

УДК 004.946

DOI: 10.26467/2079-0619-2023-26-5-30-41

Обеспечение визуальной когерентности в обучающих системах дополненной реальности с учетом авиакосмической специфики

А.Л. Горбунов¹, Юньхань Ли¹

¹*Московский государственный технический университет гражданской авиации,
г. Москва, Россия*

Аннотация: В мае 2022 года саудовская правительственная структура Saudi Arabian Military Industries приобрела обучающую платформу дополненной реальности для летчиков, в сентябре корпорация Boeing начала разработку тренажера пилота дополненной реальности, в ноябре стартовал аналогичный проект ведущего британского разработчика авиационной техники BAE Systems. Эти факты позволяют уверенно говорить о начале новой эпохи авиационных тренажеров – тренажеров с применением технологии дополненной реальности. Одно из перспективных преимуществ данной технологии – возможность безопасного моделирования опасных ситуаций в реальном мире. Необходимым условием использования этого преимущества является обеспечение визуальной когерентности сцен дополненной реальности: виртуальные объекты должны быть неотличимы от реальных. Все мировые IT-лидеры рассматривают дополненную реальность как следующую «большую волну» радикальных изменений в цифровой электронике, поэтому визуальная когерентность становится ключевым вопросом для будущего IT, а в аэрокосмических приложениях визуальная когерентность уже приобрела практическое значение. В РФ имеет место серьезное отставание в изучении проблематики визуальной когерентности в целом и для авиатренажеров дополненной реальности в частности: на момент публикации авторам удалось обнаружить в российском научном пространстве только две работы по теме, тогда как за рубежом их число уже около тысячи. Цель настоящей обзорной статьи – создать условия для купирования проблемы. Визуальная когерентность зависит от многих факторов: освещения, цветового тона, теней от виртуальных объектов на реальных, взаимных отражений, текстур виртуальных поверхностей, оптических аберраций, конвергенции и аккомодации и др. В статье анализируются публикации, посвященные методам оценки условий освещенности и цветового тона реальной сцены и переноса таковых на виртуальные объекты с использованием зондов и по отдельным изображениям, а также по рендерингу виртуальных объектов в сценах дополненной реальности, в том числе с применением нейросетей.

Ключевые слова: авиационные тренажеры, дополненная реальность, визуальная когерентность.

Для цитирования: Горбунов А.Л., Ли Ю. Обеспечение визуальной когерентности в обучающих системах дополненной реальности с учетом авиакосмической специфики // Научный Вестник МГТУ ГА. 2023. Т. 26, № 5. С. 30–41. DOI: 10.26467/2079-0619-2023-26-5-30-41

Visual coherence in augmented reality training systems considering aerospace specific features

A.L. Gorbunov¹, Yunhan Li¹

¹*Moscow State Technical University of Civil Aviation, Moscow, Russia*

Abstract: In May 2022, Saudi Arabian Military Industries, a Saudi government agency, acquired an augmented reality training platform for pilots. In September, the Boeing Corporation began the development of an augmented reality pilot simulator. In November, a similar project was launched by BAE Systems, a leading British developer of aeronautical engineering. These facts allow us to confidently speak about the beginning of a new era of aviation simulators – simulators using the augmented reality technology. One of the promising advantages of this technology is the ability to safely simulate dangerous situations in the real world. A necessary condition for using this advantage is to ensure the visual coherence of augmented reality scenes: virtual objects must be indistinguishable from real ones. All the global IT leaders consider augmented reality as the subsequent surge of radical changes in digital electronics, so visual coherence is becoming a key issue for the future of IT, and in aerospace applications, visual

coherence has already acquired practical significance. The Russian Federation lags far behind in studying the problems of visual coherence in general and for augmented reality flight simulators in particular: at the time of publication the authors managed to find only two papers on the subject in the Russian research space, while abroad their number is already approximately a thousand. The purpose of this review article is to create conditions for solving the problem. Visual coherence depends on many factors: lighting, color tone, shadows from virtual objects on real ones, mutual reflections, textures of virtual surfaces, optical aberrations, convergence and accommodation, etc. The article reviews the publications devoted to methods for assessing the conditions of illumination and color tone of a real scene and transferring them to virtual objects using various probes and by individual images, as well as by rendering virtual objects in augmented reality scenes, using neural networks.

Key words: flight simulators, augmented reality, visual coherence.

For citation: Gorbunov, A.L., Li, Yu. (2023). Visual coherence in augmented reality training systems considering aerospace specific features. Civil Aviation High Technologies, vol. 26, no. 5, pp. 30–41. DOI: 10.26467/2079-0619-2023-26-5-30-41

Введение

Зонтичный термин «расширенная реальность» (Extended Reality – XR) включает таксоны «виртуальная реальность» (Virtual Reality – VR), «дополненная реальность» (Augmented Reality – AR) и различные комбинации первого и второго. Современные авиационные и космические тренажеры фактически по умолчанию подразумевают использование VR. Преимущества такого подхода хорошо известны, поэтому мы не будем на них останавливаться, но отметим ряд существенных и, что важнее, непреодолимых недостатков, обусловленных самой природой технологии виртуальной реальности. VR – это цифровая, дискретная среда, в то время как реальный мир непрерывен. Поэтому моделирование реального мира в VR неизбежно связано с ошибками дискретизации, что снижает эффективность обучения. Однако для обучающих систем более серьезным негативным аспектом является то, что решения человека во многом основаны на подсознательном учете множества деталей реальной картины мира. Этот процесс принципиально невозможно воспроизвести с помощью чисто компьютерных технологий (например, VR) по двум причинам: 1) мы до сих пор не знаем (и вряд ли когда-либо узнаем), каков механизм работы человеческого мозга, последние спекуляции на тему искусственного интеллекта только подтверждают это и 2) детали реального мира, учитываемые при принятии решений, бесконечны по количеству, возникают случайным образом и имеют самую разную природу (визуальную, акустическую, тактильную

и др.) и поэтому фактически невозпроизводимы в VR.

Появление в 2022 году авиационных обучающих систем с AR снижает остроту проблемной ситуации, примерами могут служить разработка корпорацией Boeing пилотского тренажера дополненной реальности на основе проекта ATARS компании Red 6, параллельно стартовавший аналогичный проект ведущего британского разработчика авиационной техники BAE Systems или тренажер авиадиспетчеров из статьи [1, 2]. Все информационное богатство окружающего нас мира в AR представлено в явном виде и не требует моделирования. Но для реализации преимуществ AR, связанных с параллельным присутствием реальных и виртуальных объектов, необходимо решить проблему визуальной когерентности (Visual Coherence – VC): виртуальные объекты должны быть неотличимы от реальных.

AR является производной формой VR. AR сохраняет все возможности VR, но помимо этого как гибридная технология обладает значительными преимуществами, вытекающими из параллельного сосуществования виртуальных и реальных объектов, что и привлекает внимание разработчиков к VC. Более того, исследования [3] показывают, что среди негативных психофизиологических последствий использования устройств дополненной реальности доминирует оптический дискомфорт, который возникает из-за разницы в восприятии реальных и виртуальных объектов в одной и той же сцене из-за недостатка VC. Все лидеры IT-отрасли (Microsoft, Google, Apple, Samsung и др.) рассматривают

AR как следующую (после появления смартфонов) «большую волну» революционных изменений в цифровой электронике, поэтому проблема VC становится ключевой для IT в целом, эти лидеры проявляют растущий интерес к методам ее решения [4]. Однако в аэрокосмических приложениях проблема визуальной когерентности уже приобрела практическое значение. Важно отметить, что VC зависит от многих факторов, например освещения, теней, цветового тона, взаимных отражений, текстуры поверхности, оптических aberrаций, конвергенции и аккомодации, и, соответственно, используются различные техники AR-визуализации [5].

В РФ имеет место очевидное отставание в изучении проблематики VC, цель настоящего обзора – создать условия для купирования проблемы.

Обзор публикаций

К моменту подготовки данной статьи авторам удалось выявить только две публикации по тематике VC в российском научном информационном пространстве [1, 2], тогда как зарубежная библиография в данной области содержит уже около тысячи пунктов. Список источников в обзоре [6] включает 175 позиций; этот обзор освещает базовые подходы к достижению VC в части обеспечения корректной освещенности виртуальных объектов (за исключением основанных на нейронных сетях, рассматриваемых ниже), поэтому здесь мы кратко опишем некоторые характерные примеры, которые соответствуют двум упомянутым в [6] основным классам.

Измерение условий освещенности

Используя зонд освещенности с диффузными полосами между зеркальными сферическими квадрантами, Дебевек и др. [7] продемонстрировали, как полный динамический цветовой диапазон сцены может быть восстановлен из одной экспозиции. На основе изображения, полученного с помощью зонда, интенсивность нескольких источников света может быть оценена путем решения простой линейной системы уравнений. Результаты

были использованы для рендеринга виртуальной диффузной сферы.

Унгер и др. [8] измерили и параметризировали световое поле, накрывающее плоскую двумерную поверхность в реальных сценах. Для этого использовалась камера с объективом «рыбий глаз» с углом обзора 180° , установленная на движущемся столике. Во время рендеринга в сцене размещается простой вспомогательный объект, например сфера или рамка. Зарегистрированные данные светового поля затем превращаются в выборку путем проецирования лучей назад от вспомогательного объекта на область регистрации, где блики оцениваются посредством интерполяции для линейных и угловых координат. Сопоставление реального объекта с его виртуальным аналогом, полученное с помощью этого метода, демонстрирует высокий уровень VC.

Альхаками и Тасерян [9] оценивают направление падающего света (прямое освещение) реальной сцены с помощью прикрепленной к AR-устройству 360-градусной камеры, используя методы компьютерного зрения. Система моделирует отраженный свет от поверхностей при рендеринге виртуальных объектов. Затем определяются параметры теней для каждого виртуального объекта.

Оценка условий освещенности

Кнорр и Курц [10] предложили схему оценки условий освещения в реальном мире на основе фото человеческого лица. Метод основан на обучении модели лица на основе базы данных по лицам при известном освещении. Авторы затем восстанавливают наиболее правдоподобные условия освещения в реальном мире в базисе сферических гармоник для запечатленного лица.

Карш и др. [11] предложили полностью автоматическую систему для вставки неотражающих объектов в одиночные фото в помещении, предполагая ламбертово отражение сцены и диффузное освещение. Метод дает высококачественные результаты для многих типов сцен, однако часто не может восстановить модели освещения в сценах, где предположение о диффузности не выполняется, например в сценах на открытом воздухе.

В сценах, где карта глубины не может быть точно оценена, метод создает неточные тени, отбрасываемые виртуальными объектами на реальную сцену.

Стоит упомянуть работу [12], где описана комбинация измерения и оценки освещенности. Авторы измеряют отражательные свойства реальных объектов, используя карты глубины и цветные изображения вращающегося объекта на поворотном столе с помощью RGB-D камеры. Форма объекта реконструируется путем интеграции изображений глубины объекта, полученных с разных точек обзора. Отражательная способность объекта определяется путем оценки параметров модели отражения по восстановленным изображениям формы.

Существует большое число работ по методам передачи цвета от изображения к изображению, которые используются для выравнивания цветовых характеристик реальных и виртуальных объектов в сценах AR. В работе [13] представлен способ автоматического переноса цветовой статистики (средних значений и стандартных отклонений) с эталонного изображения на целевое. Дополнительные параметры используются для того, чтобы избежать ручной обработки, которая необходима для определения особенностей передачи цвета в случаях, когда изображения имеют сильное различие в цветовой палитре. Эти дополнительные параметры объединяют дисперсии эталонного и целевого изображений. Авторы статьи утверждают, что, хотя ручное изменение данных параметров требуется крайне редко, тем не менее иногда это необходимо. Кроме того, статистическая природа метода вызывает вопросы относительно вида и объема статистики. Также не очевидна способность метода обрабатывать некоторые типы изображений (содержащие блестящие объекты, тени). Полезное расширение этого метода предложено в [14], которое учитывает цветовой облик, вносимый как целевым, так и эталонным изображением, что приводит к смешиванию цветов.

Другой подход к передаче цвета от изображения к изображению, основанный на характеристиках человеческого восприятия,

описан в [15]. Сначала цвета изображения ассоциируются с одной из 11 основных цветных категорий. Затем строятся распределения категорий для исходного и целевого изображений. Наконец, цвета передаются между парами категорий. Для построения распределения необходимо оценить оттенок освещения для обоих изображений. Для некоторых типов изображений эта процедура требует ручной настройки. Метод очень перспективен, но он неэффективен для нефотореалистичных изображений и для изображений с несколькими локальными источниками освещения.

Сяо и Ма [16] представили алгоритм для решения проблемы достоверности передачи цвета с учетом деталей сцены. Авторы рассматривают сохранение цветового градиента как необходимое условие достоверности сцены. Они формулируют проблему передачи цвета как задачу оптимизации и решают ее в два этапа – согласование гистограмм и оптимизация с сохранением градиента. Предложена метрика для объективной оценки эффективности алгоритмов передачи цвета на основе примеров.

Недостатками методов, описанных в [13–16] (и их многочисленных аналогов) являются неуниверсальность, неавтоматический характер и невозможность передачи не только цвета, но и всех основных характеристик изображения с помощью одной процедуры. Два типичных примера: метод балансировки цвета, предложенный в [17], предлагает пользователю интерактивно определить цветное отображение, вручную указывая соответствия; подход, описанный в [18], требует специальной оценки геометрии и BRDF реальных объектов и хотя бы одного доминирующего цвета на реальных объектах.

От указанных недостатков свободен способ обеспечения VC, представленный в [1]. Несмотря на впечатляющие результаты, полученные исследователями с помощью методов, упомянутых в обзоре [6], уровень VC все еще часто недостаточен, особенно в сценах AR с реальными естественными ландшафтами в условиях рассеянного освещения, которые типичны для авиационных приложений.

Метод из статьи [1] ориентирован на такие условия и базируется на аппарате двумерных спектральных преобразований изображений. Предложена технология спектральной трансплантации, которая обеспечивает прямую передачу характеристик цвета, яркости и контраста от реального фона к виртуальным объектам. Разработан алгоритм автоматического выбора оптимального вида спектрального преобразования для использования в спектральной трансплантации.

В последние годы активно развивается направление применения нейросетевых решений для обеспечения VC, которые преимущественно используются для оценки условий освещенности и для рендеринга виртуальных объектов.

Нейросетевая оценка условий освещенности

Хорошей иллюстрацией исследовательских подходов к решению данной проблемы является обзор [19] (104 источника).

В работе [20] предложен способ рендеринга виртуальных объектов на основе одного фото реальной сцены. Результат достигается в три этапа: 1) нейросетевой классификатор условий освещения обучается автоматически аннотировать местоположение источников света на большом наборе данных карт окружающей среды низкого динамического диапазона (Low Dynamic Range – LDR); 2) полученные аннотации используются для обучения нейросети, которая предсказывает расположение источников света в сцене по одной фотографии с ограниченным полем зрения; 3) осуществляется точная настройка этой нейросети с применением небольшого набора данных карт среды высокого динамического диапазона (High Dynamic Range – HDR) для прогнозирования интенсивности освещения. Далее оценки освещенности используются для рендеринга 3D-объектов в сценах AR. Фотореалистичность результата подтверждается исследованием, в котором 105 участников отдали предпочтение данному методу в сравнении с [11].

Исследователи из Google и Принстонского университета [21] предложили сходный подход к оценке условий освещенности в 3D

по одному 2D-изображению и для конкретной точки на нем. Отличие заключается в том, что нейросеть сначала оценивает геометрию сцены, затем она трансформируется в сферическую проекцию вокруг выбранного пиксела, которая используется как вход для следующей нейросети, задача которой – генерация цветовой окраски для ненаблюдаемых фрагментов. Еще одна нейросеть генерирует картину HDR, которую можно использовать для рендеринга виртуальных объектов. Преимущество подхода – в независимости нейросетей, что упрощает их обучение.

Авторы статьи [22] адресуют свой метод приложениям для мобильных устройств, эксплуатируя наличие в них фронтальной и задней камер. Картинки этих камер используются как вход для сверточной нейросети, цель которой – прогноз низкочастотных сферических гармоник для модели освещенности, представленной в виде сферических гармонических функций, широко используемой в игровой индустрии для рендеринга виртуальных трехмерных объектов в силу своей простоты.

Для характерных в авиационных приложениях сцен AR с рассеянным светом открытых пространств интересным представляется подход, описанный в [23]. Авторы применяют сверточную нейросеть, обученную на большом наборе пар фото, – параметры освещенности для экстракта характеристик освещения из одного изображения, которые затем дополняются моделью освещенности на открытом воздухе. Недостатки метода – ограничения, накладываемые этой моделью (хорошо работает для солнечных сцен и плохо – для пасмурных), влияние на результаты конкретного обучающего набора и проблемы с воспроизведением влияния точечных источников света.

Лучше справляется с отсветами реальных бликов на виртуальных объектах метод, развиваемый авторами [24]. Сначала вид с камеры преобразуется в облако точек световых источников с центром в точке наблюдения. Затем компактная нейронная сеть используется для оценки коэффициентов сферических гармоник в точке рендеринга. Далее по коэф-

фициентам сферических гармоник восстанавливается карта освещенности HDR.

Еще один способ оценки коэффициентов сферических гармоник в точке рендеринга посредством нейросети предложен в [25]. Авторы как преимущество своего подхода заявляют отсутствие необходимости в какой-либо информации о геометрии сцены и высокую скорость (20 мс) генерации оценок для сферических гармоник пятого порядка. Однако для обучения нейросети используется специально синтезированный для конкретной сцены набор данных, что ограничивает универсальность метода.

В работе [26] являющийся сейчас сотрудником Google П. Дебевек с коллегами продолжили развитие его идеи 11-летней давности [7] об использовании специальных отражающих объектов для генерации карты HDR, но уже с применением нейросетей. Автоэнкодер обучается на массиве LDR-картинок, включающих три сферы с поверхностями разной отражательной способности, каждая из которых позволяет получить непересекающиеся данные о характере освещения. Такой подход, как утверждают авторы, делает его эффективным для генерации сцен AR и внутри помещений, и на открытом воздухе.

Другая команда из Google в сотрудничестве с исследователями из Университета Беркли предложила нейросетевой метод оценки условий освещенности по одной стереопаре LDR [27]. Предложенный метод оценивает трехмерную объемную RGB-модель сцены, включая контент за пределами наблюдаемого поля, а затем использует стандартный объемный рендеринг для оценки освещения в любой точке в пределах этого объема. Модель обучается без каких-либо реальных 3D-данных и требует только расширенного перспективного вида вблизи входной стереопары и сферической панорамы, снятой внутри сцены.

Нейросетевой рендеринг

Широкий обзор нейросетевых методов рендеринга виртуальных объектов, активно используемых в генерации AR-сцен, дан в [28] (234 источника). В качестве преамбулы публикация содержит краткое описание клас-

сических алгоритмов компьютерного зрения для рендеринга, что делает ее удобной при общем знакомстве с проблематикой.

Весьма успешно нейросети применяются для рендеринга в виде генеративно-сопоставительных сетей (Generative Adversarial Network – GAN), представленных в ставшей уже классической работе [29], которая вызвала поток публикаций по стилизовому моделированию изображений, примером может служить статья [30]. Часто процесс генерации изображений требует обеспечения возможности управления им, что первоначально делалось сопоставлением полученных данных с данными исходных изображений. Такое решение, однако, не гарантирует, что результат, сгенерированный GAN, будет выглядеть естественно. Поэтому вместо минимизации расстояния между результатами и целевыми показателями в настоящее время чаще применяются условные GAN, где оценивается соответствие статистических распределений результатов и входных данных [31].

Управление рендерингом через точку размещения виртуальной камеры в [32] основано на сверточной сети, обученной синтезировать новые изображения из шести входных. Сеть обучается с помощью большого синтетического набора данных из 1000 сцен со сложной геометрией и свойствами материалов. Метод способен синтезировать новые точки обзора для захваченных реальных данных и воспроизводить сложные эффекты, такие как перекрытия, зависящая от вида зеркальность и резкие тени.

Весьма востребован у пользователей мультимодальный нейросетевой синтез. Для получения различных выходных данных, существенно отличающихся друг от друга, нейросеть должна обладать некоторой стохастичностью или структурированной дисперсией, что, как правило, достигается с помощью вариационных автоэнкодеров, как это сделано в [33].

Классический инверсный рендеринг для виртуальных объектов в реальных сценах в нейросетевом исполнении позволяет справиться с проблемой вычислительной затратности. В [34] предложен метод, позволяющий на основе одного RGB-изображения произ-

вольной сцены в помещении создать ее полную реконструкцию путем оценки геометрии, изменяющихся в пространстве освещения и неламбертовой отражательной способности поверхностей.

Применение нейросетей для целей VC также включает решение ряда иных задач. В статье [35] перенос признаков одного изображения на другое осуществляется в два этапа, оригинальная нейросеть используется при выявлении признаков в первом изображении, вектор которых применяется в качестве фильтра другой нейросети, выявляющей признаки во втором. Авторы считают, что их решение успешно конкурирует с GAN. Обзор [36] (95 источников) содержит анализ приложений для обнаружения объектов в изображении реального мира на основе глубокого обучения в контексте дополненной реальности и анализ использования серверов или локальных устройств AR для выполнения вычислений по обнаружению объектов.

Заключение

Появление новой генерации авиатренажеров – тренажеров с использованием технологии дополненной реальности – маркирует на аэрокосмическую сферу проекцию связанного с массовым переходом на использование XR-устройств глобального тренда развития IT-технологий. Обусловленный этим трендом бум исследований по тематике VC, который проиллюстрирован настоящим обзором, требует скорейшего формирования соответствующего научного направления в РФ.

Список литературы

1. Горбунов А.Л. Визуальная когерентность в дополненной реальности // *Advanced Engineering Research*. 2023. № 2. С. 180–190. DOI: 10.23947/2687-1653-2023-23-2-180-190
2. Горбунов А.Л. Визуальная однородность сцен дополненной реальности // Запись и воспроизведение объемных изображений в кинематографе и других областях: сборник докладов IX Международной научно-прак-

тической конференции. Москва, 17–18 апреля 2017 г. М.: ВГИК им. С.А. Герасимова, 2017. С. 235–239.

3. Hughes C. The psychometrics of cyber-sickness in augmented reality / C. Hughes, C. Fidopiastis, K. Stanney, P. Bailey, E. Ruiz [Электронный ресурс] // *Frontiers in Virtual Reality*. 2020. Vol. 1. ID: 602954. DOI: 10.3389/frvir.2020.602954 (дата обращения: 05.04.2023).

4. Somanath G., Kurz D. HDR environment map estimation for real-time augmented reality [Электронный ресурс] // *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. USA: Nashville, TN, 2021. Pp. 11293–11301. DOI: 10.1109/CVPR46437.2021.01114 (дата обращения: 05.04.2023).

5. Zollmann S. Visualization techniques in augmented reality: A taxonomy, methods and patterns / S. Zollmann, T. Langlotz, R. Grasset, W.H. Lo, S. Mori, H. Regenbrecht // *IEEE Transactions on Visualization and Computer Graphics*, 2021. Vol. 27, no. 9. Pp. 3808–3825. DOI: 10.1109/TVCG.2020.2986247

6. Kronander J. Photorealistic rendering of mixed reality scenes / J. Kronander, F. Bantlerle, A. Gardner, E. Miandji, J. Unger // *Computer Graphics Forum*. 2015. Vol. 34, iss. 2. Pp. 643–665. DOI: 10.1111/cgf.12591

7. Debevec P. A single-shot light probe / P. Debevec, P. Graham, J. Busch, M. Bolas [Электронный ресурс] // *SIGGRAPH '12: Special Interest Group on Computer Graphics and Interactive Techniques Conference*. California, Los Angeles, 2012. Article No.: 10. Pp. 1–19. DOI: 10.1145/2343045.2343058 (дата обращения: 05.04.2023).

8. Unger J. Capturing and rendering with incident light fields / J. Unger, A. Wenger, T. Hawkins, A. Gardner, P. Debevec // *14th Eurographics Symposium on Rendering*, 2003. Pp. 141–149. DOI: 10.2312/EGWR/EGWR03/141-149

9. Alhakamy A., Tuceryan M. CubeMap360: Interactive global illumination for augmented reality in dynamic environment // *Proceedings of IEEE SoutheastCon*. USA, Huntsville, AL, 2019. Pp. 1–8. DOI: 10.1109/SoutheastCon42311.2019.9020588

- 10. Knorr S., Kurz D.** Real-time illumination estimation from faces for coherent rendering // 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). Germany, Munich, 2014. Pp. 349–350. DOI: 10.1109/ISMAR.2014.6948483
- 11. Karsch K., Sunkavalli K., Hadap S. и др.** Automatic scene inference for 3D object compositing // ACM Transactions on Graphics. 2014. Vol. 33, no. 3. Pp. 1–15. DOI: 10.1145/2602146
- 12. Tsunzaki S.** Reproducing material appearance of real objects using mobile augmented reality / S. Tsunzaki, R. Nomura, T. Komuro, S. Yamamoto, N. Tsumura // 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). Germany, Munich, 2018. Pp. 196–197. DOI: 10.1109/ISMAR-Adjunct.2018.00065
- 13. Reinhard E.** Real-time color blending of rendered and captured video / E. Reinhard, A.O. Akyüz, M. Colbert, C. Hughes, M. Oconor [Электронный ресурс] // Proceedings Interservice/Industry Training, Simulation and Education Conference (I/ITSEC). Orlando, Amerika Birleşik Devletleri, 2004. P. 15021. URL: <http://www.ceng.metu.edu.tr/~akyuz/files/blend.pdf> (дата обращения: 05.04.2023).
- 14. Chen W.-S., Huang M.-L., Wang C.-M.** Optimizing color transfer using color similarity measurement // 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), 2016. Pp. 1–6. DOI: 10.1109/ICIS.2016.7550799
- 15. Chang Y., Saito S., Nakajima M.** Example-Based color transformation of image and video using basic color categories // IEEE Transactions on Image Processing, 2007. Vol. 16, no. 2. Pp. 329–336. DOI: 10.1109/tip.2006.888347
- 16. Xiao X., Ma L.** Gradient-Preserving color transfer // Computer Graphics Forum. 2009. Vol. 28, iss. 7. Pp. 1879–1886. DOI: 10.1111/j.1467-8659.2009.01566.x
- 17. Oskam T.** Fast and stable color balancing for images and augmented reality / T. Oskam, A. Hornung, R.W. Sumner, M. Gross // 2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission (3DIMPVT). Zurich, Switzerland, 13–15 October 2012. Pp. 49–56. DOI: 10.1109/3DIMPVT.2012.36
- 18. Knecht M.** Adaptive camera-based color mapping for mixed-reality applications / M. Knecht, C. Traxler, W. Purgathofer, M. Wimmer // 2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2011). Switzerland, Basel, 2011. Pp. 165–168. DOI: 10.1109/ISMAR.2011.6092382
- 19. Einabadi F., Guillemaut J., Hilton A.** Deep neural models for illumination estimation and relighting: A survey // Computer Graphics Forum. 2021. Vol. 40, iss. 6. Pp. 315–331. DOI: 10.1111/cgf.14283
- 20. Gardner M., Sunkavalli K., Yumer E. и др.** Learning to predict indoor illumination from a single image // ACM Transactions on Graphics. 2017. Vol. 36, iss. 6. Article No.: 176. Pp. 1–14. DOI: 10.1145/3130800.3130891
- 21. Song S., Funkhouser T.** Neural illumination: Lighting prediction for indoor environments [Электронный ресурс] // Proceedings 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019. Pp. 6918–6926. DOI: 10.48550/arXiv.1906.07370 (дата обращения: 05.04.2023).
- 22. Cheng D.** Learning scene illumination by pairwise photos from rear and front mobile cameras / D. Cheng, J. Shi, Y. Chen, X. Deng, X. Zhang // Computer Graphics Forum. 2018. Vol. 37, iss. 7. Pp. 213–221. DOI: 10.1111/cgf.13561
- 23. Hold-Geoffroy Y.** Deep outdoor illumination estimation / Y. Hold-Geoffroy, K. Sunkavalli, S. Hadap, E. Gambaretto, J.-F. Lalonde [Электронный ресурс] // Proceedings 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). USA, Honolulu, HI, 2017. Pp. 2373–2382. DOI: 10.1109/CVPR.2017.255 (дата обращения: 05.04.2023).
- 24. Zhao Y., Guo T.** PointAR: Efficient lighting estimation for mobile augmented reality // 16th European Conference on Computer Vision (ECCV'20), 2020. Pp. 678–693. DOI: 10.48550/arXiv.2004.00006
- 25. Garon M.** Fast spatially-varying in-door lighting estimation / M. Garon, K. Sunkavalli, S. Hadap, N. Carr, J. Lalonde [Электронный ресурс] // Proceedings IEEE Conference on

Computer Vision and Pattern Recognition, 2019. Pp. 6908–6917. DOI: 10.48550/arXiv.1906.03799 (дата обращения: 05.04.2023).

26. LeGendre C., Ma W., Fyffe G. и др. Deep-light: Learning illumination for unconstrained mobile mixed reality [Электронный ресурс] // Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2019. Pp. 5918–5928. DOI: 10.48550/arXiv.1904.01175 (дата обращения: 05.04.2023).

27. Srinivasan P. Lighthouse: Predicting lighting volumes for spatially-coherent illumination / P. Srinivasan, B. Mildenhall, M. Tancik, J. Barron, R. Tucker, N. Snavely [Электронный ресурс] // Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2020. Pp. 8080–8089. DOI: 10.48550/arXiv.2003.08367 (дата обращения: 05.04.2023).

28. Tewari A., Fried O., Thies J. и др. State of the art on neural rendering // Computer Graphics Forum. 2020. Vol. 39, iss. 2. Pp. 701–727. DOI: 10.1111/cgf.14022

29. Goodfellow I., Pouget-Abadie J., Mirza M. и др. Generative adversarial nets // Proceedings of the 27th International Conference on Neural Information Processing Systems, 2014. Vol. 2. Pp. 2672–2680. DOI: 10.5555/2969033.2969125

30. Karras T., Laine S., Aila T. A style-based generator architecture for generative adversarial networks [Электронный ресурс] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). USA, Long Beach, CA, 2019. Pp. 4396–4405. DOI: 10.1109/CVPR.2019.00453 (дата обращения: 05.04.2023).

31. Bi S. Deep CG2Real: Synthetic-to-real translation via image disentanglement / S. Bi, K. Sunkavalli, F. Perazzi, E. Shechtman, V. Kim, R. Ramamoorthi // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019. Pp. 2730–2739. DOI: 10.1109/ICCV.2019.00282

32. Xu Z. Deep view synthesis from sparse photometric images / Z. Xu, S. Bi, K. Sunkavalli, S. Hadap, H. Su, R. Ramamoorthi [Электронный ресурс] // ACM Transactions on Graphics. 2019. Vol. 38, iss. 412. Article No.: 76. Pp. 1–13. DOI: 10.1145/3306346.3323007 (дата обращения: 05.04.2023).

33. Park T. Semantic image synthesis with spatially-adaptive normalization / T. Park, M. Liu, T. Wang, J. Zhu [Электронный ресурс] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019. 19 p. DOI: 10.48550/arXiv.1903.07291 (дата обращения: 05.04.2023).

34. Li Z. Inverse rendering for complex indoor scenes: Shape, spatially-varying lighting and SVBRDF from a single image / Z. Li, M. Shafiei, R. Ramamoorthi, K. Sunkavalli, M. Chandraker [Электронный ресурс] // Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2020. Pp. 2475–2484. DOI: 10.48550/arXiv.1905.02722 (дата обращения: 05.04.2023).

35. Zhan F., Yu Y., Wu R. и др. Bi-level feature alignment for semantic image translation and manipulation [Электронный ресурс] // Proceedings IEEE Conference on Computer Vision and Pattern Recognition, 2022. 18 p. DOI: 10.48550/arXiv.2107.03021 (дата обращения: 05.04.2023).

36. Ghasemi Y. Deep learning-based object detection in augmented reality: A systematic review / Y. Ghasemi, H. Jeong, S. Choi, J. Lee, K. Park [Электронный ресурс] // Computers in Industry. 2022. Vol. 139. ID: 103661. DOI: 10.1016/j.compind.2022.103661 (дата обращения: 05.04.2023).

References

1. Gorbunov, A.L. (2023). Visual coherence for augmented reality. *Advanced Engineering Research*, no. 2, pp. 180–190. DOI: 10.23947/2687-1653-2023-23-2-180-190 (in Russian)

2. Gorbunov, A.L. (2017). Visual homogeneity of augmented reality scenes. In: *Zapisi i vosproizvedeniye obyemnykh izobrazheniy v kinematografе i drugikh oblastiakh: sbornik dokladov IX Mezhdunarodnoy nauchno-prakticheskoy konferentsii*. Moscow: VGIK im. S.A. Gerasimova, pp. 235–239. (in Russian)

3. Hughes, C., Fidopiastis, C., Stanney, K., Bailey, P., Ruiz, E. (2020). The psychometrics of cybersickness in augmented reality. *Frontiers in Virtual Reality*, vol. 1, ID: 602954. DOI: 10.3389/frvir.2020.602954 (accessed: 05.04.2023).

4. **Somanath, G., Kurz, D.** (2021). HDR environment map estimation for real-time augmented reality. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, pp. 11293–11301. DOI: 10.1109/CVPR46437.2021.01114 (accessed: 05.04.2023).
5. **Zollmann, S., Langlotz, T., Grasset, R., Lo, W.H., Mori, S., Regenbrecht, H.** (2021). Visualization techniques in augmented reality: A taxonomy, methods and patterns. In: *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 9, pp. 3808–3825. DOI: 10.1109/TVCG.2020.2986247
6. **Kronander, J., Banterle, F., Gardner, A., Miandji, E., Unger, J.** (2015). Photorealistic rendering of mixed reality scenes. *Computer Graphics Forum*, vol. 34, issue 2, pp. 643–665. DOI: 10.1111/cgf.12591
7. **Debevec, P., Graham, P., Busch, J., Bolas, M.** (2012). A single-shot light probe. In *SIGGRAPH '12: Special Interest Group on Computer Graphics and Interactive Techniques Conference*, Los Angeles California, Article No.: 10, pp. 1–19. DOI: 10.1145/2343045.2343058 (accessed: 05.04.2023).
8. **Unger, J., Wenger, A., Hawkins, T., Gardner, A., Debevec, P.** (2003). Capturing and rendering with incident light fields. In: *14th Eurographics Symposium on Rendering*, pp. 141–149. DOI: 10.2312/EGWR/EGWR03/141-149
9. **Alhakamy, A., Tuceryan, M.** (2019). CubeMap360: Interactive global illumination for augmented reality in dynamic environment. In: *Proceedings of IEEE SoutheastCon*, Huntsville, AL, USA, pp. 1–8. DOI: 10.1109/SoutheastCon42311.2019.9020588
10. **Knorr, S., Kurz, D.** (2014). Real-time illumination estimation from faces for coherent rendering. In: *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Munich, Germany, pp. 349–350. DOI: 10.1109/ISMAR.2014.6948483
11. **Karsch, K., Sunkavalli, K., Hadap, S., et al.** (2014). Automatic scene inference for 3D object compositing. *ACM Transactions on Graphics*, vol. 33, no. 3, pp. 1–15. DOI: 10.1145/2602146
12. **Tsunezaki, S., Nomura, R., Komuro, T., Yamamoto, S., Tsumura, N.** (2018). Reproducing material appearance of real objects using mobile augmented reality. In: *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, Munich, Germany, pp. 196–197. DOI: 10.1109/ISMAR-Adjunct.2018.00065
13. **Reinhard, E., Akyuz, A.O., Colbert, M., Hughes, C., O'Connor, M.** (2004). Real-time color blending of rendered and captured video. In: *Proceedings Interservice/Industry Training, Simulation and Education Conference (IITSEC)*, Orlando, Amerika Birleşik Devletleri, p. 15021. Available at: <http://www.ceng.metu.edu.tr/~akyuz/files/blend.pdf> (accessed: 05.04.2023).
14. **Chen, W.-S., Huang, M.-L., Wang, C.-M.** (2016). Optimizing color transfer using color similarity measurement. In: *2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS)*, pp. 1–6. DOI: 10.1109/ICIS.2016.7550799
15. **Chang, Y., Saito, S., Nakajima, M.** (2007). Example-Based color transformation of image and video using basic color categories. In: *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 329–336. DOI: 10.1109/tip.2006.888347
16. **Xiao, X., Ma, L.** (2009). Gradient-Preserving color transfer. *Computer Graphics Forum*, vol. 28, issue 7, pp. 1879–1886. DOI: 10.1111/j.1467-8659.2009.01566.x
17. **Oskam, T., Hornung, A., Sumner, R.W., Gross, M.** (2012). Fast and stable color balancing for images and augmented reality. In: *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission (3DIMPVT)*, Zurich, Switzerland, October 13–15, pp. 49–56. DOI: 10.1109/3DIMPVT.2012.36
18. **Knecht, M., Traxler, C., Purgathofer, W., Wimmer, M.** (2011). Adaptive Camera-Based Color Mapping For Mixed-Reality Applications. In: *2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR 2011)*, Basel, Switzerland, pp. 165–168. DOI: 10.1109/ISMAR.2011.6092382
19. **Einabadi, F., Guillemaut, J., Hilton, A.** (2021). Deep neural models for illumination es-

timation and relighting: A survey. *Computer Graphics Forum*, vol. 40, issue 6, pp. 315–331. DOI: 10.1111/cgf.14283

20. Gardner, M., Sunkavalli, K., Yumer, E., et al. (2017). Learning to predict indoor illumination from a single image. *ACM Transactions on Graphics*, vol. 36, issue 6, Article No.: 176, pp. 1–14. DOI: 10.1145/3130800.3130891

21. Song, S., Funkhouser, T. (2019). Neural illumination: Lighting prediction for indoor environments. In: *Proceedings 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6918–6926. DOI: 10.48550/arXiv.1906.07370 (accessed: 05.04.2023).

22. Cheng, D., Shi, J., Chen, Y., Deng, X., Zhang, X. (2018). Learning scene illumination by pairwise photos from rear and front mobile cameras. *Computer Graphics Forum*, vol. 37, issue 7, pp. 213–221. DOI: 10.1111/cgf.13561

23. Hold-Geoffroy, Y., Sunkavalli, K., Hadap, S., Gambaretto, E., Lalonde, J. (2017). Deep outdoor illumination estimation. In: *Proceedings 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 2373–2382. DOI: 10.1109/CVPR.2017.255 (accessed: 05.04.2023).

24. Zhao, Y., Guo, T. (2020). Point AR: Efficient lighting estimation for mobile augmented reality. In: *16th European Conference on Computer Vision (ECCV'20)*, pp. 678–693. DOI: 10.48550/arXiv.2004.00006

25. Garon, M., Sunkavalli, K., Hadap, S., Carr, N., Lalonde, J. (2019). Fast spatially-varying in-door lighting estimation. In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6908–6917. DOI: 10.48550/arXiv.1906.03799 (accessed: 05.04.2023).

26. LeGendre, C., Ma, W., Fyffe, G., et al. (2019). DeepLight: Learning illumination for unconstrained mobile mixed reality. In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5918–5928. DOI: 10.48550/arXiv.1904.01175 (accessed: 05.04.2023).

27. Srinivasan, P., Mildenhall, B., Tanksik, M., Barron, J., Tucker, R., Snavely, N. (2020). Lighthouse: Predicting lighting volumes for spatially-coherent illumination. In: *Proceed-*

ings IEEE Conference on Computer Vision and Pattern Recognition, pp. 8080–8089. DOI: 10.48550/arXiv.2003.08367 (accessed: 05.04.2023).

28. Tewari, A., Fried, O., Thies, J., et al. (2020). State of the art on neural rendering. *Computer Graphics Forum*, vol. 39, issue 2, pp. 701–727. DOI: 10.1111/cgf.14022

29. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014). Generative adversarial nets. In: *Proceedings of the 27th International Conference on Neural Information Processing Systems*, vol. 2, pp. 2672–2680. DOI: 10.5555/2969033.2969125

30. Karras, T., Laine, S., Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, pp. 4396–4405. DOI: 10.1109/CVPR.2019.00453 (accessed: 05.04.2023).

31. Bi, S., Sunkavalli, K., Perazzi, F., Shechtman, E., Kim, V., Ramamoorthi, R. (2019). Deep CG2Real: Synthetic-to-real translation via image disentanglement. In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2730–2739. DOI: 10.1109/ICCV.2019.00282

32. Xu, Z., Bi, S., Sunkavalli, K., Hadap, S., Su, H., Ramamoorthi, R. (2019). Deep view synthesis from sparse photometric images. *ACM Transactions on Graphics*, vol. 38, issue 412, Article No.: 76, pp. 1–13. DOI: 10.1145/3306346.3323007 (accessed: 05.04.2023).

33. Park, T., Liu, M., Wang, T., Zhu, J. (2019). Semantic image synthesis with spatially-adaptive normalization. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 19 p. DOI: 10.48550/arXiv.1903.07291 (accessed: 05.04.2023).

34. Li, Z., Shafiei, M., Ramamoorthi, R., Sunkavalli, K., Chandraker, M. (2020). Inverse rendering for complex indoor scenes: Shape, spatially-varying lighting and SVBRDF from a single image. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2475–2484. DOI: 10.48550/arXiv.1905.02722 (accessed: 05.04.2023).

35. Zhan, F., Yu, Y., Wu, R., et al. (2022). Bi-level feature alignment for semantic image

translation and manipulation. In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, 18 p. DOI: 10.48550/arXiv.2107.03021 (accessed: 05.04.2023).

36. Ghasemi, Y., Jeong, H., Choi, S., Lee, J., Park, K. (2022). Deep learning-based

object detection in augmented reality: A systematic review. *Computers in Industry*, vol. 139, ID: 103661. DOI: 10.1016/j.compind.2022.103661 (accessed: 05.04.2023).

Сведения об авторах

Горбунов Андрей Леонидович, кандидат технических наук, доцент, доцент кафедры управления воздушным движением МГТУ ГА, a.gorbunov@mstuca.aero.

Ли Юньхань, аспирантка МГТУ ГА, antatanoe@gmail.com.

Information about the authors

Andrey L. Gorbunov, Candidate of Technical Sciences, Associate Professor, Associate Professor of the Air Traffic Management Chair, Moscow State Technical University of Civil Aviation, a.gorbunov@mstuca.aero.

Yunhan Li, Postgraduate Student, Moscow State Technical University of Civil Aviation, antatanoe@gmail.com.

Поступила в редакцию	18.04.2023	Received	18.04.2023
Одобрена после рецензирования	17.05.2023	Approved after reviewing	17.05.2023
Принята в печать	21.09.2023	Accepted for publication	21.09.2023