

МЕТОД ФОРМАЛИЗАЦИИ И ОЦЕНКИ ВЫЧИСЛИТЕЛЬНОЙ НАГРУЗКИ В АДАПТИВНЫХ СИСТЕМАХ 3D-ВИЗУАЛИЗАЦИИ

Синицын А.В., Алпатова М.В., Иерусалимов И.Д.

Федеральное государственное бюджетное образовательное учреждение высшего образования «МИРЭА – Российский технологический университет», e-mail: sinicyn@mirea.ru

В работе рассматривается задача адаптивного управления уровнем детализации в 3D Gaussian Splatting (3DGS) при ограниченном вычислительном бюджете. Предложен метод покадрового отбора гауссиан на основе оценки их значимости в экранном пространстве, при котором управление вычислительной нагрузкой формализуется как задача оптимизации с учётом бюджетного ограничения. Метод объединяет оценку экранной стоимости примитивов и гибридный предиктор, учитывающий пространственные и временные характеристики изображения. Разработан воспроизводимый экспериментальный контур для унифицированного сбора метрик производительности и качества, включая частоту кадров, время рендеринга и показатели качества изображения. Проведено сравнение предложенного подхода с градиентными вариантами отбора на сценах различной сложности. Экспериментальные результаты показывают, что использование экранной стоимости в качестве критерия отбора обеспечивает более устойчивый компромисс между качеством изображения и производительностью по сравнению с градиентными методами и сохраняет применимость подхода в системах реального времени. Предложенная постановка задачи, воспроизводимый экспериментальный контур и унифицированный набор метрик могут быть использованы при формировании стандартизируемых процедур сравнительной оценки методов адаптивного рендеринга.

Ключевые слова: 3D Gaussian Splatting, уровень детализации, адаптивный рендеринг, LOD, оптимизация, экранное пространство, перцептивные метрики.

A METHOD FOR FORMALIZING AND EVALUATING COMPUTATIONAL LOAD IN ADAPTIVE 3D VISUALIZATION SYSTEMS

Sinicyn A.V., Alpatova M.V., Ierusalimov I.D.

Federal State Budgetary Educational Institution of Higher Education “MIREA – Russian Technological University”, e-mail: sinicyn@mirea.ru

This paper investigates adaptive level-of-detail control in 3D Gaussian Splatting (3DGS) under limited computational budgets. We propose a frame-wise Gaussian selection method that estimates primitive importance in screen space and formulates computational load control as a budget-constrained optimization problem. The method combines a screen-space cost estimate with a hybrid predictor that captures both spatial and temporal image characteristics. To ensure consistent evaluation, we develop a reproducible experimental framework for collecting unified performance and quality metrics, including frame rate, rendering time, and image quality measures. The proposed approach is evaluated against gradient-based selection strategies on scenes of varying complexity. The results show that using screen-space cost as the selection criterion yields a more robust trade-off between image quality and performance than gradient-based alternatives while remaining suitable for real-time applications. The proposed problem formulation, reproducible experimental setup, and unified set of metrics can be used in the development of standardized procedures for the comparative evaluation of adaptive rendering methods.

Keywords: 3D Gaussian Splatting, level of detail, adaptive rendering, LOD, optimization, screen space, perceptual metrics.

Введение

Метод 3D Gaussian Splatting (3DGS) широко применяется в задачах нейронного рендеринга для синтеза изображений новых ракурсов. Представление сцены в виде набора анизотропных трёхмерных гауссиан позволяет отказаться от ресурсоёмкого процесса дискретного отбора точек вдоль луча — сэмплирования, характерного для методов на основе нейронных полей излучения (Neural Radiance Fields, NeRF), и обеспечивает высокую скорость визуализации при сохранении качества изображения.

Вместе с тем производительность 3DGS напрямую зависит от количества используемых примитивов. Современные сцены могут содержать миллионы гауссиан, что приводит к существенному росту вычислительных затрат и ограничивает применение метода в интерактивных системах, особенно на устройствах с ограниченными вычислительными ресурсами.

Существующие подходы к ускорению 3DGS преимущественно ориентированы на статические оптимизации на уровне сцены: сжатие атрибутов, удаление избыточных примитивов, построение пространственных иерархий. Такие методы снижают среднюю вычислительную стоимость, однако выполняются на этапе подготовки модели и не позволяют динамически регулировать вычислительную нагрузку в зависимости от текущего ракурса и условий рендеринга. В классической компьютерной графике аналогичная задача решается с помощью методов уровня детализации (Level of Detail, LOD), обеспечивающих динамическое изменение сложности сцены в зависимости от условий наблюдения. Однако прямое перенесение этой идеи на представления типа 3DGS затруднено отсутствием явной топологической структуры и независимым характером вклада отдельных примитивов в итоговое изображение.

В этих условиях возникает задача покадрового управления количеством обрабатываемых гауссиан, позволяющего гибко регулировать вычислительную нагрузку в зависимости от ракурса и условий рендеринга при допустимой потере визуального качества. Для корректного сопоставления различных подходов требуется также единая процедура измерения производительности и качества изображения, обеспечивающая воспроизводимость результатов при одинаковых параметрах эксперимента.

Целью работы является разработка и экспериментальное исследование метода покадрового отбора гауссиан в 3D Gaussian Splatting, позволяющего адаптировать вычислительную нагрузку на основе оценки значимости примитивов в экранном пространстве.

Обзор существующих исследований

Предлагаемая работа находится на стыке исследований по снижению вычислительной стоимости 3D Gaussian Splatting (3DGS), методам уровня детализации (Level of Detail, LOD) в компьютерной графике и подходам к оценке визуальной значимости. В данном разделе рассматриваются ключевые публикации, формирующие контекст настоящего исследования.

Метод 3D Gaussian Splatting, предложенный Kerbl и соавт. [1], основан на явном представлении сцены в виде набора анизотропных гауссиан. Такой подход обеспечивает синтез изображений новых ракурсов с высоким качеством при интерактивных частотах кадров. Развитие этого направления привело к появлению множества работ, посвящённых повышению эффективности представления и рендеринга. К таким работам относятся методы сжатия атрибутов гауссиан, включая векторное квантование и энтропийное кодирование [2], а также структурные улучшения, нацеленные на упорядочение примитивов и ускорение растеризации [3]. Систематизация современных направлений развития метода представления сцен на основе проецирования и смешивания гауссиан (гауссова сплэттинга), включая вопросы производительности, качества и областей применения, представлена в обзоре [4].

Значительная часть исследований посвящена статическим оптимизациям, выполняемым после этапа обучения модели. Методы прореживания и объединения гауссиан позволяют уменьшить число примитивов и снизить стоимость рендеринга [5]. Альтернативный подход основан на построении пространственных иерархий и структур отсечения, что уменьшает количество примитивов, участвующих в обработке для заданного положения камеры [6–8]. Развиваются и гибридные решения, сочетающие оптимизацию структуры представления с эффективными реализациями на графических процессорах [9].

Общей чертой указанных работ является то, что основная оптимизация выполняется заранее, на уровне сцены или модели в целом. В отличие от них, в настоящей работе рассматривается динамическое управление вычислительной нагрузкой: покадровый отбор примитивов, выполняемый непосредственно во время рендеринга без изменения структуры обученной модели. Такой подход ориентирован не на предварительное упрощение сцены, а на гибкое регулирование вычислительной сложности в зависимости от текущего ракурса и условий визуализации.

Методы уровня детализации являются фундаментальным инструментом снижения вычислительной сложности в компьютерной графике [8]. Для полигональных моделей традиционно используются заранее подготовленные представления различной сложности, выбор которых зависит от расстояния до наблюдателя, параметров проекции или допустимой ошибки в экранном пространстве. Идеи иерархического представления и пространственного отсечения применяются также к воксельным и иным пространственным структурам данных [6, 7].

В задачах нейронного рендеринга развиваются подходы адаптивного распределения вычислений. Например,

AdaNeRF изменяет плотность сэмплирования вдоль лучей в зависимости от ожидаемого вклада в итоговое изображение [11]. В смежных работах используются многомасштабные представления признаков для перераспределения вычислений между различными областями сцены [10].

Однако прямое перенесение классических схем LOD на 3D Gaussian Splatting затруднено отсутствием у гауссиан явной топологической структуры. Вклад каждого примитива определяется не только его положением в пространстве, но и параметрами проекции в текущем кадре. По этой причине для 3DGS практический интерес представляет не только предварительное упрощение модели, но и кадровый отбор примитивов по критериям их вклада в изображение. Рассматриваемый в статье подход следует именно этой логике: управление вычислительной нагрузкой осуществляется на уровне текущего кадра на основе признаков экранного пространства.

При адаптивном рендеринге сокращение числа обрабатываемых примитивов неизбежно связано с потерей части информации, поэтому существенное значение имеет выбор критериев, позволяющих оценить вклад отдельных областей изображения в воспринимаемое качество. В качестве стандартных метрик в работах по компьютерному зрению и рендерингу широко используются PSNR, SSIM и LPIPS [16]. Эти показатели позволяют количественно оценивать различие между результатом рендеринга и эталонным изображением, хотя их чувствительность к различным типам искажений неодинакова.

Отдельное направление исследований связано с выделением значимых областей изображения и предсказанием визуального внимания. Для этой цели предложен ряд моделей обнаружения значимых областей в изображениях и видеопоследовательностях [17–20]. Для анализа временной согласованности применяются также подходы, учитывающие движение и межкадровые различия, в том числе на основе оптического потока и ошибок варпинга [21, 22]. Вместе с тем использование сложных обучаемых моделей значимости в контуре рендеринга реального времени не всегда допустимо из-за их вычислительной стоимости.

В этой связи представляют интерес более простые признаки, вычисляемые непосредственно в экранном пространстве. К ним относятся градиенты изображения, позволяющие выделять границы и области с высокой локальной изменчивостью [23, 24], а также межкадровая разность, отражающая появление новых или изменяющихся областей [25]. Близкие по смыслу эвристики применяются и в задачах анализа видеоданных, где требуется быстро выделять области, наиболее значимые с точки зрения восприятия или последующей обработки [26].

Настоящая работа опирается на указанные направления и рассматривает задачу кадрового отбора гауссиан как задачу оптимизации при ограничении вычислительного бюджета с использованием приближённой оценки значимости примитивов в экранном пространстве. Тем самым объединяются идеи адаптивного управления вычислениями, уровня детализации и оценки визуальной значимости применительно к представлению сцены в форме 3D Gaussian Splatting.

Методы

Предлагаемый метод представляет собой кадровый этап обработки, выполняемый перед финальным проходом растеризации 3DGS. Его цель — выбор подмножества гауссиан для рендеринга, максимизирующего предсказанное качество изображения в экранном пространстве при соблюдении ограничения на время кадра B . Метод включает три основных этапа: предварительный рендер, вычисление значимости примитивов и отбор с учётом вычислительного бюджета. При необходимости после отбора может выполняться этап постобработки.

Пусть $G = \{g_i\}_{i=1}^N$ — полный набор из N гауссиан обученной модели 3DGS. Оператор рендеринга $R(G; v)$ формирует изображение I для заданного ракурса v , при этом время выполнения обозначается как $T(G)$.

Для заданного ракурса v_t и заданного вычислительного бюджета B , определяемого через прокси-оценку стоимости примитивов, требуется найти подмножество $G'_t \subseteq G$, максимизирующее визуальное качество:

$$\max_{G'_t \subseteq G} Q(I_t, I_t^{gt}) \quad \text{при условии} \quad \sum_{g_i \in G'_t} c(g_i) \leq B, \quad \#(1)$$

где $I_t = R(G'_t; v_t)$, I_t^{gt} — эталонное изображение, Q — функция качества изображения (например, комбинация PSNR, SSIM), а $c(g_i)$ — приближённая оценка вычислительной стоимости гауссианы в экранном пространстве.

Следует отметить, что $c(g_i)$ используется как приближённая оценка вычислительной нагрузки, поэтому фактическое время кадра может отличаться от заданного бюджета B . В связи с этим задача рассматривается как задача оптимизации при бюджетном ограничении с приближённым контролем вычислительной нагрузки.

Поскольку эталонное изображение I_t^{gt} недоступно во время рендеринга, задача оптимизации заменяется

использованием приближённой оценки значимости $S(g_i, v_t)$, после чего выбираются гауссианы с наибольшими значениями этой величины при соблюдении бюджетного ограничения.

Таким образом, задача сводится к дискретной оптимизации вида:

$$\max_{G'_t \subseteq G} \sum_{g_i \in G'_t} S(g_i, v_t) \quad \text{при условии} \quad \sum_{g_i \in G'_t} c(g_i) \leq B, \#(2)$$

где $c(g_i)$ — оценка вычислительной стоимости примитива. Данная постановка эквивалентна NP-трудной задаче о рюкзаке, что обосновывает использование приближённых алгоритмов.

Оценка значимости определяется как комбинация пространственной и временной составляющих:

$$S(g_i, v_t) = w_s S_{\text{grad}}(g_i, v_t) + w_t S_{\text{temp}}(g_i, v_t), \#(3)$$

где w_s и w_t — скалярные веса.

Пространственная составляющая. Сначала выполняется предварительный рендер изображения пониженного разрешения I_t^{prev} с использованием небольшого фиксированного подмножества гауссиан или приближённого рендерера. Далее вычисляется величина градиента изображения в экранном пространстве $\|\nabla I_t^{\text{prev}}\|$. Использование подмножества гауссиан позволяет дополнительно снизить вычислительную нагрузку на этапе предварительного рендера без существенного ухудшения оценки значимости.

Для каждой гауссианы g_i её центр проецируется на плоскость изображения, после чего формируется мягкая маска растеризации M_i . Пространственная значимость определяется как:

$$S_{\text{grad}}(g_i) = \sum_{p \in \text{pixels}} M_i(p) \|\nabla I_t^{\text{prev}}(p)\|. \#(4)$$

Предложенная составляющая повышает приоритет гауссиан, вносящих вклад в границы и высокочастотные детали изображения.

Временная значимость (temporal surprise). Для учёта динамики или появления новых областей выполняется сравнение текущего предварительного изображения I_t^{prev} с предыдущим изображением, приведённым к текущему ракурсу $\tilde{I}_{t-1}^{\text{prev}}$. Карта временных различий определяется следующим образом:

$$D_t(p) = |I_t^{\text{prev}}(p) - \tilde{I}_{t-1}^{\text{prev}}(p)|. \#(5)$$

Тогда временная значимость вычисляется следующим образом:

$$S_{\text{temp}}(g_i) = \sum_{p \in \text{pixels}} M_i(p) D_t(p). \#(6)$$

Данная составляющая повышает приоритет гауссиан, расположенных в областях, изменяющихся во времени или становящихся видимыми.

В работе используются значения $w_s = 1.0$ и $w_t = 0.5$, обеспечивающие баланс между сохранением пространственной структуры и учётом межкадровых изменений. Эти значения выбраны эмпирически на основе предварительных экспериментов. Анализ чувствительности показал, что в диапазоне $w_t \in [0.3, 0.7]$ качественные и временные характеристики метода изменяются незначительно.

Дополнительно проведён качественный анализ вклада отдельных составляющих предиктора значимости. Использование только пространственной составляющей (S_{grad}) приводит к более агрессивному снижению качества изображения при жёстких ограничениях бюджета, тогда как добавление временной составляющей (S_{temp}) способствует большей устойчивости процесса отбора и уменьшению изменчивости набора выбранных гауссиан между кадрами. При этом влияние временной составляющей на статические метрики качества изображения остаётся ограниченным. Этот вывод согласуется с результатами экспериментов, где добавление временной составляющей не приводит к заметному улучшению статических метрик, но влияет на поведение метода в последовательности кадров.

Для каждой гауссианы вводится приближённая оценка вычислительной стоимости $c(g_i)$, пропорциональная

площади её проекции в экранном пространстве, то есть числу пикселей, затрагиваемых при растеризации.

Гауссианы сортируются по убыванию отношения значимости к стоимости:

$$\frac{S(g_i, v_t)}{c(g_i)}. \#(7)$$

Далее выполняется жадный отбор примитивов до тех пор, пока выполняется ограничение:

$$\sum_{g_i \in G'_t} c(g_i) \leq B. \#(8)$$

Использование жадного алгоритма обусловлено требованиями работы в реальном времени и его практической применимостью к задачам данного типа.

Предварительный рендер выполняется на пониженном разрешении (например, 256×256) с использованием фиксированного подмножества гауссиан. Это позволяет снизить вычислительные затраты на этапе оценки значимости.

Использование пониженного разрешения не оказывает существенного влияния на оценку градиентов и временных различий, поскольку указанные признаки в рассматриваемой постановке сохраняют информативность при масштабировании изображения.

При принятой схеме реализации основные этапы алгоритма характеризуются следующими оценками вычислительной сложности:

- предварительный рендер: $O(N')$, где $N' \ll N$;
- вычисление значимости: $O(N)$;
- сортировка: $O(N \log N)$;
- отбор: $O(N)$.

В рамках данной оценки наиболее затратным этапом является сортировка. На практике итоговое время работы определяется также затратами на предварительный проход, вычисление значимости и организацию отбора, поэтому фактическая применимость метода в интерактивных сценариях оценивается по результатам экспериментов.

Эксперименты

Цель экспериментального исследования заключается в сравнительном анализе стратегий покaдрового отбора гауссиан в 3D Gaussian Splatting в условиях ограниченного вычислительного бюджета.

Эксперимент направлен на ответы на следующие исследовательские вопросы:

1. Существует ли устойчивая связь между заданным вычислительным бюджетом и фактической вычислительной нагрузкой при использовании метода Cost-Budget?
2. Сохраняется ли качество изображения при уменьшении доступного вычислительного бюджета?
3. Как соотносятся между собой методы Cost-Budget, Gradient и Gradient + Temporal?
4. Влияет ли учёт временной компоненты на устойчивость покaдрового отбора?

Набор сцен. Для проверки обобщающей способности подхода использовались три сцены различной сложности:

- objaverse_composite40 — композитная сцена высокой сложности, содержащая 1 999 999 гауссиан;
- xbox_composite20 — сцена средней сложности, содержащая 419 711 гауссиан;
- xbox_0005 — компактная сцена малой сложности, содержащая 18 168 гауссиан.

Такой набор покрывает три характерных режима работы: очень плотное представление, среднюю по размеру сцену и компактный объектный пример. Это позволяет оценить устойчивость методов как на тяжёлых сценах с высокой стоимостью рендеринга, так и на относительно простых случаях, где качество близко к исходному даже при существенном прореживании. В экспериментальном исследовании использованы сцены, представленные в формате 3D Gaussian Splatting и полученные из открытых наборов данных, распространяемых через платформу HuggingFace¹. Компактные сцены с небольшим числом примитивов (например, xbox_0005, xbox_composite20) сформированы на основе синтетических объектов из коллекций ModelNet и ShapeNet, адаптированных к представлению в виде гауссиан [1,27]. Для оценки поведения метода в

¹ <https://huggingface.co/datasets/ashawkey/ShapeNetGaussian>

условиях высокой вычислительной нагрузки использованы крупномасштабные сцены (например, objaverse_composite40), сформированные на основе набора Objaverse² [28]. Использование сцен различной сложности позволяет провести всесторонний анализ соотношения качества изображения и производительности, а также оценить устойчивость предложенного метода при изменении вычислительного бюджета.

В экспериментах рассматривались следующие методы:

1. Full — полный рендер без отбора гауссиан;
2. Cost-Budget — предложенный метод отбора на основе экранной стоимости;
3. Gradient — отбор на основе пространственного градиента изображения;
4. Gradient + Temporal — метод отбора, учитывающий пространственный градиент и межкадровые различия.

Метод Cost-Budget рассматривается как основной объект исследования, тогда как методы Gradient и Gradient + Temporal используются для сравнительного анализа различных критериев отбора примитивов.

Протокол оценки. Во всех экспериментах использовались одинаковые параметры рендеринга: разрешение 1280×720 , длина траектории 60 кадров, из которых первые 10 кадров рассматривались как этап прогрева. Отбор выполнялся с периодом $sel_every = 8$. Для обеспечения воспроизводимости использовалось фиксированное значение $seed = 12345$.

Эксперименты проводились в двух режимах:

- режим фиксированной доли примитивов (ratio mode), в котором доля сохраняемых гауссиан задавалась значениями 1.0, 0.5, 0.25 и 0.125;
- режим целевого времени кадра (time mode) в котором использовались ориентиры 16.7 мс, 11.1 мс и 8.3 мс.

Использование единых параметров обеспечивает воспроизводимость результатов и корректность сравнительного анализа.

Для оценки эффективности методов использовались следующие метрики:

- FPS — частота кадров, характеризующая производительность;
- среднее время кадра (frame time);
- PSNR, SSIM — метрики качества изображения;
- число выбранных примитивов (selected count) — количество гауссиан, отобранных для финального рендера.

Техническая реализация. Все методы были реализованы в едином экспериментальном контуре на базе модифицированного растеризационного контура 3D Gaussian Splatting. Для каждого прогона сохранялись:

- агрегированные CSV-файлы с итоговыми показателями;
- покадровые CSV-файлы;
- результаты визуализации;
- временные показатели предварительного прохода, отбора и финального рендера.

Такая организация эксперимента обеспечила сопоставимость результатов между методами и позволила исключить ошибки, характерные для предыдущей версии экспериментальной инфраструктуры.

Эксперименты проводились на ноутбуке со следующими характеристиками: центральный процессор Intel (архитектура Intel64, семейство 6, модель 158), 16 ГБ оперативной памяти и графический процессор NVIDIA GeForce GTX 1660 Ti с 6 ГБ видеопамяти.

Вычисления выполнялись с использованием графического ускорения (GPU) в среде PyTorch 2.4.1 с поддержкой CUDA 12.4. Использование GPU позволило обеспечить выполнение рендеринга и алгоритмов отбора примитивов в интерактивном режиме.

Все методы сравнивались в идентичных аппаратных и программных условиях, что исключает влияние вычислительной платформы на результаты и обеспечивает корректность сравнительного анализа.

Следует отметить, что фактическое время кадра систематически превышает заданное целевое значение. Это связано с использованием приближённой оценки стоимости примитивов и наличием дополнительных накладных расходов, возникающих на этапах предварительного прохода, сортировки и отбора. Таким образом, предлагаемый метод обеспечивает приближённое, а не строгое управление вычислительной нагрузкой.

Для количественной оценки точности управления вычислительным бюджетом используется относительное отклонение фактического времени кадра от заданного бюджета:

² <https://huggingface.co/datasets/allenai/objaverse>

$$\epsilon = \frac{T_{actual} - T_{target}}{T_{target}} \cdot \#(9)$$

где T_{actual} — фактическое время кадра, T_{target} — заданный вычислительный бюджет.

Результаты анализа представлены на рисунке 1, где показано соответствие между заданным вычислительным бюджетом и фактическим временем кадра.

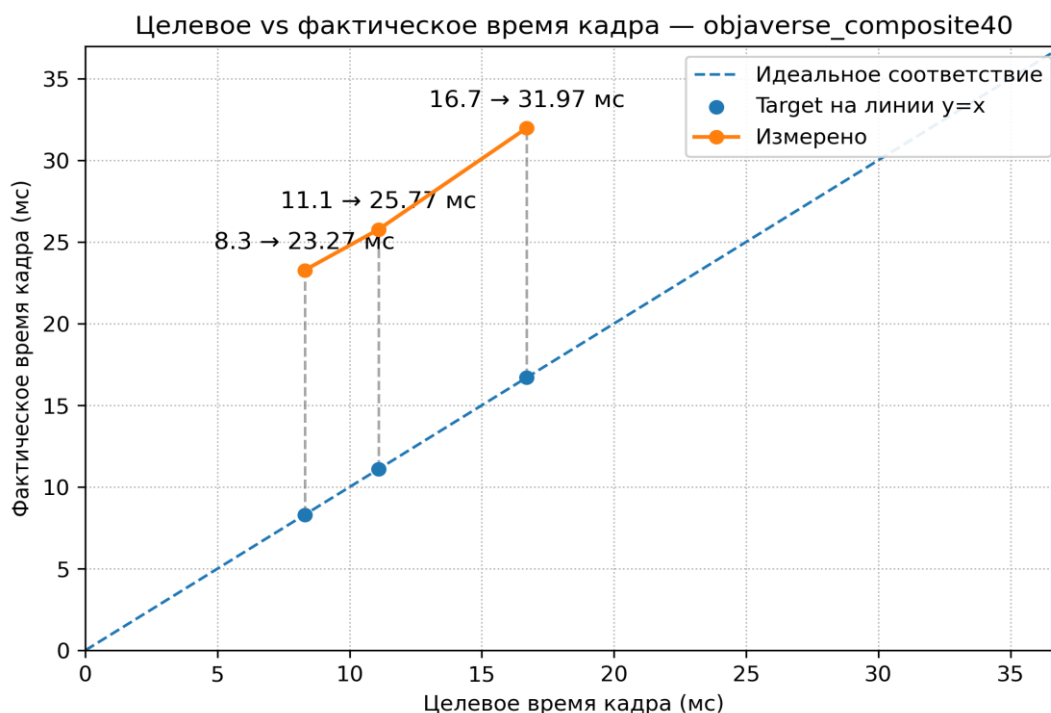


Рисунок 1 – Соответствие между заданным бюджетом и фактическим временем кадра для сцены objaverse_composite40

Как показано на рисунке 1, пунктирная линия соответствует идеальному случаю $T_{actual} = T_{target}$, тогда как реальные измерения располагаются выше неё. Вертикальные отрезки показывают величину отклонения фактического времени кадра от заданного бюджета. Наблюдается систематическое превышение заданного времени, усиливающееся при уменьшении бюджета, что указывает на приближённый характер используемой прокси-оценки вычислительной стоимости.

Анализ показал, что для сложных сцен наблюдается систематическое превышение заданного времени кадра. В частности, для сцены *objaverse_composite40* отклонение составляет от 91% до 180% в зависимости от значения бюджета, что наглядно отражается расстоянием между точками фактического времени и линией идеального соответствия.

Для сцены средней сложности (*xbox_composite20*) отклонение находится в диапазоне от 27% до 132%, тогда как для компактной сцены *xbox_0005* различия между заданным и фактическим временем незначительны.

При этом сохраняется монотонная зависимость между заданным и фактическим временем кадра, что позволяет использовать предложенный метод для относительного управления вычислительной нагрузкой.

Полученные результаты подтверждают, что используемая прокси-оценка стоимости примитивов приводит к приближённому, а не точному контролю времени кадра. Дополнительно наблюдается устойчивая зависимость между оценкой стоимости примитивов и фактическим временем рендеринга, что подтверждает пригодность используемой прокси-метрики для относительного управления вычислительной нагрузкой.

Результаты

Режим фиксированной доли примитивов. Сначала был исследован режим фиксированной доли примитивов, в котором ограничение задаётся как доля экранной стоимости полного набора гауссиан. Результаты приведены в таблице 1.

Из таблицы 1 следует, что уменьшение доли сохраняемых примитивов сопровождается сокращением числа выбранных гауссиан и, как правило, снижением времени финального рендера. На сценах средней и высокой сложности это в большинстве случаев приводит к росту производительности, однако для отдельных конфигураций наблюдаются отклонения от общего тренда, связанные с накладными расходами

предварительного прохода, сортировки и отбора.

Таблица 1 – Сравнение методов при фиксированной доле примитивов.

Scene	Method	Ratio	FPS (frames/s)	Total (ms)	Render (ms)	Selected	PSNR
objaverse_composite40	Full	1.0	22.2	45.03	24.96	1999999	≥60
objaverse_composite40	Cost-Budget	1.0	22.5	44.47	26.33	1999999	≥60
objaverse_composite40	Cost-Budget	0.5	12.3	81.05	14.27	781672	44.29
objaverse_composite40	Cost-Budget	0.25	41.8	23.94	6.04	334140	37.16
objaverse_composite40	Cost-Budget	0.125	46.8	21.36	4.57	134334	33.05
xbox_composite20	Full	1.0	47.2	21.17	7.08	419711	≥60
xbox_composite20	Cost-Budget	1.0	49.5	20.19	7.06	419711	≥60
xbox_composite20	Cost-Budget	0.5	58.4	17.13	4.37	206515	55.80
xbox_composite20	Cost-Budget	0.25	65.6	15.24	2.82	106973	44.52
xbox_composite20	Cost-Budget	0.125	68.3	14.65	2.12	59211	38.79
xbox_0005	Full	1.0	44.0	22.73	1.88	18168	≥60
xbox_0005	Cost-Budget	1.0	45.4	22.05	2.00	18168	≥60
xbox_0005	Cost-Budget	0.5	43.7	22.89	1.58	8719	53.53
xbox_0005	Cost-Budget	0.25	43.3	23.09	1.40	4460	38.59
xbox_0005	Cost-Budget	0.125	44.2	22.61	1.25	2591	30.48

Для наиболее сложной сцены *objaverse_composite40* при переходе от доли 1.0 к 0.125 число выбранных гауссиан уменьшается с 1 999 999 до 134 334, а время финального рендера — с 26.33 до 4.57 мс. При этом зависимость общего времени кадра и частоты кадров оказывается немонотонной: при доле 0.5 зафиксированы 12.3 FPS и 81.05 мс, тогда как при дальнейшем уменьшении доли до 0.25 и 0.125 производительность возрастает до 41.8 и 46.8 FPS при общем времени кадра 23.94 и 21.36 мс соответственно. Это показывает, что итоговая производительность определяется не только числом отобранных примитивов, но и накладными расходами алгоритма.

Для сцены средней сложности *xbox_composite20* наблюдается более согласованная картина. При уменьшении доли от 1.0 до 0.125 производительность возрастает с 49.5 до 68.3 FPS, общее время кадра уменьшается с 20.19 до 14.65 мс, а число выбранных гауссиан сокращается с 419 711 до 59 211. При этом время финального рендера снижается с 7.06 до 2.12 мс.

Для компактной сцены *xbox_0005* число выбранных гауссиан также последовательно уменьшается — с 18 168 до 2 591, а время финального рендера снижается с 2.00 до 1.25 мс. Однако итоговые значения FPS меняются незначительно: от 45.4 при доле 1.0 до 44.2 при доле 0.125. Это указывает на то, что для компактных сцен эффект отбора примитивов по общему времени кадра выражен слабее.

Режим целевого времени кадра. В режиме целевого времени кадра сравниваются методы Cost-Budget, Gradient и Gradient + Temporal при значениях 16.7, 11.1 и 8.3 мс, рассматриваемых как ориентиры для настройки вычислительного бюджета. Результаты приведены в таблице 2.

Как видно из таблицы 2, при одинаковых целевых настройках методы достигают различных фактических значений времени кадра. По этой причине приведённые результаты следует интерпретировать как сравнение поведения методов при одинаковых настройках управляющего параметра, а не как строгое сопоставление при равном времени выполнения.

Следует отметить, что заданные значения времени кадра рассматриваются как ориентиры для настройки вычислительного бюджета, а не как жёсткие ограничения, поскольку используемая приближённая оценка стоимости примитивов и накладные расходы алгоритма приводят к систематическому отклонению фактического времени от целевого значения. В результате сравнение методов при одинаковых целевых значениях не соответствует строго одинаковым вычислительным условиям.

Наиболее сложной является сцена *objaverse_composite40*. Для неё метод Cost-Budget во всех трёх режимах показывает наилучшие значения как по производительности, так и по PSNR среди сравниваемых методов. В режиме 16.7 мс в таблице указаны 31.3 FPS, 31.97 мс и PSNR ≥ 60 дБ при 1 006 349 выбранных гауссиан, тогда как методы Gradient и Gradient + Temporal дают 23.7 и 23.3 FPS при PSNR 35.18 и 34.87 дБ соответственно. При переходе к целевым значениям 11.1 и 8.3 мс метод Cost-Budget сохраняет преимущество: PSNR составляет 39.51 и 36.08 дБ при 38.4 и 42.9 FPS, тогда как для Gradient эти значения равны 32.72 и 31.36 дБ при 25.8 и 29.8 FPS, а для Gradient + Temporal — 32.63 и 31.34 дБ при 25.9 и 28.2 FPS.

На сцене средней сложности *xbox_composite20* различия по качеству изображения выражены слабо: во

всех режимах для всех методов в таблице указано $PSNR \geq 60$ дБ. В этих условиях сравнение смещается в сторону производительности. Метод Cost-Budget во всех трёх режимах показывает наибольшую частоту кадров — 51.7, 52.1 и 54.0 FPS — при общем времени кадра 19.35, 19.20 и 18.53 мс, тогда как для Gradient эти значения составляют 48.0, 50.5 и 52.0 FPS, а для Gradient + Temporal — 47.7, 49.3 и 50.3 FPS.

Таблица 2 – Сравнение методов при целевом времени кадра.

Scene	Method	Target (ms)	FPS (frames/s)	Total (ms)	Render (ms)	Selected	PSNR
objaverse_composite40	Cost-Budget	16.7	31.3	31.97	14.63	1006349	≥ 60
objaverse_composite40	Gradient	16.7	23.7	42.21	22.27	1524685	35.18
objaverse_composite40	Grad + Temp	16.7	23.3	42.85	22.00	1490787	34.87
objaverse_composite40	Cost-Budget	11.1	38.4	26.02	9.35	630227	39.51
objaverse_composite40	Gradient	11.1	25.8	38.78	18.33	1151559	32.72
objaverse_composite40	Grad + Temp	11.1	25.9	38.58	18.22	1130486	32.63
objaverse_composite40	Cost-Budget	8.3	42.9	23.30	6.56	391245	36.08
objaverse_composite40	Gradient	8.3	29.8	33.60	14.52	830820	31.36
objaverse_composite40	Grad + Temp	8.3	28.2	35.40	14.69	829500	31.34
xbox_composite20	Cost-Budget	16.7	51.7	19.35	6.54	382421	≥ 60
xbox_composite20	Gradient	16.7	48.0	20.85	6.53	385024	≥ 60
xbox_composite20	Grad + Temp	16.7	47.7	20.98	6.45	385599	≥ 60
xbox_composite20	Cost-Budget	11.1	52.1	19.20	6.20	344479	≥ 60
xbox_composite20	Gradient	11.1	50.5	19.81	5.85	343850	≥ 60
xbox_composite20	Grad + Temp	11.1	49.3	20.30	5.69	347815	≥ 60
xbox_composite20	Cost-Budget	8.3	54.0	18.53	5.53	305082	≥ 60
xbox_composite20	Gradient	8.3	52.0	19.22	5.42	308896	≥ 60
xbox_composite20	Grad + Temp	8.3	50.3	19.90	5.51	305585	≥ 60
xbox_0005	Cost-Budget	16.7	46.8	21.35	1.87	16480	≥ 60
xbox_0005	Gradient	16.7	27.5	36.36	2.75	16581	≥ 60
xbox_0005	Grad + Temp	16.7	42.8	23.36	1.92	16510	≥ 60
xbox_0005	Cost-Budget	11.1	48.8	20.51	1.79	16316	≥ 60
xbox_0005	Gradient	11.1	25.8	38.81	2.41	15875	≥ 60
xbox_0005	Grad + Temp	11.1	41.4	24.14	2.00	15829	≥ 60
xbox_0005	Cost-Budget	8.3	50.0	20.00	1.68	14775	≥ 60
xbox_0005	Gradient	8.3	27.1	36.89	3.14	12959	≥ 60
xbox_0005	Grad + Temp	8.3	45.3	22.08	1.85	15049	≥ 60

Для компактной сцены *xbox_0005* различия по качеству изображения также выражены слабо, поскольку во всех исследованных режимах в таблице приведены значения $PSNR \geq 60$ дБ. При этом и здесь метод Cost-Budget показывает наилучшие показатели производительности: 46.8, 48.8 и 50.0 FPS против 27.5, 25.8 и 27.1 FPS для Gradient и 42.8, 41.4 и 45.3 FPS для Gradient + Temporal. Это позволяет рассматривать компактную сцену прежде всего как пример сравнения методов по производительности, а не по качеству изображения.

Таким образом, по данным таблицы 2 метод Cost-Budget на исследованных сценах не только формирует иной компромисс между качеством изображения и производительностью, но и в представленных режимах не уступает сравниваемым вариантам по производительности. Вместе с тем эти результаты следует трактовать как сравнение при одинаковых целевых настройках вычислительного бюджета, а не при строго одинаковом фактическом времени выполнения.

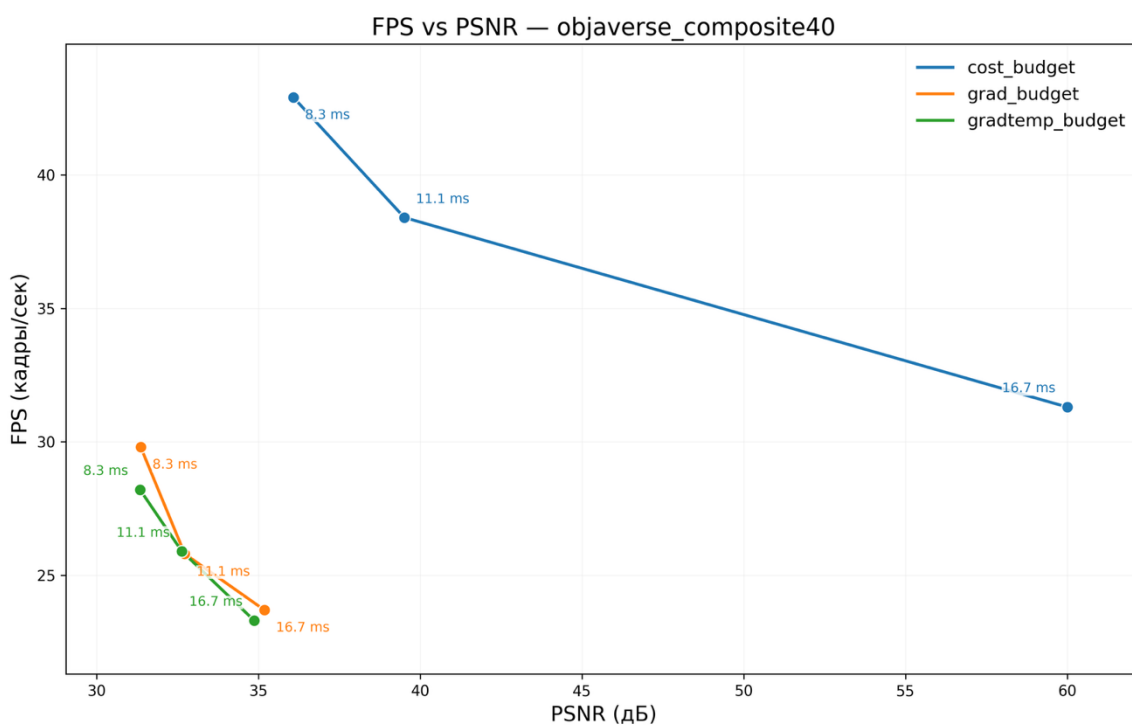


Рисунок 2 – Граница Парето в пространстве «качество–производительность» для различных методов отбора гауссиан. Точка Full соответствует максимальному качеству и вычислительной стоимости.

Как видно из рисунка 2, после приведения графика в соответствие с табличными данными точки метода Cost-Budget для сцены *objaverse_composite40* располагаются выше и правее точек методов Gradient и Gradient + Temporal. Это означает, что в представленных режимах метод Cost-Budget обеспечивает одновременно более высокое качество изображения и более высокую частоту кадров.

Градиентные методы формируют точки с меньшими значениями PSNR и FPS. Добавление временной компоненты не приводит к улучшению положения на границе Парето по сравнению с методом Gradient: соответствующие точки остаются близкими по качеству изображения и производительности.

Таким образом, в рамках представленных результатов метод Cost-Budget демонстрирует преимущество на сложной сцене как по качеству изображения, так и по производительности. Качественный анализ последовательностей рендеринга при этом указывает на возможное повышение устойчивости отбора при использовании временной компоненты, однако количественная оценка этого эффекта требует отдельного исследования.

На основании проведённого эксперимента можно сделать следующие выводы:

1. Предложенный метод позволяет приближённо управлять вычислительной нагрузкой за счёт изменения числа выбираемых гауссиан.
2. Уменьшение числа выбранных примитивов в большинстве исследованных конфигураций приводит к снижению стоимости финального рендера, однако итоговое время кадра зависит также от накладных расходов алгоритма.
3. На сложных сценах различия между стратегиями отбора выражены заметнее, чем на компактных сценах;

4. В режиме целевого времени кадра на исследованных сценах метод Cost-Budget не уступает сравнимым вариантам по производительности, а на сложной сцене *objaverse_composite40* также показывает более высокие значения PSNR.

5. Предлагаемый способ задания бюджета позволяет говорить о приближённом, а не точном контроле времени кадра.

Ограничения

Несмотря на полученные результаты, предложенный подход имеет ряд ограничений, которые необходимо учитывать при интерпретации результатов и дальнейшем развитии метода. Используемые метрики качества изображения (PSNR, SSIM) обладают ограниченной чувствительностью к локальным структурным и временным искажениям, что может влиять на точность количественной оценки визуального качества. В частности, данные метрики не всегда отражают субъективное восприятие изменений изображения, особенно в динамических сценах.

Экспериментальное исследование проводилось на ограниченном наборе сцен, что может снижать обобщаемость полученных результатов. Для более полной оценки поведения метода требуется расширение экспериментальной базы, включая сцены с различной геометрической структурой и динамическими объектами.

Выбор параметров метода, таких как веса пространственной и временной составляющих, выполнялся эмпирически. Несмотря на устойчивое поведение на рассматриваемых сценах, их универсальность для других условий требует дополнительного анализа. Модель временной значимости в текущей реализации в большей степени учитывает изменения, обусловленные движением камеры, и в меньшей степени отражает динамику объектов внутри сцены. Предложенный метод использует приближённое решение задачи отбора примитивов, что обусловлено требованиями работы в реальном времени. В результате оптимальность получаемого решения не гарантируется, однако получаемое решение позволяет практически исследовать компромисс между качеством изображения и вычислительной стоимостью. В работе не представлен детальный количественный анализ показателей временной устойчивости отбора, что требует дальнейшего исследования.

Перечисленные ограничения определяют направления дальнейших исследований, включая разработку более чувствительных метрик качества, расширение набора тестовых сцен и совершенствование методов оценки значимости примитивов.

Заключение

В работе рассмотрена задача покадрового адаптивного управления уровнем детализации в 3D Gaussian Splatting в условиях ограниченного вычислительного бюджета. Предложен метод отбора гауссиан, основанный на оценке их значимости в экранном пространстве с использованием пространственных и временных признаков.

Разработанный подход реализует приближённое динамическое управление вычислительной нагрузкой на уровне отдельных кадров и может быть интегрирован в существующие пайплайны 3DGS без изменения структуры представления сцены. Предложенная формализация задачи и используемая эвристика отбора позволяют рассматривать метод как пригодный для интерактивных сценариев рендеринга.

Проведённое экспериментальное исследование показало, что использование экранной стоимости в качестве критерия отбора позволяет уменьшать стоимость финального рендера в режиме фиксированной доли примитивов и формирует устойчивые различия по сравнению с градиентными вариантами отбора в режиме целевого времени кадра. На сложной сцене *objaverse_composite40* по данным таблицы 2 метод Cost-Budget показывает более высокие значения PSNR при более высокой частоте кадров по сравнению с методами Gradient и Gradient + Temporal. При этом добавление временной компоненты влияет на характер покадрового отбора, однако количественная оценка этого эффекта требует дополнительного исследования.

Практическим результатом работы является формирование воспроизводимого экспериментального контура, включающего унифицированные метрики, протокол измерений и процедуру сравнения методов. В этом смысле работа вносит вклад в формализацию и унификацию процедур оценки адаптивных методов рендеринга, задавая состав измеряемых показателей, режимы испытаний и принципы сопоставления результатов.

Представленные результаты и предложенная экспериментальная схема могут быть использованы при формировании требований к процедурам сравнительной оценки и методическим рекомендациям для испытаний систем интерактивной 3D-визуализации и нейронного рендеринга.

Список литературы

1. Kerbl B., Kopanas G., Leimkühler T., Drettakis G. 3D gaussian splatting for real-time radiance field rendering // *ACM Transactions on Graphics (TOG)*. 2023. Vol. 42, no. 4. P. 1–14.
2. Niedermayr S., Stumpfegger J., Westermann R. Compressed 3D gaussian splatting for accelerated novel view synthesis / *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. 2024. P. 10349–10358.
3. Lu T., Yu M., Xu L., Xiangli Y., Wang L., Lin D., Dai B. Scaffold-GS: Structured 3D gaussians for view-adaptive rendering / *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. 2024. P. 5656–5665.
4. Петрищев К. С., Морозкин Н. К., Семенов В. А., Тарлапан О. А., Шуткин В. Н. Обзор современных методов гауссова сплэттинга в приложениях компьютерной графики // *Труды Института системного программирования РАН*. — 2025. — Т. 37, № 6. — С. 201–226. — [https://doi.org/10.15514/ISPRAS-2025-37\(6\)-59](https://doi.org/10.15514/ISPRAS-2025-37(6)-59).
5. Wu T., Yuan Y., Zhang L., Yang J., Cao Y., Yan L., Gao L. Recent advances in 3D gaussian splatting // *Computational Visual Media*. 2024. Vol. 10. P. 1–25.
6. Laine S., Karras T. Efficient sparse voxel octrees / *Proceedings of the ACM SIGGRAPH symposium on interactive 3D graphics and games (I3D)*. 2010. P. 55–63.
7. Wald I., Boulos S., Shirley P. Ray tracing deformable scenes using dynamic bounding volume hierarchies // *ACM Transactions on Graphics (TOG)*. 2007. Vol. 26, no. 1. P. 6:1–6:8.
8. Luebke D., Reddy M., Cohen J.D., Varshney A., Watson B., Huebner R. Level of detail for 3D graphics / *Morgan Kaufmann*, 2003.
9. Franke L., Fink L., Stamminger M. VR-splatting: Foveated radiance field rendering via 3D gaussian splatting and neural points // *Proceedings of the ACM on Computer Graphics and Interactive Techniques*. 2025.
10. Müller T., Evans A., Schied C., Keller A. Instant neural graphics primitives with a multiresolution hash encoding // *ACM Transactions on Graphics (TOG)*. 2022. Vol. 41, no. 4. P. 1–15.
11. Klingner M., Martin-Brualla R., Wetzstein G. AdaNeRF: Adaptive sampling for real-time rendering of neural radiance fields / *European conference on computer vision (ECCV)*. 2022. P. 254–270.
12. Luebke D., Erikson C. View-dependent simplification of arbitrary polygonal environments / *Proceedings of the 24th annual conference on computer graphics and interactive techniques (SIGGRAPH)*. 1997. P. 199–208.
13. Cater K., Chalmers A. Perceptually based rendering: Concepts and applications // *Proceedings of the Eurographics*. 2003.
14. Wang L., Shi X., Liu Y. Foveated rendering: A state-of-the-art survey // *Computational Visual Media*. 2023. Vol. 9, no. 1. P. 1–32.
15. Семёнов В.А., Шуткин В.Н., Золотов В.А., Морозов С.В. Расширение метода иерархических уровней детализации для динамических сцен с детерминированным характером событий / *Труды Международной конференции GraphiCon*. Брянск: ИПМ им. М.В. Келдыша, 2019. P. 37–41.
16. Zhang R., Isola P., Efros A.A., Shechtman E., Wang O. The unreasonable effectiveness of deep features as a perceptual metric // *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018. P. 586–595.
17. Wang W., Shen J., Shao L. Video salient object detection via fully convolutional networks // *IEEE Transactions on Image Processing*. 2017. Vol. 27, no. 1. P. 38–49.
18. Wei L., Zhao S., Bourahla O.F., Li X., Wu F. End-to-end video saliency detection via a deep contextual spatiotemporal network // *IEEE Transactions on Neural Networks and Learning Systems*. 2020.
19. Zhang K., Chen Z., Liu S. A spatial-temporal recurrent neural network for video saliency prediction // *IEEE Transactions on Image Processing*. 2020.
20. Huang L., Yan P., Li G., Wang Q., Lin L. Attention embedded spatio-temporal network for video salient object detection // *IEEE Access*. 2019.
21. Maczyta L., Bouthemy P., Le Meur O. CNN-based temporal detection of motion saliency in videos // *Pattern Recognition Letters*. 2019.
22. Jian M., Lu X., Yu X., Ju Y., Yu H. Flow-edge-net: Video saliency detection based on optical flow and edge-weighted balance loss // *IEEE Transactions*. 2023.
23. Wang W., Shen J., Shao L. Consistent video saliency using local gradient flow optimization and global refinement // *IEEE Transactions on Image Processing*. 2015. Vol. 24, no. 11. P. 4185–4196.
24. Droste R., Jiao J., Noble J.A. Unified image and video saliency modeling / *European conference on computer vision (ECCV)*. Springer, 2020. P. 472–489.
25. Bhattacharya S., Venkatesh K.S., Gupta S. Background estimation and motion saliency detection using total variation-based video decomposition // *Signal, Image and Video Processing*. Springer, 2017. Vol. 11, no. 4. P. 699–706.
26. Culibrk D., Mirkovic M., Zlokolica V., Pokric M. Salient motion features for video quality assessment // *IEEE Transactions on Image Processing*. 2010. Vol. 19, no. 4. P. 948–963.
27. Wu Z., Song S., Khosla A., Yu F., Zhang L., Tang X., Xiao J. 3D ShapeNets: A deep representation for volumetric shapes / *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*. 2015. P. 1912–1920.
28. Deitke M., Schwenk D., Salvador J., et al. Objaverse: A Universe of Annotated 3D Objects // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. — 2023. — Pp. 13142–13153. — https://openaccess.thecvf.com/content/CVPR2023/papers/Deitke_Objaverse_A_Universe_of_Annotated_3D_Objects

References

1. Kerbl B., Kopanas G., Leimkühler T., Drettakis G. 3D gaussian splatting for real-time radiance field rendering // *ACM Transactions on Graphics (TOG)*. 2023. Vol. 42, no. 4. P. 1–14.
2. Niedermayr S., Stumpfegger J., Westermann R. Compressed 3D gaussian splatting for accelerated novel view synthesis / *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. 2024. P. 10349–10358.
3. Lu T., Yu M., Xu L., Xiangli Y., Wang L., Lin D., Dai B. Scaffold-GS: Structured 3D gaussians for view-adaptive rendering / *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (CVPR)*. 2024. P. 5656–5665.
4. Petrishchev K.S., Morozkin N.K., Semenov V.A., Tarlapan O.A., Shutkin V.N. A survey of modern gaussian splatting methods in computer graphics applications // *Proceedings of the Institute for System Programming of the Russian Academy of Sciences*. 2025. Vol. 37, no. 6. P. 201–226.
5. Wu T., Yuan Y., Zhang L., Yang J., Cao Y., Yan L., Gao L. Recent advances in 3D gaussian splatting // *Computational Visual Media*. 2024. Vol. 10. P. 1–25.
6. Laine S., Karras T. Efficient sparse voxel octrees / *Proceedings of the ACM SIGGRAPH symposium on interactive 3D graphics and games (I3D)*. 2010. P. 55–63.
7. Wald I., Boulos S., Shirley P. Ray tracing deformable scenes using dynamic bounding volume hierarchies // *ACM Transactions on Graphics (TOG)*. 2007. Vol. 26, no. 1. P. 6:1–6:8.
8. Luebke D., Reddy M., Cohen J.D., Varshney A., Watson B., Huebner R. *Level of detail for 3D graphics* / Morgan Kaufmann, 2003.
9. Franke L., Fink L., Stamminger M. VR-splatting: Foveated radiance field rendering via 3D gaussian splatting and neural points // *Proceedings of the ACM on Computer Graphics and Interactive Techniques*. 2025.
10. Müller T., Evans A., Schied C., Keller A. Instant neural graphics primitives with a multiresolution hash encoding // *ACM Transactions on Graphics (TOG)*. 2022. Vol. 41, no. 4. P. 1–15.
11. Klingner M., Martin-Brualla R., Wetzstein G. AdaNeRF: Adaptive sampling for real-time rendering of neural radiance fields / *European conference on computer vision (ECCV)*. 2022. P. 254–270.
12. Luebke D., Erikson C. View-dependent simplification of arbitrary polygonal environments / *Proceedings of the 24th annual conference on computer graphics and interactive techniques (SIGGRAPH)*. 1997. P. 199–208.
13. Cater K., Chalmers A. *Perceptually based rendering: Concepts and applications* // *Proceedings of the Eurographics*. 2003.
14. Wang L., Shi X., Liu Y. Foveated rendering: A state-of-the-art survey // *Computational Visual Media*. 2023. Vol. 9, no. 1. P. 1–32.
15. Semenov V. A., Shutkin V. N., Zolotov V. A., Morozov S. V. Extension of the hierarchical level-of-detail method for dynamic scenes with deterministic event behavior // *Proceedings of the International Conference GraphiCon*. — Bryansk, 2019. — Pp. 37–41. — <https://doi.org/10.30987/graphicon-2019-1-37-41>.
16. Zhang R., Isola P., Efros A.A., Shechtman E., Wang O. The unreasonable effectiveness of deep features as a perceptual metric // *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2018. P. 586–595.
17. Wang W., Shen J., Shao L. Video salient object detection via fully convolutional networks // *IEEE Transactions on Image Processing*. 2017. Vol. 27, no. 1. P. 38–49.
18. Wei L., Zhao S., Bourahla O.F., Li X., Wu F. End-to-end video saliency detection via a deep contextual spatiotemporal network // *IEEE Transactions on Neural Networks and Learning Systems*. 2020.
19. Zhang K., Chen Z., Liu S. A spatial-temporal recurrent neural network for video saliency prediction // *IEEE Transactions on Image Processing*. 2020.
20. Huang L., Yan P., Li G., Wang Q., Lin L. Attention embedded spatio-temporal network for video salient object detection // *IEEE Access*. 2019.
21. Maczyta L., Bouthemy P., Le Meur O. CNN-based temporal detection of motion saliency in videos // *Pattern Recognition Letters*. 2019.
22. Jian M., Lu X., Yu X., Ju Y., Yu H. Flow-edge-net: Video saliency detection based on optical flow and edge-weighted balance loss // *IEEE Transactions*. 2023.
23. Wang W., Shen J., Shao L. Consistent video saliency using local gradient flow optimization and global refinement // *IEEE Transactions on Image Processing*. 2015. Vol. 24, no. 11. P. 4185–4196.
24. Droste R., Jiao J., Noble J.A. Unified image and video saliency modeling / *European conference on computer vision (ECCV)*. Springer, 2020. P. 472–489.
25. Bhattacharya S., Venkatesh K.S., Gupta S. Background estimation and motion saliency detection using total variation-based video decomposition // *Signal, Image and Video Processing*. Springer, 2017. Vol. 11, no. 4. P. 699–706.
26. Culibrk D., Mirkovic M., Zlokolica V., Pokric M. Salient motion features for video quality assessment // *IEEE Transactions on Image Processing*. 2010. Vol. 19, no. 4. P. 948–963.
27. Wu Z., Song S., Khosla A., Yu F., Zhang L., Tang X., Xiao J. 3D ShapeNets: A deep representation for volumetric shapes / *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*. 2015. P.

1912–1920

28. Deitke M., Schwenk D., Salvador J., et al. Objaverse: A Universe of Annotated 3D Objects // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). — 2023. — Pp. 13142–13153. — https://openaccess.thecvf.com/content/CVPR2023/papers/Deitke_Objaverse_A_Universe_of_Annotated_3D_Objects_CVPR_2023_paper.pdf