

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

Статья поступила в редакцию: 24.08.2026

Статья принята к публикации: 14.09.2026

Статья опубликована: 28.09.2026

УДК 004.8+004.9

DOI: 10.32743/UniTech.2026.150.9.23412

Моделирование процесса формирования визуального контента в интеллектуальных системах искусственного интеллекта на основе Prompt Engineering

Ганиев Акмал Абдухалилович

ст. преп. кафедры управления промышленностью и цифровых технологий

Международный университет Нордик

Республика Узбекистан, г. Ташкент

ORCID: 0009-0006-9402-1508

E-mail: g.akmal@nordicuniversity.org

Modeling the Process of Visual Content Generation in Intelligent AI Systems Based on Prompt Engineering

Ganiev Akmal Abdukhalilovich

senior Lecturer

Department of Industrial Management and Digital Technologies Nordic International University

Republic of Uzbekistan, Tashkent

Аннотация

В статье рассматривается процесс формирования визуального контента в интеллектуальных системах искусственного интеллекта на основе технологий Prompt Engineering. Актуальность исследования обусловлена необходимостью повышения точности преобразования естественно-языковых пользовательских запросов в визуальный контент с учетом заданных характеристик, контекста и требований к изображению. Проведен анализ существующих подходов к генерации изображений с использованием больших языковых моделей и современных генеративных нейронных сетей. Предложена авторская концептуальная модель интеллектуального формирования визуального контента, предусматривающая последовательные этапы получения исходного запроса, его семантического анализа, структуризации, оптимизации и последующей оценки сформированного изображения. В качестве структурной основы промпта выделены четыре

Библиографическое описание: Ганиев А.А. Моделирование процесса формирования визуального контента в интеллектуальных системах искусственного интеллекта на основе Prompt Engineering // Universum: технические науки : электрон. научн. журн. 2026. 9(150). URL: <https://7universum.com/ru/tech/archive/item/23412>

взаимосвязанных компонента: содержание визуальной сцены (S), визуальные характеристики (V), пользовательские требования (U) и контекст формирования (C). Разработан алгоритм интеллектуальной оптимизации промпта, предусматривающий проверку полноты и согласованности его компонентов, уточнение параметров и итерационную корректировку запроса на основе полученного результата. Для оценки эффективности предложенного подхода использованы критерии семантического соответствия, полноты визуального описания, стилевой согласованности, детализации и устойчивости результатов при повторной генерации. Экспериментальная проверка проведена с использованием генеративных систем ChatGPT Images 2.0, Pixmind, Qwen 3.0 Pro и FLUX.2 [klein]. Полученные результаты подтверждают положительное влияние структуризации и оптимизации промптов на качество формируемого визуального контента. Предлагаемый подход может быть использован при разработке интеллектуальных систем поддержки проектирования, цифрового дизайна, образовательных платформ и мультимедийных приложений.

Abstract

The article examines the process of visual content generation in intelligent artificial intelligence systems based on Prompt Engineering technologies. The relevance of the study is determined by the need to improve the accuracy of transforming natural-language user queries into visual content while taking into account specified characteristics, contextual information, and user requirements. Existing approaches to image generation using large language models and modern generative neural networks are analyzed. An original conceptual model for intelligent visual content generation is proposed, incorporating sequential stages of obtaining the initial query, performing semantic analysis, structuring and optimizing the prompt, and subsequently evaluating the generated image. Four interconnected components are identified as the structural basis of the prompt: visual scene content (S), visual characteristics (V), user requirements (U), and generation context (C). An algorithm for intelligent prompt optimization is developed, including verification of the completeness and consistency of its components, parameter refinement, and iterative query correction based on the generated result. The effectiveness of the proposed approach is evaluated using the criteria of semantic correspondence, completeness of visual description, stylistic consistency, image detail, and stability of results during repeated generation. Experimental verification was conducted using ChatGPT Images 2.0, Pixmind, Qwen 3.0 Pro, and FLUX.2 [klein]. The obtained results confirm the positive effect of prompt structuring and optimization on the quality of generated visual content. The proposed approach can be applied in the development of intelligent design support systems, digital design tools, educational platforms, and multimedia applications.

Ключевые слова: искусственный интеллект, Prompt Engineering, генеративный искусственный интеллект, визуальный контент, генерация изображений, диффузионные модели, интеллектуальные системы, моделирование.

Keywords: Artificial Intelligence, Prompt Engineering, Generative AI, Visual Content, Image Generation, Diffusion Models, Intelligent Systems, Modeling.

Введение

В последние годы искусственный интеллект (ИИ) стал одной из ключевых технологий цифровой трансформации, оказывая существенное влияние на обработку информации, автоматизацию интеллектуальной деятельности и создание цифрового контента [7, 9]. Особое развитие получили генеративные модели искусственного интеллекта, основанные на использовании больших языковых моделей и диффузионных нейронных сетей, позволяющих автоматически создавать текстовые и графические материалы по естественно-языковым запросам пользователя [2, 6, 10]. Современные системы, такие как ChatGPT Images 2.0, Pixmind, Qwen 3.0 Pro, FLUX.2 [klein] активно применяются в цифровом дизайне, промышленном проектировании, образовании и мультимедийных технологиях [3].

Одним из ключевых факторов, определяющих качество работы генеративных моделей, является технология Prompt Engineering, обеспечивающая структурированное формирование текстовых запросов для получения требуемого результата [4]. Качество создаваемого визуального контента во многом определяется корректностью построения промпта, полнотой контекстной информации и точностью описания объекта генерации [5].

Несмотря на значительное количество исследований в области генеративного искусственного интеллекта, основное внимание уделяется совершенствованию архитектур нейронных сетей и методов генерации изображений [2, 3, 6, 10, 8]. При этом вопросы моделирования процесса преобразования текстового промпта в визуальный контент, включая этапы семантического анализа, структурной обработки запроса и оценки качества результатов генерации, остаются недостаточно изученными. Отсутствие комплексной модели взаимодействия пользователя с интеллектуальной системой определяет актуальность настоящего исследования.

Ранее автором были рассмотрены отдельные аспекты применения технологий искусственного интеллекта и методов Prompt Engineering при разработке цифрового образовательного контента и интеллектуальных информационных решений [1]. Настоящая работа является продолжением этих исследований и направлена на разработку модели формирования визуального контента в интеллектуальных системах искусственного интеллекта.

Целью исследования является разработка модели процесса формирования визуального контента в интеллектуальных системах искусственного интеллекта на основе технологии Prompt Engineering, обеспечивающей повышение качества интерпретации текстовых запросов и эффективности генерации изображений.

Задачи

1. провести анализ современных методов формирования визуального контента в интеллектуальных системах искусственного интеллекта, исследовать особенности генеративных моделей и определить роль технологии Prompt Engineering в процессе обработки текстовых запросов;
2. исследовать существующие подходы к структурированию текстовых промптов, выявить факторы, влияющие на точность интерпретации пользовательских запросов и качество генерируемых изображений;

3. разработать концептуальную модель процесса формирования визуального контента на основе технологии Prompt Engineering, включающую этапы семантического анализа, структурной обработки промпта, генерации изображения и оценки качества результата;

4. предложить алгоритм интеллектуальной обработки текстовых промптов, обеспечивающий повышение эффективности взаимодействия пользователя с генеративной системой искусственного интеллекта;

5. разработать математическую модель оценки качества формируемого визуального контента с учетом семантического соответствия, полноты описания, стилевой согласованности, детализации изображения и устойчивости результата при повторной генерации;

6. провести экспериментальное исследование предложенной модели с использованием современных генеративных систем искусственного интеллекта и выполнить сравнительный анализ полученных результатов по показателям качества визуального контента.

Материалы и методы

Объектом исследования является процесс формирования визуального контента в интеллектуальных системах искусственного интеллекта на основе технологии Prompt Engineering. Предмет исследования включает методы структурирования и семантической обработки текстовых промптов, а также механизмы преобразования естественно-языковых запросов в изображения с использованием современных генеративных моделей.

Материалами исследования послужили научные публикации по генеративному искусственному интеллекту, Prompt Engineering, большим языковым и диффузионным моделям, а также результаты экспериментального применения систем ChatGPT Images 2.0, Pixmind, Qwen 3.0 Pro и FLUX.2 [klein].

Методологическую основу составили системный и сравнительный анализ, концептуальное моделирование, семантический анализ и структурное проектирование. Предлагаемая модель предусматривает последовательную обработку запроса: семантический анализ, оптимизацию промпта, генерацию изображения и оценку результата.

Эффективность модели оценивалась по пяти критериям: семантическому соответствию, полноте визуального описания, стилевой согласованности, детализации и устойчивости результата при повторной генерации. Сравнение проводилось при сопоставимых условиях и одинаковых исходных заданиях до и после оптимизации промптов.

Таблица 1.

Методы исследования

Метод	Назначение
Системный анализ	Исследование структуры процесса формирования визуального контента
Семантический анализ	Интерпретация текстовых промптов
Концептуальное моделирование	Разработка авторской модели процесса
Сравнительный анализ	Сопоставление генеративных AI-систем
Экспериментальный метод	Оценка качества результатов генерации

Результаты и обсуждение

2.1. Концептуальная модель процесса формирования визуального контента

Результаты исследования позволили сформировать концептуальную модель преобразования естественно-языкового пользовательского запроса в визуальный контент. В отличие от прямой схемы «текстовый запрос – изображение», предлагаемая модель предусматривает последовательные этапы семантической и структурной обработки, оптимизации и оценки результата.

Анализ текстовых промптов показал, что их структура и полнота существенно влияют на качество генерируемого визуального контента. Недостаточно структурированный запрос может приводить к неполному представлению объектов, нарушению их взаимного расположения и несоответствию заданному стилю. С учетом выявленных особенностей процесс формирования визуального контента представлен следующей последовательностью:



Рисунок 1. Процесс генерации визуального контента

На первом этапе формируется исходный пользовательский запрос, после чего выполняется его семантический анализ с выделением основных объектов, атрибутов, контекста, пространственных отношений и визуальных ограничений.

На этапе структуризации выделенная информация преобразуется в упорядоченную структуру промпта, после чего выполняется его оптимизация путем уточнения параметров, устранения неоднозначностей и дополнения необходимой контекстной информации. Оптимизированный промпт передается генеративной модели для формирования визуального контента.

Полученный результат оценивается по критериям семантического соответствия, полноты, стилевой согласованности, детализации и устойчивости при повторной генерации. При выявлении несоответствий выполняется корректировка промпта и повторная генерация. Таким образом, предложенная модель имеет итерационный характер:

$$P_i \rightarrow G(P_i) \rightarrow K \rightarrow P_{i+1}$$

где P_i — текущий промпт, $G(P_i)$ — результат его обработки генеративной моделью, K_i — оценка качества полученного визуального результата, а P_{i+1} — скорректированный промпт, сформированный с учетом выявленных несоответствий.

Важным элементом предложенной модели является обратная связь между результатом генерации и последующей обработкой текстового запроса. Она обеспечивает последовательное уточнение промпта с учетом выявленных несоответствий и повышение соответствия визуального результата исходному пользовательскому замыслу.

Таким образом, качество визуального контента определяется характеристиками исходного запроса, структурой и качеством семантической обработки промпта, а также возможностями генеративной модели. Prompt Engineering в предложенной модели выступает промежуточным интеллектуальным механизмом, связывающим пользовательское намерение с процессом визуального синтеза.

2.2. Структурная модель Prompt Engineering

Для повышения точности формирования визуального контента промпт рассматривается как структурированный набор взаимосвязанных компонентов, отражающих содержание изображения, его визуальные характеристики, требования пользователя и контекст. Такой подход позволяет перейти от свободного текстового описания к формализованной структуре. В рамках исследования структура промпта представляется следующим образом:

$$P=\{S,V,U,C\}$$

где (P) — структурированный промпт; (S) — содержание или основной объект визуальной сцены; (V) — визуальные характеристики; (U) — пользовательские требования и ограничения; (C) — контекст формирования визуального контента.

Компонент (S) (Subject) определяет содержание изображения: основные объекты, их действия и взаимное расположение.

Компонент (V) (Visual characteristics) характеризует визуальное представление: художественный стиль, цветовую гамму, освещение, композицию, перспективу и детализацию.

Компонент (U) (User requirements) отражает требования и ограничения пользователя, включая формат, соотношение сторон, степень реалистичности и наличие заданных элементов.

Компонент (C) (Context) определяет контекст формирования изображения: место и время действия, назначение контента и окружающую среду.

Таким образом, структурированный промпт может рассматриваться как совокупность четырех взаимосвязанных компонентов:

$$P=\{S,V,U,C\}$$

Компоненты S , V , U и C взаимосвязаны: изменение одного компонента может потребовать корректировки других. Поэтому качество изображения определяется не только полнотой отдельных компонентов, но и согласованностью их содержания.

Предлагаемая структура создает основу для интеллектуальной оптимизации промпта, позволяя выявлять отсутствующие или противоречивые параметры и формировать уточненный запрос перед его передачей генеративной модели.

Таким образом, модель $P=\{S,V,U,C\}$ выступает промежуточным уровнем между естественно-языковым запросом и генеративной моделью, обеспечивая структурированное представление пользовательского намерения и автоматизированную оптимизацию промпта.

2.3. Алгоритм интеллектуальной оптимизации промпта

На основе разработанной структурной модели предложен алгоритм интеллектуальной оптимизации текстового промпта, направленный на повышение точности интерпретации пользовательского запроса и качества визуального контента. Алгоритм предусматривает последовательную обработку содержательных, визуальных, пользовательских и контекстных характеристик запроса. Алгоритм включает следующие основные этапы:

1. Получение исходного запроса — прием естественно-языкового описания требуемого визуального контента.
2. Семантический анализ — определение объектов, атрибутов, действий, взаимосвязей и контекста, а также выявление неоднозначностей.
3. Декомпозиция запроса — распределение выделенной информации по компонентам S , V , U и C .
4. Проверка полноты и согласованности — выявление отсутствующих параметров и логических противоречий между компонентами.
5. Дополнение и уточнение параметров — устранение выявленных неопределенностей с сохранением исходного пользовательского намерения.
6. Формирование оптимизированного промпта — объединение обработанных компонентов в единый структурированный запрос.
7. Генерация и оценка результата — передача промпта генеративной модели и оценка полученного изображения по установленным критериям качества.
8. Корректировка промпта — при выявлении несоответствий корректируются соответствующие компоненты и выполняется повторная генерация. Обобщенно алгоритм может быть представлен следующей последовательностью:

$$Q \rightarrow O\{S,V,U,C\} \rightarrow P^* \rightarrow G(P^*) \rightarrow K$$

где Q — исходный пользовательский запрос; S,V,U,C — компоненты структурного представления запроса; P^* — оптимизированный промпт; $G(P^*)$ — результат генерации; K — оценка качества визуального контента.

В отличие от непосредственной передачи исходного запроса генеративной модели, предложенный алгоритм предусматривает предварительную проверку и структурную обработку информации, что позволяет снизить влияние неоднозначных формулировок и повысить полноту промпта. Проверка согласованности компонентов позволяет выявлять потенциальные противоречия и отсутствующие параметры до начала генерации.

Таким образом, алгоритм обеспечивает переход от свободного естественно-языкового описания к структурированному и оптимизированному промпту и формирует итерационный процесс, в котором результаты генерации используются для последующей корректировки запроса. Разработанный алгоритм служит основой для экспериментальной проверки предложенного подхода и сравнительного анализа результатов генерации исходных и оптимизированных промптов.

2.4. Математическая модель оценки качества визуального контента

Для количественной оценки эффективности предложенного подхода определены критерии соответствия визуального контента исходному пользовательскому запросу. Качество результата оценивается по пяти взаимосвязанным критериям: семантическому соответствию, полноте визуального описания, стилевой согласованности, детализации и устойчивости результата при повторной генерации. На их основе интегральный показатель качества определяется следующим образом:

$$K = w_1 K_s + w_2 K_c + w_3 K_{st} + w_4 K_d + w_5 K_r$$

где K — интегральная итоговая оценка качества визуального контента; K_s — показатель семантического соответствия изображения исходному промпту; K_c — показатель полноты представления заданных элементов; K_{st} — показатель стилевой согласованности; K_d — показатель детализации изображения; K_r — показатель устойчивости результата при повторной генерации; $w_1, w_2, \dots, w_5$ — весовые коэффициенты, определяющие вклад соответствующих критериев в итоговую оценку качества.

Для обеспечения сопоставимости критериев оценки значения показателей K_s, K_c, K_{st}, K_d и K_r принимаются в диапазоне от 0 до 1. Значение, близкое к 0, соответствует низкой степени выполнения соответствующего критерия, тогда как значение, близкое к 1, характеризует высокий уровень соответствия установленным требованиям.

Весовые коэффициенты w_1, w_2, w_3, w_4, w_5 определяют относительную значимость каждого критерия при формировании итоговой оценки качества. В рамках настоящего исследования для всех пяти критериев установлены одинаковые весовые коэффициенты, что

позволяет обеспечить сопоставимость результатов различных генеративных систем и исключить дополнительное влияние субъективного определения приоритетности отдельных критериев. Поэтому принимается $w_1 = w_2 = w_3 = w_4 = w_5 = 0,2$.

Таким образом, интегральный показатель К объединяет характеристики визуального результата в единую оценку качества, при этом более высокое значение К свидетельствует о большем соответствии изображения исходному пользовательскому запросу.

Для проверки эффективности предложенного подхода далее проводится сравнительный анализ результатов генерации исходных и оптимизированных промптов при одинаковых условиях и установленных критериях оценки.

2.5. Экспериментальное исследование и сравнительный анализ

Для экспериментальной проверки предложенной модели проведено сравнительное исследование с использованием систем ChatGPT Images 2.0, Pixmind, Qwen 3.0 Pro и FLUX.2 [klein]. Эксперимент выполнялся при одинаковых условиях и идентичных исходных заданиях. Оптимизированные промпты формировались в соответствии со структурной моделью $P=\{S,V,U,C\}$.

В качестве базового тестового задания использовался следующий пользовательский запрос:

«Создать фотореалистичное изображение современной университетской аудитории, в которой преподаватель проводит занятие по искусственному интеллекту для студентов. На экране отображается схема нейронной сети, аудитория имеет современный технологичный интерьер, естественное освещение, реалистичные цвета, горизонтальная композиция».

Для визуального сопоставления на рисунке 2 представлены три повторные генерации по исходному и три по оптимизированному промпту, полученные с использованием FLUX.2 [klein]. Сравнение позволяет оценить изменения детализации, полноты представления объектов, стилевой согласованности и устойчивости генерации после оптимизации промпта

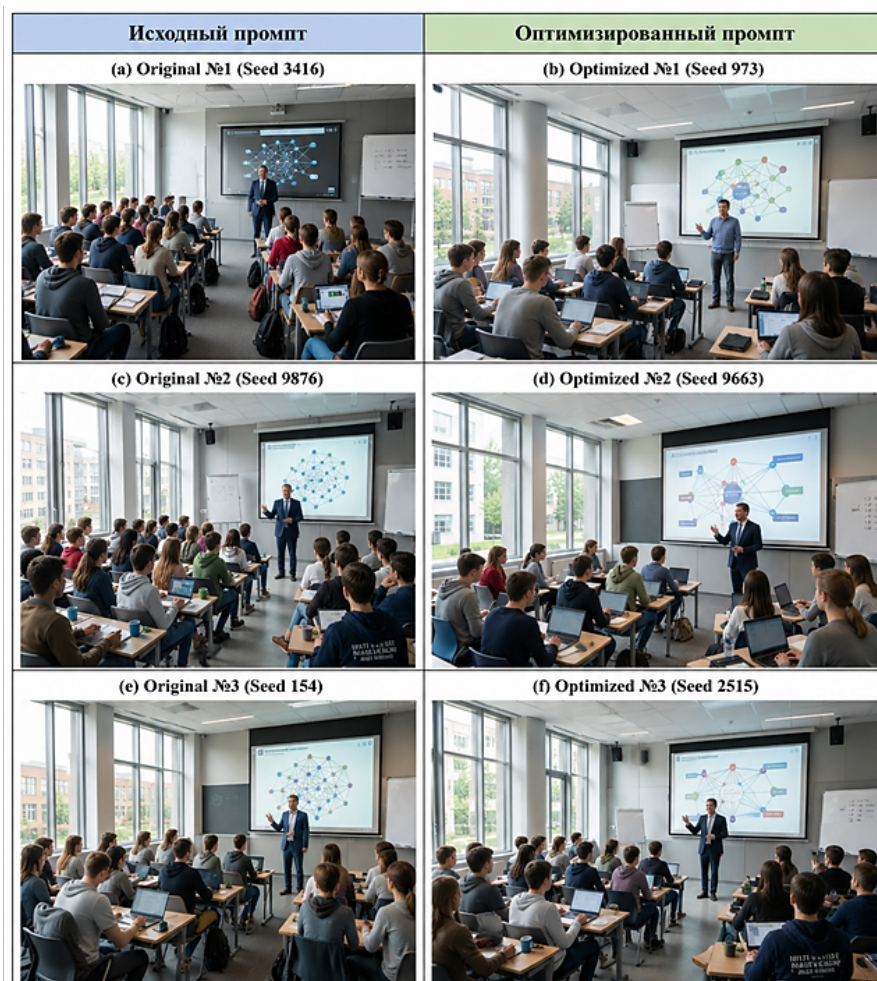


Рисунок 2. Сравнение результатов генерации FLUX.2 [klein] по исходному и оптимизированному промптам

Для дальнейшей обработки исходный пользовательский запрос был декомпозирован в соответствии с четырьмя компонентами структурной модели $P=\{S,V,U,C\}$. Результаты структуризации представлены в таблице 2.

Таблица 2.

Структуризация исходного пользовательского запроса

Компонент	Содержание
S – Содержание	Современная университетская аудитория, преподаватель, студенты, проведение занятия по искусственному интеллекту, схема нейронной сети на экране
V – Визуальные характеристики	Фотореалистичный стиль, современный технологичный интерьер, естественное освещение, реалистичные цвета, горизонтальная композиция
U – Пользовательские требования	Фотореалистичное представление, четкое отображение основных объектов, горизонтальный формат изображения
C – Контекст	Образовательная среда университета, учебное занятие по искусственному интеллекту, использование цифровых технологий

На основе проведенной структуризации был сформирован оптимизированный вариант промпта:

«Фотореалистичное изображение современной университетской аудитории, в которой преподаватель проводит занятие по искусственному интеллекту для студентов. На большом экране отображается схема нейронной сети с четко различимыми узлами и связями. Студенты находятся за учебными столами с ноутбуками и внимательно слушают преподавателя. Аудитория имеет современный технологичный интерьер, естественное освещение, реалистичную цветовую гамму, высокую детализацию и горизонтальную композицию. Широкий угол обзора, четкое расположение основных объектов, реалистичная перспектива».

Для оценки результатов использовались пять критериев: семантическое соответствие, полнота визуального описания, стилевая согласованность, детализация и устойчивость при повторной генерации. Для единообразия оценки применялась пятиуровневая шкала, представленная в таблице 3. Оценка проводилась автором исследования на основе визуального сопоставления результатов генерации с исходным заданием.

Таблица 3.

Шкала оценки критериев качества визуального контента

Значение показателя	Уровень соответствия
0,0–0,2	Очень низкий
>0,2–0,4	Низкий
>0,4–0,6	Средний
>0,6–0,8	Высокий
>0,8–1,0	Очень высокий

В рамках исследования всем пяти критериям присвоены одинаковые весовые коэффициенты $w_1=w_2=w_3=w_4=w_5=0,2$, что обеспечивает сопоставимость результатов различных генеративных систем. При этом сумма весовых коэффициентов составляет 1, а интегральный показатель K определяется как среднее значение пяти нормированных критериев:

$$K = 0,2K_s + 0,2K_c + 0,2K_{st} + 0,2K_d + 0,2K_r$$

Эксперимент проводился в два этапа: сначала генеративным системам передавался исходный промпт, затем он подвергался семантическому анализу и структуризации по компонентам S,V,U,C, после чего формировался оптимизированный промпт и выполнялась повторная генерация.



Рисунок 3. Сравнение результатов генерации исходного и оптимизированного промптов

Для визуального сопоставления результатов ChatGPT Images 2.0 на рисунке 3 представлены три генерации по исходному и три по оптимизированному промпту. Сравнение позволяет оценить изменения детализации, представления ключевых объектов и соответствия заданным характеристикам

Как видно на рисунке 3, использование оптимизированного промпта в ChatGPT Images 2.0 обеспечивает более полное представление основных элементов сцены — преподавателя, студентов, экрана с визуализацией нейронной сети и современной учебной среды, а также повышает детализацию визуального результата.

Для обеспечения сопоставимости во всех системах использовались одинаковые исходные задания и структура требований. При сравнении учитывались визуальные характеристики и степень сохранения исходного пользовательского замысла после оптимизации промпта. Для дополнительной проверки предложенного подхода аналогичный эксперимент проведен в системе Pixmind



Рисунок 4. Сравнение результатов генерации Pixmind по исходному и оптимизированному промптам

На рисунке 4 представлены три генерации по исходному и три по оптимизированному промпту. Сравнение позволяет оценить влияние оптимизации на полноту представления объектов, композиционную согласованность и соответствие заданному визуальному контексту.

Для дополнительной визуальной проверки предложенного подхода аналогичный эксперимент был проведен с использованием системы Qwen 3.0 Pro. На рисунке 5 представлены три повторные генерации по исходному промпту и три повторные генерации по оптимизированному промпту. Сопоставление результатов позволяет оценить влияние структуризации и оптимизации промпта на полноту представления объектов, композиционную согласованность, детализацию и соответствие заданному визуальному контексту.



Рисунок 5. Сравнение результатов генерации Qwen 3.0 Pro по исходному и оптимизированному промптам

В рамках сравнительного анализа сопоставлялись результаты, полученные при использовании исходного и семантически и структурно оптимизированного промптов. Такой подход позволяет оценить влияние предложенной модели $P=\{S,V,U,C\}$ на качество формируемого визуального контента [1].

Результаты оценивались по пяти критериям: семантическому соответствию, полноте визуального описания, стиливой согласованности, детализации и устойчивости при повторной генерации. На их основе формировалась интегральная оценка качества, представленная в таблице 4.

Таблица 4.
Сравнительная оценка результатов генерации исходных и оптимизированных промптов

Генеративная система	Вариант промпта	K_s	K_c	K_{st}	K_d	K_r	K
ChatGPT Images 2.0	Исходный	0,7	0,6	0,8	0,7	0,8	0,72
ChatGPT Images 2.0	Оптимизированный	0,9	0,9	0,9	0,9	0,9	0,90
Pixmind	Исходный	0,84	0,83	0,86	0,79	0,85	0,83
Pixmind	Оптимизированный	0,88	0,87	0,89	0,83	0,90	0,87
Qwen 3.0 Pro	Исходный	0,85	0,84	0,82	0,85	0,80	0,83
Qwen 3.0 Pro	Оптимизированный	0,94	0,95	0,94	0,96	0,90	0,94
FLUX.2 [klein]	Исходный	0,90	0,85	0,90	0,90	0,90	0,89
FLUX.2 [klein]	Оптимизированный	0,95	0,95	0,90	0,90	0,90	0,92

Результаты экспериментального исследования показывают, что применение оптимизированных промптов привело к повышению интегрального показателя качества K во всех исследуемых генеративных системах. Для ChatGPT Images 2.0 показатель K увеличился с 0,72 до 0,90, что соответствует приросту на 25,0 %. Для Pixmind показатель увеличился с 0,83 до 0,87, или на 4,8 %. В системе Qwen 3.0 Pro показатель K повысился с 0,83 до 0,94, что соответствует приросту на 13,3 %. Для FLUX.2 [klein] показатель увеличился с 0,89 до 0,92, или на 3,4 %. Наибольший прирост интегрального показателя наблюдается в системе ChatGPT Images 2.0 (+25,0 %), тогда как наименьший прирост зафиксирован в системе FLUX.2 [klein] (+3,4 %). В целом полученные результаты подтверждают положительное влияние структуризации и оптимизации промптов на качество формируемого визуального контента.

Заключение

В результате исследования разработана модель формирования визуального контента на основе технологии Prompt Engineering, включающая семантический анализ, структуризацию, оптимизацию промпта, генерацию и оценку результата. Предложена структурная модель промпта $P=\{S,V,U,C\}$ и интегральный показатель качества K , основанный на пяти критериях оценки.

Экспериментальная проверка с использованием ChatGPT Images 2.0, Pixmind, Qwen 3.0 Pro и FLUX.2 [klein] показала повышение показателя K при применении оптимизированных промптов: соответственно на 25,0 %, 4,8 %, 13,3 % и 3,4 %. Полученные результаты свидетельствуют о положительном влиянии структуризации и оптимизации промптов на качество формируемого визуального контента. Предложенная модель может применяться в интеллектуальных системах проектирования, цифрового дизайна, образования и мультимедийных приложениях.

Список литературы

1. Ганиев А.А. Моделирование визуализации на основе промптов в системах искусственного интеллекта // Научные основы и современные проблемы интеграции медиа и искусственного интеллекта: сборник докладов Международной научно-практической конференции. – 2026.
2. Cao P., Zhou F., Song Q., Yang L. Controllable Generation With Text-to-Image Diffusion Models: A Survey // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2026. – Т. 48, № 4. – Р. 4771–4791.
3. Chen Z., Zhang L., Weng F., Pan L., Lan Z. Tailored Visions: Enhancing Text-to-Image Generation with Personalized Prompt Rewriting // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024.
4. Gu J., Han Z., Chen S., Beirami A., He B., Zhang G., Liao R., Qin Y., Tresp V., Torr P.H.S. A Systematic Survey of Prompt Engineering on Vision-Language Foundation Models. – 2023. – 15 P.

5. Mahajan S., Rahman T., Yi K.M., Sigal L. Prompting Hard or Hardly Prompting: Prompt Inversion for Text-to-Image Diffusion Models // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024.
6. Mo W., Zhang T., Bai Y., Su B., Wen J.-R., Yang Q. Dynamic Prompt Optimizing for Text-to-Image Generation // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024.
7. Rombach R., Blattmann A., Lorenz D., Esser P., Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022.
8. Yao J. et al. PromptCoT: Align Prompt Distribution via Adapted Chain-of-Thought // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024.
9. Zhang C., Zhang C., Zhang M., Kweon I.S., Kim J. Text-to-image Diffusion Models in Generative AI: A Survey. – 2023. – 14 P.
10. Zhang L., Rao A., Agrawala M. Adding Conditional Control to Text-to-Image Diffusion Models // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). – 2023.

References

1. Ganiev A.A. Modeling Prompt-Based Visualization in Artificial Intelligence Systems // Nauchnye osnovy i sovremennyye problemy integratsii media i iskusstvennogo intellekta: sbornik dokladov Mezhdunarodnoj nauchno-prakticheskoy konferentsii. – 2026. – P. 95–98. (In Russ.)
2. Cao P., Zhou F., Song Q., Yang L. Controllable Generation With Text-to-Image Diffusion Models: A Survey // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2026. – Vol. 48, No. 4. – P. 4771–4791.
3. Chen Z., Zhang L., Weng F., Pan L., Lan Z. Tailored Visions: Enhancing Text-to-Image Generation with Personalized Prompt Rewriting // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024.
4. Cao P., Zhou F., Song Q., Yang L. Controllable Generation With Text-to-Image Diffusion Models: A Survey // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2026. – Vol. 48, No. 4. – P. 4771–4791.
5. Mahajan S., Rahman T., Yi K.M., Sigal L. Prompting Hard or Hardly Prompting: Prompt Inversion for Text-to-Image Diffusion Models // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024.
6. Mo W., Zhang T., Bai Y., Su B., Wen J.-R., Yang Q. Dynamic Prompt Optimizing for Text-to-Image Generation // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024.

7. Rombach R., Blattmann A., Lorenz D., Esser P., Ommer B. High-Resolution Image Synthesis with Latent Diffusion Models // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2022.
8. Yao J. et al. PromptCoT: Align Prompt Distribution via Adapted Chain-of-Thought // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). – 2024.
9. Zhang C., Zhang C., Zhang M., Kweon I.S., Kim J. Text-to-image Diffusion Models in Generative AI: A Survey. – 2023. – 14 P.
10. Zhang L., Rao A., Agrawala M. Adding Conditional Control to Text-to-Image Diffusion Models // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). – 2023.