

УДК 004.946
<https://doi.org/10.31854/3034-2201-2026-4-2-C05>
EDN: FKQFIV

Цветовая согласованность виртуальных объектов в видеосистемах смешанной реальности: обзор и классификация методов

Бучатский А. Н., Кузьмина Д. Д. ✉

Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

Постановка задачи: цветовая согласованность встроенных виртуальных объектов с реальным окружением — одна из нерешенных задач видеосистем смешанной реальности. Существующие подходы либо применяют единое глобальное преобразование без учета геометрии освещения, либо требуют вычислительных ресурсов, недоступных в режиме реального времени. **Целью работы** является систематизация и критический анализ методов переноса цвета, оценки освещения и гармонизации изображений применительно к видеосистемам смешанной реальности, включая новейшие подходы 2024–2026 гг. **Используемые методы:** аналитический обзор научных публикаций; сравнительная классификация методов по принципу работы, вычислительным требованиям и области применимости. **Новизна:** впервые методы переноса цвета и гармонизации изображений для видеосистем смешанной реальности (включая генеративные методы на основе диффузионных моделей и мягкую гармонизацию методами оптимального транспорта) систематизированы в едином обзоре, охватывающем период с 2001 по 2026 гг. **Результат:** предложены классификация по девяти группам методов и подход к решению задачи цветовой согласованности объектов на основе семантической сегментации и раздельной коррекции яркостной и хроматических компонент; проведен сравнительный анализ метрик оценки цветового расхождения (CIEDE2000, SSIM, PSNR) и метрик гармонизации (fMSE, LPIPS). **Теоретическая / Практическая значимость:** полученная классификация позволяет обоснованно выбирать метод цветовой коррекции при проектировании видеосистем смешанной реальности с учетом требований к качеству и доступным вычислительным ресурсам.

Ключевые слова: смешанная реальность, цветовая коррекция, перенос цвета, гармонизация изображений, CIEDE2000, цветовое пространство CIELAB, генеративные модели, оптимальный транспорт, семантическая сегментация

Актуальность исследования

Системы смешанной реальности находят применение в медицине, промышленном проектировании, образовании и архитектурной визуализации. Ключевое требование к таким системам — визуальная согласованность виртуальных и реальных элементов сцены: встроенный объект должен восприниматься как органичная часть окружения, а не как посторонний элемент, наложенный поверх него. Цветовая рассогласованность разрушает это единство. Виртуальный объект, синтезированный при нейтральных условиях освещения (стандартный источник D65), выглядит чужеродно при наложении на сцену с теплым, холодным или насыщенным цветным световым полем.

Библиографическая ссылка на статью:

Бучатский А. Н., Кузьмина Д. Д. Цветовая согласованность виртуальных объектов в видеосистемах смешанной реальности: обзор и классификация методов // Вестник СПбГУТ. 2026. Т. 4. № 2. С. 5. EDN: FKQFIV

Reference for citation:

Buchatsky A., Kuzmina D. Color Transfer Methods in Mixed Reality Systems: A Review // Herald of SPbSUT. 2026. Vol. 4. Iss. 2. P. 5. EDN: FKQFIV

На протяжении двух последних десятилетий для решения этой задачи было предложено большое число подходов: от классических статистических преобразований до генеративных методов на основе диффузионных моделей. Параллельно сформировалось самостоятельное направление — глубокая гармонизация изображений (*от англ. Deep Image Harmonization*), — рассматривающее встраивание объекта как задачу выравнивания цветового домена переднего плана и фона. Наиболее свежие разработки 2024–2026 гг. привнесли два новых подхода: генеративную оценку освещения и мягкую гармонизацию, ориентированную непосредственно на работу в режиме дополненной реальности (AR, *аббр. от англ. Augmented Reality*) [1, 2]. Настоящая статья систематизирует рассмотренные подходы и анализирует метрики оценки качества встраивания виртуальных объектов.

Постановка задачи

Задача цветовой коррекции виртуального объекта в системе смешанной реальности формулируется следующим образом. Даны реальная сцена, заданная набором пикселей в цветовом пространстве CIELAB, и виртуальный объект, синтезированный при нейтральных условиях освещения. Требуется преобразовать цветовые характеристики объекта так, чтобы минимизировать перцептивное расхождение с реальным окружением при следующих ограничениях: необходимо сохранять исходный цветовой облик объекта; преобразование должно работать в реальном времени, а результат — допускать объективную оценку метриками, коррелирующими с восприятием человека.

Предложенные в литературе методы принципиально различаются по двум ключевым характеристикам: глубине моделирования светового поля сцены и вычислительной сложности. Общая классификация рассматриваемых подходов представлена на рисунке 1.



Рис. 1. Классификация методов цветовой коррекции в видеосистемах смешанной реальности

Статистические методы переноса цвета

Первым и наиболее широко цитируемым методом является описанный в 2001 г. подход [3], суть которого состоит в согласовании среднего значения и стандартного отклонения по каждому из каналов пространства CIELAB. Преобразование имеет вид:

$$v'_x = (v_x - \mu_v) \cdot \left(\frac{\sigma_r}{\sigma_v} \right) + \mu_r,$$

где v_x — значение компоненты пикселя виртуального объекта; μ_v и σ_v — среднее и стандартное отклонение объекта; μ_r и σ_r — соответствующие параметры реальной сцены. Метод прост в реализации и вычислительно эффективен, что обеспечило ему широкое распространение. Принципиальным ограничением является единое глобальное преобразование для всего объекта без учета пространственной неоднородности освещения и геометрии сцены.

Методы адаптации баланса белого [4] решают смежную задачу нормировки цветовых компонентов изображения при изменении условий освещения. Алгоритм Gray World основан на допущении о равномерном распределении цветов; White Patch предполагает нейтральность максимального значения в каждом канале. Нейросетевой метод Deep White-Balance [4] обучен предсказывать матрицу трансформации из признаков изображения. При всей эффективности для цветовой нормализации снимка эти методы не предназначены для согласования встроенного объекта с конкретным местом сцены: они корректируют общий баланс, не затрагивая локальные условия освещения.

Нейросетевые методы оценки освещения

Авторы статьи [5] в 2017 г. предложили первую архитектуру, предсказывающую полную панораму с расширенным динамическим диапазоном (HDR, *аббр. от англ.* High Dynamic Range) в виде равноугольной карты освещения из единственного снимка интерьера. Карта затем используется для рендеринга виртуального объекта с корректными параметрами светового поля. Метод DeepLight [6] (2019) адаптирует эту идею для мобильной смешанной реальности: нейросеть предсказывает пятимерное HDR-освещение из одного обычного снимка и работает с интерактивной частотой кадров на смартфоне. При обучении используются видеозаписи зеркальных сфер с различными характеристиками отражения, которые раскрывают разные аспекты освещения в одном кадре.

Метод EMLight [7] (2021) аппроксимирует световое поле набором сферических гауссиан: нейросеть предсказывает их параметры — направление, интенсивность и ширину, — что позволяет моделировать несколько источников освещения одновременно и обеспечивает более реалистичное затенение.

Общий недостаток нейросетевых методов — зависимость от соответствия тестовых сцен обучающей выборке, а также значительная вычислительная нагрузка при прямом проходе через сеть, затрудняющая работу в реальном времени на мобильных устройствах без отдельного графического процессора.

Оценка освещения на основе генеративных моделей

Принципиально иной подход к оценке освещения предложен в [8] (метод DiffusionLight, 2024). Вместо прямой регрессии к параметрам освещения используется генеративная (диффузионная) модель, обученная «дорисовывать» хромированный шар-зонд в произвольную точку изображения. Шар отражает все окружающее световое поле, и после развертки его отражения получается HDR-карта освещения для данной точки сцены. Дополнительное дообучение с использованием адаптера LoRA обеспечивает генерацию шаров с различными уровнями экспозиции, необходимыми для HDR-слияния.

Ключевое достоинство данного метода — значительно лучшая способность к обобщению на «естественных» сценах (In-the-Wild): генеративная модель, обученная на миллиардах обычных изображений, несет богатые сведения о естественном освещении, которых недостаточно в специализированных наборах HDR-изображений. На практике метод превзошел нейросетевые аналоги как для помещений, так и для уличных сцен.

Методы на основе физических моделей освещения

Физические модели представляют световое поле аналитическими функциями. Разложение по сферическим гармоникам низкого порядка [9] дает компактное описание рассеянного освещения: уже девяти коэффициентов нулевого и первого порядков достаточно для реалистичного рендеринга объектов в диффузных условиях. Модель Фонга и физически корректные методы рендеринга (PBR, *аббр. от англ. Physically Based Rendering*) учитывают зеркальные отражения и параметры материалов поверхности. Устройства с лидаром (LiDAR) позволяют получать точные карты глубины в реальном времени, что значительно упрощает построение геометрии сцены.

На практике проблема применения PBR связана с тем, что точная трехмерная геометрия сцены и оптические свойства поверхностей в большинстве случаев заранее не известны, а без них корректный рендеринг невозможен. Восстановить эти данные по видеопотоку без специализированного оборудования можно лишь приближенно — нейросетевая оценка глубины справляется с задачей, но по точности уступает прямым измерениям.

Глубокая гармонизация изображений

Моделировать физические закономерности освещения сложно и дорого; значительно продуктивнее выглядит другой путь: показать нейросети тысячи примеров удачно и неудачно вписанных объектов и дать ей самой разобраться, как должны выглядеть цвета. Именно эта идея породила направление Deep Image Harmonization — подход, который отказывается от явных физических моделей в пользу обучаемого выравнивания цветового домена. Метод DoveNet [10] стал одной из первых убедительных реализаций этой идеи: авторы ввели в архитектуру дискриминатор верификации домена — специальный блок, штрафующий модель всякий раз, когда статистики цветовых признаков переднего плана расходятся с фоном. Сеть, по сути, учится не просто менять цвет, а делать передний план и фон неразличимыми с точки зрения цветового домена. Параллельно был опубликован набор данных iHarmony4 — более 73 000 синтетических составных изображений, — который впервые дал исследователям единую репрезентативную базу для достоверного сравнения методов гармонизации.

Трансформерный метод CDT (*аббр. от англ. Collaborative Dual Transformations*) [11] применяет модуль совместных двойственных преобразований, позволяющий независимо обрабатывать область переднего плана и все изображение целиком, что улучшает передачу локального контекста освещения. Существенное ограничение обоих методов — обучение на синтетических парах изображений, что порождает разрыв домена при работе с реальными составными изображениями, в которых несоответствие освещения носит иной характер.

Гармонизация в пространстве внутренних составляющих изображения

Принципиально иную архитектуру предложили в 2023 г. авторы работы [12]. Изображение разлагается на внутренние составляющие: альбедо (albedo, собственный цвет поверхности) и затенение (shading, вклад освещения). Согласование цвета выполняется в пространстве альбедо при помощи самообучаемой сети, предсказывающей параметры базовых операций редактирования (экспозиция, насыщенность, баланс белого). Затем по картам нормалей и затенения оценивается простая глобальная модель освещения сцены, на основе которой генерируется корректное затенение для вставленного объекта.

Ключевое преимущество по сравнению с методами глубокой гармонизации состоит в способности моделировать реальные эффекты освещения (направленный свет, обводящая подсветка), а не просто выравнивать цветовой баланс. Обучение в самообучаемом режиме на наборах данных для семантической сегментации снижает зависимость от пар составных изображений.

В контексте смешанной реальности разложение на альбедо и затенение меняет саму логику цветовой коррекции. Работая в пространстве альбедо, можно затронуть только собственный цвет поверхности объекта, — не касаясь вклада освещения, — и наоборот. Именно это давало сбой в статистических методах: они изменяли яркостные и хроматические характеристики объекта одновременно, что приводило к его перекрашиванию вместо передачи световой атмосферы сцены.

Мягкая гармонизация методом оптимального транспорта

Метод MKL-Harmonizer [13] (аббр. от англ. Monge – Kantorovich Linear – линейное отображение Монжа – Канторовича) создавался с конкретной практической целью – работать внутри цепочки обработки AR-приложения, без сложных вычислений. Вместо нейросетевой регрессии авторы обратились к теории оптимального транспорта: компактный кодировщик находит 12 параметров MKL, которое переводит цветное распределение объекта как можно ближе к распределению фона. При этом, в отличие от методов глубокой гармонизации, никаких синтетических пар изображений для обучения не нужно, что исключает разрыв между синтетическими обучающими и реальными данными.

Проверка на смартфонах Pixel 4 и Pixel 7 подтвердила, что метод работает прямо на устройстве (on-Device) и укладывается в близкое к реальному времени. На специально собранном наборе данных реальных AR-композиций с точными масками объектов MKL-Harmonizer показал лучший суммарный результат среди всех сравниваемых подходов. Однако линейная природа транспортного отображения не позволяет учесть геометрию освещения, и в сценах с выраженным направленным светом точность гармонизации заметно падает.

Семантическая сегментация и отдельная коррекция яркостной и хроматических компонент

Чтобы виртуальный объект располагался в сцене естественно, необходимо определить ее пространственную структуру, т. е. положение опорных поверхностей (пола, стен) и удаленность областей сцены от камеры. Для этого в связке могут работать две нейросети: MiDaS [14] строит карту относительных глубин, а SAM [15] выделяет семантически однородные области. Вместе они позволяют автоматически найти подходящую горизонтальную поверхность и предположить масштаб объекта через среднюю глубину выбранного участка.

Суть предлагаемого авторами настоящей статьи подхода заключается в том, что вместо изменения цвета объекта целиком яркость и цветность задаются по-разному. Компонента яркости L^* пространства CIELAB подстраивается под сцену полностью, а хроматические компоненты a^* и b^* – лишь частично, благодаря чему объект должен перенимать световую атмосферу окружения, сохраняя свой исходный цвет. При этом алгоритм принимает на вход три набора данных: RGB-изображение реальной сцены, бинарную маску области размещения виртуального объекта и синтезированное изображение этого объекта в нейтральных условиях освещения (источник D65). Результатом работы является скорректированное изображение объекта, цветовые характеристики которого приведены в соответствие со световой атмосферой реальной сцены. Предполагается, что алгоритм будет допускать работу в режиме реального времени без специализированного аппаратного обеспечения.

Сопоставление рассмотренных подходов показывает, что чем полнее метод учитывает световое поле и геометрию сцены, тем выше его вычислительные затраты. Статистические методы и методы адаптации баланса белого работают быстро, но не учитывают пространственную неоднородность освещения. Нейросетевые, генеративные и физические методы описывают освещение точнее, однако требуют значительных ресурсов или дополнительного оборудования. Методы гармонизации занимают промежуточное положение, но их точность зависит от состава обучающих данных. Принцип работы и основные ограничения каждой группы методов приведены в таблице 1.

Таблица 1. Сравнительный анализ методов переноса цвета для систем смешанной реальности

Метод / подход	Принцип работы	Основные ограничения
Глобальный перенос цвета [3]	Согласование среднего и стандартного отклонения по каналам L^* , a^* , b^* цветового пространства CIELAB	Единое преобразование для всего объекта; пространственная неоднородность освещения игнорируется
Адаптация баланса белого (Gray World, White Patch, Deep White Balance) [4]	Нормировка цветовых компонент на основе статистических допущений о нейтральных областях или обученной модели	Предназначены для нормализации съемки, а не для согласования встроенного объекта с конкретной сценой

Метод / подход	Принцип работы	Основные ограничения
Нейросетевая оценка освещения (DeepLight, EMLight и др.) [5–7]	Предсказание HDR-карты освещения или параметрической модели из одного RGB-кадра при помощи сверточных или трансформерных сетей	Высокая вычислительная нагрузка; чувствительность к расхождению обучающих и тестовых данных
Диффузионная оценка освещения (DiffusionLight) [8]	Генеративная модель «дорисовывает» хромированный шар-зонд в указанную точку сцены; шар разворачивается в HDR-карту освещения	Итеративный процесс диффузии медленнее прямой регрессии; требует дополнительного этапа развертки шара
Физические модели освещения (сферические гармоники, модель Фонга, PBR) [9]	Аппроксимация светового поля аналитическими функциями; рендеринг с учетом свойств материалов и геометрии поверхности	Требуется точная трехмерная геометрия сцены; работа в реальном времени без специализированного оборудования (LiDAR) затруднена
Глубокая гармонизация изображений (DoveNet, CDT) [10, 11]	Нейросеть выравнивает цветовой домен вставленного переднего плана и фона; набор данных iHarmony4	Разрыв домена (domain gap) при работе с реальными составными изображениями; геометрия освещения не моделируется явно
Гармонизация в пространстве внутренних компонент (альbedo / затенение) [12]	Разложение на составляющие альbedo и затенение; согласование в пространстве альbedo с последующим восстановлением затенения по оцененной модели освещения	Точность ограничена качеством разложения на внутренние составляющие; вычислительно затратнее статистических методов
Мягкая гармонизация методом оптимального транспорта (MKL-Harmonizer) [13]	Компактный кодировщик предсказывает параметры MKL между цветовыми распределениями объекта и фона; работа прямо на устройстве	Линейная модель ограничивает точность при сложном направленном освещении; геометрия сцены не учитывается
Метод, использующий семантическую сегментацию и отдельную коррекцию яркостной и хроматических компонент (MiDaS + SAM + CIELAB, в разработке) [3, 14, 15]	Яркостная компонента L^* и цветовые a^* , b^* корректируются в разной степени; освещение анализируется с учетом положения объекта в сцене; масштаб и позиция подбираются автоматически по карте глубины	Результат зависит от того, насколько точно сегментирована сцена; степень коррекции хроматических компонент требует экспериментального обоснования

Метрики оценки цветового расхождения

Выбор метрики существенно влияет на интерпретацию результатов и возможность сравнения методов между собой. Классическая метрика $\Delta E \cdot ab$ (CIE76) определяет цветовое расхождение как евклидово расстояние в пространстве CIELAB. Ее принципиальный недостаток — перцептивная неравномерность: одинаковое числовое расстояние воспринимается человеком по-разному в зависимости от области пространства. Расхождение в три единицы в синей области менее заметно визуально, чем то же расхождение в нейтральной области. Метрика ΔE_{94} частично исправляет это введением весовых коэффициентов, однако формула не охватывает синюю область и считается частично устаревшей.

Наилучшую корреляцию с субъективными оценками наблюдателей обеспечивает стандартизированная формула CIEDE2000 (ΔE_{00}) [16], где введены весовые функции SL , SC , SH и поправочный член RT , учитывающий взаимодействие цветности и оттенка в синей области:

$$\Delta E_{00} = \left(\frac{\Delta L'}{S_L} \right)^2 + \left(\frac{\Delta C'}{S_C} \right)^2 + \left(\frac{\Delta H'}{S_H} \right)^2 + R_T \cdot \left(\frac{\Delta C'}{S_C} \right) \cdot \left(\frac{\Delta H'}{S_H} \right),$$

где $\Delta L'$, $\Delta C'$, $\Delta H'$ — нормированные разности яркости, цветности и оттенка соответственно. При $\Delta E_{00} < 1$ цветовое расхождение практически неразлично невооруженным взглядом, при $\Delta E_{00} > 2$ оно отчетливо воспринимается наблюдателем. Именно это свойство делает CIEDE2000 наиболее подходящей метрикой для задач смешанной реальности, где сцены могут включать как нейтральные источники, так и насыщенные цветные подсветки [2].

Структурный индекс SSIM [17] оценивает близость яркостных распределений и текстурных паттернов двух изображений. Он полезен для оценки естественности встраивания, однако в практике работы с системами смешанной реальности может вводить в заблуждение: при агрессивной цветовой коррекции, переводящей объект, например, из насыщенно-красного в синий, яркостная структура существенно меняется, и SSIM фиксирует ухудшение, тогда как CIEDE2000 и PSNR при этом улучшаются. Именно поэтому для всестороннего анализа целесообразно совместное применение всех трех метрик. В задачах гармонизации изображений, где объект занимает лишь часть кадра, стандартом стала метрика fMSE (аббр. от англ. foreground Mean Squared Error – среднеквадратическая ошибка, вычисляемая исключительно по пикселям переднего плана). В отличие от обычной MSE по всему изображению fMSE не «разбавляется» неизменным фоном и точнее отражает качество коррекции самого встроенного объекта [10]. Наконец, метрика LPIPS (аббр. от англ. Learned Perceptual Image Patch Similarity – обучаемая мера перцептивного сходства фрагментов изображения) [18] оценивает перцептивное сходство через разность активаций предобученной нейросети-классификатора и демонстрирует наилучшую среди метрик общего назначения корреляцию с оценками наблюдателей. По сравнению с SSIM она лучше переносит агрессивную цветовую коррекцию и точнее реагирует на структурные искажения, но за счет запуска нейросетевого классификатора, что на практике осложняет использование метрики в системах реального времени. Сводная характеристика рассмотренных метрик приведена в таблице 2.

Таблица 2. Сравнительный анализ метрик оценки цветового расхождения

Метрика	Физический смысл	Достоинства	Ограничения
$\Delta E \cdot ab$ (CIE76)	Евклидово расстояние в пространстве CIELAB	Простота вычисления; широкая поддержка в инструментах	Неравномерная перцептивная чувствительность в разных областях пространства
ΔE_{94}	Взвешенное расстояние с учетом насыщенности	Лучше отражает зрительное восприятие цветовых различий, чем CIE76	Не учитывает нелинейность в синей области; формула частично устарела
CIEDE2000 (ΔE_{00})	Нормированное цветовое расхождение с весовыми функциями SL , SC , SH и поправкой RT	Наилучшая корреляция с оценками человека; стандартизована МКО; порог $\Delta E_{00} < 1$ означает практически неразличимое расхождение	Вычислительно сложнее; требует перевода в CIELAB
SSIM	Структурное сходство яркостных распределений двух изображений	Учитывает текстурные особенности; значение 1 соответствует полному совпадению	Может снижаться при сильной цветовой коррекции даже при видимом улучшении
PSNR	Отношение пикового значения сигнала к среднеквадратической ошибке, дБ	Широко принята как эталонная; удобна для сравнения методов	Слабо коррелирует с субъективным восприятием цвета
fMSE	Среднеквадратическая ошибка, вычисляемая только по пикселям переднего плана (области встроенного объекта)	Точнее отражает качество коррекции самого объекта, не «разбавляя» результат неизменным фоном; стандарт набора данных iHarmony4	Требуется маски объекта; для систем без явной сегментации необходима дополнительная разметка
LPIPS	Перцептивное сходство двух изображений через разность активаций предобученной нейросети-классификатора (AlexNet или VGG)	Наилучшая корреляция с субъективными оценками наблюдателей среди метрик общего назначения; стандарт в задачах гармонизации изображений	Требуется запуск нейросетевого классификатора; вычислительно затратнее пиксельных метрик; менее интерпретируема, чем CIEDE2000

Выводы

Данная работа охватывает методы переноса цвета и гармонизации изображений в системах смешанной реальности за период с 2001 по 2026 гг. Все рассмотренные подходы распределены по девяти группам, в том числе два новейших направления: генеративная оценка освещения на основе диффузионных моделей и мягкая гармонизация методами оптимального транспорта.

Статистические методы и методы адаптации баланса белого отличаются высокой скоростью работы, однако не учитывают геометрию освещения. Нейросетевые методы (DeepLight, EMLight и др.) в теории точнее, но требуют значительных вычислительных ресурсов. Метод DiffusionLight обеспечивает принципиально лучшую способность к обобщению за счет использования богатых сведений об освещении, накопленных генеративной моделью. Методы глубокой гармонизации (DoveNet, CDT) хорошо работают в границах обучающего распределения, однако теряют точность из-за разрыва домена на реальных составных изображениях.

Наиболее концептуально значимым следует признать метод гармонизации во внутреннем пространстве изображения: разложение на альbedo и затенение позволяет разделить собственный цвет объекта и вклад освещения, устраняя главный недостаток статистических методов — «перекрашивание» вместо передачи световой атмосферы. С практической точки зрения наибольший интерес представляет метод MKL-Harmonizer, впервые обеспечивший гармонизацию непосредственно на мобильном устройстве в реальном времени.

Для объективной оценки качества встраивания целесообразно совместное применение метрик CIEDE2000, SSIM и PSNR. Метрика CIEDE2000 наилучшим образом отражает перцептивное цветовое расхождение и рекомендуется как основной инструмент сравнительного анализа. Предложенный авторами подход к созданию метода цветовой согласованности виртуальных объектов в видеосистемах смешанной реальности на основе семантической сегментации и отдельной коррекции яркостной и хроматических компонент планируется реализовать в виде программного алгоритма, допускающего работу в режиме реального времени, с последующей интеграцией в конвейеры обработки видео в системах смешанной реальности. Перспективным направлением дальнейших исследований является экспериментальная верификация разрабатываемого метода с количественной оценкой по метрике CIEDE2000.

Литература

1. Бучатский А. Н., Кузьмина Д. Д. Управление цветом в современных медиатехнологиях: актуальные методы и практики // Актуальные проблемы инфотелекоммуникаций в науке и образовании: материалы XIII Международной научно-технической и научно-методической конференции (АПИНО 2024, Санкт-Петербург, 27–28 февраля 2024 г.): сб. статей. СПб.: СПбГУТ, 2024. С. 580–585. EDN: WANRTA
2. Бучатский А. Н., Кузьмина Д. Д. Проблема восприятия метамерных пар при наложении виртуальных объектов на реальные: новые подходы к повышению качества визуализации // 65-я научно-техническая конференция профессорско-преподавательского состава, научных работников и аспирантов (HTK ППС 2025, Санкт-Петербург, 17–21 февраля 2025 г.): сб. статей. СПб.: СПбГУТ, 2025. С. 37–42. EDN: UJKVBB
3. Reinhard E., Adhikhmin M., Gooch B., Shirley P. Color Transfer between Images // IEEE Computer Graphics and Applications. 2001. Vol. 21. Iss. 5. PP. 34–41. DOI: 10.1109/38.946629
4. Afifi M., Brown M. S. Deep White-Balance Editing // Proceedings of the IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR, Seattle, USA, 13–19 June 2020). 2020. PP. 1397–1406. DOI: 10.1109/CVPR.42600.2020.00147
5. Gardner M.-A., Sunkavalli K., Yumer E. Shen X., Gambaretto E., et al. Learning to Predict Indoor Illumination from a Single Image // ACM Transactions on Graphics. 2017. Vol. 36. Iss. 6. P. 176. DOI: 10.1145/3130800.3130891
6. LeGendre C., Ma W.-C., Fyffe G., Flynn J., Charbonnel L., et al. DeepLight: Learning Illumination for Unconstrained Mobile Mixed Reality // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR, Long Beach, USA, 15–20 June 2019). 2019. PP. 5911–5921. DOI: 10.1109/CVPR.2019.00607
7. Zhan F., Zhang Ch., Yu Y., Chang Y., Lu S., et al. EMLight: Lighting Estimation via Spherical Distribution Approximation // Proceedings of the AAAI Conference on Artificial Intelligence (2–9 February 2021, Held Virtually). 2021. Vol. 35. Iss. 4. PP. 3287–3295. DOI: 10.1609/aaai.v35i4.16440

8. Phongthawee P., Chinchuthakun W., Sinsunthithet N., Raj A., Jampaniet V., et al. DiffusionLight: Light Probes for Free by Painting a Chrome Ball // Proceedings of the IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR, Seattle, USA, 16–22 June 2024). 2024. PP. 98–108. DOI: 10.1109/CVPR52733.2024.00018
9. Ramamoorthi R., Hanrahan P. An Efficient Representation for Irradiance Environment Maps // Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'01). New York: ACM Press, 2001. PP. 497–500. DOI: 10.1145/383259.383317
10. Cong W., Zhang J., Niu L., Liu L., Linget Zh., et al. DoveNet: Deep Image Harmonization via Domain Verification // Proceedings of the IEEE / CVF Conference on Computer Vision and Pattern Recognition (Seattle, USA, 13–19 June 2020). 2020. PP. 8394–8403. DOI: 10.1109/CVPR42600.2020.00842
11. Cong W., Tao X., Niu L., Liang J., Gao X., et al. High-Resolution Image Harmonization via Collaborative Dual Transformations // Proceedings of the IEEE / CVF Conference on Computer Vision and Pattern Recognition (CVPR, New Orleans, USA, 18–24 June 2022). 2022. PP. 18449–18458. DOI: 10.1109/CVPR52688.2022.01792
12. Careaga C., Miangoleh S. M. H., Aksoy Y. Intrinsic Harmonization for Illumination-Aware Compositing // Proceedings of Computer Graphics and Interactive Techniques-Asia Conference (SIGGRAPH Asia, Sydney, Australia, 12–15 December 2023). 2023. P. 100. DOI: 10.1145/3610548.3618178
13. Larchenko M., Guskov D., Lobashev A., Derevyanko G. Lightweight Optimal-Transport Harmonization on Edge Devices // Proceedings of the AAAI Conference on Artificial Intelligence. 2026. Vol. 40. Iss. 7. PP. 5827–5835. DOI: 10.1609/aaai.v40i7.37504. EDN: NYIMRG
14. Ranftl R., Bochkovskiy A., Koltun V. Vision Transformers for Dense Prediction // Proceedings of the IEEE / CVF International Conference on Computer Vision (ICCV, Montreal, Canada, 10–17 October 2021). 2021. PP. 12159–12168. DOI: 10.1109/ICCV48922.2021.01196
15. Kirillov A., Mintun E., Ravi N., Mao H., Rolland C., et al. Segment Anything // Proceedings of the IEEE / CVF International Conference on Computer Vision (ICCV, Paris, France, 01–06 October 2023). 2023. PP. 3992–4003. DOI: 10.1109/ICCV51070.2023.00371
16. Sharma G., Wu W., Dalal E. N. The CIEDE2000 Color-Difference Formula: Implementation Notes, Supplementary Test Data, and Mathematical Observations // Color Research and Application. 2005. Vol. 30. Iss. 1. PP. 21–30. DOI: 10.1002/col.20070
17. Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P. Image Quality Assessment: From Error Visibility to Structural Similarity // IEEE Transactions on Image Processing. 2004. Vol. 13. Iss. 4. P. 600–612. DOI: 10.1109/TIP.2003.819861
18. Zhang R., Isola P., Efros A. A., Shechtman E., Wang O. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric // Proceedings of the IEEE / CVF Conference on Computer Vision and Pattern Recognition (Salt Lake City, USA, 18–23 June 2018). 2018. PP. 586–595. DOI: 10.1109/CVPR.2018.00068. EDN: FZQHAF

Статья поступила 16 мая 2026 г.
Одобрена после рецензирования 24 июня 2026 г.
Принята к публикации 12 октября 2026 г.

Информация об авторах

Бучатский Александр Николаевич — кандидат технических наук, доцент кафедры цифрового телевидения и метрологии Санкт-Петербургского государственного университета телекоммуникаций. E-mail: abuchatsky@sut.ru

Кузьмина Дарья Денисовна — аспирант кафедры цифрового телевидения и метрологии Санкт-Петербургского государственного университета телекоммуникаций. E-mail: kuzmina.dd@sut.ru

<https://doi.org/10.31854/3034-2201-2026-4-2-C05>
EDN: FKQFIV

Color Transfer Methods in Mixed Reality Systems: A Review

A. Buchatsky, D. Kuzmina ✉

The Bonch-Bruevich Saint Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Background. One of the unsolved problems in mixed reality systems is the color consistency of embedded virtual objects with their real surroundings. Existing methods either apply a single global transformation without accounting for lighting geometry, or demand computational resources unavailable in real time. **Aim.** To systematize and critically analyze existing color transfer, illumination estimation and image harmonization methods for mixed reality, including the latest approaches of 2024–2026. **Methods.** Analytical review of scientific publications; comparative classification by operating principle, computational requirements and scope of applicability. **Results.** A nine-group classification is proposed; an approach to the color consistency of virtual objects based on semantic segmentation and separate correction of lightness (L^*) and chromatic (a^* , b^*) components is proposed, intended for subsequent software implementation; a comparative analysis of color difference and image quality metrics is conducted, including CIEDE2000, SSIM, PSNR, fMSE and LPIPS. **Novelty.** For the first time, color transfer and image harmonization methods for video-based mixed reality systems are systematized in a unified review spanning 2001 to 2026. **Theoretical and practical significance.** The proposed classification supports justified selection of color correction methods in mixed reality system design, balancing quality and computational constraints.

Key words: mixed reality, color correction, color transfer, image harmonization, CIEDE2000, CIELAB color space, generative models, optimal transport, semantic segmentation

Information about Authors

Buchatsky Alexander – PhD in Engineering, Associate Professor of the Department of Digital Television and Metrology (The Bonch-Bruevich Saint Petersburg State University of Telecommunications).
E-mail: abuchatsky@sut.ru

Kuzmina Daria – Postgraduate Student of the Department of Digital Television and Metrology (The Bonch-Bruevich Saint Petersburg State University of Telecommunications). E-mail: kuzmina.dd@sut.ru