introduced industry-standardized operating procedures (SOPs) for design to improve efficiency. Through the requirements table, agent roles, and conference rooms, we visualize the operating logic of AI agents for designer participants.

II. SYSTEM DESIGN

We have built the DesignGPT system, a multi-agent collaboration system that incorporates a design thinking knowledge base and the design process, aiming to provide designers with a new method of collaborating with generative AI agents for product design. The system integrates the large language model GPT4 [10] and the image generative model Stable Diffusion [4], providing design inspiration through graphic dialogue and supporting designers' innovative design in the conceptual stage. As shown in Fig. 1.

A. Core components

The core components of DesignGPT include Requirements Import Form, Role Definition and Selection, and Meeting
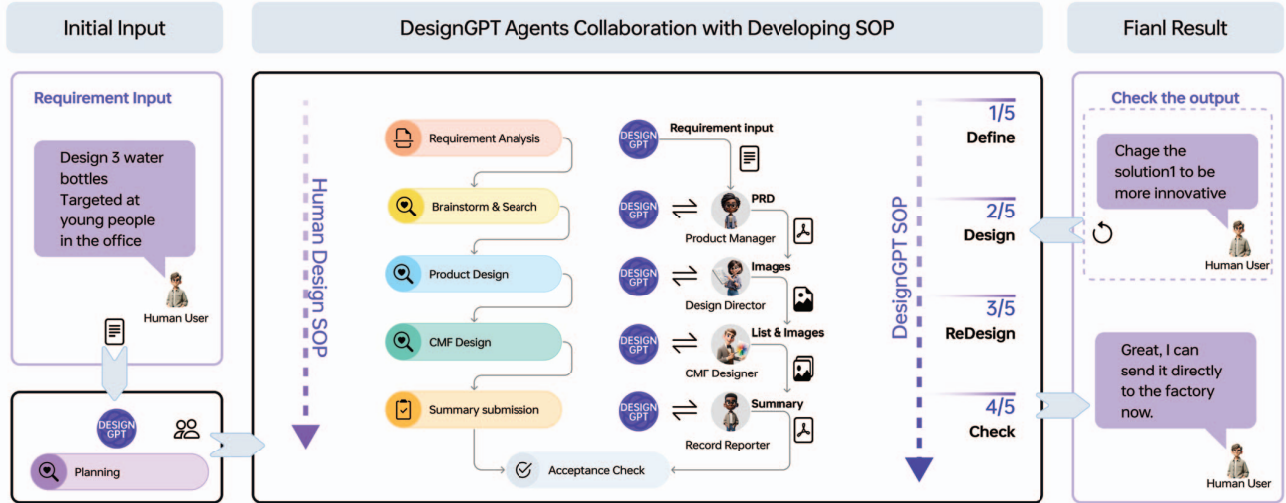
Fig. 1. System schematic. The system defines the job roles and workflow in the design company. Initially, the designer user needs to input design requirements, such as designing three cups for young people in the office. The DesignGPT system initializes multiple employees in the design company, allowing participants to select roles to start meetings. Employees complete the work in the order of design SOP , and finally output the results. The designer user can check the output of the design process in real time, and intervene by typing midway to make Scheme 1 more innovative.

Room.

1) Requirements Import Form: It mainly presents an interface for filling in design requirements, and users fill in design requirements here. The design requirement will be recorded as an initial prompt, and the design requirement will be set as the meeting topic.

2) Role Definition and Selection: The DesignGPT system initializes multiple employees in the design company, allowing participants to select roles to start meetings. The roles include virtual users, bosses, Product Manager , design directors, CMF designers, scoring recorder, and technical staff. The roles are clearly defined and undertake tasks such as requirement proposal, requirement analysis, design proposal, detailed design, and technical iteration. Virtual users and mainly Product Manager, Design Director, CMF Designers and scoring recorder participate in the final design. In order to facilitate the differentiation of roles among designer participants, we named them ABCDE in order, using random avatars. Product Manager A is responsible for setting the product direction, analyzing the market, and defining product features; Design Director B is responsible for organizing designers and finding design solutions; CMF Designer C is responsible for creating color and style schemes for various scenarios; Recording ReporterD is responsible for Document meetings and assessing proposal feasibility.

3) Meeting room: After filling in the requirements, participants enter the meeting room interface, select employees, and press the Join Meeting Room button to prepare for the meeting. The meeting room chat format is similar to common group chat software. A chat bar is set below the system, and the right side of the chat bar allows uploading pictures to perform picture description and modify color scheme tasks.

Role definitions, imported requirements, and historical records will be input to LLM together. Users can supplement the initial design ideas or propose new design requirements according to the selected employee responsibilities. This idea and the initial filled design requirements will be considered in the meeting. After the user enters the content, the agent employees summoned in the meeting room will start the discussion according to their responsibilities, gradually complete the demand analysis, design proposal, detailed design, evaluation report work, and finally output the design plan. During the discussion process, users can output messages through the chat bar to modify the meeting direction.

B. Generative AI settings

1) Image Generation AI settings: Images are an important part of the design plan. Therefore, the integration of image generation AI needs to ensure the quality of images and the ease of interaction of the system interface. We chose the image generation model based on Stable Diffusion 1.5. Image generation is completed by the role of the design director according to the user's design needs. Sketch is a tool for designers to express design ideas. The process from sketch to final concept map is often time-consuming and laborious, including modifications to product color, material, and process, which we abbreviate as CMF design. The system sets up the role of CMF designer as an agent, and uses Controlnet's canny model to generate design plans. Its input is a picture, and AI automatically extracts the line draft. The line draft is combined with the CMF-related prompt words provided by the agent to generate different CMF design plans.

2) Text Generation AI settings: The text generative model under the system framework can be replaced at will. In order to
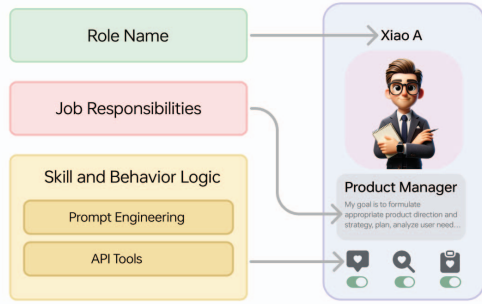
Fig. 2. The role card encompasses information such as the role's name (e.g., Xiao A), job responsibilities (e.g., Product Manager), as well as skills and behavior logic, which include prompt engineering and API tools.

ensure smooth communication between AI agents and smooth collaboration with designers, we ultimately selected the high-performing model, GPT-4.

The text input to the large language model is called prompt. Our system has designed a module to reconstruct the prompt words input to GPT4. The prompt words generally include role definitions, current skills, output format restrictions, user input, historical records, etc.

The market and user information required for design analysis is updated quickly. Due to the training data for GPT-4 may lag behind the current time, the system has integrated the function of searching for information on the network, and the information obtained from the search will be added as additional text prompts. This function is mainly reflected in the market analysis operation of the Product Manager agent role.

Different roles have different strategies for reconstructing prompt words. For example, in the drawing skills of the design director, the prompt words ultimately used for image generation are comprehensively generated by GPT4 considering the format of Stable Diffusion prompt words, historical records, initial requirements, and current plans.

If only the text model is used, the agent cannot recognize the image, so we use the BLIP [16] model to convert the image uploaded by the user into a text description, and cascade to realize the agent's recognition of the image. Like the network search information function, the text description recognized by BLIP is also added as an additional text prompt to the prompt word submitted to GPT4.

C. Composition of roles

The concept of roles helps designers understand and guides LLM to generate actions that match role characteristics. Therefore, roles under the DesignGPT framework consist of personal information, job responsibilities, and skills. Roles have basic abilities, such as reflection, Knowledge Base, meeting presentation, and status management. Different roles have different skills.

The personal information sets the basic information for the role, such as the name "Xiao A". The job responsibilities

determine the professional knowledge and possible behavioral logic of the role. For example, the goal of Product Manager is to develop appropriate product direction and strategy, planning; competing product analysis, product definition, scene, character portrait; responsible for user research, analyzing user requests, and providing user data support for design.

Skills and logic include the ability of the role and the actions to be followed, which are composed of two parts: prompt engineering and api tools. For example, CMF designer Xiao C needs to import the table for CMF analysis according to the demand, combine the CMF scheme, output the "CMF scheme list", and utilize a CMF API tool to produce CMF images according to the scheme list.

III. USER STUDY

In order to investigate whether the DesignGPT multi-agent collaboration system will change the design process of individual designers and affect the overall design results, and to evaluate the degree of assistance provided by the system in the design process, we conducted an online experiment using inter-group design.

A. Participants

We recruited 30 design learners for the study, each with an average design learning duration of 2.81 years and an age range between 21 and 24, with an average age of 23. All were college students with similar design proficiency, possessing design experience, computer literacy, exposure to generative AI such as ChatGPT for design.

B. Procedure

Participants were randomly divided into two groups and designed using one strategy:

- **Strategy 1:** allows designers to create using the AI tools provided by the experiment, specifically the web version of Stable Diffusion (SD) and GPT4.
- **Strategy 2:** allows designers to create web pages using the DesignGPT system, which utilizes the same SD and GPT APIs as Strategy 1.

The final output is a design scheme of a product concept. Participants are required to design a "human-centered future intelligent product" in 30 minutes, create a complete design plan, mainly including requirement analysis, concept proposal, specific design presentation, and finally present it in a simple document layout. During the experiment, participants evaluate the product design process themselves, and the system will fully record and export the process document.

C. Evaluation indicators

In this experiment, we focus on mapping the effectiveness of the design process through the quality of the final plan, so we choose to evaluate the design plan in three dimensions: novelty [17], completeness, and feasibility [18]. Novelty measures the uniqueness of the design plan, and the evaluation criteria are the novelty compared to existing products or concepts on the market and other concepts generated in this experiment.
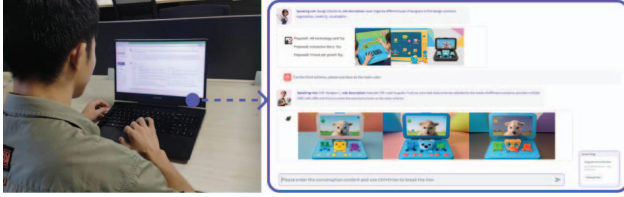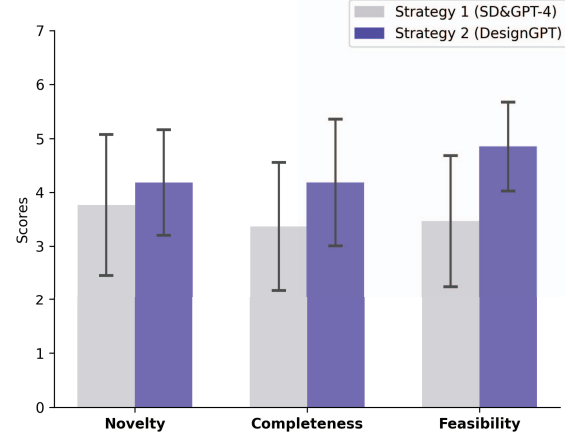
206

Fig. 3. A designer is using DesignGPT system.

Fig. 4. Mean with SD Plot of the designing results. The plot shows the average scores of the participants' design schemes evaluated by experts in the three dimensions of novelty, completeness and feasibility under the two strategies

Feasibility measures the viability of the design plan in technology and the market, as well as the closeness to meeting the requirements of the design summary. completeness measures whether a design idea is fully expressed, and the judgment is based on the richness of details in terms of function, form, material, structure, color, etc.

This evaluation uses a seven-point scale, with 1 point representing the lowest score and 7 points representing the highest score. The evaluators need to score the design scheme in terms of novelty, completeness, and feasibility by viewing the integrated design scheme content and design conclusions.

After the design task, we conducted a 10-15 minute semi-structured interview with each participant about the designer's reaction to the system. The content includes the degree of support and contribution of AI to the design scheme.

IV. EXPERIMENTAL RESULTS

Based on Consensual Assessment Technique (CAT) [19], we invited two experts with 6 years of design experience to independently score each design scheme. The reliability among judges is high, with a significant intraclass correlation coefficient (ICC) of 0.90 ($p < 0.001$).

A. quantitative analysis

As shown in Fig. 4, in terms of three indicators, the mean and median scores of Strategy 2 (DesignGPT) are higher than those of Strategy 1 (SD&GPT-4). Compared with Strategy 1, Strategy 2 group has an improvement of 11.06%, 24.26%, and 39.90% in novelty, completeness, and feasibility, respectively.Utilizing an independent samples one-tailed t-test ($p = 0.05$), our analysis reveals that Strategy 2 demonstrates statistically significant advantages over Strategy 1 in terms of completeness and feasibility. Strategy 2 participants' schemes have gained higher overall recognition. The score variance and Standard Deviation of Strategy 2 are slightly lower than those of Strategy 1, indicating that the scheme scores produced by the DesignGPT collaborative system are more concentrated and the evaluators' evaluations are more consistent.

B. Qualitative Analysis

According to our interviews, the users gave high praise for the interactive form of the DesignGPT. P12 (Strategy 2) pointed out that DesignGPT deduces the scheme from the perspective of multiple roles, breaking through the problem of having a relatively single perspective in conceiving the scheme in the past. P5 (Strategy 2) pointed out that the AI

role proficiently uses the design process and design expression, including background analysis, requirement import form, concept screening, scheme expression and scheme scoring, etc.

Regarding the level of support and contribution provided by AI in design thinking and program output, Strategy 1 ranges from 5% to 90%, and Strategy 2 ranges from 30% to 80%. Compared to Strategy 2, Strategy 1 seems to be more likely to have completely AI-led design or designer-led design. Users of Strategy 1 tend to characterize the assistance provided by the experiment as utilizing AI tools or systems, whereas users of Strategy 2 are more inclined to describe it as collaborating with AI.

V. CONCLUSIONS

Our research developed DesignGPT, a multi-agent-based design collaboration system that supports designers to collaborate with AI agents to complete innovative designs during the conceptual stage. We first conducted an interview study and invited design learners who have tried AI design tools to experiment to understand how designers use AI tools for design and what challenges they encounter during the process. We created a chat room style interactive website that allows users to input requirements and form collaborative AI agent teams, establishing a new workflow of generative AI with designers. Our results show that the DesignGPT system can effectively support designers' design process. We believe that this research has important theoretical and practical significance for understanding the role of AI in the design process.

REFERENCES

[1] Y.-c. Tang, J.-j. Huang, M.-t. Yao, J. Wei, W. Li, Y.-x. He, and Z.-j. Li, "A review of design intelligence: progress, problems, and challenges," *Frontiers of Information Technology & Electronic Engineering*, vol. 20, no. 12, pp. 1595–1617, 2019.

[2] X. Zhang, X. Mao, Y. Yin, C. Chai, and T. Zhang, "Melting your models: An integrated ai-based creativity support tool for inspiration evolution," in *2022 15th International Symposium on Computational Intelligence and Design (ISCID)*. IEEE, 2022, pp. 97–101.

[3] H. Vartiainen and M. Tedre, "Using artificial intelligence in craft education: crafting with text-to-image generative models," *Digital Creativity*, vol. 34, no. 1, pp. 1–21, 2023.

[4] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 10 684–10 695.

[5] P. P. Ray, "Chatgpt: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope," *Internet of Things and Cyber-Physical Systems*, 2023.

[6] Y.-M. Fang, "The role of generative ai in industrial design: Enhancing the design process and learning."

[7] L. Sun, Y. Zhang, Z. Zhou, and Z. Zhou, "Current situation and development trend of intelligent product design under the background of human-centered ai," *Packaging Engineering*, vol. 41, no. 02, pp. 1–6, 2020.

[8] D. Wang, E. Churchill, P. Maes, X. Fan, B. Shneiderman, Y. Shi, and Q. Wang, "From human-human collaboration to human-ai collaboration: Designing ai systems that can work together with people," in *Extended abstracts of the 2020 CHI conference on human factors in computing systems*, 2020, pp. 1–6.

[9] M. Wooldridge and N. R. Jennings, "Intelligent agents: Theory and practice," *The knowledge engineering review*, vol. 10, no. 2, pp. 115–152, 1995.

[10] S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, P. Lee, Y. T. Lee, Y. Li, S. Lundberg *et al.*, "Sparks of artificial general intelligence: Early experiments with gpt-4," *arXiv preprint arXiv:2303.12712*, 2023.

[11] T. et al., "Auto-gpt," https://github.com/Significant-Gravitas/Auto-GPT, 2023.

[12] G. Li, H. A. A. K. Hammoud, H. Itani, D. Khizbullin, and B. Ghanem, "Camel: Communicative agents for" mind" exploration of large scale language model society," *arXiv preprint arXiv:2303.17760*, 2023.

[13] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein, "Generative agents: Interactive simulacra of human behavior," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, ser. UIST '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3586183.3606763

[14] S. Hong, X. Zheng, J. Chen, Y. Cheng, C. Zhang, Z. Wang, S. K. S. Yau, Z. Lin, L. Zhou, C. Ran *et al.*, "Metagpt: Meta programming for multi-agent collaborative framework," *arXiv preprint arXiv:2308.00352*, 2023.

[15] C. Qian, X. Cong, C. Yang, W. Chen, Y. Su, J. Xu, Z. Liu, and M. Sun, "Communicative agents for software development," *arXiv preprint arXiv:2307.07924*, 2023.

[16] J. Li, D. Li, C. Xiong, and S. Hoi, "Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation," in *International Conference on Machine Learning*. PMLR, 2022, pp. 12 888–12 900.

[17] L. Chen, P. Wang, H. Dong, F. Shi, J. Han, Y. Guo, P. R. Childs, J. Xiao, and C. Wu, "An artificial intelligence based data-driven approach for design ideation," *Journal of Visual Communication and Image Representation*, vol. 61, pp. 10–22, 2019.

[18] P.-A. Verhaegen, D. Vandevenne, J. Peeters, and J. R. Duflou, "Refinements to the variety metric for idea evaluation," *Design Studies*, vol. 34, no. 2, pp. 243–263, 2013.

[19] T. M. Amabile, "The social psychology of creativity: A componential conceptualization." *Journal of personality and social psychology*, vol. 45, no. 2, p. 357, 1983.