ОБЗОР ОБОРУДОВАНИЯ ДЛЯ СЪЕМКИ 3D ИЗОБРАЖЕНИЙ С ПОСЛЕДУЮЩЕЙ ПЕРЕДАЧЕЙ ЧЕРЕЗ СЕТИ СВЯЗИ

Б. О. Паньков¹, М. А. Маколкина^{1*}

¹Санкт-Петербургский государственный университет телекоммуникаций
им. проф. М. А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

*Адрес для переписки: makolkina@list.ru

Аннотация—Относительно недавнее возникновение общедоступных потребительских датчиков RGB-D позволило раскрыть большое количество функциональных возможностей для взаимодействия с 3D-технологией. К примеру, такие устройства, как Microsoft Kinect v1, Microsoft Kinect v2 и Intel Realsense F200, на сегодняшний день являются распространенными сенсорами RGB-D, использующими технологии структурированного света и времени пролета для получения информации о пиксельной глубине. Статья содержит обширное сопоставление RGB-D устройств на предмет разрешающей способности RGB и Depth камер, задержки (времени, необходимого для построения карты глубины и её обработки), а также сравниваются углы обзора, необходимые USB-интерфейсы, размеры датчиков и т. д. Главная цель статьи – предоставить полноценное и наглядное сравнение RGB-D устройств, которые потенциально могут быть применены для захвата трехмерных изображений с последующей их передачей сетями связи.

Ключевые слова—передача трехмерных изображений, дополненная реальность, сети связи 2030, голография, RGB-D, Intel Realsense F200, Kinect v1/v2, 3D.

Информация о статье

УДК 621.391.

Язык статьи – русский.

Поступила в редакцию 19.11.2021, принята к печати 10.12.2021.

Ссылка для цитирования: Паньков Б. О., Маколкина М. А. Обзор оборудования для съемки 3D изображений с последующей передачей через сети связи // Информационные технологии и телекоммуникации. 2021. Том 9. № 3. С. 48–71. DOI 10.31854/2307-1303-2021-9-3-56-71.

OVERVIEW OF EQUIPMENT FOR CAPTURING 3D IMAGES WITH SUBSEQUENT TRANSMISSION THROUGH A COMMUNICATION NETWORK

B. Pankov¹, M. Makolkina^{1*}

¹The Bonch-Bruевич Saint-Petersburg State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

*Corresponding author: makolkina@list.ru

Abstract—The relatively recent advent of the publicly available consumer RGB-D sensors has enabled a wide range of functionality to interact with 3D technology. For example, devices such as Microsoft Kinect v1, Microsoft Kinect v2, and Intel Realsense F200 are currently available RGB-D sensors using Structured light and Time of Flight technologies for getting information about pixel depth. The article contains an extensive comparison of RGB-D devices in terms of the resolution of RGB and Depth cameras, latency (the time required to build a depth map and its processing), and also compares viewing angles, required USB interfaces, sensor sizes, etc. The main goal of the article is to provide a complete and visual comparison of devices that can potentially be used to capture three-dimensional images with their subsequent transmission by communication networks.

Keywords—transmission of three-dimensional images, augmented reality, communication networks 2030, holography, RGB-D, Intel Realsense F200, Kinect v1 / v2, 3D.

Article info

Article in Russian.

Received 19.11.2021, accepted 10.12.2021.

For citation: Pankov B., Makolkina M.: Overview of Equipment for Capturing 3D Images with Subsequent Transmission Through a Communication Network // Telecom IT. 2021. Vol. 9. Iss. 3. pp. 56–71 (in Russian). DOI 10.31854/2307-1303-2021-9-3-56-71.

Введение

На сегодняшний день развитие сферы ИТ протекает с огромной скоростью, что, безусловно, влечет за собой совершенствование способов передачи данных. Кроме всего прочего, уже существует документ, планирующий стандарт связи шестого поколения или сети связи 2030 [1]. Вследствие чего в скором времени следует ожидать формирования новейших высокоперспективных систем и технологий, которые предъявят иные, гораздо более существенные запросы по отношению к коммуникационным сетям, чем существующие на сегодняшний момент.

Одними из наиболее перспективных приложений для сетей 2030 являются НТС (*Holographic Type Communications*) или голографические коммуникации, а также технология телеприсутствия, основанная на использовании аватаров, способных воспроизводить действия управляющего ими человека [2]. Голографическая модель будет включать в себя данные о пространстве, позиции и точке наблюдения, а также около 43 различных перспектив для каждого пикселя кадра. Следовательно, объем передаваемой информации вырастет в разы, что повлечет за собой необходимость в резком повышении пропускной способности сетей будущего.

По прогнозам, к 2030 году инженерия достигнет понимания процесса построения высококачественных голограмм в полном объеме, вследствие чего возникнет вопрос о возможности использования повседневной коммуникации пользователей посредством трехмерных изображений.

Однако сети нынешнего поколения не в силах предоставить пропускную способность, задержку и качество передачи данных, которые в ближайшее время понадобятся для реализации голографической связи [3]. Поэтому необходимо начать планировать концепцию этой технологии уже сейчас, а именно требуется протестировать и провести сравнительный анализ оборудования RGB-D (*Red-Green-Blue-Depth*), которое впоследствии может быть применено для оцифровки трехмерных объектов. Также важной задачей является определение ключевых принципов извлечения данных глубины из каждой точки кадра, что позволит осуществить программную генерацию анаглифных голографических 3D стереоизображений.

Устройства RGB-D присваивают каждому пикселю одновременно два значения: цвет и глубина, вследствие чего производится цифровизация реальных трехмерных объектов и различных сред. Одним из первых потребительских датчиков RGB-D были датчики на основе PrimeSense, использующие технологию определения дальности структурированного света (*Structured light – SL*), которая и реализована в Microsoft Kinect v1 и Intel Realsense F200. В 2015 году Microsoft выпустила Kinect v2, в котором для определения параметров глубины используется принцип времени пролета (*Time of Flight – ToF*). Представленные датчики RGB-D способны захватывать данные высокой глубины с достаточной частотой кадров по приемлемой цене. Доступность этих устройств повысила интерес к использованию 3D-технологии в самых разных сферах, начиная от здравоохранения, 3D-сканирования, моделирования и измерения объектов, робототехники, навигации и т. д. Все 3 датчика используют проекцию инфракрасного света (*Near*

Infrared – NIR) на длинах волн порядка 750–1400 нм для получения данных о глубине.

Сравнение устройств RGB-D

Kinect V1

Kinect v1 (*Kinect* для *Xbox 360*) был разработан компанией Microsoft и впервые представлен на выставке E3 в 2009 году под кодовым названием Project Natal [4]. В 2011 году была выпущена первая версия Kinect Software Development Kit (SDK), вследствие чего популярность Kinect среди пользователей резко возросла, и он стал активно применяться во многих исследовательских и научных проектах, связанных с технологией 3D.

Kinect v1 состоит из нескольких основных элементов (рис. 1):

1) VGA-камера, регистрирующая видимый диапазон. Представляет из себя обычную RGB-камеру разрешением 640×480 пикселей и частотой 30 кадров в секунду. RGB-камера используется для распознавания лиц, съемки видео (видео-чата), фотоснимков, обработки мимики и прочих «тонких моментов».

2) Лазерный проектор инфракрасного диапазона, излучающий структурированный ИК-свет. Данный компонент обеспечивает так называемое «прощупывание» пространства сеткой из точек.

3) Инфракрасная камера (CMOS-сенсор) – это интегральная схема, которая содержит в себе фотодетекторную матрицу. Монохромный CMOS-датчик осуществляет регистрацию отраженного от поверхности структурированного света (формат спроецированной ИК матрицы).

4) Специальные микрофоны, обеспечивающие локализацию источника звука, подавление окружающего шума, и работающие на частоте дискретизации 16 кГц.

5) Сервопривод автоматического наклона устройства в вертикальной плоскости.

Рис. 1. Основные составляющие элементы устройства Microsoft Kinect v1

Сенсор глубины включает в себя два компонента: ИК проектор и CMOS-датчик, которые функционируют одновременно и обеспечивают возможность захвата трехмерных изображений [5]. Kinect v1 осуществляет получение данных о глубине на основании метода, называемого 3D-сканирование со структурированным светом (*Structured light* – SL).

Kinect v1 стал первым сенсором в потребительском сегменте, в котором была применена реализация SL. Принцип работы большинства структурированных световых сканеров основан на проекции узкой полосы света на трехмерную поверхность. Спроецированная световая полоса претерпевает искажения вследствие разного преломления луча на неоднородном рельефе 3D объекта, и, благодаря этому, возможна точная геометрическая идентификация любой формы [6].

В системе Kinect вместо проектирования световых линий видимого диапазона, инфракрасный проектор распространяет узор ИК лучей (называемый закодированным инфракрасным изображением или сеткой), который затем отражается от объемной поверхности и принимается цифровой матрицей CMOS-сенсора (рис. 2). Соответственно, чем ярче свет зарегистрировал CMOS, тем ближе расположен сканируемый объект. Однако камера не в состоянии различать области, находящиеся слишком близко к сенсору, так как наблюдается слишком яркое поверхностное отражение.

Захваченное закодированное инфракрасное изображение передается на процессорную микросхему PrimeSense, на которой сравниваются эталонное («плоское») относительное положение точек, содержащееся в памяти устройства, и полученные искаженные координаты. Эталон получают на этапе калибровки при производстве чипа. Когда точка проецируется на область, находящуюся ближе или дальше от запрограммированного в чипе положения проекционной плоскости, ее положение, соответственно, будет смещено. Благодаря вычислению смещения каждого пикселя изображения на основе процедуры корреляции, осуществляется построение матрицы несоответствия, из значений которой Kinect определяет действительное расстояние от плоскости устройства до любой точки сканируемой поверхности [7].

Рис. 2. Закодированное ИК-изображение Kinect v1

Рис. 3. Карта глубины (слева), воссозданная из закодированного инфракрасного шаблона (справа). Изображения получены с помощью Kinect v1

ный пиксель на 2D-изображении перед дисплеем (белый) или за плоскостью экрана (черный).

Результат вычисления переводится в карту глубины (изображение, хранящее расстояние от каждого пикселя до плоскости камеры) с VGA-разрешением, содержащую 2048 различных значений, представленных графически уровнями серого цвета от белого (2048) до черного (0). Карта глубины (рис. 3) указывает, нужно ли отображать определенный

Таким образом, сенсор глубины можно рассматривать как устройство, возвращающее координаты (x, y, z) для каждого пикселя изображения и формирующее облако точек, на основании которого можно построить трехмерную картину.

Kinect V2.0

Kinect v2 (рис. 4) был представлен в 2014 году, и отличался существенно лучшей производительностью и разрешением кадра по сравнению с Kinect первой версии. В Kinect v2 была использована технология измерения времени пролета (*Time of Flight* – ToF) вместо принципа структурированного света в первой версии, что позволило применять Kinect v2 на открытом воздухе при солнечном свете [8]. Наглядное сопоставление функциональных возможностей Kinect v1 и v2 представлено в таблице (см. ниже).

Рис. 4. Основные составляющие элементы устройства Microsoft Kinect v2

Технология ToF в Kinect v2 (рис. 5) реализуется на принципе непрерывной модуляции сигнала и измерении смещения фазы световой волны, в результате чего процесс получения данных о глубине можно разложить на несколько последовательных этапов:

- 1) ИК-излучатель проецирует непрерывную модулированную волну в ИК-диапазоне (невидимый для человеческого глаза);
- 2) Непрерывная волна отражается от объемной поверхности;
- 3) Объектив инфракрасной камеры регистрирует отраженную волну;
- 4) Фазометр измеряет фазовый сдвиг $\Delta\varphi$ принятого сигнала.
- 5) Расстояние до объекта (данные о глубине) рассчитывается по формуле:

$$d = \frac{\Delta\varphi}{4\pi f} \cdot c,$$

где $\Delta\varphi$ – фазовый сдвиг, f – частота модуляции сигнала, c – скорость света (стабильна и в воздухе составляет около 300 тыс. км/с), d – расстояние до объекта.

Луч ИК-света, испускаемый датчиком, возвращается быстро, если объект съемки находится близко к объективу, соответственно, и фазовый сдвиг $\Delta\varphi$ будет небольшой. Чем дальше расположен объект или его часть, тем больше времени

уйдет на то, чтобы луч достиг его, отразился и вернулся обратно, что увеличивает смещение фазы принятой волны. Аналогичным образом должно быть измерено расстояние для каждого пикселя, составляющего изображение. Это формирует детальное представление об особенностях формы сканируемого объекта.

б) Устройство получает карту глубины в диапазоне от 50 до 450 см с разрешением 512×424 пикселей и цветное (RGB) изображение 1080p (1920×1080 пикселей).

Рис. 5. Принцип функционирования технологии Time of Flight – ToF в Kinect v2

Таблица

Сравнительная характеристика устройств RGB-D

Устройство		Microsoft Kinect v1	Microsoft Kinect v2	Intel RealSense F200
Технология расчета глубины		Структурированный свет, SL	Время пролета, ToF	Структурированный свет, SL
Камера глубины (Depth)	Разрешающая способность / FPS	320×240 30	512×424 30	640×480 30
	Поле зрения (FOV)	$58,5^\circ$ – горизонталь, $46,6^\circ$ – вертикаль	$70,6^\circ$ – горизонталь, 60° – вертикаль	73° – горизонталь, 59° – вертикаль
Цветная камера (RGB)	Разрешающая способность / FPS	640×480 30	1920×1080 30	1920×1080 30
	Поле зрения (FOV)	62° – горизонталь, $48,6^\circ$ – вертикаль	$84,1^\circ$ – горизонталь, $53,8^\circ$ – вертикаль	73° – горизонталь, 59° – вертикаль
Диапазон распознавания глубины, м		0,8–4	0,5–4,5	0,4–2,8
Интерфейс		USB 2.0	USB 3.0	USB 3.0
Поддерживаемая операционная система		Windows, ROS	Windows, ROS	Windows, Linux
Габаритные размеры (Ш×В×Г), мм		$305 \times 76,2 \times 63,5$	$249 \times 66 \times 67$	$110 \times 37,5 \times 12,5$
SDK		Kinect SDK v1.8	Kinect SDK v2.0	LibRealsense
Захват аудио		Набор из четырёх микрофонов с частотой дискретизации 16 кГц	Набор из четырёх микрофонов с частотой дискретизации 48 кГц	Набор из двух микрофонов с частотой дискретизации 16 кГц
Задержка, мс		~ 90 с обработкой	~ 60 с обработкой	~ 60 с обработкой

Устройство	Microsoft Kinect v1	Microsoft Kinect v2	Intel RealSense F200
Автоматический наклон	По вертикали	Нет	Нет
Возможность работы с SDK	Python, C, C++, C#, Java, JavaScript	Python, C, C++, C#, Java, JavaScript, Actionscript	C#, Unity, Java, Processing interface
Трехмерное сканирование объекта	Да	Да	Да
Способность вывода инфракрасного (IR) потока	Нет	Да	Да
Отслеживание траектории движения тела	Да	Да	Нет
Управление компенсацией экспозиции (возможность снимать изображение при малой яркости освещения)	Нет	Да	Нет
Трекинг и выражение лица, распознавание жестов рук	Да	Да	Да
Внешний источник питания	Да	Да	Нет
Потребляемая мощность, Вт	12	15	1,7

Инфракрасное изображение от Kinect v2 (и, следовательно, карта глубины) намного более независимы от окружающего света, чем инфракрасное изображение от Kinect v1, и поэтому предоставляют данные лучшего качества в большем количестве различных сцен¹.

Kinect v2 может функционировать практически в полной темноте, распознавать более мелкие объекты по сравнению с Kinect v1, а также различать одновременно до шести фигур, считывая при этом до 2 Гб данных в секунду [9]. Kinect v2 позволяет распознавать трехмерные движения 25 суставов человеческого скелета и способен определять: сидит пользователь или стоит. Также Kinect v2 может выделять такие жесты рук, как открытая кисть; кисть, сжатая в кулак; количество отогнутых пальцев и т. д.

После выхода Kinect появилось множество других устройств для сбора данных RGB-D, одним из которых является Intel RealSense F200.

Realsense F200

Камера Intel RealSense намного меньше громоздкого Kinect, независимо от версии v1 или v2, поэтому ее можно использовать для небольших мобильных роботов и БПЛА. Несмотря на то, что RealSense F200 является сравнительно маленьким устройством, оно также позволяет захватывать изображения RGB с разрешением 1080p FullHD, а карты глубины с разрешением 640×480 пикселей.

Существуют две версии исполнения камеры Intel RealSense: интегрированная и внешняя. Интегрированная версия располагается в ноутбуках или в новых планшетах [10]. Возможности камеры RealSense способны обеспечить видеочат, управление жестами и 3D-сканирование среды с помощью камеры RGB и датчика

¹ Kinect 2 для Windows – эволюция бесконтактного контроллера. URL: <https://clck.ru/YgVt3>

глубины. Автономная (внешняя) Intel RealSense (рис. 6) – это легкая, небольшая камера с питанием от USB.

Устройство включает в себя:

- 1) Сенсор глубины (инфракрасная камера + ИК лазерный проектор).
 - 2) Цветную (RGB) камеру разрешением 1920×1080 пикселей с частотой 30 кадров в секунду.
 - 3) Стереомикрофон с частотой дискретизации 16 кГц.
- В инфракрасном режиме камера Intel RealSense способна работать со скоростью 60 кадров в секунду.

Рис. 6. Основные составляющие элементы устройства Intel RealSense F200

Intel RealSense F200 реализует построение карты глубины посредством технологии структурированного света (*Structured light – SL*), что аналогично системе, развернутой в Kinect v1. То есть ИК лазерный проектор 1-го класса (полностью безопасные лазеры) проецирует инфракрасные лучи с длиной волны 860 нм (кодированное ИК-изображение), которые, отражаясь от поверхности, принимаются цифровой матрицей инфракрасной камеры². Затем производится измерение смещения каждого пикселя по отношению к эталонному («плоскому») положению точек, что в дальнейшем используется для вычисления истинного расстояния от сенсора до любой точки объекта и последующего построения карты глубины.

Общая сравнительная характеристика устройств RGB-D

На рис. 7 представлено несколько RGB, Depth и IR изображений, полученных с помощью Kinect v2 (рис. 7 (а, б, в)) и RealSense F200 (рис. 7 (г, д, е)) на одной и той же сцене. По изображениям видно, что качество данных RGB-D, предоставляемых Kinect v2, значительно превосходит RealSense F200. Kinect v2 может собирать данные RGB с лучшей цветопередачей, контрастностью и четкостью изображения, по сравнению с RealSense, а также получаемые Kinect данные глубины более плотные и содержат меньшее количество зашумленных точек [11]. Так что в случае отсутствия ограничений по габаритам устройства, Microsoft Kinect v2 является лучшим выбором из-за высокого качества захватываемых данных. Однако Kinect – это камеры, требующие внешний источник питания, что затрудняет

² Сравнение камер переднего обзора Intel RealSense SR300 и F200. URL: <https://habr.com/ru/company/intel/blog/303508/>

их использование на мобильных платформах или в сценариях, где источник питания ограничен [12]. Поэтому если требуется использовать датчик RGB-D на переносном устройстве или в ограниченных габаритах, то Intel RealSense F200 способен удовлетворить требования благодаря низкому весу и небольшому размеру.

Рис. 7. Кадры RGB, Depth и IR, сделанные на Kinect v2 и RealSense F200 в одной и той же сцене; Kinect v2: а) цветное (RGB) изображение; б) карта глубины (Depth); в) инфракрасный поток (IR). RealSense F200: г) цветное изображение; д) карта глубины; е) инфракрасный поток

Функциональные возможности выбранных для данной работы RGB-D датчиков и их технические характеристики представлены в таблице (см. выше). Вся информация была получена из инструкций по эксплуатации и с веб-сайтов производителей.

Принцип получения объемного изображения

Для получения объемного изображения необходимо пространственную карту глубины, полученную сенсором, объединить с данными RGB, передаваемыми цветной камерой (рис. 8). Результатом является цветное изображение в виде облака точек (координаты x , y , z для каждого пикселя изображения), в котором каждому значе-

Рис. 8. Облако точек, отображающее объект в истинных цветах

нию глубины будет присвоено значение цвета³. Соответственно, объект, представленный облаком точек, будет отображен в истинных цветах.

Для удобства анализа облако точек часто отображается в цветах радуги (наложение цветов радуги). Значения глубины в ближнем диапазоне представлены в цветах от красного до желтого, более удаленные точки – в цветах от зеленого до синего.

Генерация анаглифных 3D-изображений из одного 2D-изображения в сочетании с информацией о глубине

Анаглиф — формат 3D-изображения, стереоэффект в котором образуется вследствие наложения определенного цветового фильтра на ракурс каждого глаза по отдельности⁴. Данная технология осуществляет кодирование изображения, предназначенного для левого глаза, в голубой (синий) цвет, а для правого, соответственно, в красный. Затем оба ракурса совмещаются друг с другом, в результате чего образуется «сдвоенная» красно-голубая картинка. Таким образом, используя специальные анаглифические очки, предоставляется возможность наблюдения объемного 3D-изображения.

Происходит это вследствие того, что анаглифные 3D очки, благодаря разным цветам линз, отфильтровывают цвета. То есть, ракурс для левого глаза (голубое изображение) виден только через красный левый фильтр очков, в котором «задерживается» правое изображение, кодированное в красный цвет. Аналогично для голубого правого светофильтра очков, в котором «задерживается» левый ракурс, кодированный в голубой цвет. В итоге левый и правый глаз видят разные изображения, и мозг человека воспринимает эти два различных изображения как одно объёмное. Анаглиф считается одним из самых экономичных и простых методов трехмерной визуализации стереоизображений [13].

Обычно анаглифные стереоизображения генерируются простым синтезом пары изображений, снятых с бинокулярной камеры, дистанция между линзами которой соответствует межзрачковому расстоянию у человека. С другой стороны, анаглифные стереоизображения могут быть сгенерированы из одного 2D RGB-изображения в сочетании с информацией о глубине каждого пикселя, полученной сенсором глубины (инфракрасная камера + ИК лазерный проектор).

Межзрачковое расстояние у человека в среднем равняется 6,5 см. В результате один глаз наблюдает объект чуть с левой стороны, а второй глаз захватывает его же немного справа [14]. То есть, объект регистрируется каждым глазом человека под определёнными отличающимися углами. Вследствие этого, возникают два слегка отличных друг от друга ракурса для левого и правого глаза, которые располагаются рядом, но не накладываются друг на друга. Другими словами, формируется стереопара, которую затем мозг преобразовывает в одно объёмное изображение.

³ RGB-D: Цветные 3D-изображения в виде облака точек. URL:

<https://www.baslerweb.com/ru/produkty/kamery/3d-kamery/blaze-rgb-d/>

⁴ Форматы 3D фото и видео. Анаглиф. URL: <https://vokrug3d.ru/3d-foto/formaty-3d-izobrazhenij.html>

Соответственно, необходимо получить стереопару (левое и правое изображения) на основе 2D RGB и карты глубины. Существуют два подхода генерации стереоизображения:

1) Исходное 2D-изображение обрабатывается как центральное изображение, относительно которого создаются два новых стереовида.

2) В качестве ракурса для одного глаза используется исходное изображение, и лишь для второго глаза генерируется смещенное изображение, что позволяет минимизировать затраты на преобразование и обработку.

В анаглифном стереоизображении существует очень важный параметр, называемый горизонтальным сдвигом между аналогичными объектами, расположенными на левом и правом кадрах. Этот сдвиг, характеризующий величину взаимного смещения сопряженных точек в обоих изображениях, называется горизонтальным параллаксом. Именно его величина и параметр знака («+» или «-») определяют формат восприятия объемного изображения.

На рис. 9 представлены примеры стереопар, обладающих различными по знаку значениями горизонтальных параллакса. Существуют также понятие нулевого параллакса $P_r = 0$, когда точки правого и левого ракурсов полностью накладываются друг на друга. В этом случае изображение будет восприниматься в плоскости экрана. Примером проявления нулевого параллакса на рис. 9 служит исходное RGB-изображение Kinect. Однако при отображении объектов, обладающих положительным параллаксом ($P_r > 0$), точки правого (красного) ракурса будут размещаться на экране правее сопряженных точек левого ракурса, и, соответственно, объект будет наблюдаться за плоскостью экрана (углубленным в него). И, наоборот, точки правого изображения у объектов, обладающих отрицательным параллаксом ($P_r < 0$), будут располагаться левее сопряженных точек левого (голубого) изображения, а объекты будут восприниматься перед экраном (выпуклая картинка). Однако необходимо учитывать, что для изображений, отображаемых в предэкранных и заэкранных областях, должны обязательно соблюдаться предельно допустимые значения горизонтальных параллакса.

Рис. 9. Процесс получения стереопары с отрицательным и положительным параллаксом

Кроме того, если наличие горизонтального параллакса является важным и неотъемлемым условием при формировании 3D изображения, то существование вертикального параллакса в стереографическом изображении абсолютно недопустимо, и он всегда должен равняться нулю $P_B = 0$.

Во время создания стереопары пиксели исходного изображения смещаются влево или вправо в зависимости от карты глубины (значения d), максимального выбранного параллакса (P) и положения поверхности экрана ($d_{\text{норм}}$).

Смещение параллакса вычисляется для каждого пикселя изображения по следующей формуле:

$$P_{\Gamma} = \frac{L}{d}, \quad (1)$$

где d – значение глубины для отдельно выбранного пикселя (расстояние от плоскости камеры до пикселя объекта), L – среднее межзрачковое расстояние, P – значение параллакса.

Соответственно, чем дальше расположен предмет, тем меньше его параллактическое смещение (степень объемности фигуры). Значение параллакса изменяет свой знак в зависимости от параметра $d_{\text{норм}}$ (значение, при котором наблюдается нулевой параллакс – отсутствует смещение), заложенного в программе. Процесс идентификации положительного и отрицательного параллакса показан на рис. 10.

Рис. 10. Процесс определения характера параллакса, применяемого к каждому пикселю

Данный алгоритм состоит из несколько шагов:

1) Программа сопоставляет каждый RGB-пиксель и принадлежащее ему значение глубины.

2) Программа вычисляет из формулы (1) значение параллакса P_r (величину смещения) для каждого пикселя.

3) Каждое значение глубины d сопоставляется с нормированным значением $d_{\text{норм}}$, которое заложено в программе и характеризует нулевой параллакс (изображение находится на плоскости экрана).

4) Если значение глубины пикселя d равняется нормированному значению $d_{\text{норм}}$, то параллакс остается равным нулю ($P_r=0$) и правое (красное) изображение не смещается.

5) Если значение d меньше $d_{\text{норм}}$, то параллакс принимает отрицательное значение ($-P_r$) и правое (красное) изображение смещается влево.

6) Если значение d больше $d_{\text{норм}}$, то параллакс принимает положительное значение ($+P_r$) и правое (красное) изображение смещается вправо по отношению к голубому ракурсу.

Пример формирования анаглифного стереоизображения для трехмерной визуализации представлен на рис. 11, на котором изображены исходное 2D изображение при освещении белым светом (рис. 11 (а)), карта глубины (рис. 11 (б)) и анаглифное стереоизображение, созданное описанным выше способом (рис. 11 (в)). Трехмерный эффект можно увидеть при наблюдении через очки с красно-голубыми фильтрами.

Рис. 11. Процесс формирования анаглифного стереоизображения на основе данных глубины

Заключение

В данной статье был осуществлен сравнительный анализ нескольких наиболее распространенных на сегодняшний день устройств RGB-D, а именно Microsoft Kinect v1/v2 и Intel RealSense F200. В ходе выполнения исследования было определено, что качество данных RGB-D, получаемых устройством Kinect v2, значительно превосходит показатели RealSense F200 и Kinect v1. Выявлено, что Kinect v2 осуществляет захват данных с меньшим количеством зашумленных точек и более плотными показателями глубины, по отношению к остальным устройствам. Этот фактор является безусловным плюсом в применении Kinect v2 при работе с анаглифами.

В случае необходимости захвата трехмерной поверхности объектов на открытом воздухе, Kinect v2 также превосходит F200 и Kinect v1 по качеству и точ-

ности сформированной карты глубины. В дополнение к этому, Kinect второй версии способен регистрировать depth-значения однородно текстурированных поверхностей, чем не могут похвастаться остальные рассматриваемые устройства. В условиях помещения Kinect v2 также выигрывает в четкости и детализированности сформированной 3D модели, что также является весомым преимуществом при работе с трехмерной графикой.

Кроме того, результаты проведенного исследования показывают, что анаглифные стереоизображения успешно генерируются с помощью одного 2D RGB-изображения и информации о глубине, что указывает на достоверность описанного в статье метода формирования стереопары.

В ходе исследования выявлено, что визуальное качество анаглифа пропорционально зависит от точности предоставляемой информации о глубине. Из этого можно сделать вывод, что при работе с голографией необходимо использовать устройства с наилучшими параметрами и разрешениями RGB и Depth-камер.

Исследование выполнено в рамках исполнения ПНИ по государственному заданию СПбГУТ на 2021 год.

Литература

1. Кучерявый А. Е., Киричек Р. В., Маколкина М. А., Парамонов А. И., Дунайцев Р. А., Пирмагомедов Р. Я., Бородин А. С., Владыко А. Г., Мутханна А. С. А., Выборнова А. И., Владимиров С. С., Гришин И. В. Новые перспективы научных исследований в области сетей связи на 2021–2024 годы // Информационные технологии и телекоммуникации. 2020. Том 8. №3. С. 1–19. DOI 10.31854/2307-1303-2020-8-3-1-19.
2. Кучерявый А. Е., Бородин А. С., Киричек Р. В. Сети связи 2030 // Электросвязь. 2018. № 11. С. 52–55.
3. Галактионов М. А., Маколкина М. А., Киричек Р. В. Обзор протоколов сетей связи шестого поколения и сетей 2030 // 76-я научно-техническая конференция Санкт-Петербургского НТО РЭС им. А.С. Попова, посвященная Дню радио: сб. докладов; СПбГЭТУ «ЛЭТИ». Санкт-Петербург. 2021. С. 183–186.
4. Ларионов П. Хроники Kinect [Электронный ресурс]. URL: <https://xboxunion.ru/article/hroniki-kinect-chast-1-rozhdenie/> (дата обращения 17.10.21).
5. Антонов А. Сенсор Microsoft Kinect [Электронный ресурс]. URL: <http://robotosha.ru/robotics/sensor-microsoft-kinect.html> (дата обращения 19.10.21).
6. Хисматуллина В. Структурированный свет в Kinect [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/strukturirovannyi-svet-v-kinect> (дата обращения 20.10.21).
7. P. Henry, M. Krainin, E. Herbst, et al. "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor International Journal of Robotics Research, vol. 31, p. 647–663, May 2014.
8. Lj Shantj, Zhou Kang, Mu Qj, et al. TOF camera real-time high-precision depth error compensation method [J]. Infrared and laser engineering, 2019, 48 (12): 1213004–1213004 (10).
9. Lachat E, Macher H, Landes T, et al. Assessment and calibration of an RGB-D camera (Kinect v2 Sensor) towards a potential use for close range 3D modeling [J]. Remote Sensing, 2015, 7 (10):13070–13097.
10. Попов Д., Демидов Д., Зотов В. Обзор технологии Intel RealSense [Электронный ресурс]. URL: <https://cyberleninka.ru/article/n/obzor-tehnologii-intel-realsense> (дата обращения 21.10.21).
11. L. C. Chin, S. N. Basah, S. Yaacob, "Accuracy and Reliability of Optimum Distance for High Performance Kinect Sensor", Biomedical Engineering, IEEE 2nd International Conference, Penang, 2015, pp. 1–7.
12. M. S. A. Manap, R. Sahak, A. Zabidi, I. Yassin, "Object Detection using Depth Information from Kinect Sensor", Signal Processing and its application, IEEE 11th International Colloquium, Malasiya, 2015, pp. 160–163.

13. Диковицкий В., Менькова А. Повышение качества визуализации стереоизображений за счёт изменения угла конвергенции [Электронный ресурс]. URL: <https://clck.ru/Ygaxx> (дата обращения 28.10.21).

14. Потыльцева П. Стереокинематограф с точки зрения физики [Электронный ресурс]. URL: <https://school-science.ru/2/11/29701> (дата обращения 02.11.21).

References

1. Koucheryavy A., Kirichek R., Makolkina M., Paramonov A., Dunaytsev R., Pirmagomedov R., Borodin A., Vladyko A., Muthanna A., Vybornova A., Vladimirov S., Grishin I.: Novel Prospects of Scientific Research in the Field of Telecommunications Networks for the 2021–2024 Years // *Telecom IT*. 2020. Vol. 8. Iss. 3. pp. 1–19 (in Russian). DOI 10.31854/2307-1303-2020-8-3-1-19.
2. Koucheryavy A. Eu., Borodin A. S., Kirichek R. V. Network 2030 // *Electrosvyaz*. 2018. No 11. pp. 52–55.
3. Galaktionov M. A., Makolkina M. A., Kirichek R. V. Overview of the protocols of the sixth generation communication networks or networks 2030 // *76 SPbNTORES*. Saint-Petersburg. 2021. pp. 183–186.
4. Larionov P. Hroniki Kinect [Elektronnyj resurs]. URL: <https://xboxunion.ru/article/hroniki-kinect-chast-1-rozhdenie/> (data obrashcheniya 17.10.21).
5. Antonov A. Sensor Microsofr Kinect [Elektronnyj resurs]. URL: <http://robotosha.ru/robotics/sensor-microsoft-kinect.html> (data obrashcheniya 19.10.21).
6. Hismatullina V. Strukturirovannyj svet v Kinect [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/strukturirovannyj-svet-v-kinect> (data obrashcheniya 20.10.21).
7. P. Henry, M. Krainin, E. Herbst, et al. "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor International Journal of Robotics Research, vol. 31, p. 647–663, May 2014.
8. Lj Shantj, Zhou Kang, Mu Qj, et al. TOF camera real-time high-precision depth error compensation method [J]. *Infrared and laser engineering*, 2019, 48 (12): 1213004–1213004 (10).
9. Lachat E, Macher H, Landes T, et al. Assessment and calibration of an RGB-D camera (Kinect v2 Sensor) towards a potential use for close range 3D modeling [J]. *Remote Sensing*, 2015, 7 (10):13070–13097.
10. Popov D., Demidov D., Zotov V. Obzor tekhnologii Intel RealSense [Elektronnyj resurs]. URL: <https://cyberleninka.ru/article/n/obzor-tehnologii-intel-realsense> (data obrashcheniya 21.10.21).
11. L. C. Chin, S. N. Basah, S. Yaacob, "Accuracy and Reliability of Optimum Distance for High Performance Kinect Sensor", *Biomedical Engineering*, IEEE 2nd International Conference, Penang, 2015, pp. 1–7.
12. M. S. A. Manap, R. Sahak, A. Zabidi, I. Yassin, "Object Detection using Depth Information from Kinect Sensor", *Signal Processing and its application*, IEEE 11th International Colloquium, Malasiya, 2015, pp. 160–163.
13. Dikovickij V., Men'kova A. Povyshenie kachestva vizualizacii stereioizobrazhenij za schyot izmeneniya ugla konvergencii [Elektronnyj resurs]. URL: <https://clck.ru/Ygaxx> (data obrashcheniya 28.10.21).
14. Potyl'ceva P. Stereokinematograf s točki zreniya fiziki [Elektronnyj resurs]. URL: <https://school-science.ru/2/11/29701> (data obrashcheniya 02.11.21).

Панков Богдан Олегович

магистрант Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича, bogdan.usman48@yandex.ru

Pankov Bogdan O.

Undergraduate, The Bonch-Bruевич Saint-Petersburg State University of Telecommunications, bogdan.usman48@yandex.ru

Маколкина Мария Александровна

доктор технических наук, доцент, профессор кафедры Санкт-Петербургского государственного университета телекоммуникаций им. проф. М. А. Бонч-Бруевича, makolkina@list.ru

Makolkina Maria A.

Doctor of Engineering Sciences, Docent, Professor at the Department, The Bonch-Bruевич Saint-Petersburg State University of Telecommunications, makolkina@list.ru