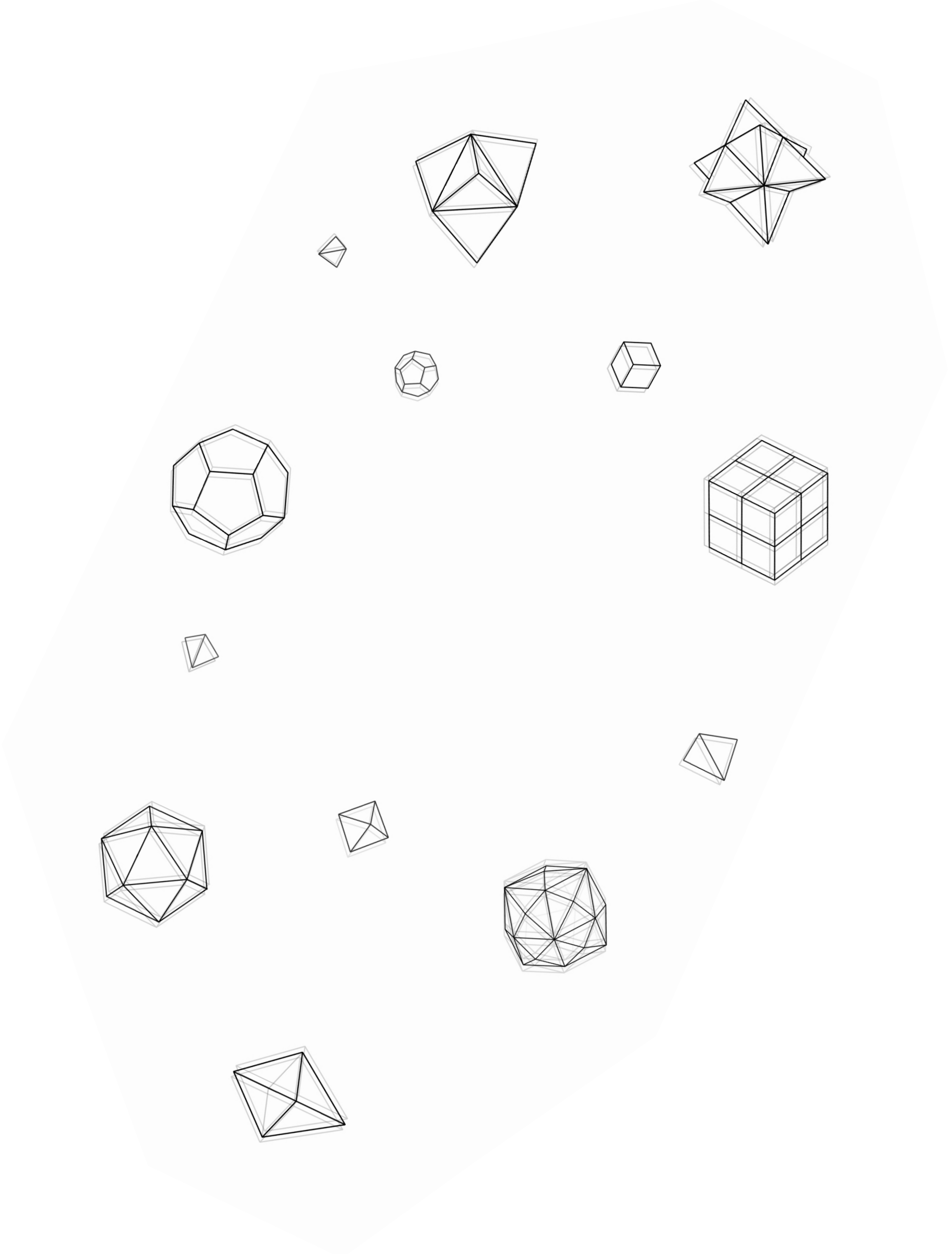
MACHINE VISION

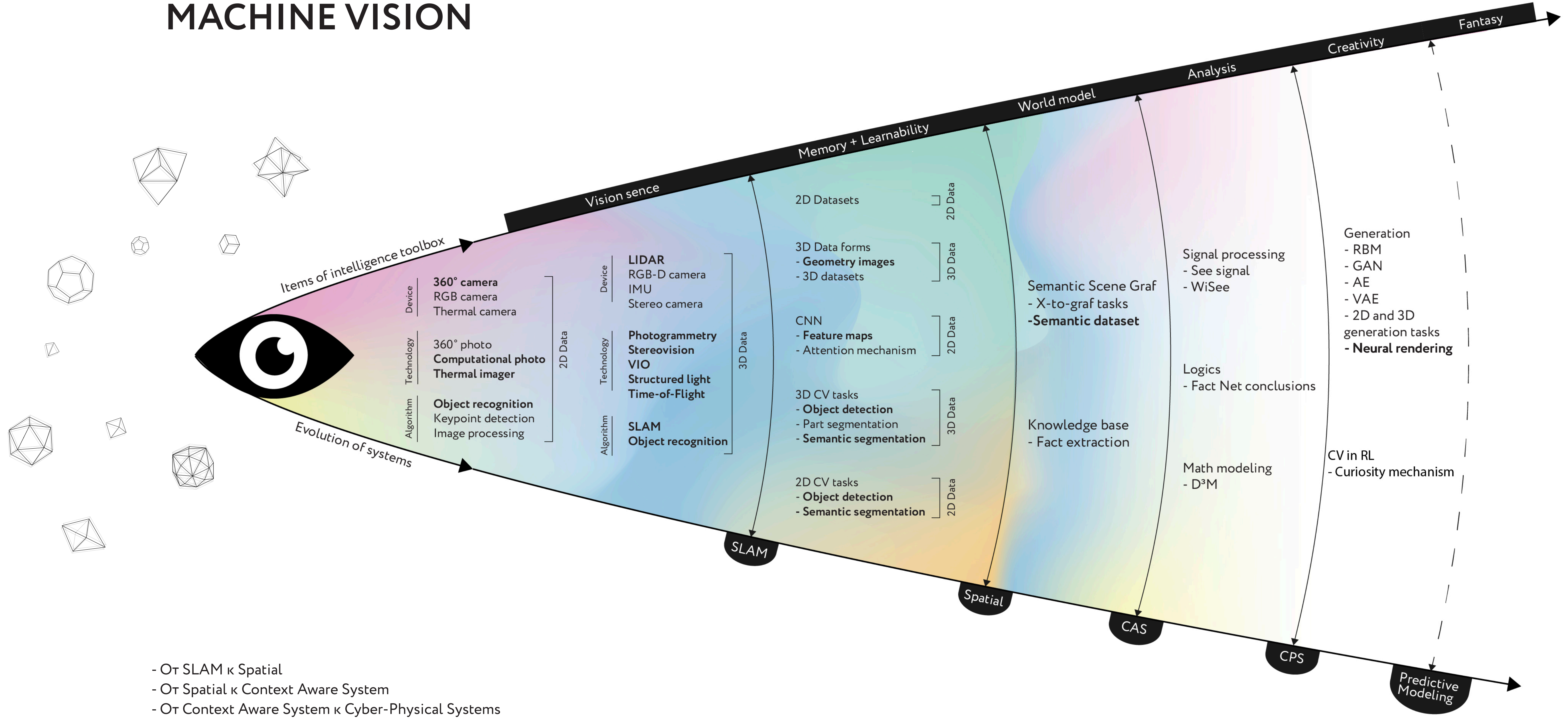
PHYGITALISM

v. 0.2
10.2020
RU



СОДЕРЖАНИЕ

MACHINE VISION



- От SLAM к Spatial
- От Spatial к Context Aware System
- От Context Aware System к Cyber-Physical Systems
- От Cyber-Physical Systems к Predictive Modeling

www.phygitalism.com

Область градиента - Sensor Fusion
Алгоритмы и технологии рассматриваемые в данной версии исследования выделены **жирным** шрифтом

MACHINE VISION (ОПИСАНИЕ СХЕМЫ)

Схема **Machine Vision** - наше видение эволюции систем компьютерного зрения, попытка составить таксономию этих систем. Схема представляет из себя систему концентрических секторов: каждый новый сектор - это новый этап развития систем компьютерного зрения и систем, использующих компьютерное зрение как составную часть.

Внутри каждого сектора отображены рассматриваемые технологии, алгоритмы и устройства, характерные для данного этапа развития технологий. Иногда логическое деление также производится по типам данных (2D или 3D).

Степень развития систем компьютерного зрения мы определяем **двумя базовыми характеристиками (две оси - верхняя и нижняя).**

Первая характеристика (нижняя ось) - это сложность системы, частью которой является компьютерное зрение: сначала компьютерное зрения - это SLAM система, целью которой является позиционирование и восстановление только геометрических характеристик, далее рассматриваются SPATIAL системы, которые способны анализировать более общие пространственные характеристики пространства за счет использования памяти и продвинутых алгоритмов. Следующие ступени чуть более условны в порядке своего следования и расставлены в таком порядке в основном по причине упорядоченности вдоль другой оси. После SPATIAL систем рассматриваются контекстно ориентированные системы (Context Aware Systems / CAS), позволяющие помимо пространственных характеристик восстанавливать семантическую информацию об объектах.

Помимо контекста, с помощью технологий машинного зрения можно делать предсказания о физических характеристиках объектов и о физических характеристиках процессов в которых эти объекты участвуют, в этом смысле, компьютерное зрение становится часть кибер-физических систем (Cyber-physical systems / CPS). В тот момент, когда технологии компьютерного зрения начинают позволять не только анализировать существующие формы, но и создавать новые, происходит естественное вливание данных технологий в системы, которые мы называем термином Predictive modelling (Pr.Mod.). Pr.Mod. системы включают в себя все рассмотренные выше системы и целью их является расширить способности и горизонты человеческого воображения и стать новым лучшим способом взаимодействия между физическим миром (природой), цифровым миром (компьютером) и человеком.

Вторая характеристика (верхняя ось) - это свойства и качества мыслящей системы, составные кирпичики любого интеллектуального агента. Первый кирпичик - это частный сенсорный механизм - зрение (Vision sense). Далее подключается память (Memory), позволяющая осуществлять процесс самообучения (Learnability). К памяти добавляется возможность строить в своем “сознании” модели окружающего мира (World models) и на основе их делать предсказания о процессах, которые протекают в этом мире (Analyses). После составных частей необходимых для анализа, появляются части отвечающие за синтез: способность к творчеству на основе ранее изученной информации (Creativity) и способность к созданию собственных миров со своими правилами и особенностями, внутри которых интеллектуальный агент может представлять и себя (Fantasy).

MACHINE VISION

Данная книга - это структурированное визуальное исследование компьютерного зрения и графики. Мы рассматриваем понятие **vision** в интегрированном подходе: не просто зрение (eyesight), но и видение (vision) и предвидение (foresight), синергию этих понятий.

Таким образом **machine vision** - это, по аналогии с человеческим восприятием, процесс познания машиной окружающего мира, с использованием инструментов интеллекта от зрения до фантазии. Рассматривая компьютерное зрение через такую призму, мы одновременно пытаемся предвидеть как слабый искусственный интеллект может развиться в сильный.

Эта книга, следуя концепции **Art-Science-Business**, рассматривает область компьютерного зрения как предмет исследования одновременно в сферах искусства, науки и бизнеса.

Art
Мы хотели преподнести алгоритмическую эстетику современного компьютерного зрения, как отправной базис для направления в искусстве, которое мы называем Фиджитализм. Фиджитализм или Фиджитал-арт - это направление в искусстве, которое сочетает в себе физические и цифровые взаимодействия / миры, и одной из лучших метафор для объяснения этой связи являются технологии компьютерного зрения и сканирования.

Science
Нашей целью было сформировать схематичное представление того, как развиваются технологии компьютерного зрения для широкого круга читателей: от новичков до специалистов CV/ML. Базовая схема и визуальный способ повествования были составлены таким образом, чтобы можно было, во-первых, окинуть взглядом все существующие на сегодняшний день технологии и системы, связанные с компьютерным зрением, а, во-вторых, попытаться предсказать путь их дальнейшего развития.

Business
Компьютерное зрение - важная составная часть многих технологических систем в современном мире: от приложений на смартфон до систем машинного зрения промышленных роботов. Благодаря знанию технологии, бизнес эффективнее организовывает процессы и улучшает ключевые показатели. Мы хотим, чтобы благодаря данной книги бизнес лучше стал понимать те, открывающиеся возможности, которые дают новые технологии перед ними.

СТРУКТУРА КНИГИ

Название технологии / алгоритма

Краткое описание технологии / алгоритма

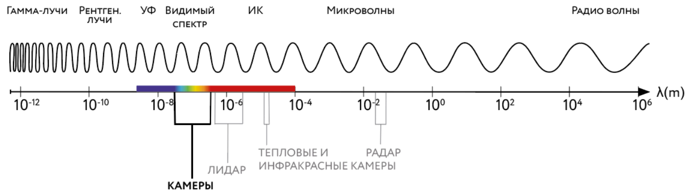
Электромагнитный спектр и зоны его использования сенсорами технологии / алгоритма

Технологические события за определенный промежуток времени, влияющие на развитие технологии / алгоритма

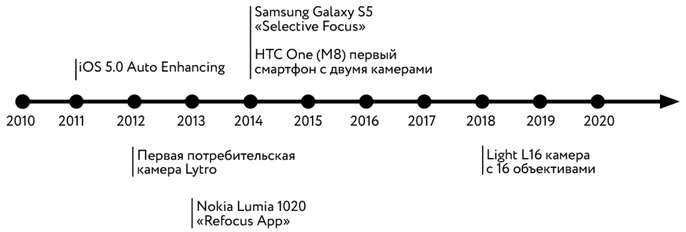
ВЫЧИСЛИТЕЛЬНАЯ ФОТОГРАФИЯ

Набор методов компьютерной визуализации, улучшающих или расширяющих возможности цифровой фотографии, при использовании которых получается обычная фотография, которая не могла технически быть снята на данную камеру традиционным способом.

Изображение 1
Электромагнитный спектр и зоны его использования сенсорами технологии
Длина волны, λ(m)

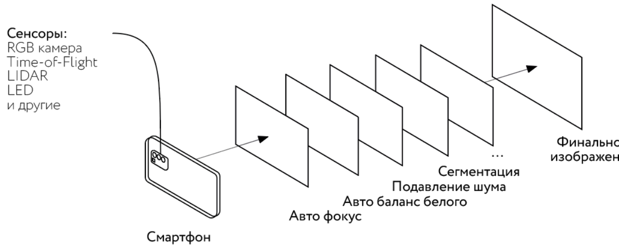


Таймлайн
Технологические события
2010 - 2020

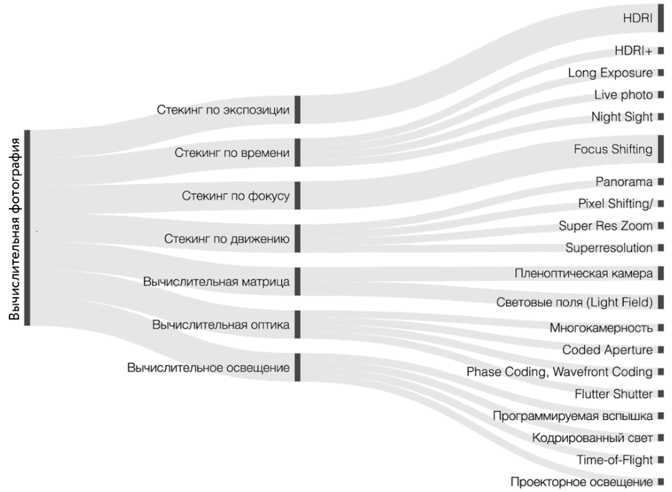


16

Изображение 2
Схема работы вычислительной фотографии



Сенсоры используемые в данной техноогии и упрощенная схема работы технологии / алгоритма

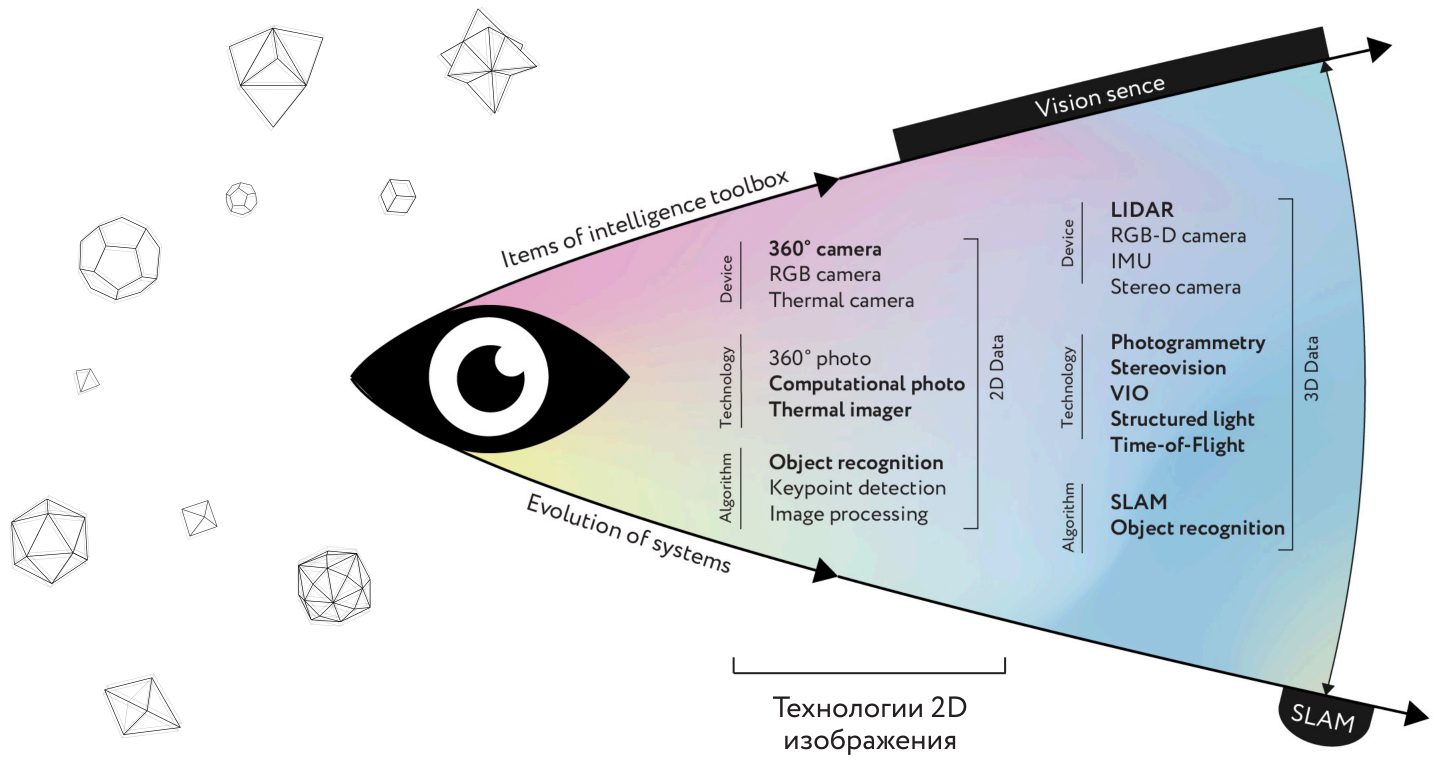


Изображение 3
Основные алгоритмы вычислительной фотографии

Дополнительная информация (таксономия алгоритмов, программного обеспеченья и т.д.)

17

ЗРЕНИЕ

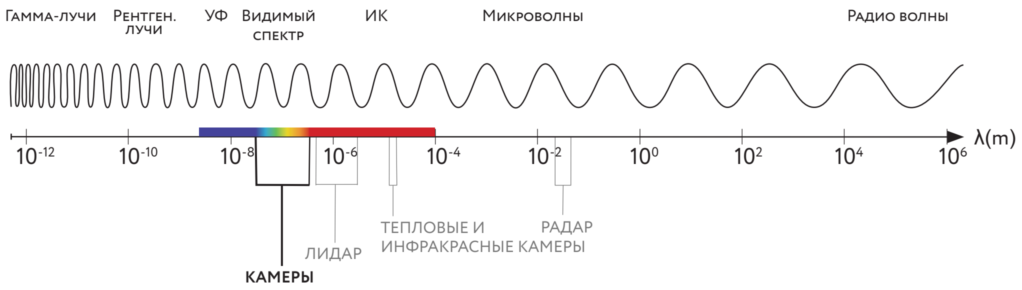


Vision sence
Items of intelligence toolbox
Evolution of systems
Device
360° camera
RGB camera
Thermal camera
Technology
360° photo
Computational photo
Thermal imager
Algorithm
Object recognition
Keypoint detection
Image processing
2D Data
Device
LIDAR
RGB-D camera
IMU
Stereo camera
Technology
Photogrammetry
Stereovision
VIO
Structured light
Time-of-Flight
Algorithm
SLAM
Object recognition
3D Data
SLAM
Технологии 2D изображения

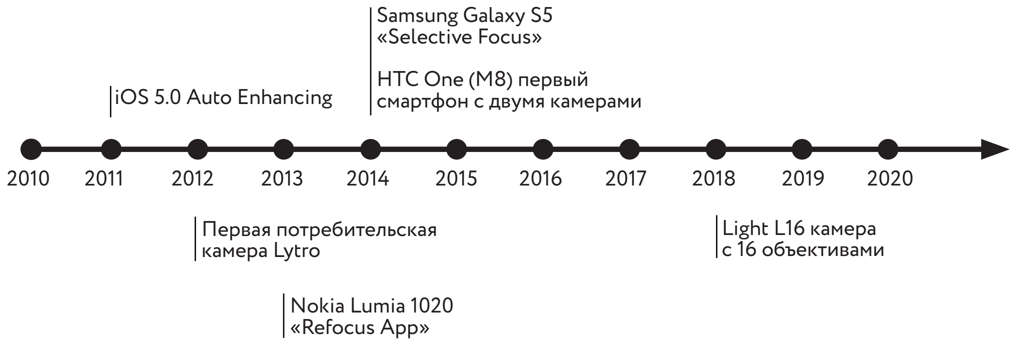
ВЫЧИСЛИТЕЛЬНАЯ ФОТОГРАФИЯ

Набор методов компьютерной визуализации, улучшающих или расширяющих возможности цифровой фотографии, при использовании которых получается обычная фотография, которая не могла технически быть снята на данную камеру традиционным способом.

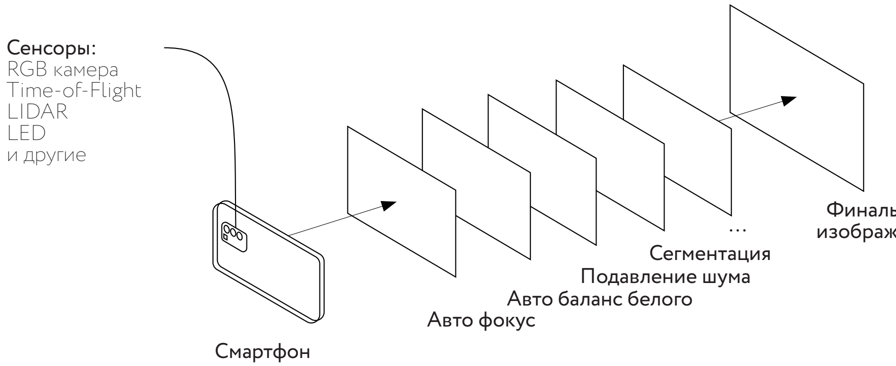
Изображение 1
Электромагнитный спектр и зоны его использования сенсорами технологии
Длина волны, λ(m)

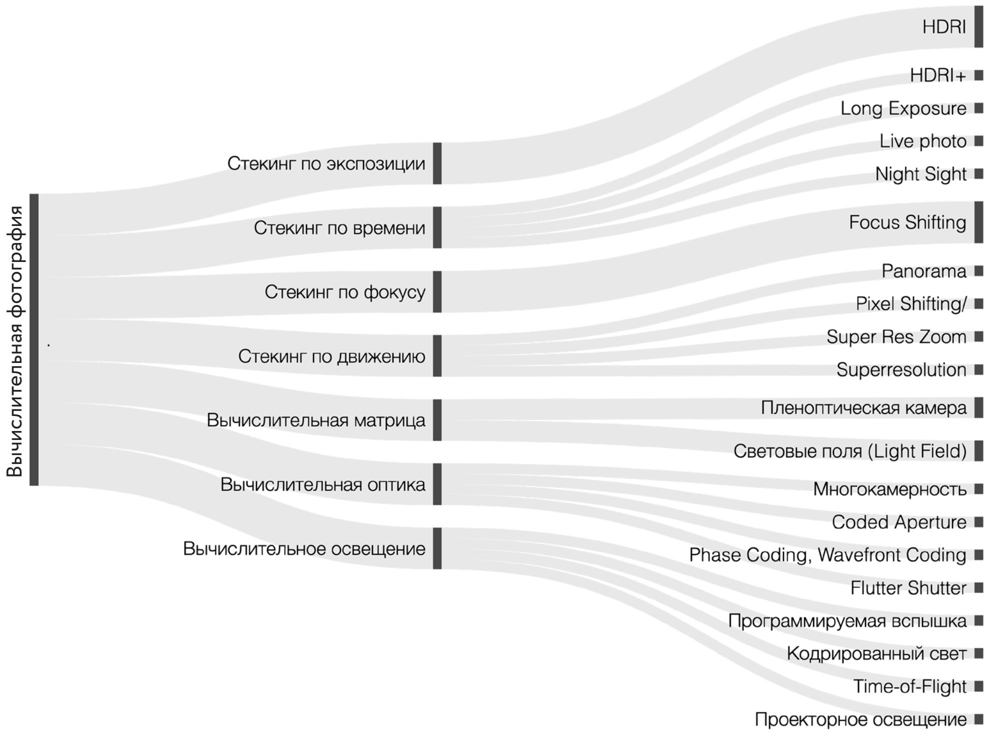


Таймлайн
Технологические события
2010 - 2020



Изображение 2
Схема работы вычислительной фотографии



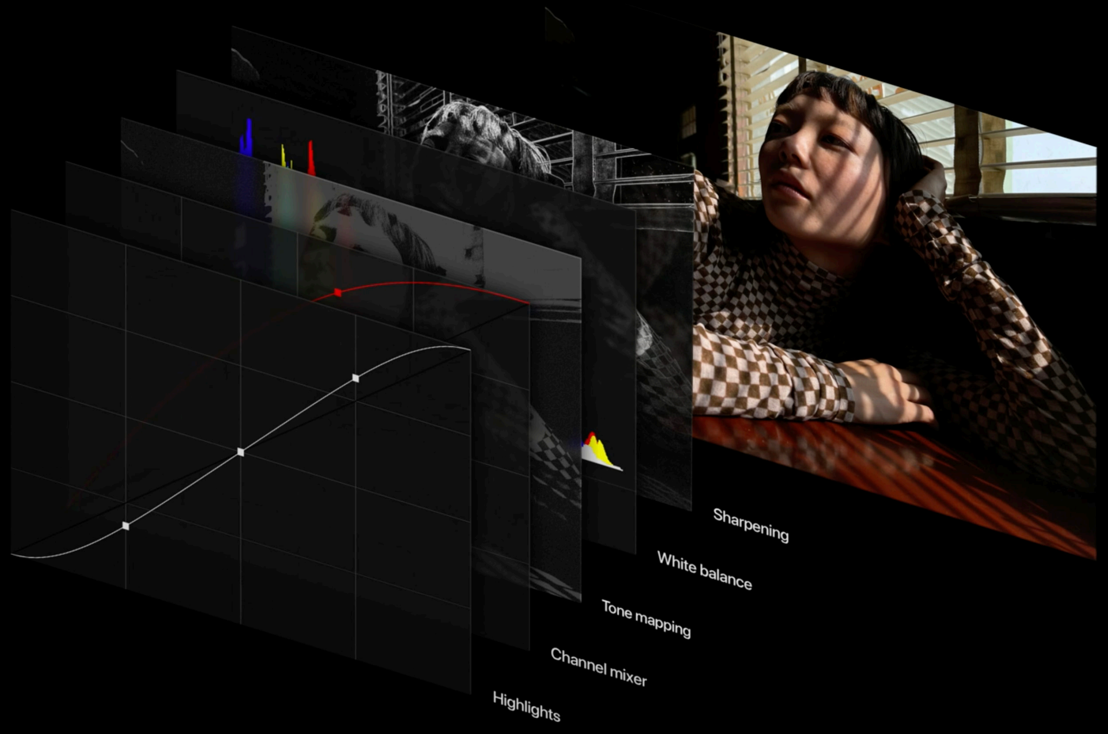


Стекинг - представляет собой метод цифровой обработки изображений, который объединяет несколько изображений, чтобы получить более качественное результирующее изображение.

Изображение 3
Основные алгоритмы вычислительной фотографии

Изображение 4 Презентация Apple возможностей камеры смартфона iPhone XS.

Новый пайплайн Apple, который берет компоненты обработки, выполняемые в CPU, GPU и Neural Engine, и объединяет их в новый файл Deep Image, вычисляемый во время захвата без какой-либо задержки срабатывания затвора.



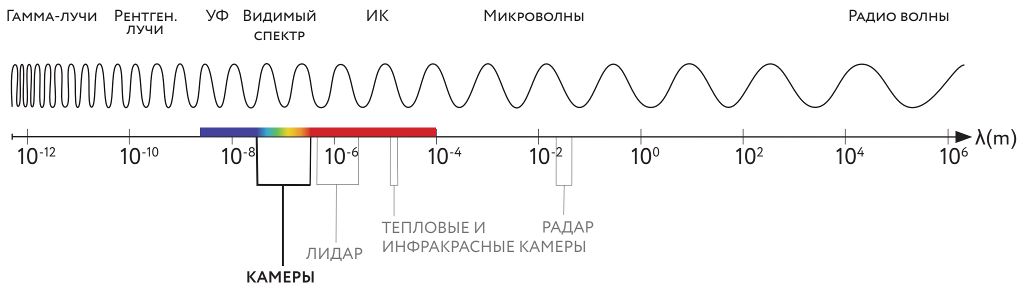
Изображение 5 и 5.1 Презентация Apple возможностей камеры смартфона iPhone 12 Pro и iPhone Pro Max, которые поддерживают формат Apple ProRAW- он объединяет в себе разработанные в Apple технологии мультикадровой съёмки, возможности вычислительной фотографии и удобство формата RAW.

Изображение 6 Apple ProRAW - пример обработки изображения.

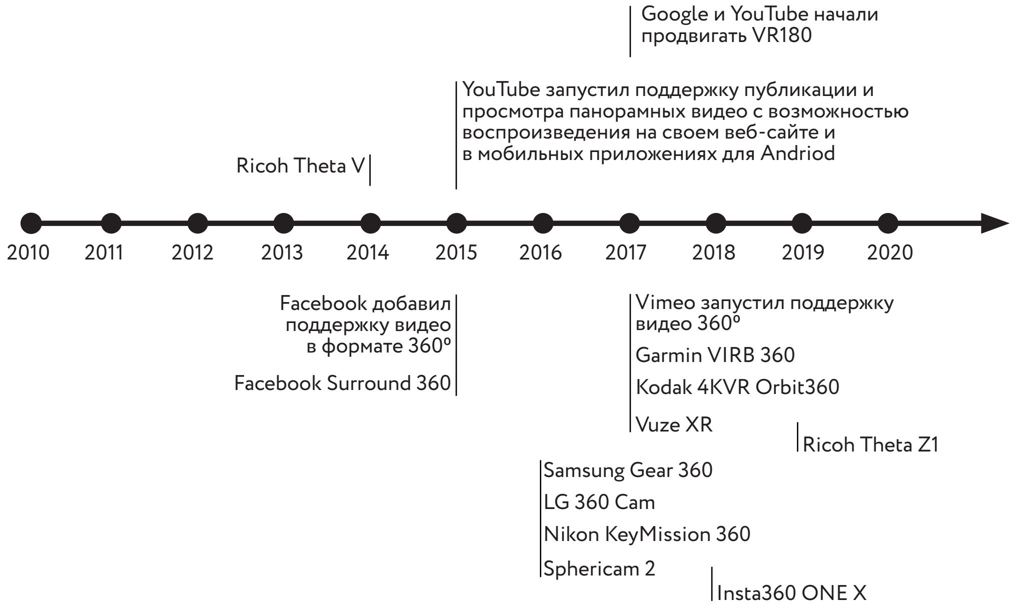
КАМЕРА 360º

В фотографии всенаправленная камера, также известная как камера 360°, представляет собой камеру, поле обзора которой покрывает примерно всю сферу или, по крайней мере, полный круг в горизонтальной плоскости.

Изображение 7
Электромагнитный спектр и зоны его использования сенсорами технологии
Длина волны, λ(m)

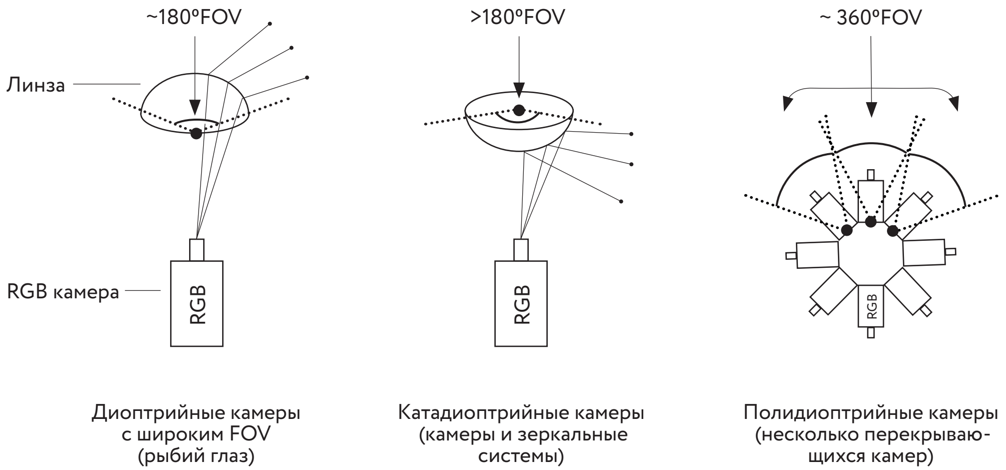


Таймлайн
Технологические события 2010 - 2020

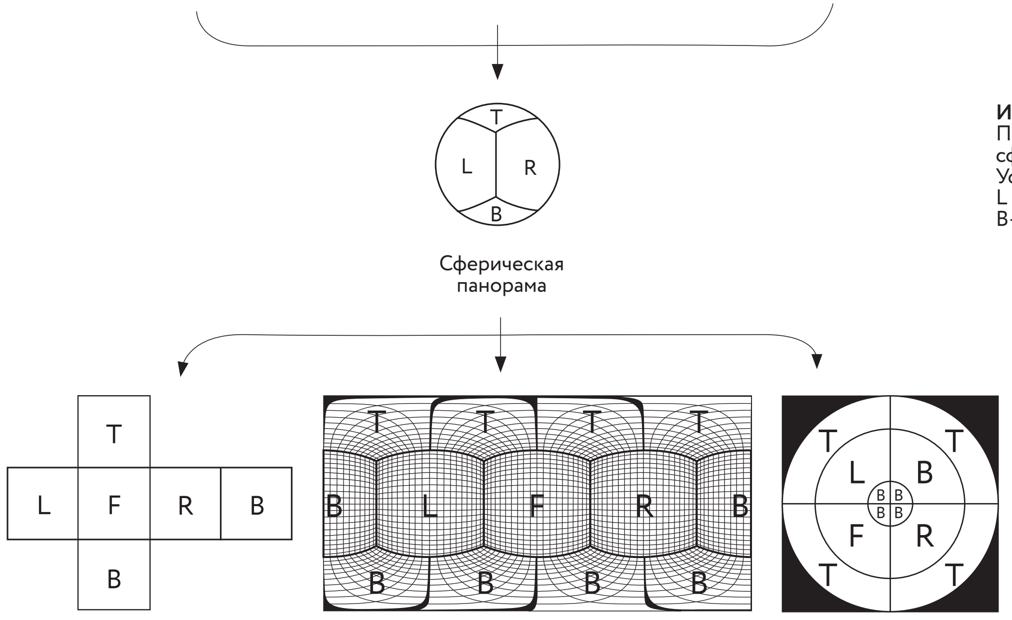


КАМЕРЫ 360º

Изображение 8
Схемы работы камер 360º



Изображение 9
Примеры развертки сферической панорамы.
Условные обозначения:
L - лево, F - фронт, R- право, B- тыл, T - зенит, B - надир

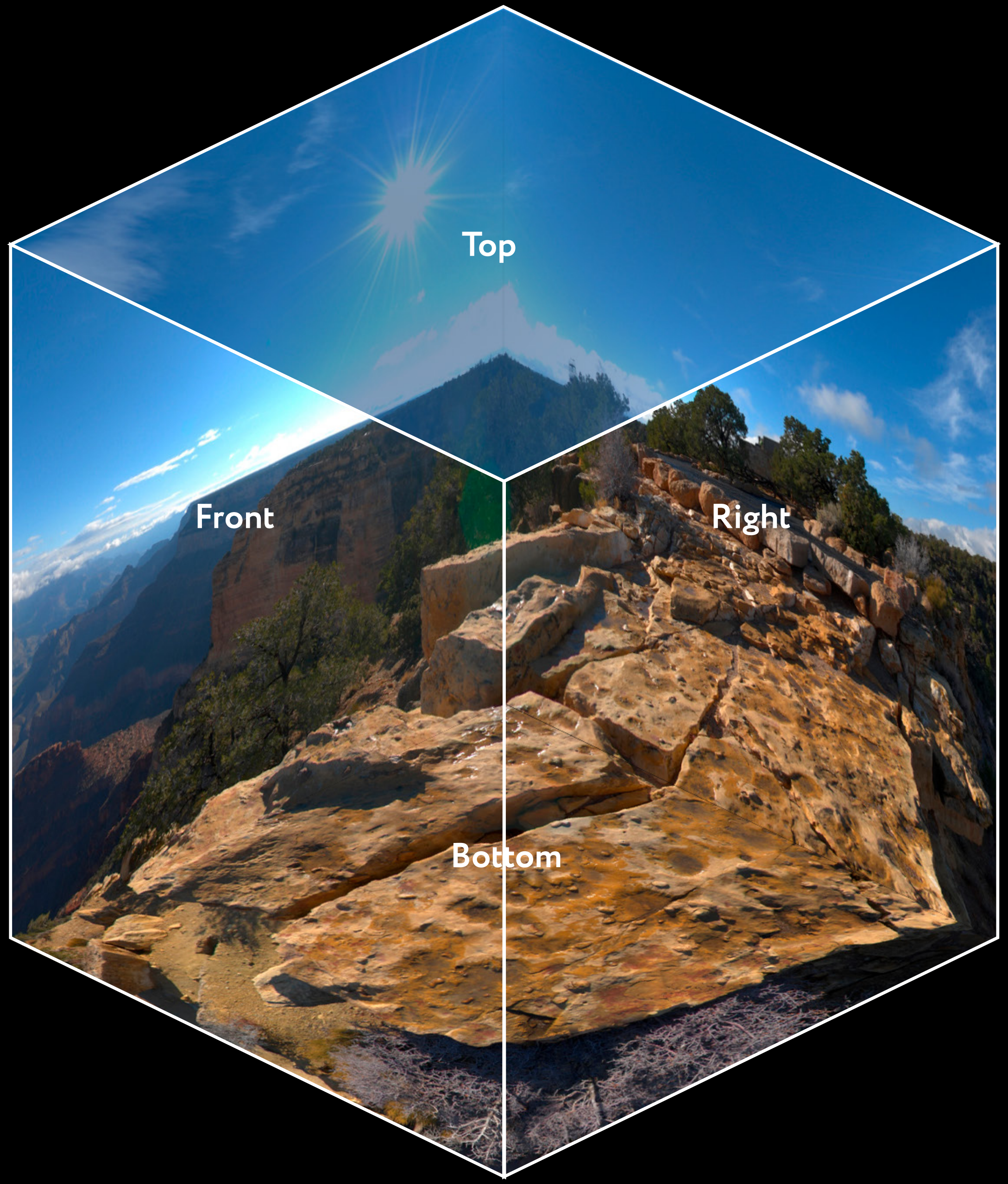


Панорама в кубической проекции

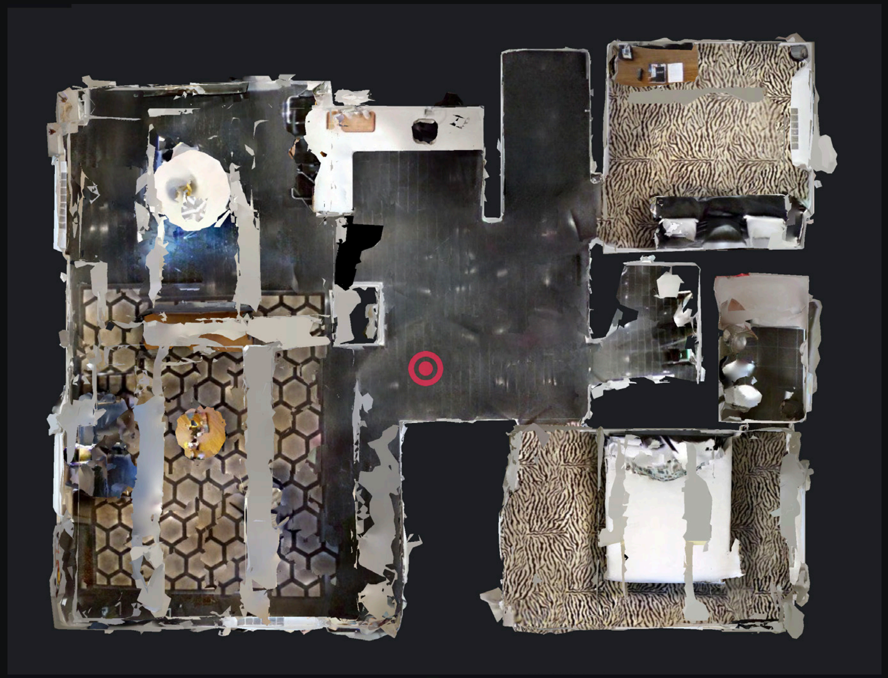
Сферическая панорама в эквидистантной проекции

Панорама "Маленькая планета"

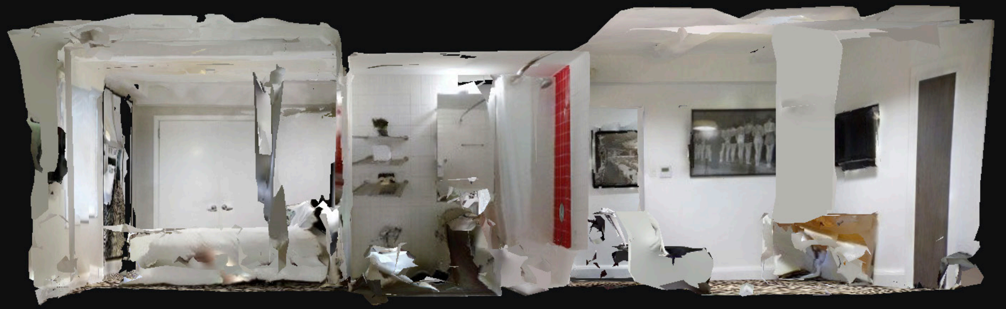
Изображение 10 Сферическая панорама в проекции "маленькая планета", в эквидистантной и кубической проекции.



Изображение 11 Развертка сферической панорамы в кубе.



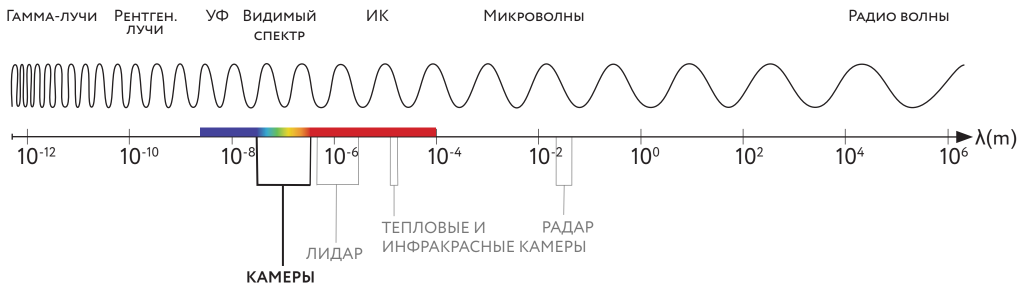
Изображение 12 Matterport предлагает пользователям возможность загружать фотографии в формате 360 ° и превращать их в 3D-модели.



ФОТОГРАММЕТРИЯ

Наука и технология получения надежной информации о физических объектах и окружающей среде посредством процесса записи, измерения и интерпретации фотографических изображений и структур изображений электромагнитного излучения и других явлений.

Изображение 13
Электромагнитный спектр и зоны его использования сенсорами технологии
Длина волны, λ(m)



Таймлайн
Технологические события
1070- 2020

1970 1972 1980 1990 2000 2010 2020

МИССИЯ APOLLO 17
Фредерик Дойл, фото-группа Apollo Orbital Science Photo Team.
14000 точек, фотографии из 4-х разных миссий, сделаны серии карт Луны.

GPS
Новый инструмент позиционирования. Глобальная система определения местоположения (GPS) - это глобальная навигационная спутниковая система (GNSS), которая предоставляет информацию о местоположении и времени при любых погодных условиях, в любом месте на Земле или рядом с ней, где есть беспрепятственная прямая видимость для четырех или более спутников GPS.

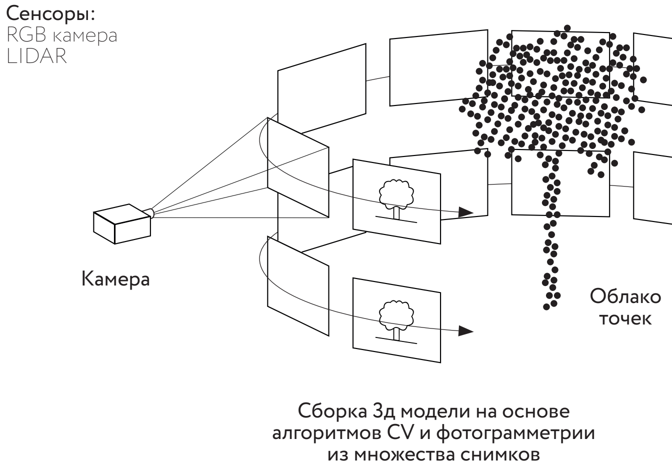
LANDSAT
Программа Landsat - это самая продолжительная программа по получению спутниковых изображений Земли. 23 июля 1972 года был запущен Earth Resources Technological Satellite. В конечном итоге он был переименован в Landsat.

ГИС
Новый инструмент пространственного анализа. Географическая информационная система (ГИС) - это компьютерная система для сбора, хранения, проверки и отображения данных, связанных с позициями на поверхности Земли.

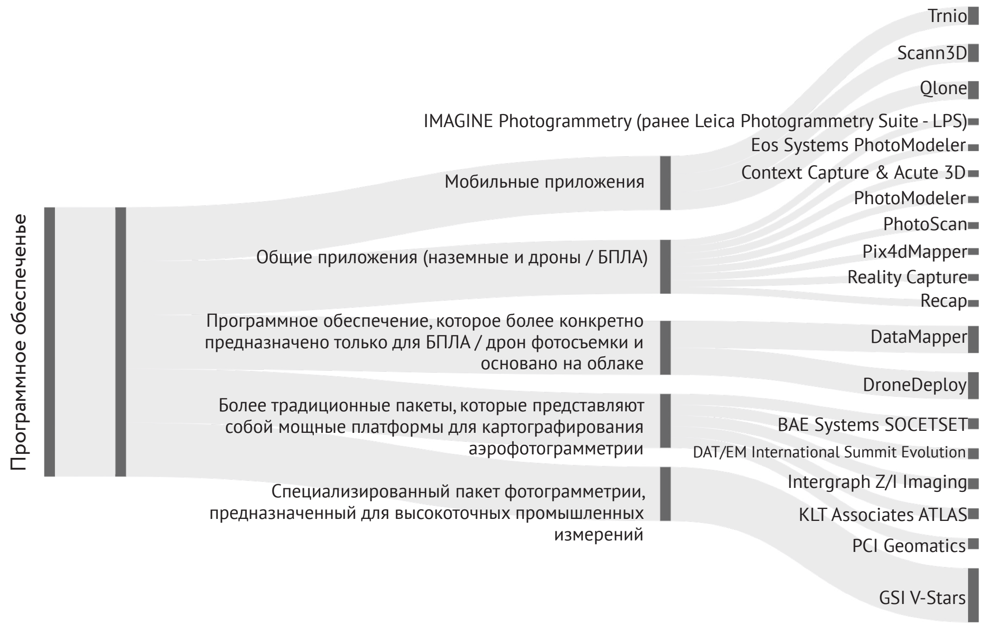
НАБОР ДЛЯ ФОТОГРАММЕТРИИ PYTHON
osm-bundler разрабатывает программное обеспечение с открытым исходным кодом для восстановления трехмерной геометрии из набора фотографий

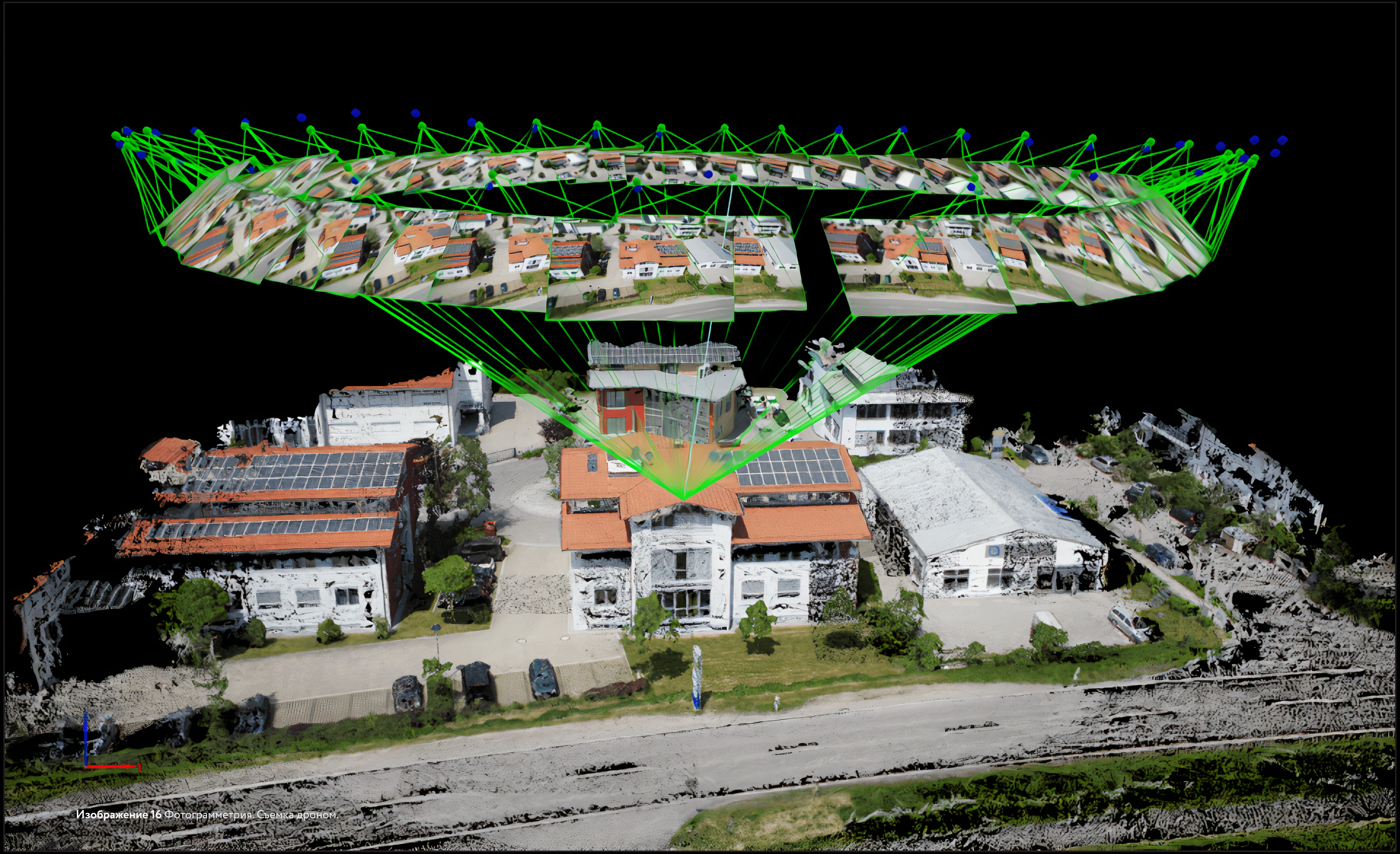
ISPRS
ISP меняет свое название на Международное Общество Фотограмметрии и Дистанционного Зондирования.

Изображение 14
Схема работы фотограмметрии

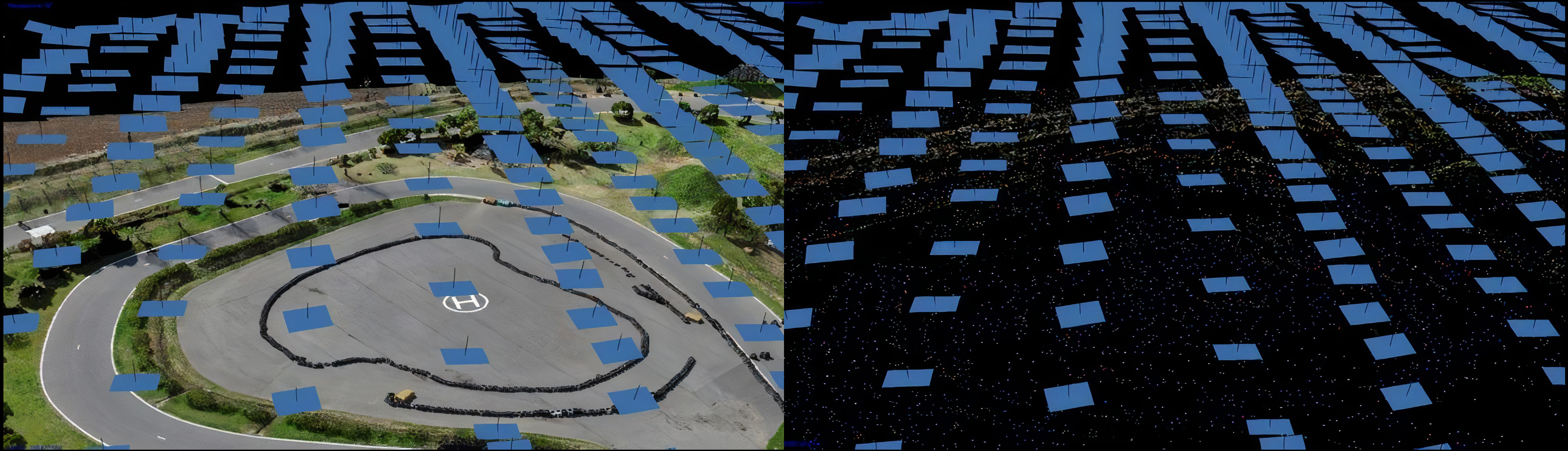


Изображение 15
Виды программного обеспеченья

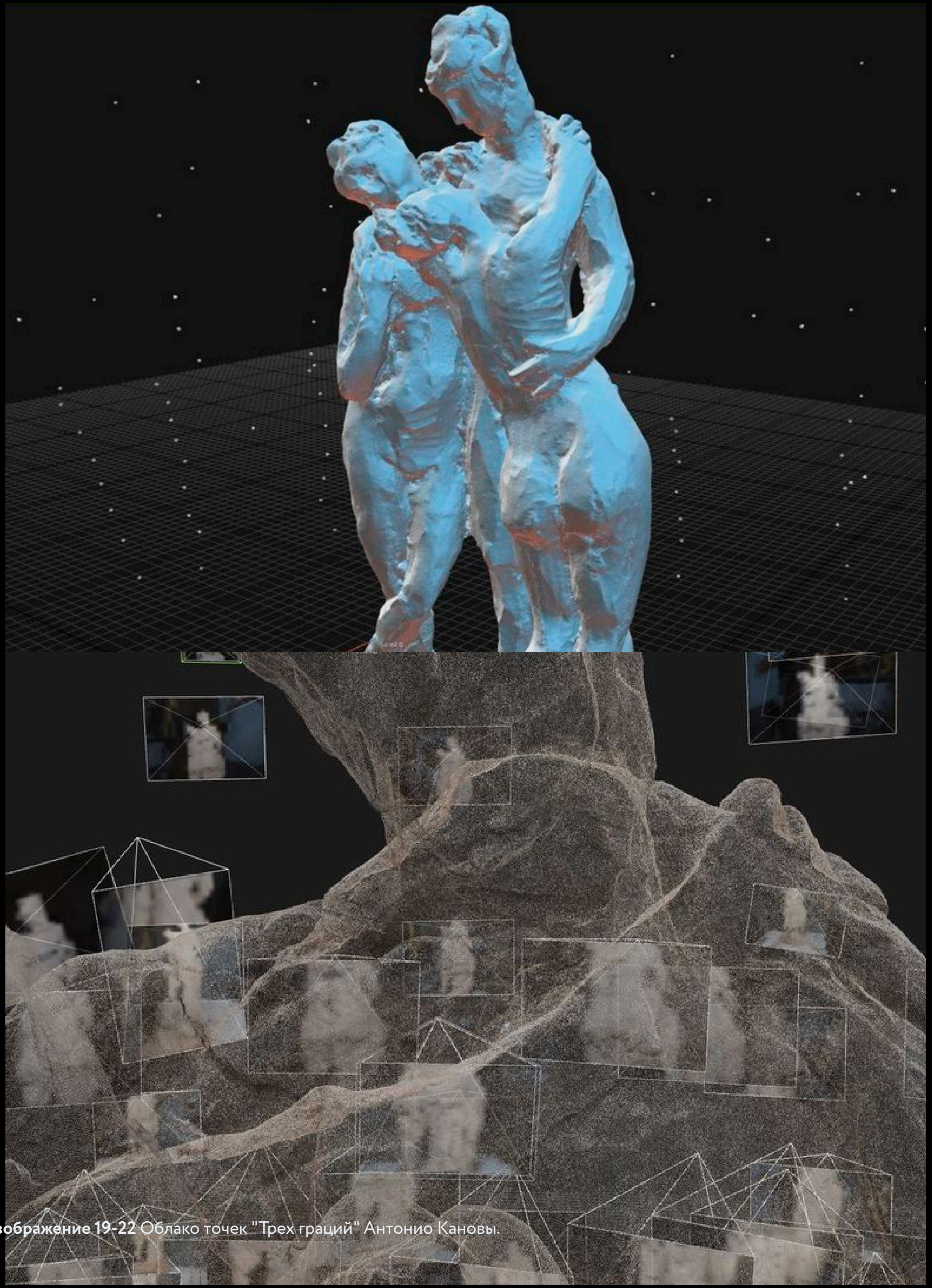


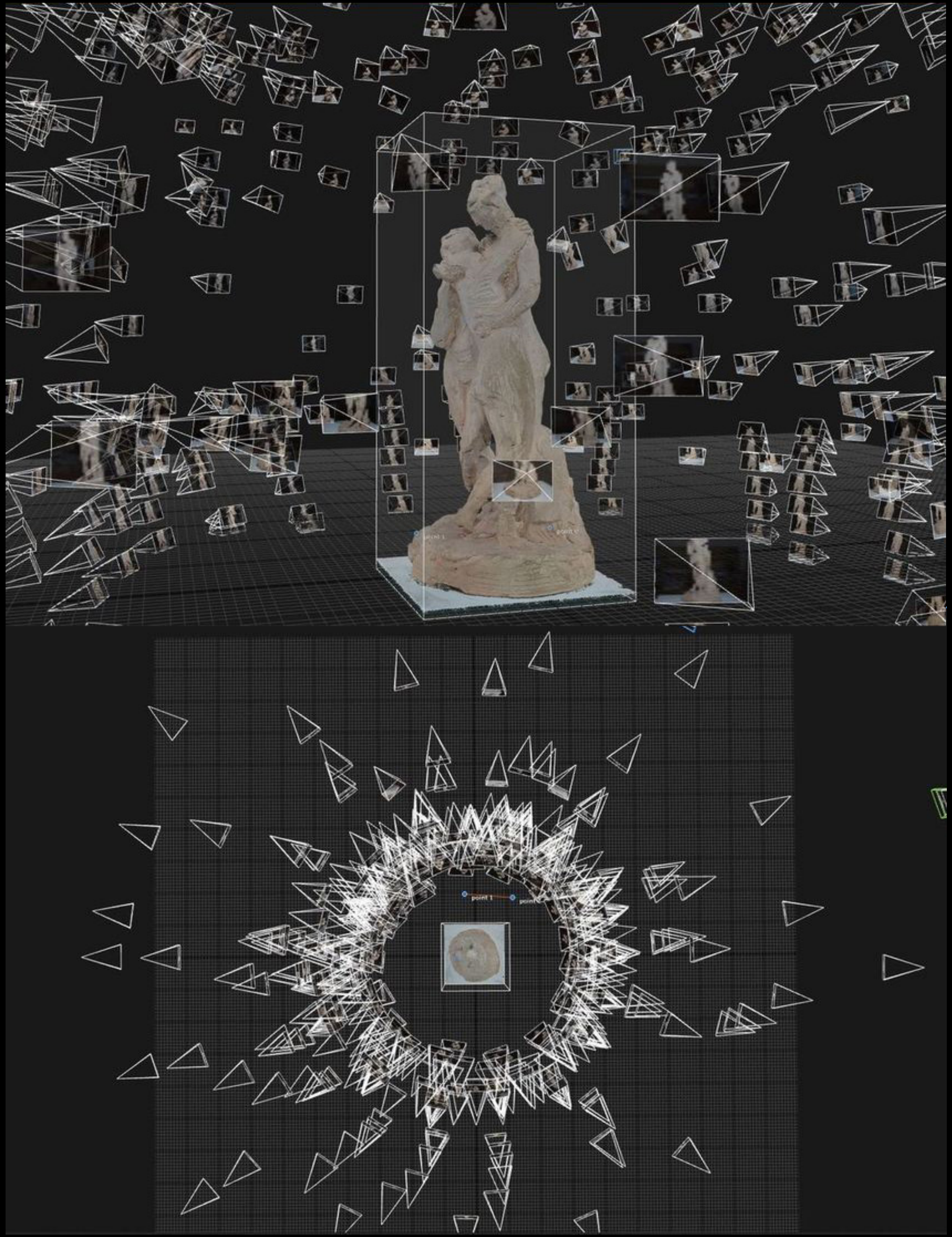


Изображение 16 Фотограмметрия. Съемка дроном.



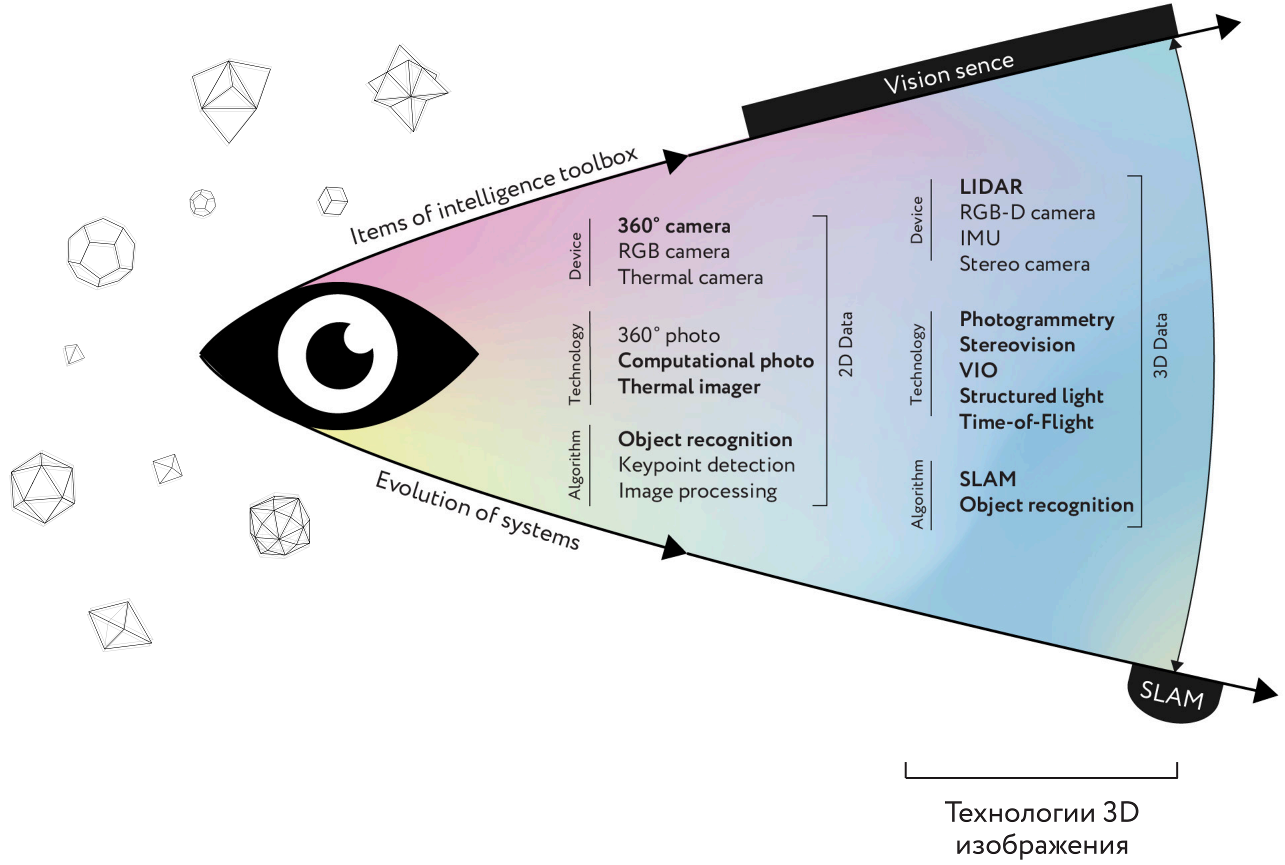
Изображение 17-18 Расчет облака точек из данных фотографии и создание поверхности и изображения.





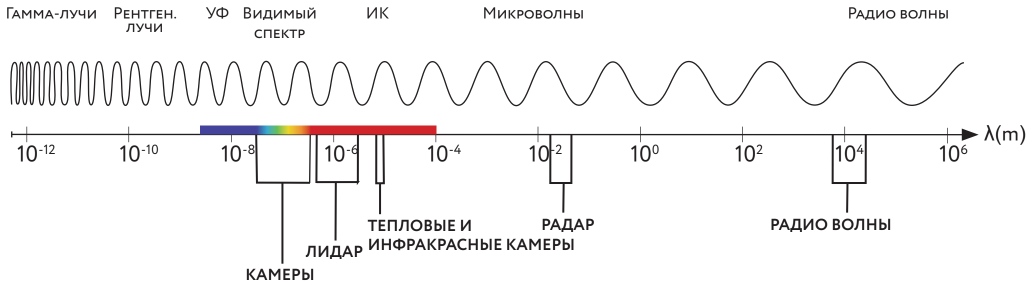
Изображение 19-22 Облако точек "Трех граций" Антонио Кановы.

ЗРЕНИЕ

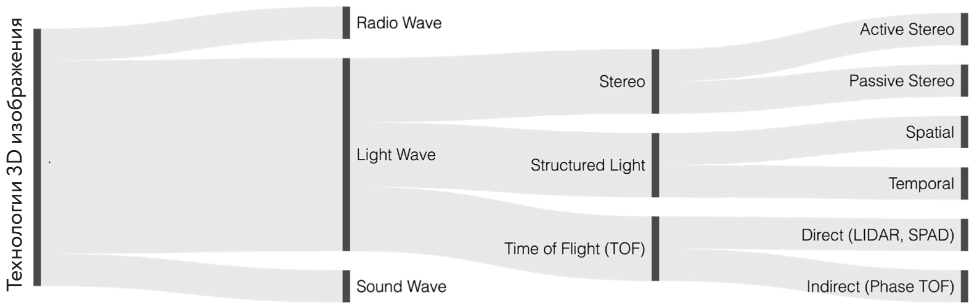


ТЕХНОЛОГИИ 3D ИЗОБРАЖЕНИЯ

Технологии 3D изображения можно широко классифицировать на три области: Стереозрение (Stereovision), Структурированный свет (Structured Light), Time-of-Flight камеры (TOF).



Изображение 23
Электромагнитный спектр и зоны его использования сенсорами 3D технологий Длина волны, λ(m)

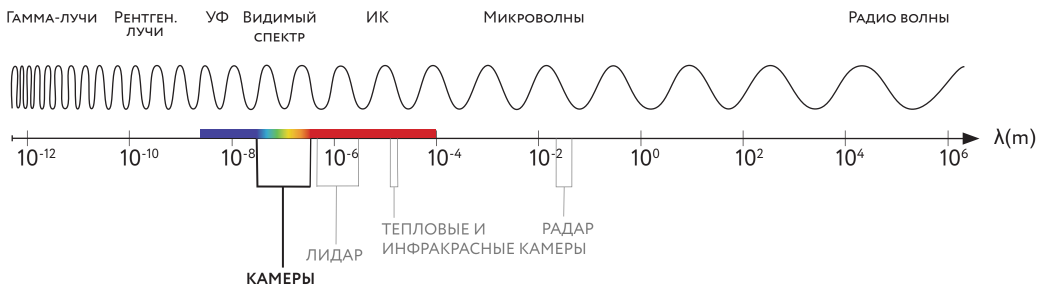


Изображение 24
Классификация технологий 3D изображения

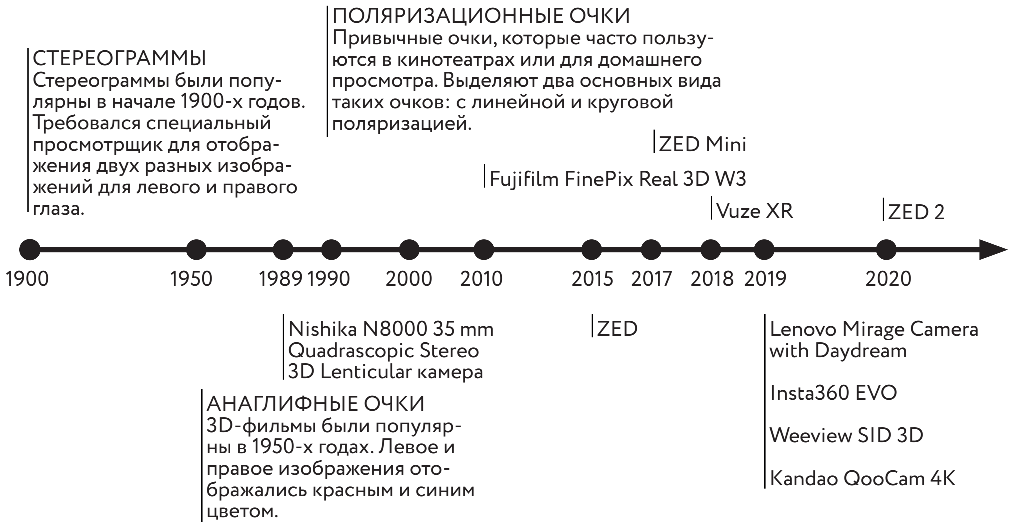
СТЕРЕОЗРЕНИЕ

Это извлечение трехмерной информации из цифровых изображений. Сравнивая информацию о сцене с двух точек обзора, можно извлечь трехмерную информацию, исследуя относительное положение объектов на двух панелях. Это похоже на биологический процесс стереопсис — субъективное ощущение глубины пространства при бинокулярном зрении, обусловленное сетчаточной горизонтальной диспарантностью.

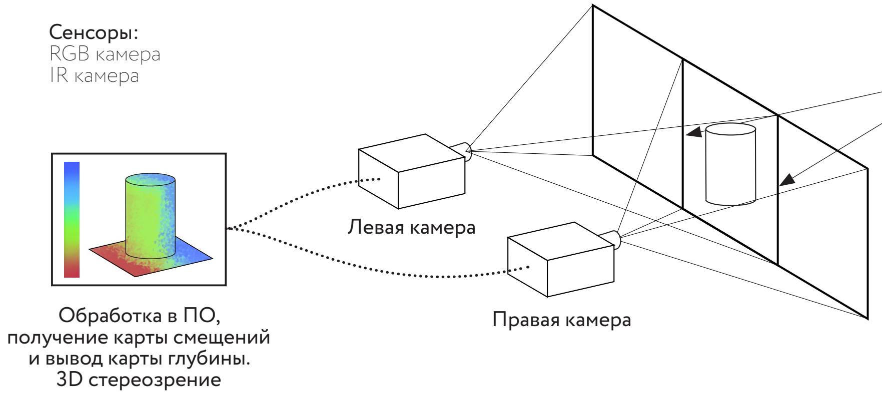
Изображение 25
Электромагнитный спектр и зоны его использования сенсорами технологии
Длина волны, λ(m)

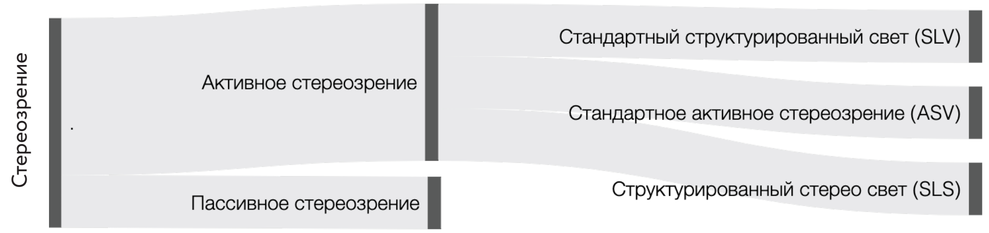


Таймлайн
Технологические события 1900- 2020



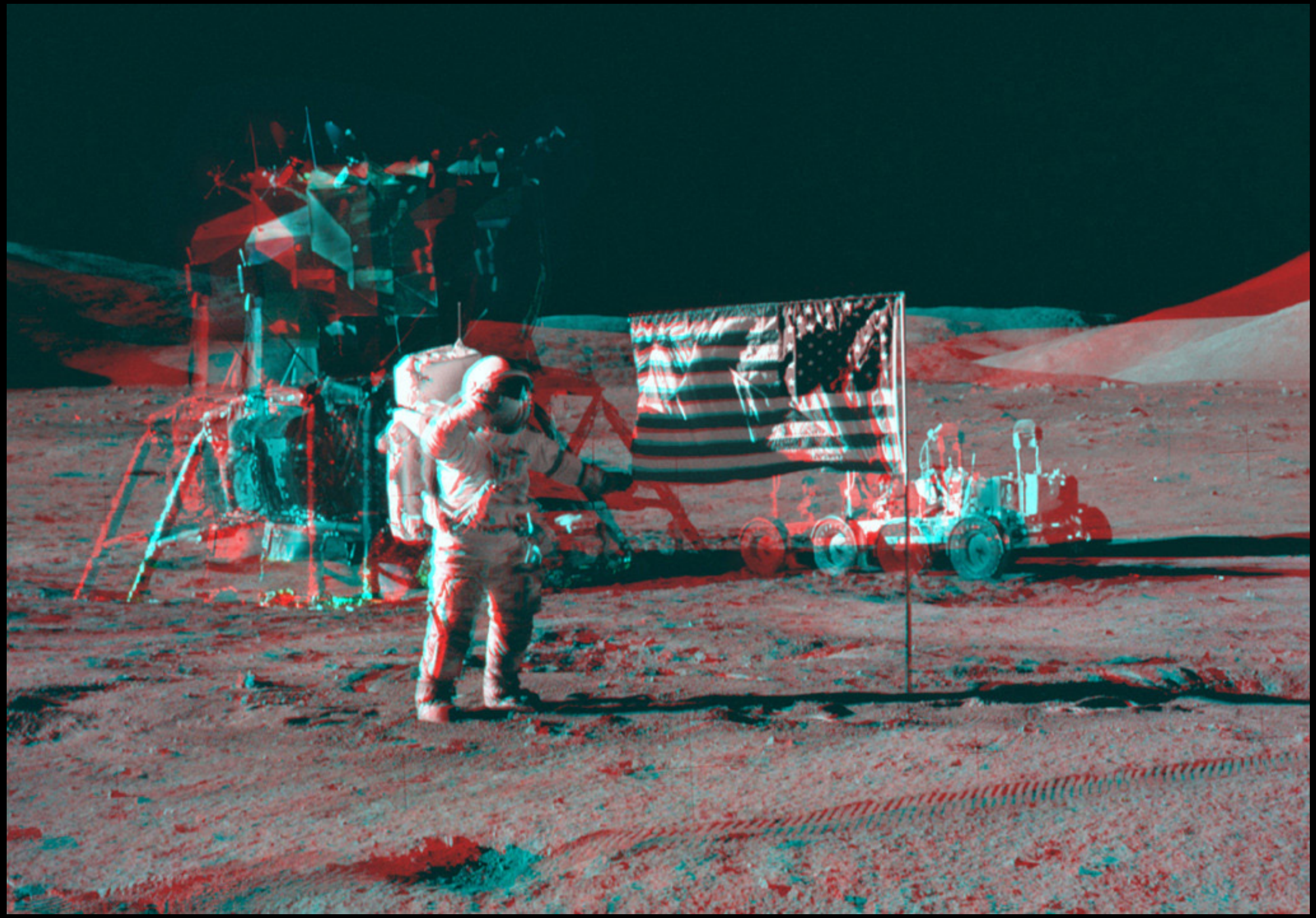
Изображение 26
Схема работы Стерео зрения



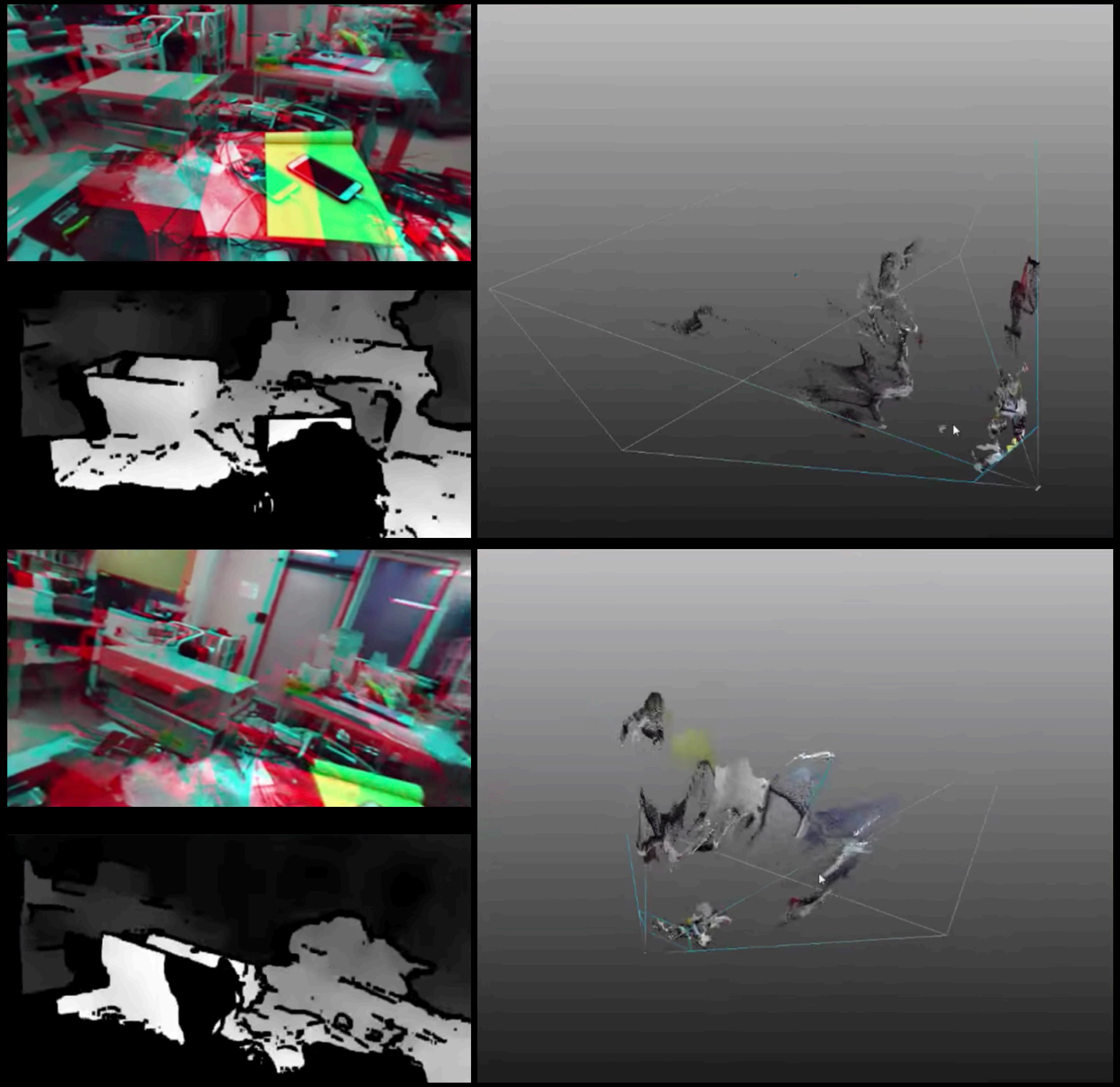


Изображение 27
Виды Стереозрения

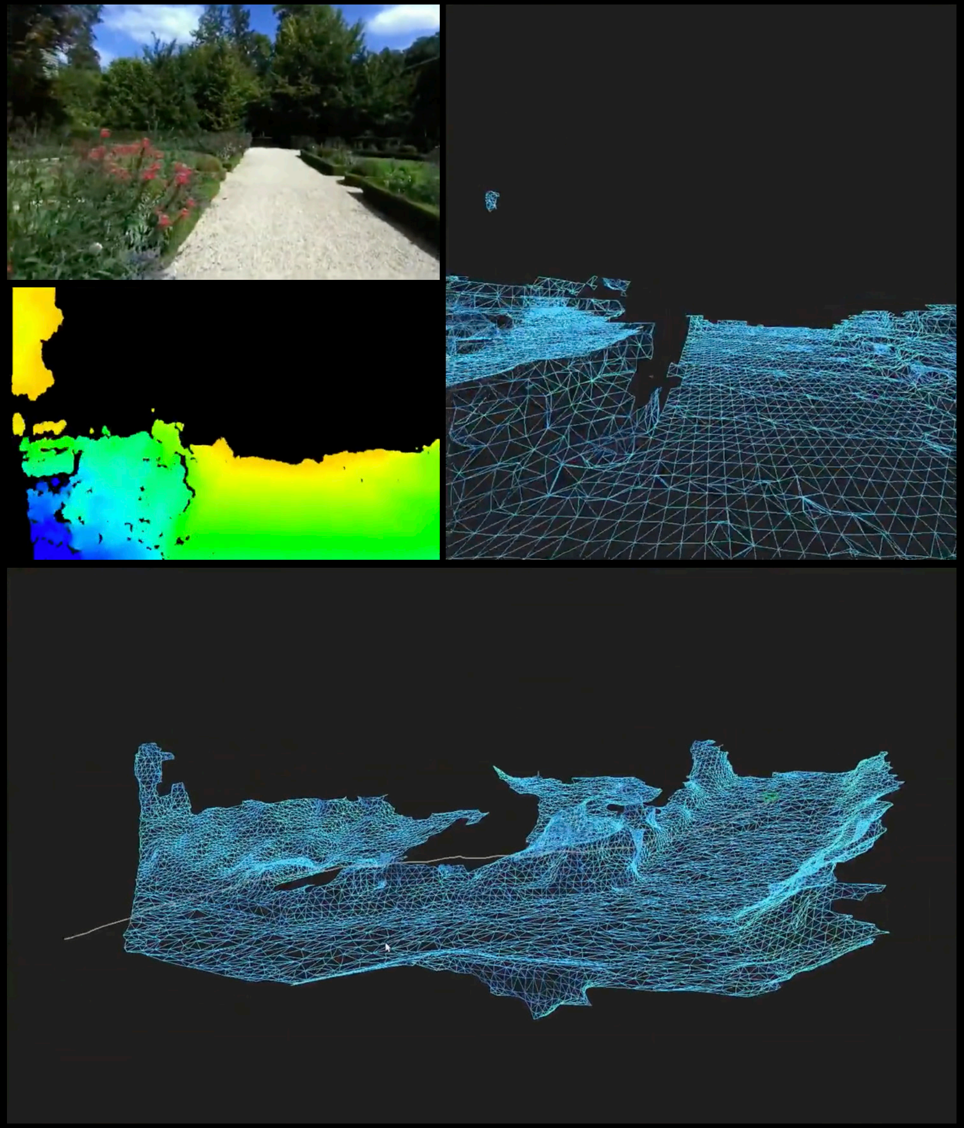
Изображение 28 Стерео монохромное изображение анаглифировано для фильтров красного (левый глаз) и голубого (правый глаз). Для правильного просмотра этого изображения рекомендуется использовать красные голубые 3D-очки.



Изображение 29 Красно-синий анаглиф астронавта Аполлона-17 Джина Сернана на Луне, 11 декабря 1972 года.



Изображение 30 Камера ZED - тест стереозрения.

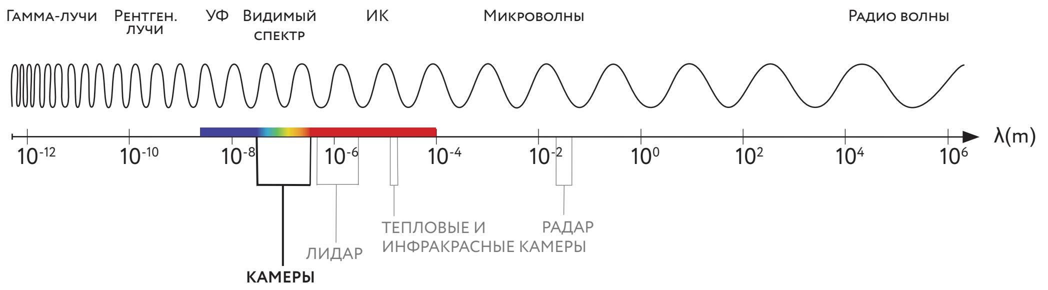


Изображение 31 3D-мэппинг на Jetson Nano с помощью камеры ZED

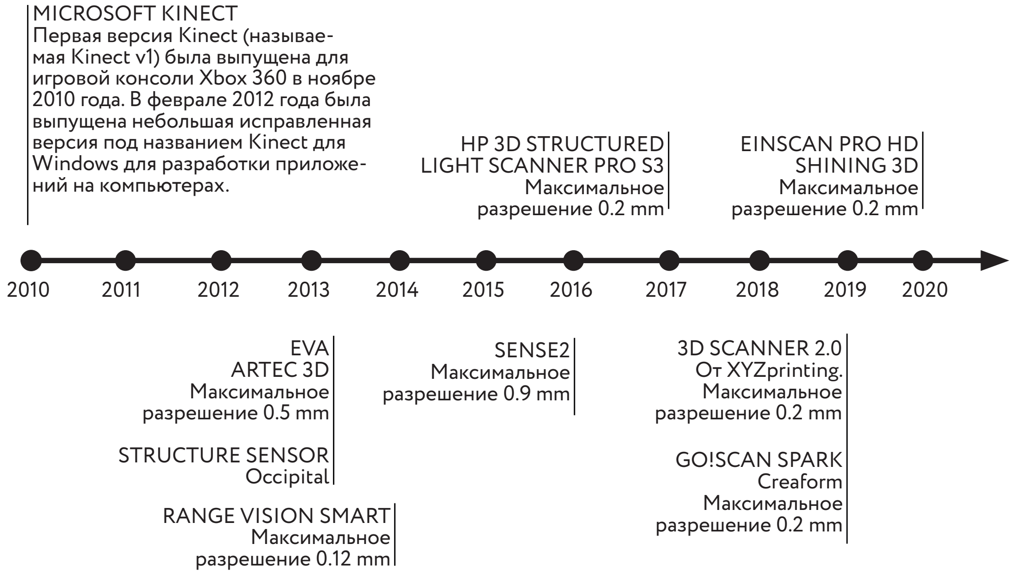
СТРУКТУРИРОВАННЫЙ СВЕТ

Структурированный свет - это процесс проецирования паттерна (часто сетки или горизонтальных полос) на сцену. То, как они деформируются при столкновении с поверхностями, позволяет системам зрения вычислять глубину и информацию о поверхности объектов в сцене, как это используется в 3D-сканерах со структурированным светом .

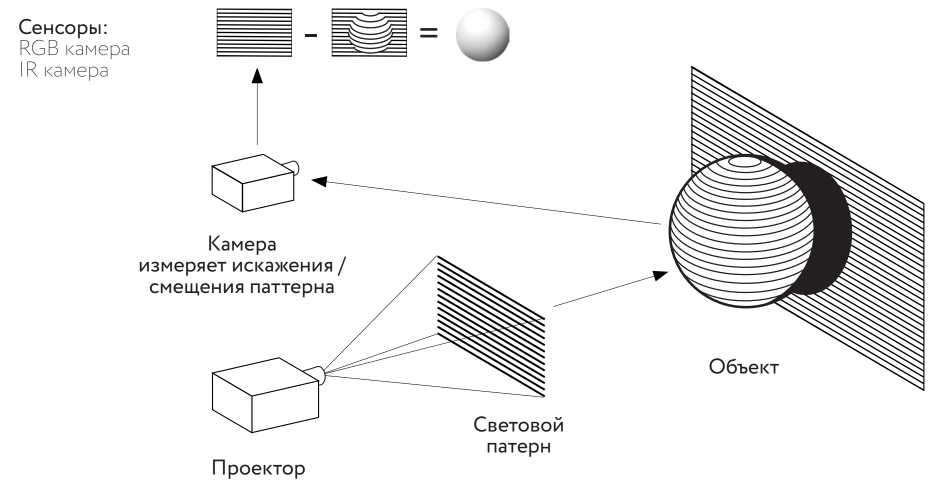
Изображение 32
Электромагнитный спектр и зоны его использования сенсорами технологии
Длина волны, λ(m)



Таймлайн
Технологические события 2010- 2020

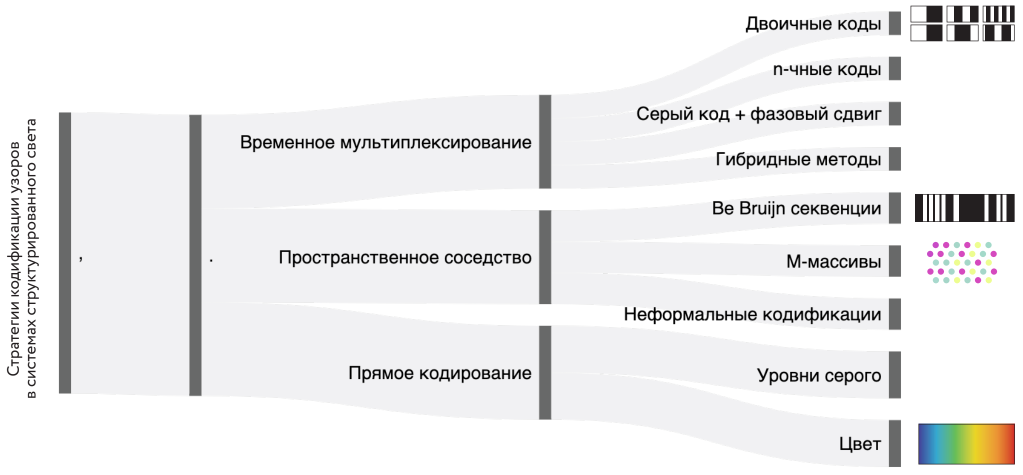


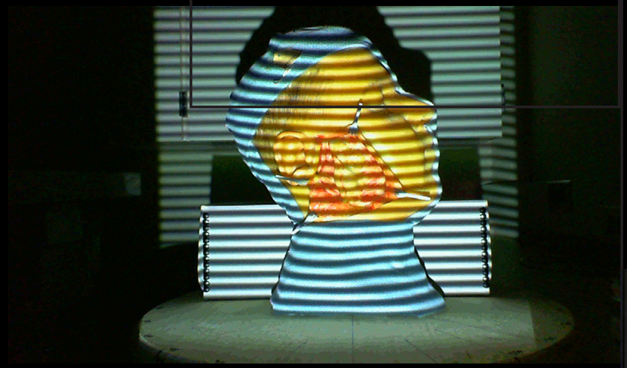
Изображение 33
Схема работы Структурированного света



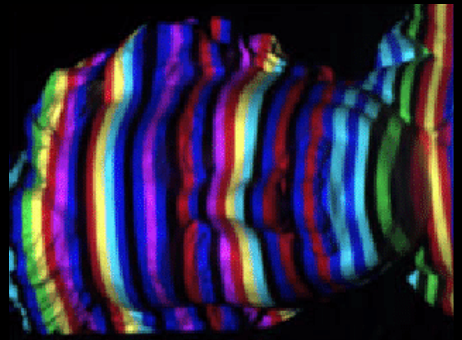
Изображение 34

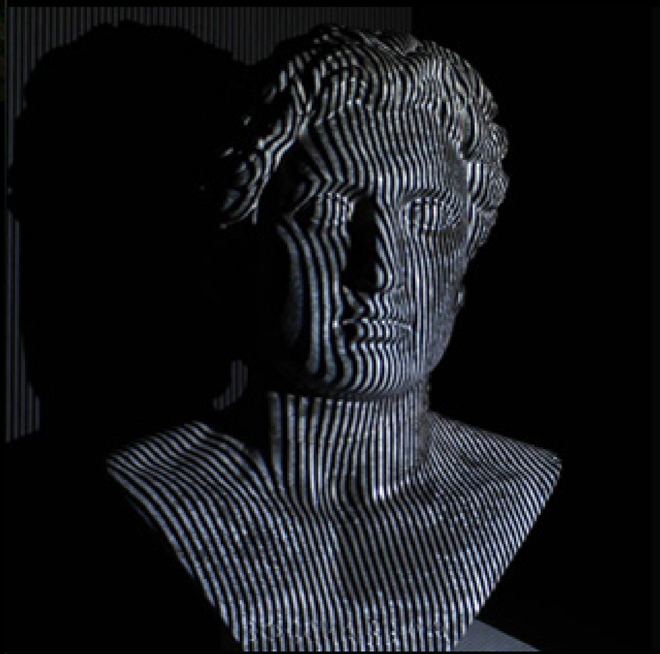
Стратегии кодификации узоров в системах структурированного света

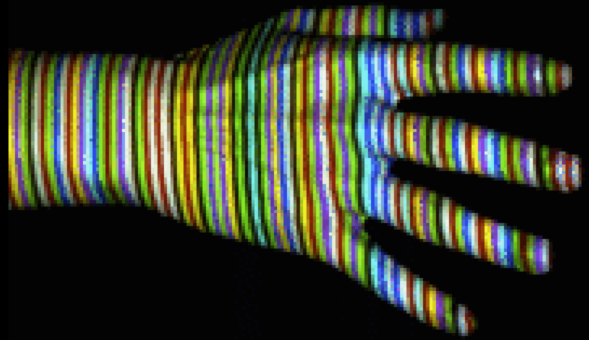


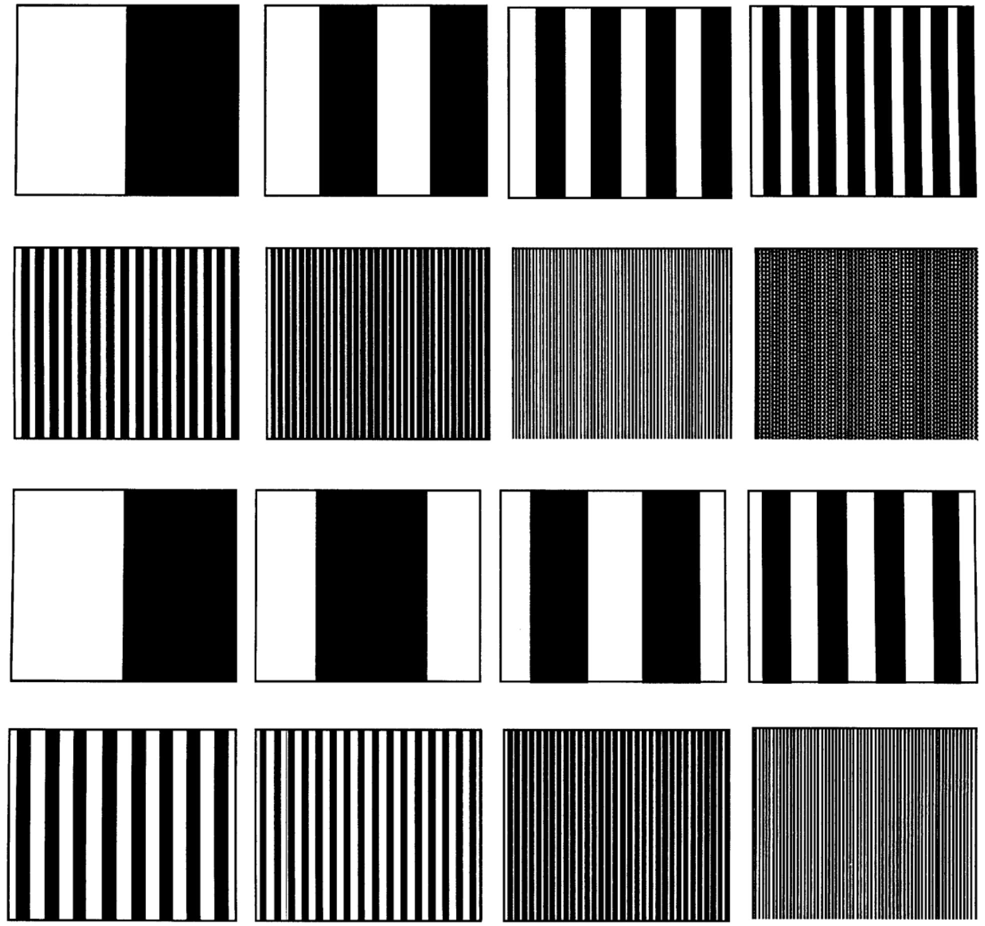


Изображение 35 Узоры структурированного света.

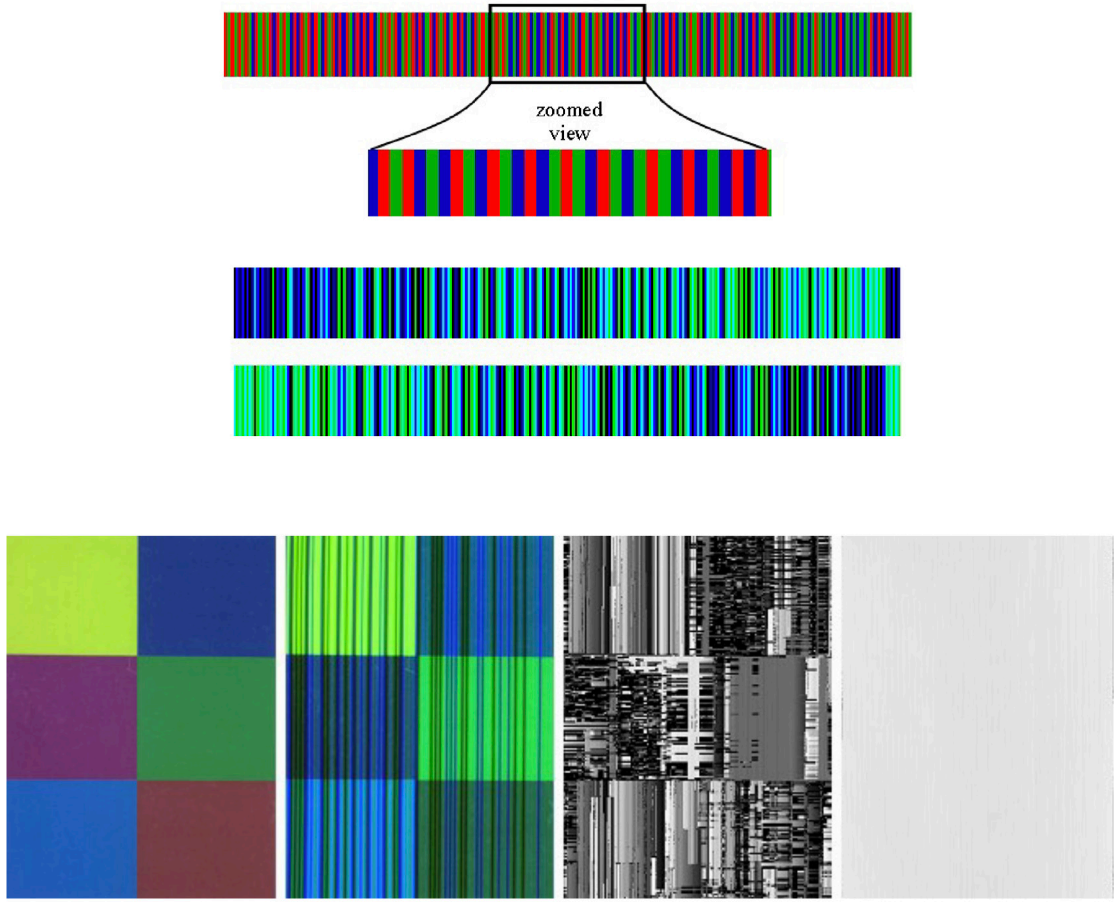




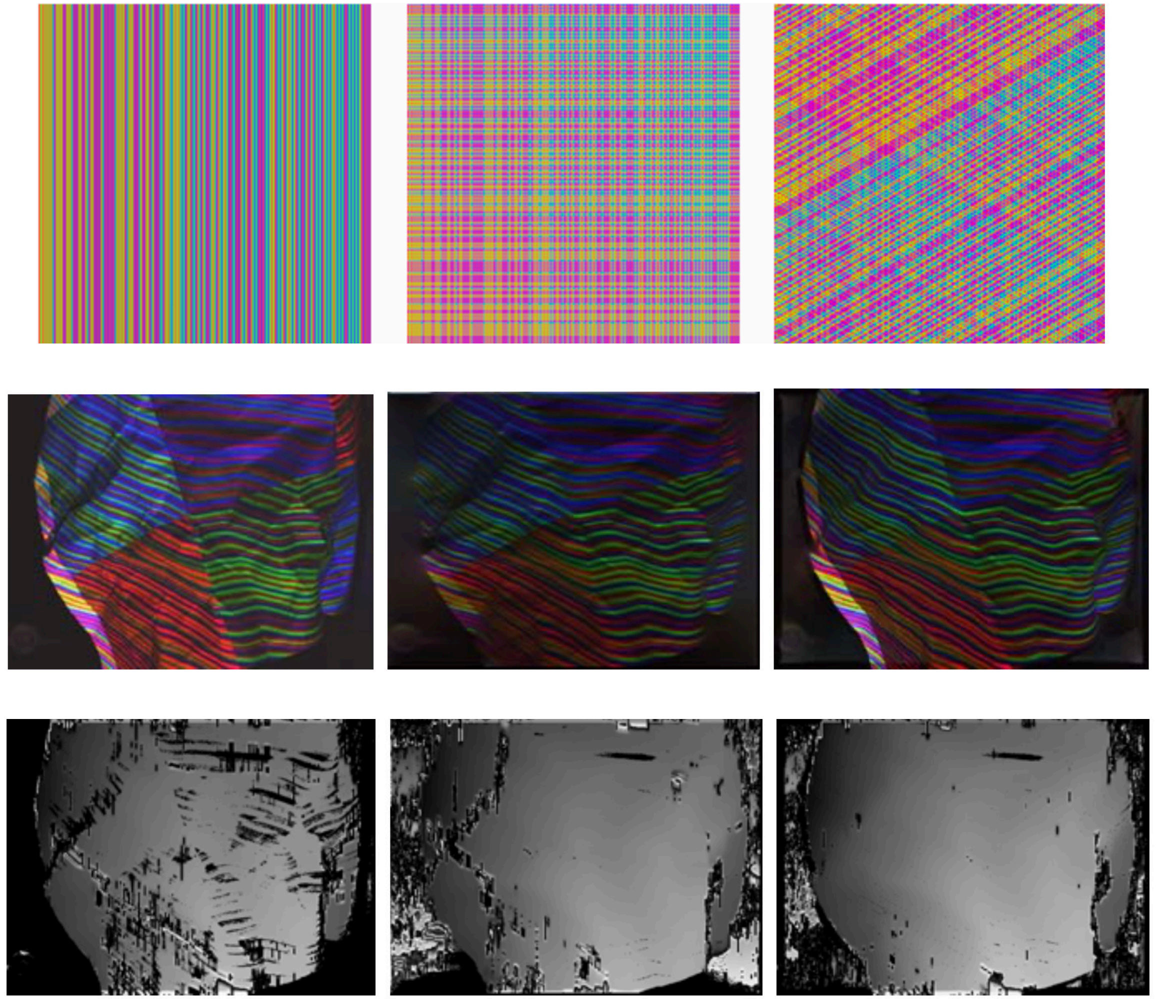




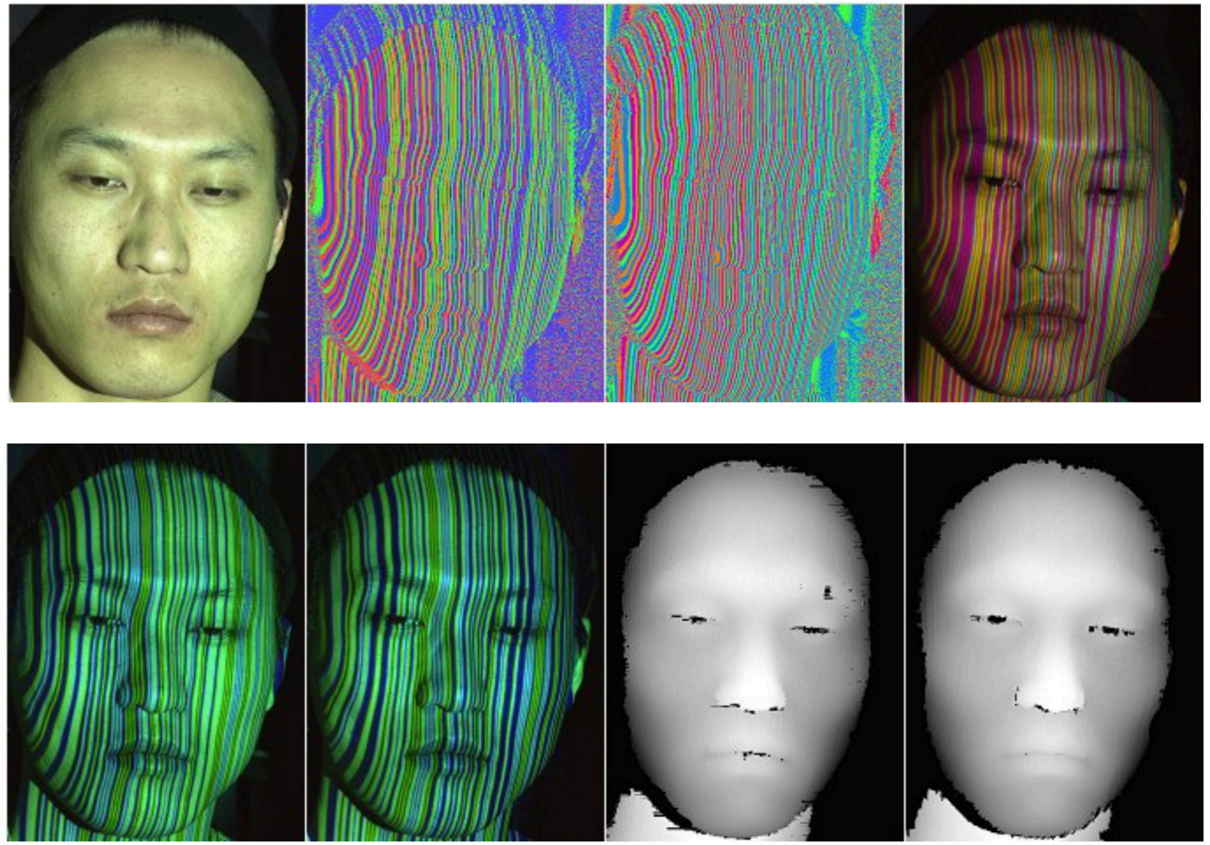
Изображение 36 Узоры структурированного света.



Изображение 37 Высококонтрастный узор из цветных полос для быстрой визуализации в диапазоне структурированного света.



Изображение 38 Структурированный свет с цветными полосами, устойчивый к цвету поверхности и неоднородности.

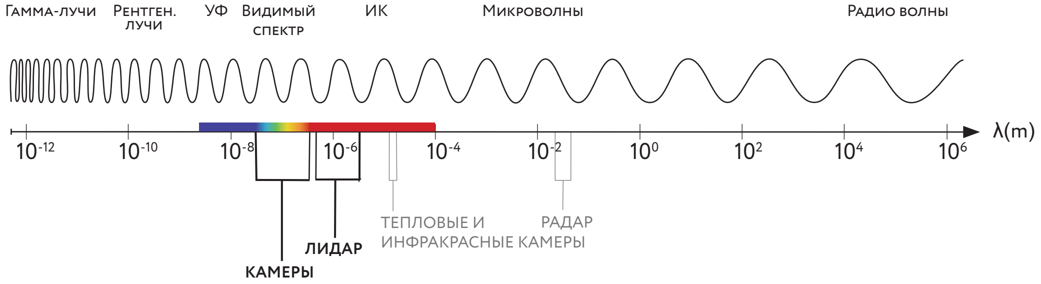


Изображение 39 Высококонтрастный узор из цветных полос для быстрой визуализации в диапазоне структурированного света.

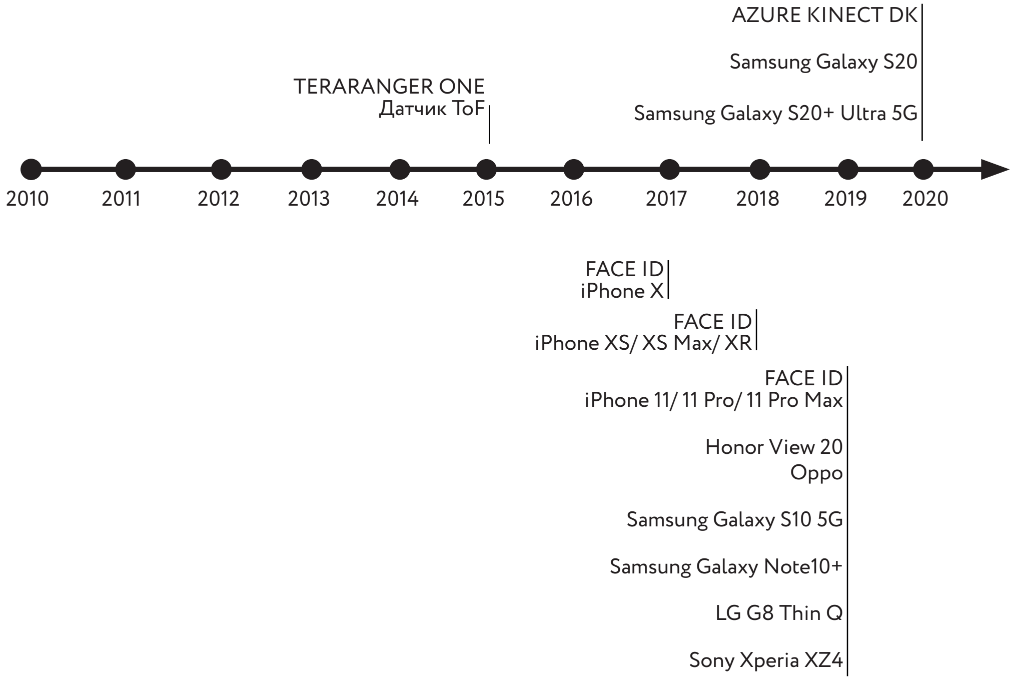
TIME-OF-FLIGHT CAMERA

Time-of-flight камера (ToF) — видеокамера, формирующая так называемое дальностное изображение (дальностный портрет). Используется для создания изображений, которые в качестве пикселей содержат оценки расстояний от экрана до конкретных точек наблюдения. ToF делится на Direct ToF (SPAD, LIDAR) Indirect ToF (Phase Tof).

Изображение 40
Электромагнитный спектр и зоны его использования сенсорами технологии
Длина волны, λ(m)

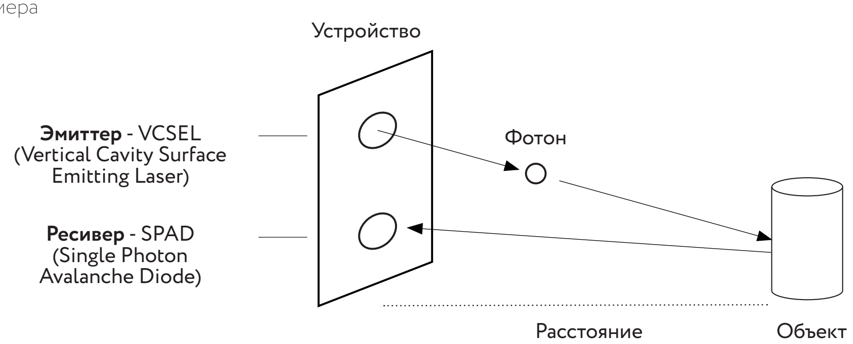


Таймлайн
Технологические события
2010- 2020

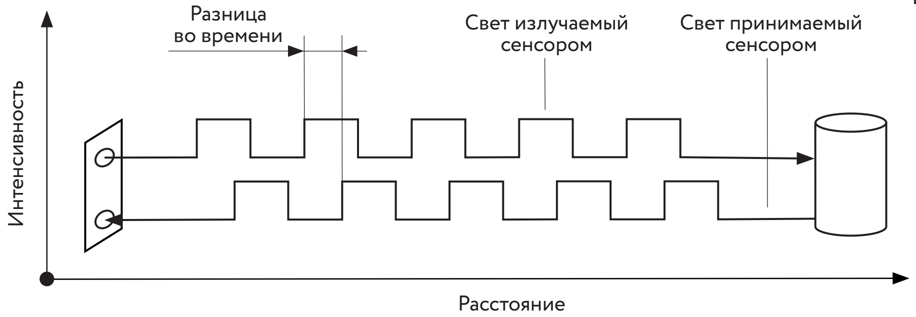


Сенсоры:
RGB камера
LIDAR

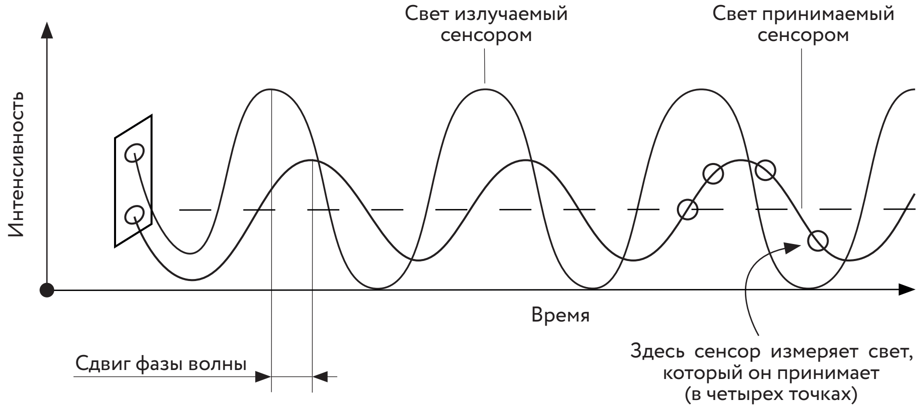
Изображение 41
Схема работы TOF камеры
Величина дистанции = (Время прохождения фотона)/ 2 х скорость света



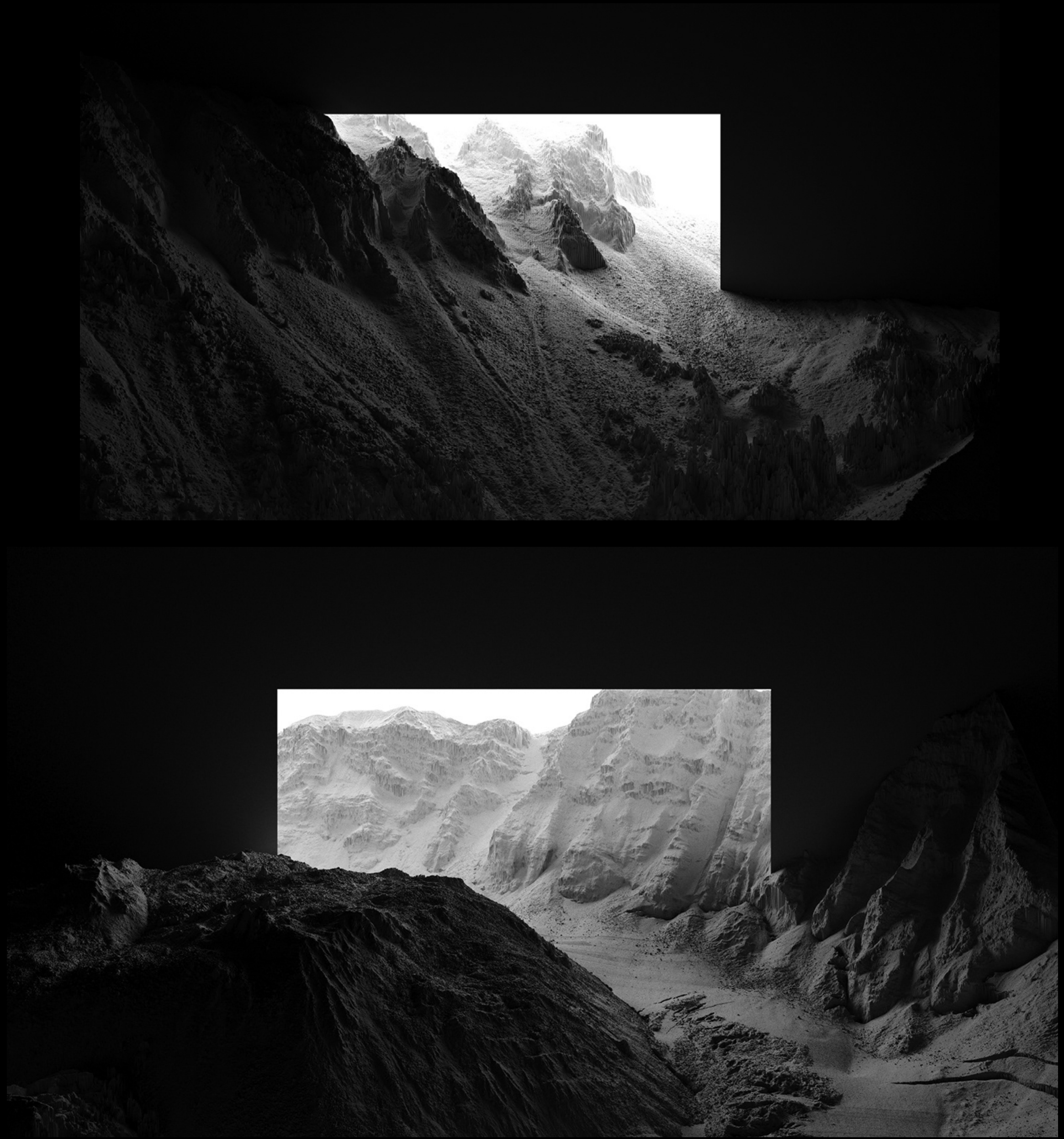
Изображение 42
Схема работы
DIRECT TOF (SPAD, LIDAR)



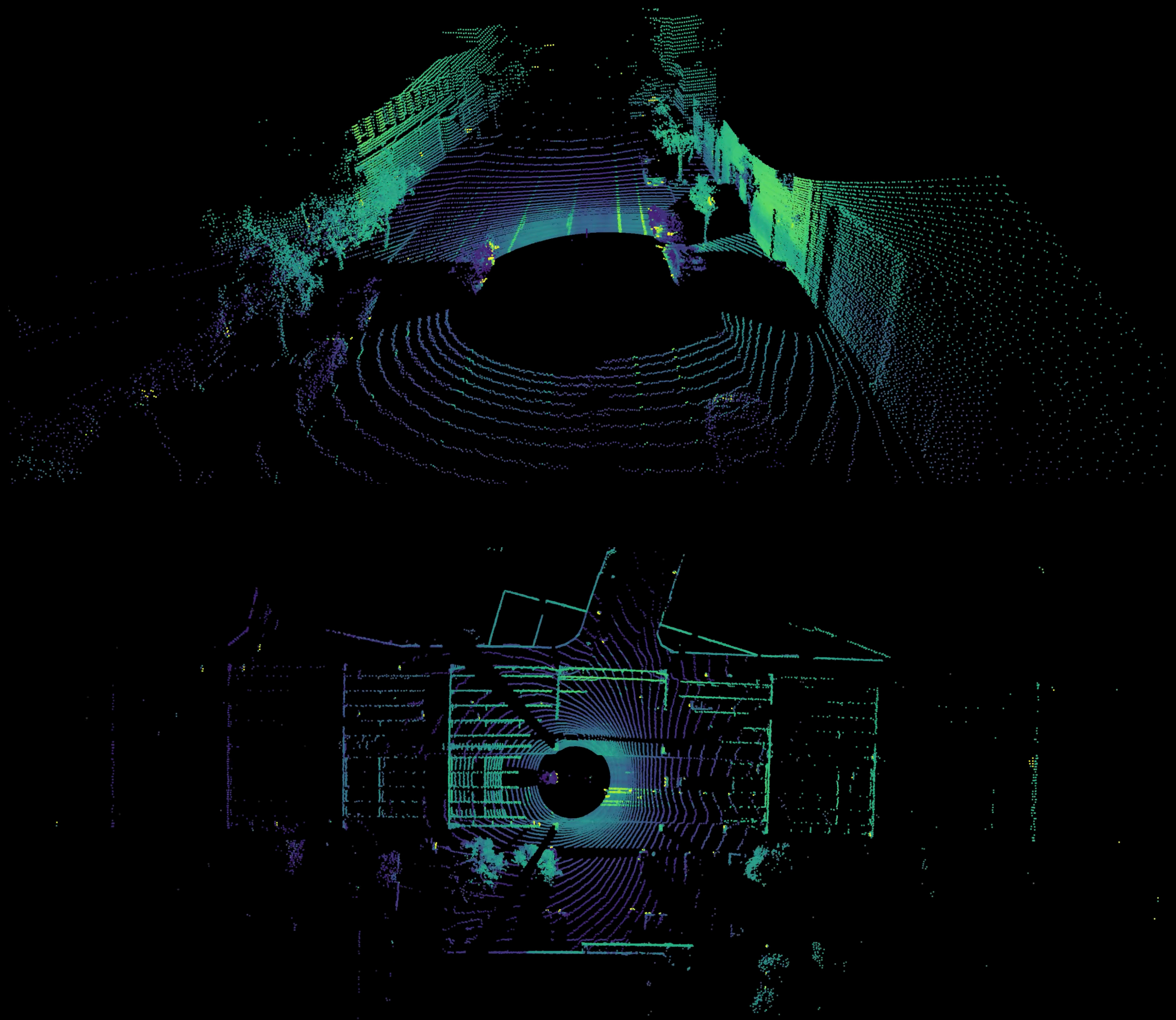
Изображение 43
Схема работы
INDIRECT TOF (PHASE TOF)



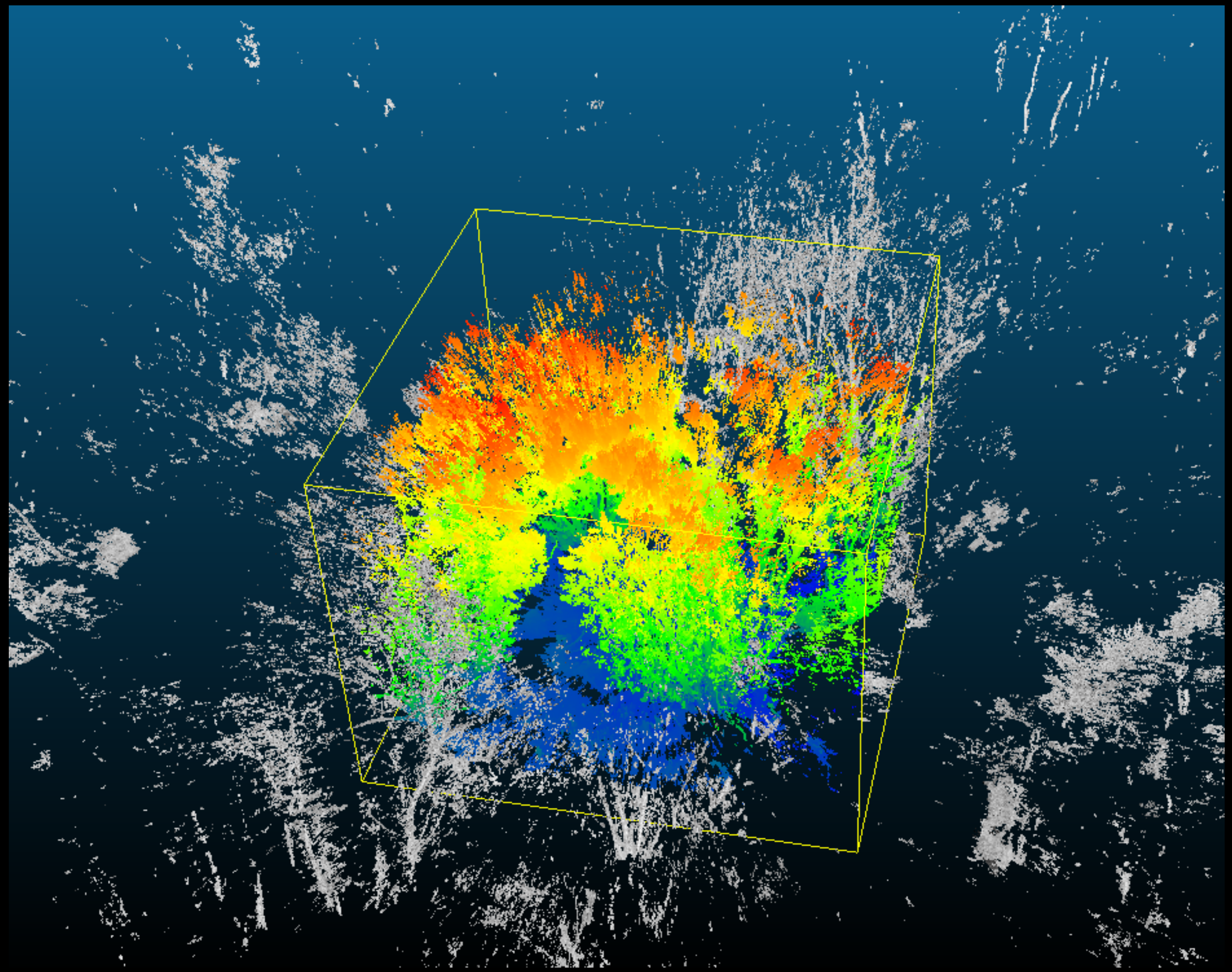
Изображение 44 Фотосессия на инфракрасную времяпролетную камеру Kinect.



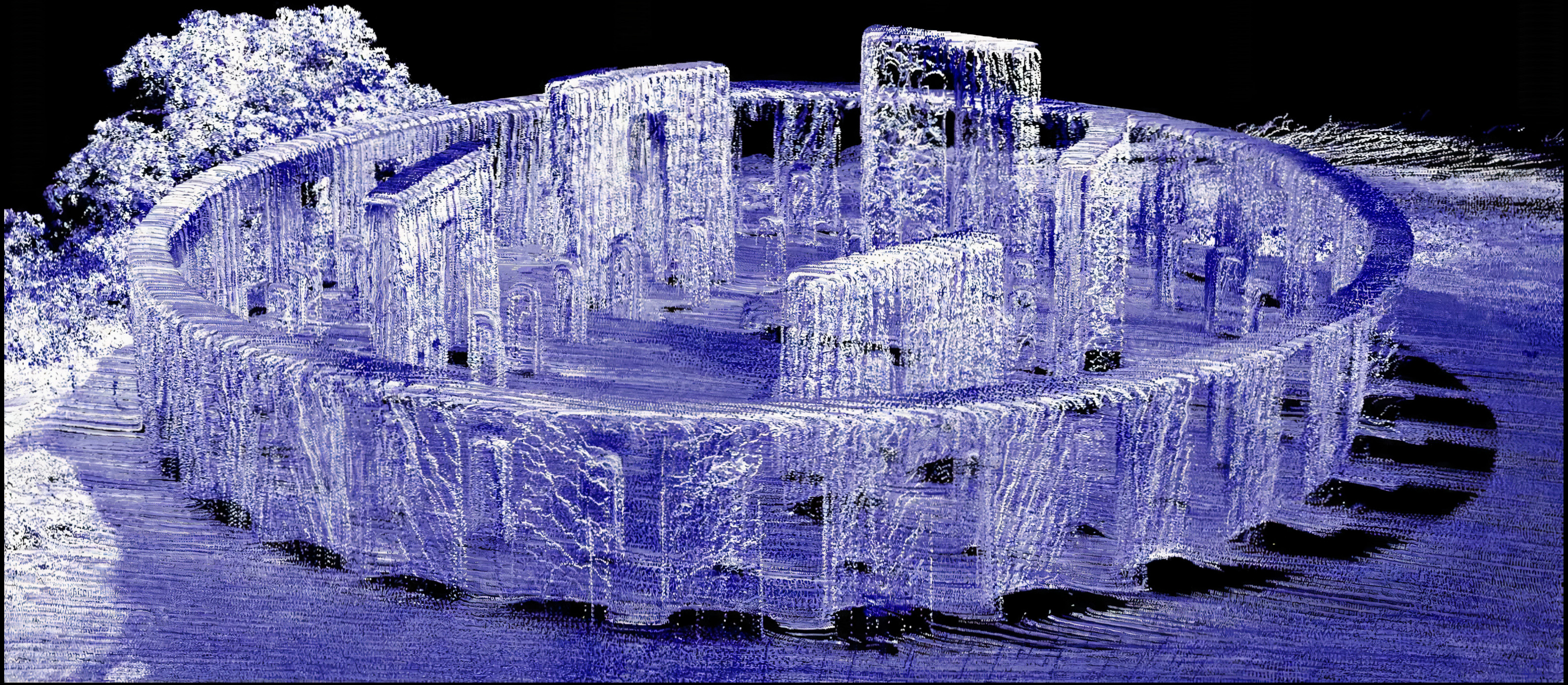
Изображение 45 Лидарные снимки в художественной интерпритации.



Изображение 46 Облако точек лидарного датчика Ouster OS1-64 и двухмерные изображения окружающей среды, сигнала и дальности.

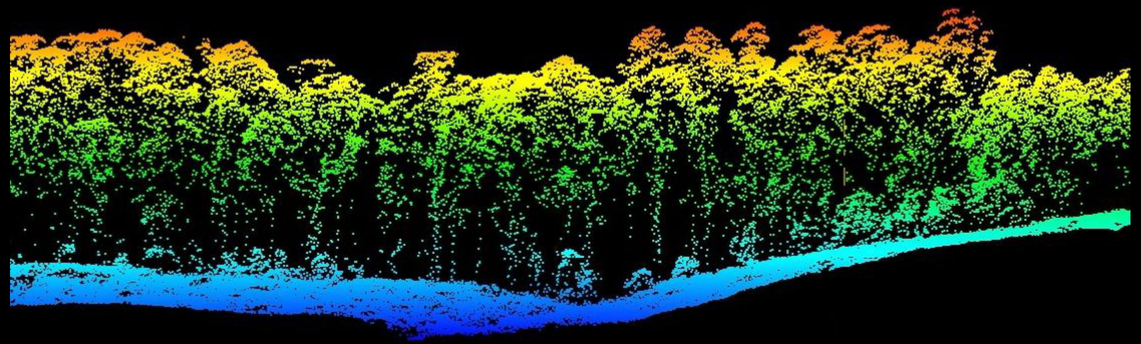


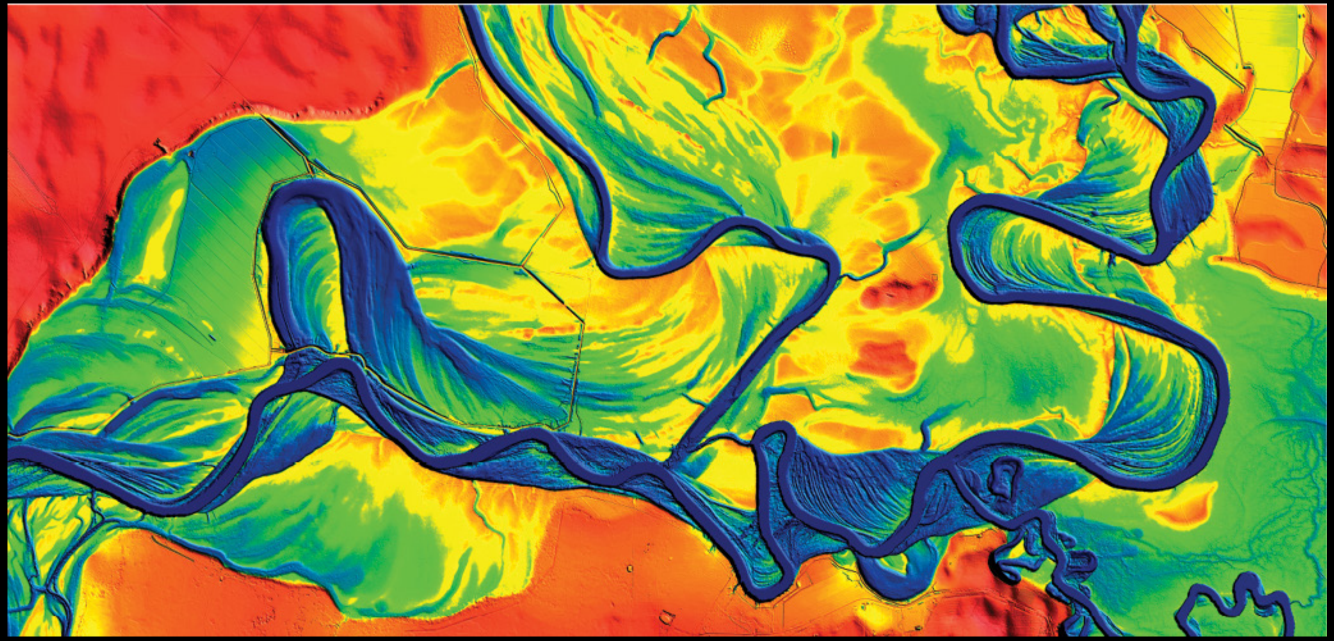
Изображение 47 Трехмерное облако точек прибрежной растительности, полученное с помощью наземного лазерного сканера.



Изображение 48 Детальное облако точек LiDAR военного памятника в Мэрихилле, Вашингтон, США.

Изображение 49 Детальное облако точек LiDAR.



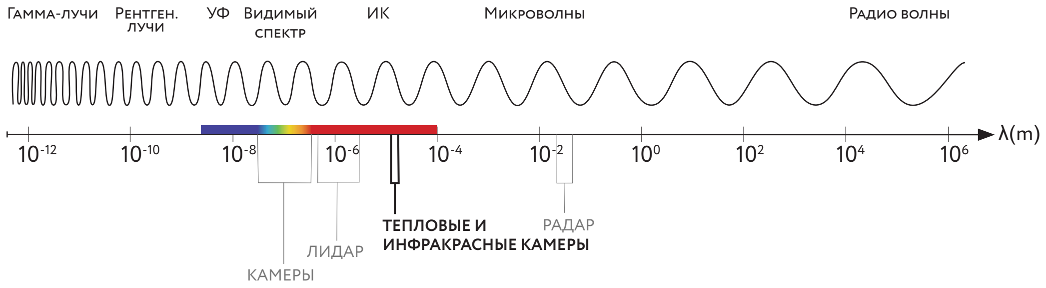


Изображение 50 Лидарные снимки земной поверхности.

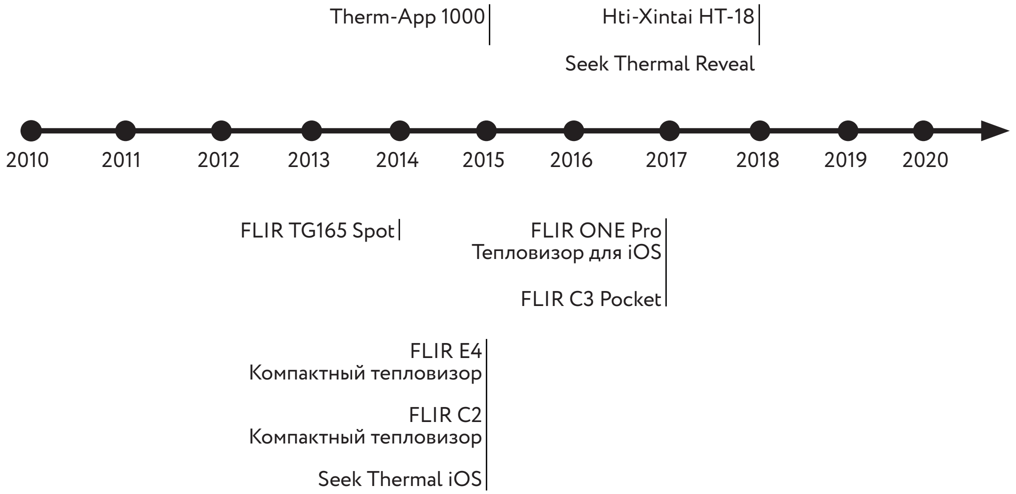
ТЕРМОГРАФИЧЕСКИЕ КАМЕРЫ

Термографические камеры (также называется инфракрасная камера или тепловизионная камера или тепловизор) является устройством , которое создает изображение с помощью инфракрасного излучения, по аналогии с общей камерой, которая формирует изображение с использованием видимого света.

Изображение 51
Электромагнитный спектр
и зоны его использования
сенсорами технологии
Длина волны, λ(m)

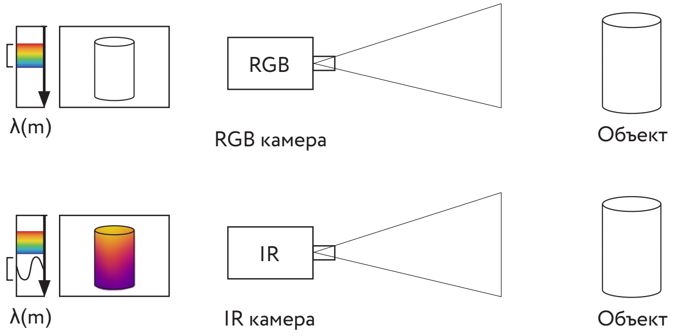


Таймлайн
Технологические события
2010- 2020

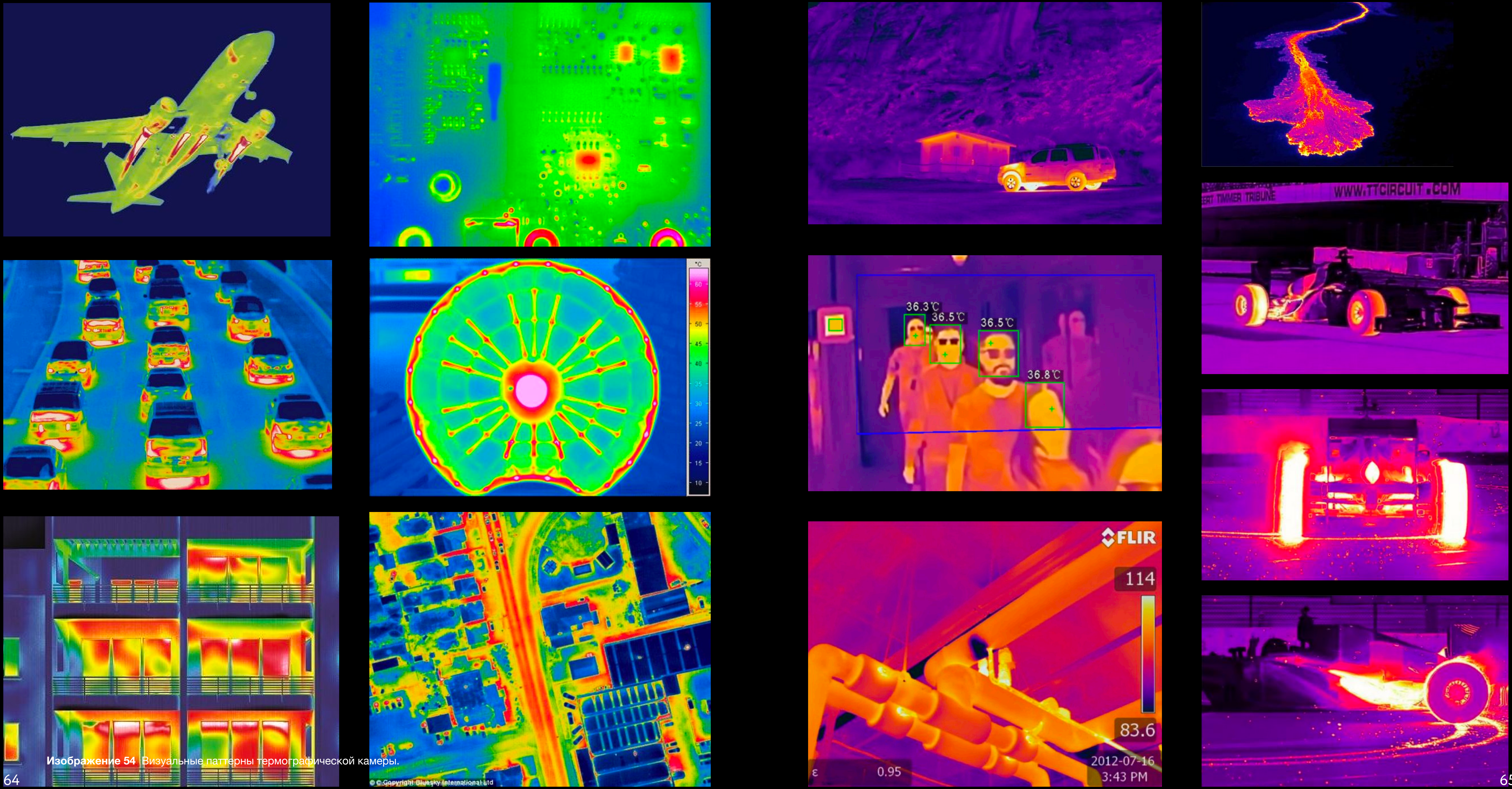


Сенсоры:
IR камера

Изображение 52
Схема работы
Термографической камеры



Изображение 53
Классификация
термографических камер



Изображение 54 Визуальные паттерны термографической камеры.

ВИЗУАЛЬНО-ИНЕРЦИАЛЬНАЯ ОДОМЕТРИЯ (VISUAL INERTION ODOMETRY)

В робототехнике и компьютерном зрении VIO (визуальная инерциальная одометрия) - это процесс определения положения и ориентации робота путем анализа связанных изображений камеры. Составными частями VIO являются CV (компьютерное зрение) и IMU (инерциальная единица измерения). IMU - это электронное устройство, которое может измерять и сообщать угловую скорость, положение, направление ходьбы и удельную силу тела. IMU содержит акселерометры и гироскопы, иногда магнитометры. IMU является технической основой для VIO (Visual Inertial Odometry), где распознование объектов и фотограмметрия алгоритмической основой. В состав IMU часто входит: **акселерометр** - показывает ускорение относительно собственных осей X, Y и Z. Это помогает определить направление к центру Земли. **Гироскоп** - показывает скорость вращения относительно собственных осей X, Y и Z. **Магнитометр/компас** - определяет напряженность магнитного поля относительно собственных осей. Это помогает определить направление на Север. **Барометр** - измеряет атмосферное давление и помогает вычислить высоту над уровнем моря.

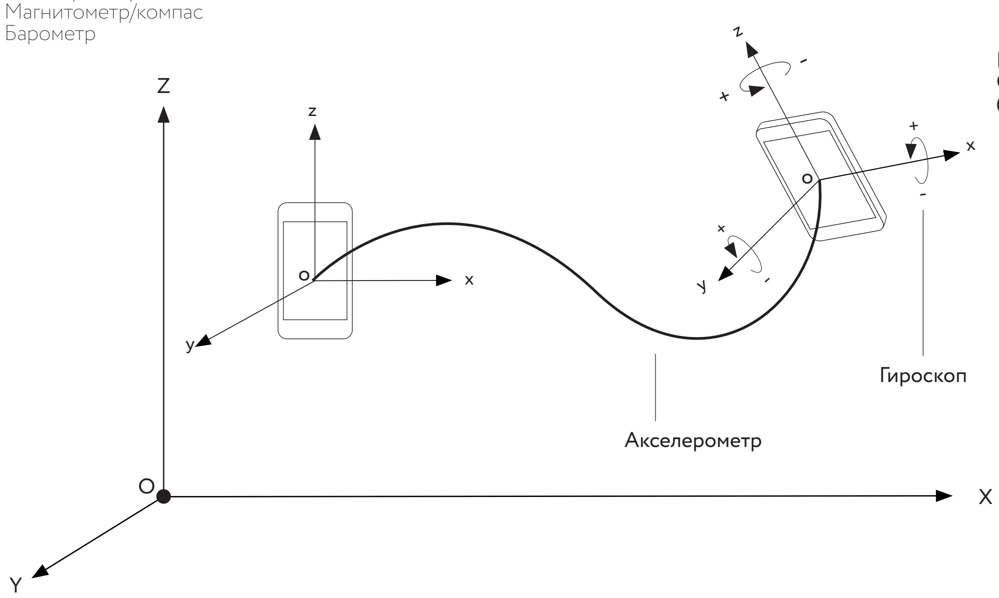
Алгоритмы/ Виды:

• 6-осевое семейство IMU (3-осевой гироскоп + 3-осевой акселерометр)

• 7-осевое семейство IMU (3-осевой гироскоп + 3-осевой акселерометр + датчик барометрического давления)

• 9-осевое семейство IMU (3-осевой гироскоп, 3-осевой акселерометр и 3-осевой компас)

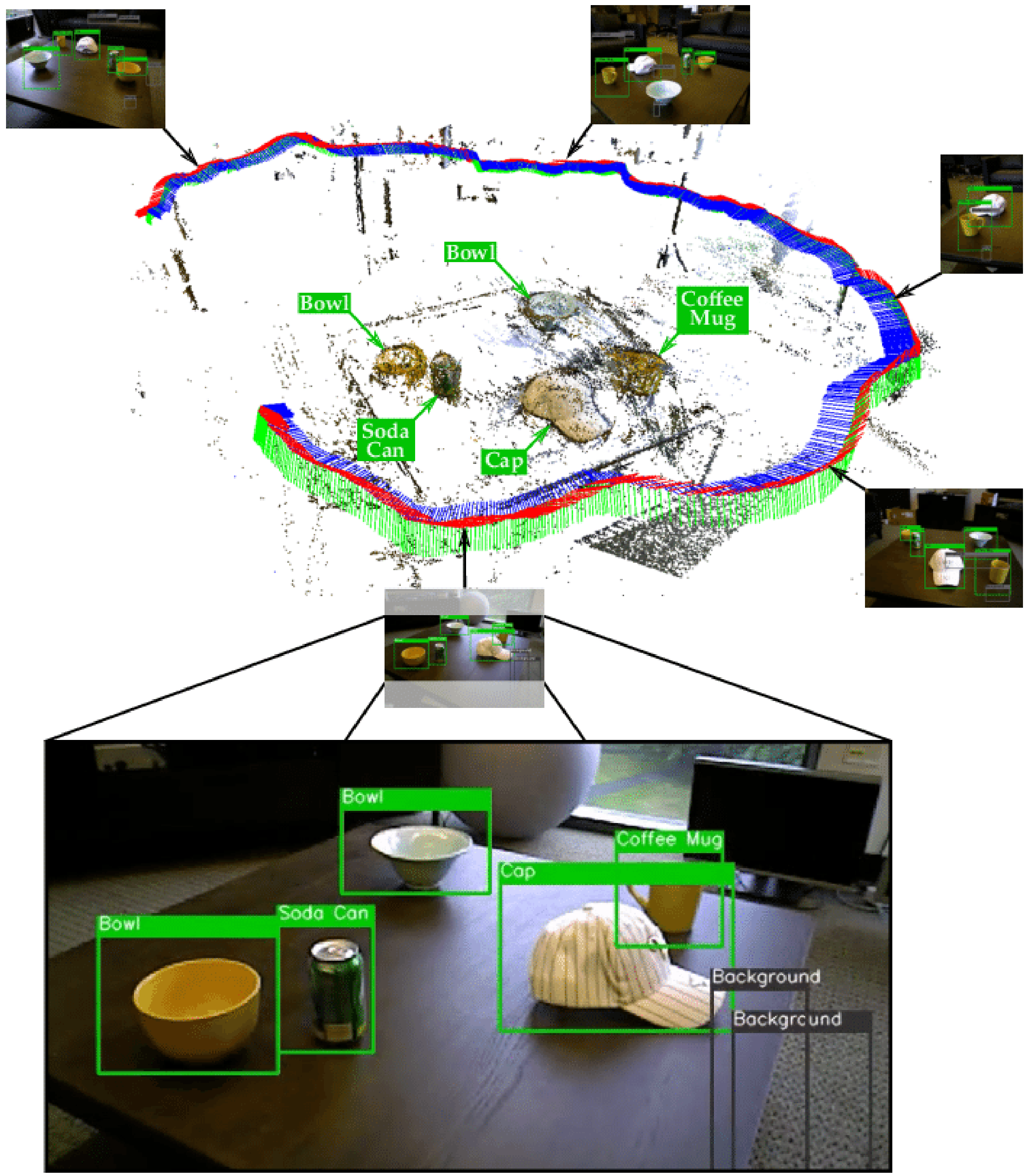
Сенсоры:
Гироскоп
Акселерометр
Магнитометр/компас
Барометр



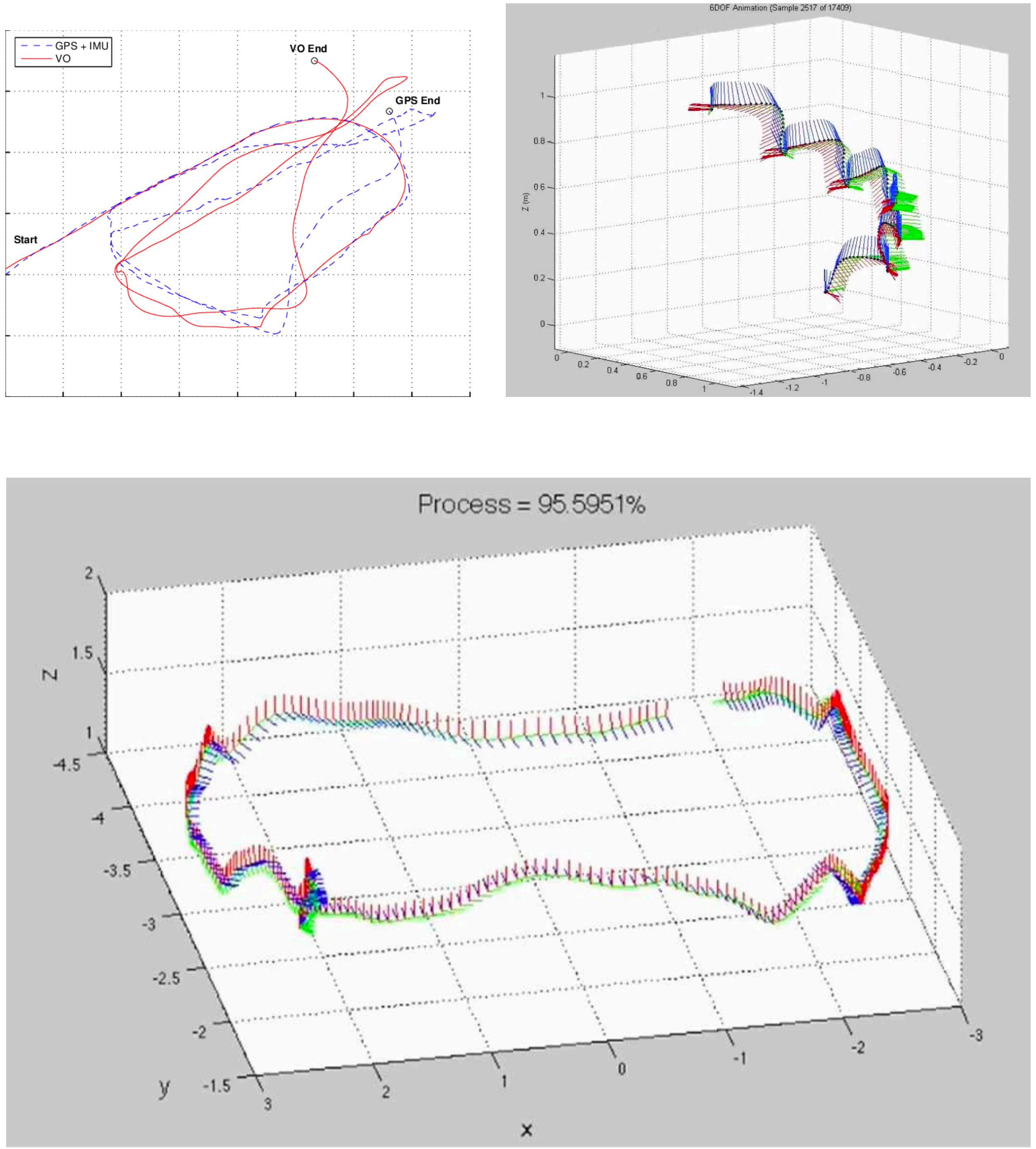
Изображение 55
Схема работы IMU
(6 степеней свободы)

Изображение 56
Производительность IMU
в разных сценариях применения

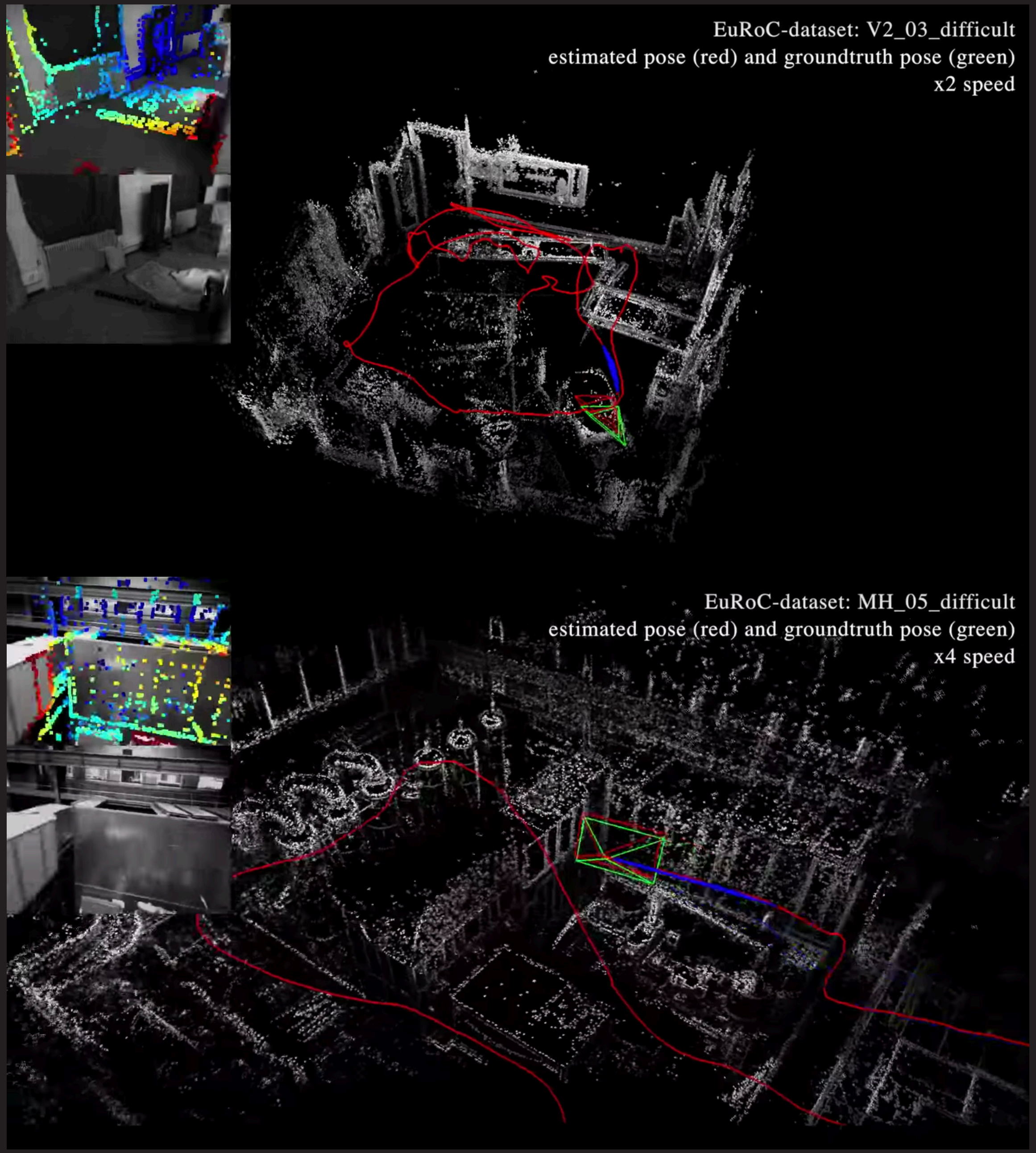
		Акселерометр	Гироскоп	
	Навигация	<100 µg	<0.01°/h	Высокая производительность
	Курсовертикаль	100 µg <x<3mg	0.05°/h<x<25°/h	
	Стабилизация полета	~1 mg	~25°/h	
	Стабилищация	100 µg <x<1mg	0.001°/h<x<10°/h	
	Навигация и захват цели	100 µg <x<500µg	0.002°/h<x<0.05°/h	
	Позиционирование	~1 mg	~20°/h	Низская производительность
	Стабилизация	~10 mg	~100°/h	
	Ориентация или стабидизация	~100 mg	~360°/h	



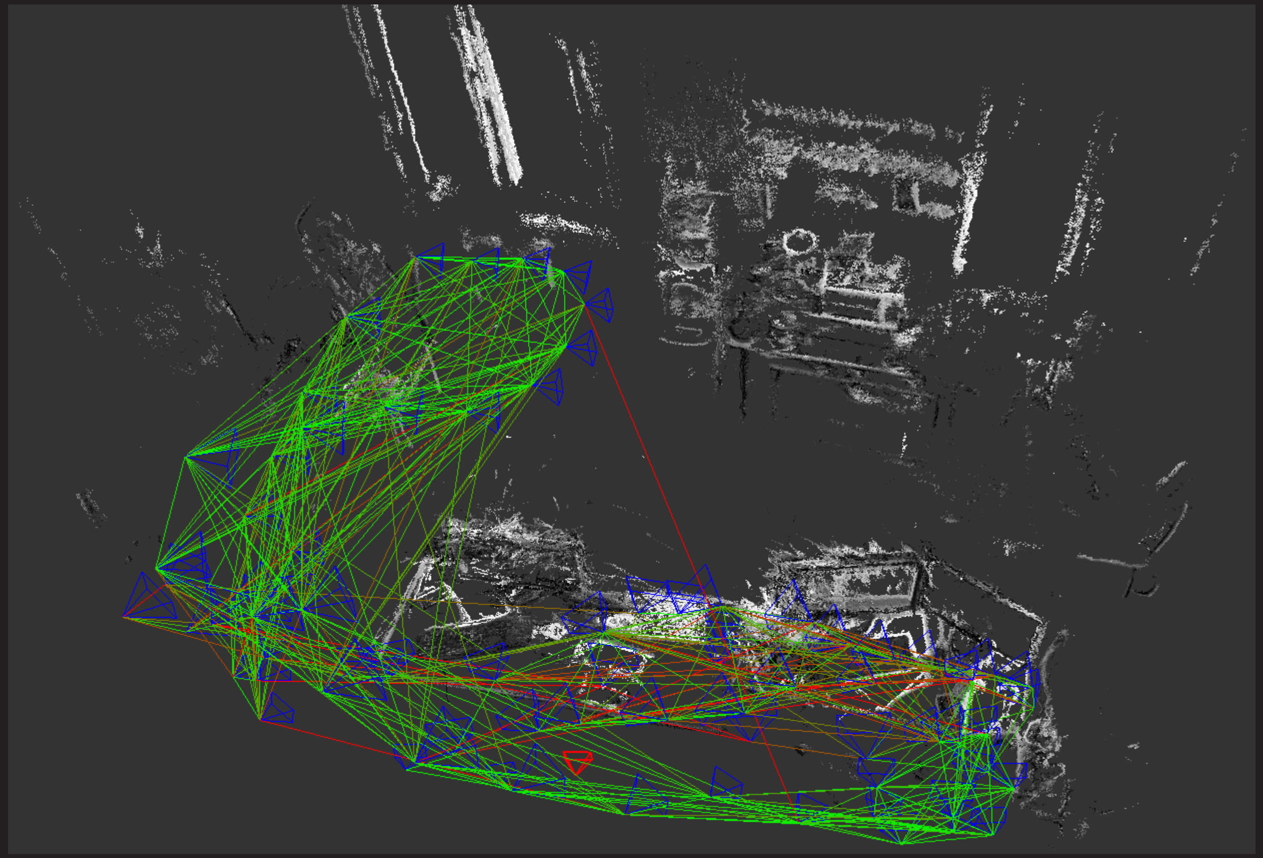
Изображение 57 Визуальные паттерны термографической камеры.



Изображение 58 2D-3D трекинг с помощью IMU.



Изображение 59 Набор данных комнаты LSD-SLAM.



Изображение 60 Набор данных комнаты LSD-SLAM.

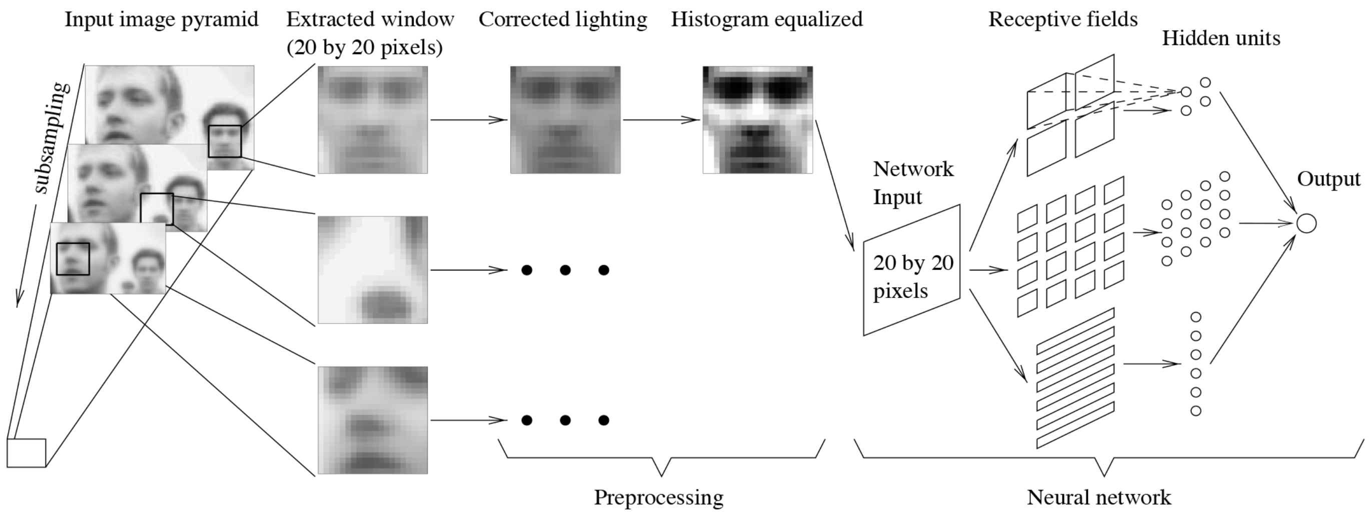
РАСПОЗНАВАНИЕ ОБЪЕКТОВ

Распознование объектов - это распространенная проблема компьютерного зрения, которая связана с идентификацией и локализацией объектов определенных классов на изображении. Интерпретация локализации объекта может выполняться различными способами, включая создание ограничивающего прямоугольника вокруг объекта или маркировку каждого пикселя изображения, содержащего объект (так называемая сегментация).

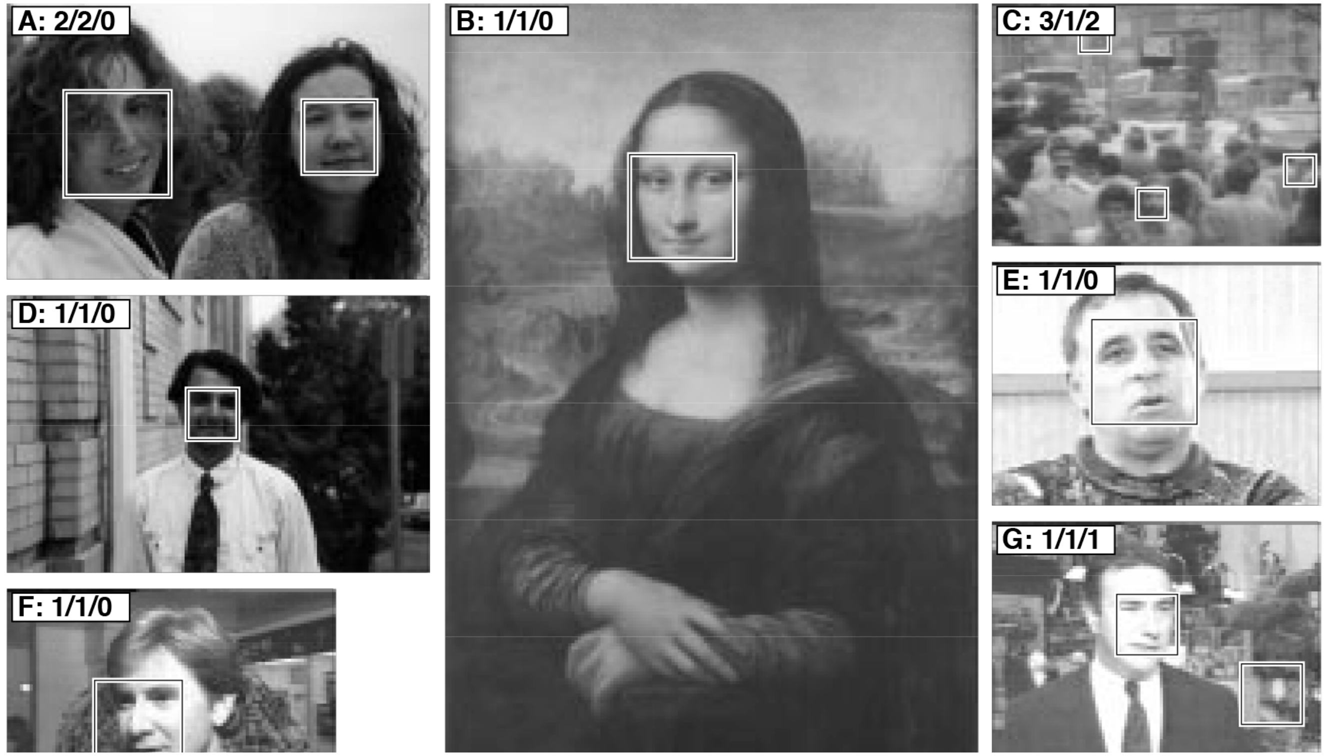
Обнаружение объектов изучалось еще до того, как сверточные нейронные сети (CNN) стали популярными в области компьютерного зрения. Хотя сверточные нейронные сети (CNN) способны автоматически извлекать более сложные и лучшие функции, взгляд на традиционные методы в худшем случае может быть небольшим обходным путем и в лучшем случае вдохновением.

Обнаружение объектов, до глубокого обучения, представляло собой несколько этапов, начиная с обнаружения краев и выделения признаков с использованием таких методов, как SIFT, HOG, SURF и т. д. Эти изображения затем сравнивались с существующими шаблонами объектов, обычно на разных уровнях, для обнаружения и локализации объектов, присутствующих в образе.

Алгоритм выделения признаков прописывали иссследователи, поэтому у алгоритма нет памяти, а память заменяет логичекие правила прописанные человеком.

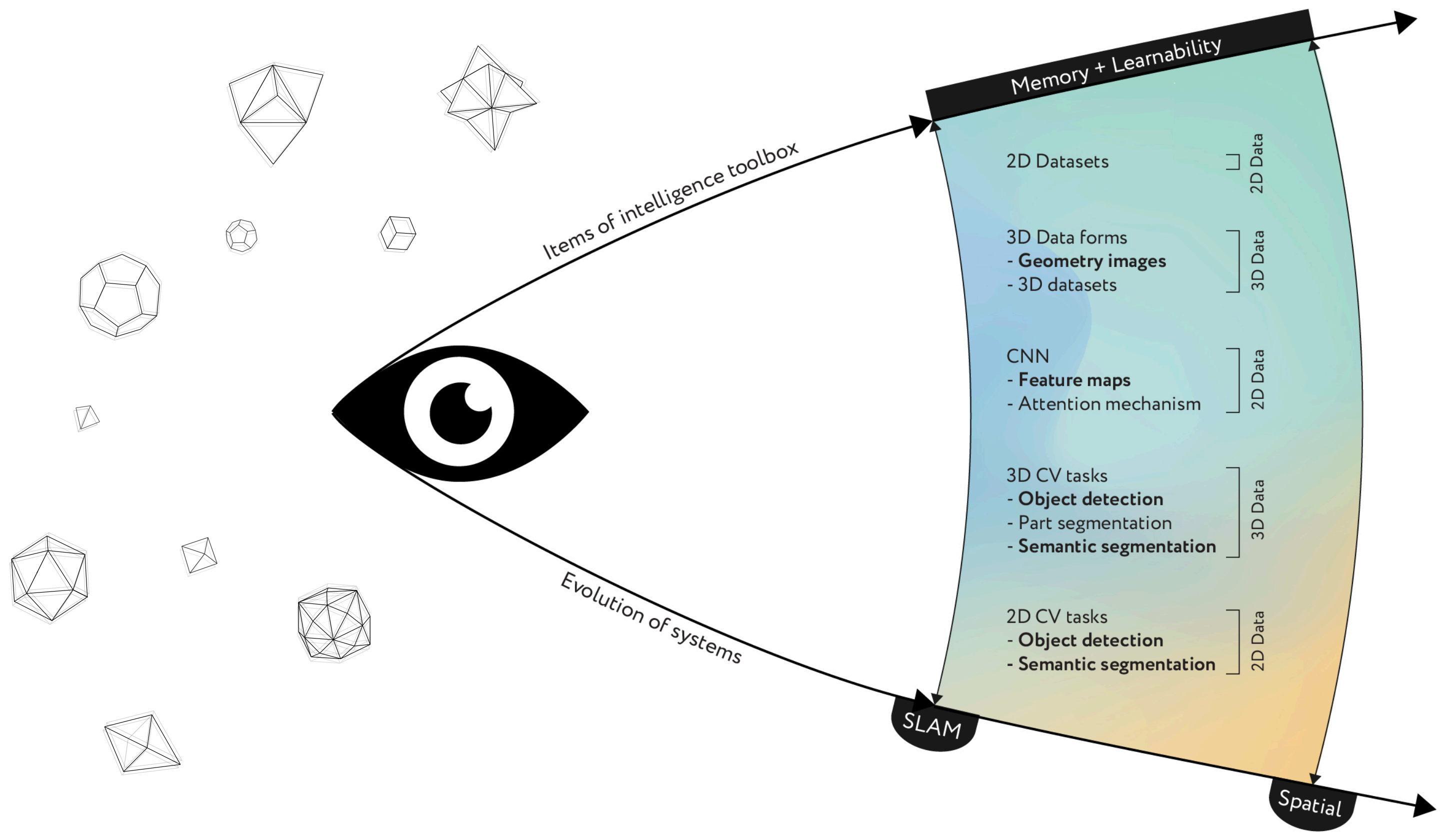


A neural network for face detection (Rowley, Baluja, and Kanade 1998a) c 1998 IEEE.



Изображение 61 Результаты обнаружения лиц, полученные Rowley, Baluja и Kanade (1998a) c 1998 IEEE.

ПАМЯТЬ + ОБУЧАЕМОСТЬ



GEOMETRY IMAGES

Geometry images - захватывает геометрию как простой массив n × n значений [x, y, z]. Другие атрибуты поверхности, такие как нормали и цвета, хранятся в виде дополнительных квадратных изображений, разделяя тот же домен, что и геометрия. Поскольку геометрия и атрибуты имеют одну и ту же параметризацию, сама параметризация является неявной - «координаты текстуры» отсутствуют. Более того, эта параметризация полностью использует область текстуры (без потери места). Геометрические изображения можно кодировать с использованием традиционного алгоритма сжатия изображений, например кодеров на основе вейвлетов. Также геометрические изображения идеально подходят для аппаратного рендеринга. Они могут быть переданы в графический конвейер в сжатом виде, как и изображения текстур. И они исключают дорогостоящие структуры на основе указателей, такие как списки индексированных вершин.

Изображение 62 Geometry Images
(Xianfeng Gu, Steven J. Gortle, Hugues Hoppe, 2002)

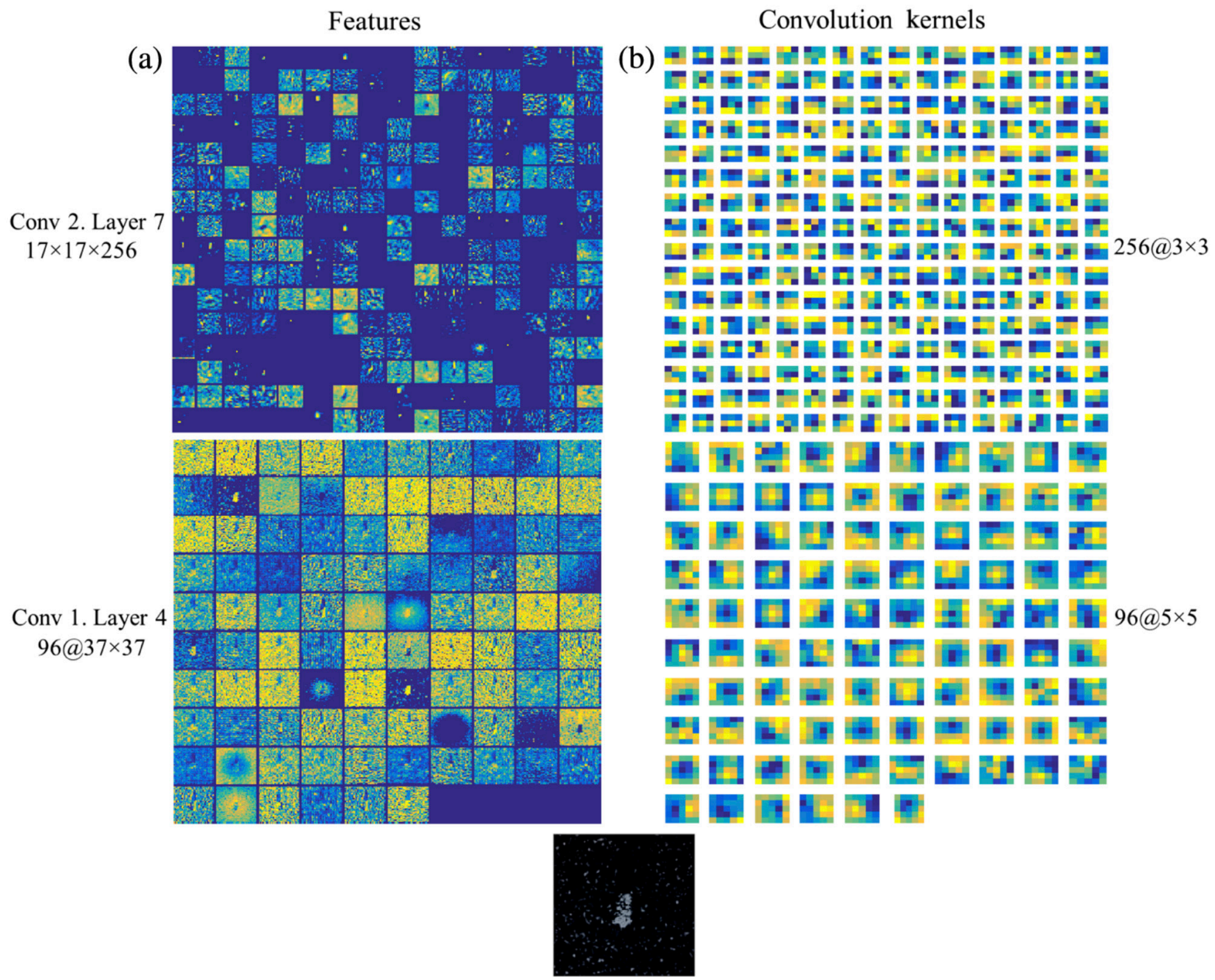
КАРТЫ (ХАРАКТЕРНЫХ) ПРИЗНАКОВ

Сверточные нейронные сети (CNN), имеют внутренние структуры, которые предназначены для работы с данными двухмерного изображения и, как таковые, сохраняют пространственные отношения для того, что было изучено моделью. В частности, двумерные фильтры, изученные моделью, могут быть проверены и визуализированы, чтобы обнаружить типы функций, которые модель будет обнаруживать, а карты активации, выводимые сверточными слоями, могут быть проверены, чтобы точно понять, какие особенности были обнаружены для данного входа.

В сверточной нейронной сети элементы в скрытом слое сегментируются на «карты характеристик», где элементы карты функций совместно используют матрицу весов или, проще говоря, ищут одну и ту же функцию. Скрытые блоки на карте объектов уникальны тем, что они связаны с различными блоками нижнего уровня. Таким образом, для первого скрытого слоя единицы на карте объектов будут связаны с разными областями входного изображения.

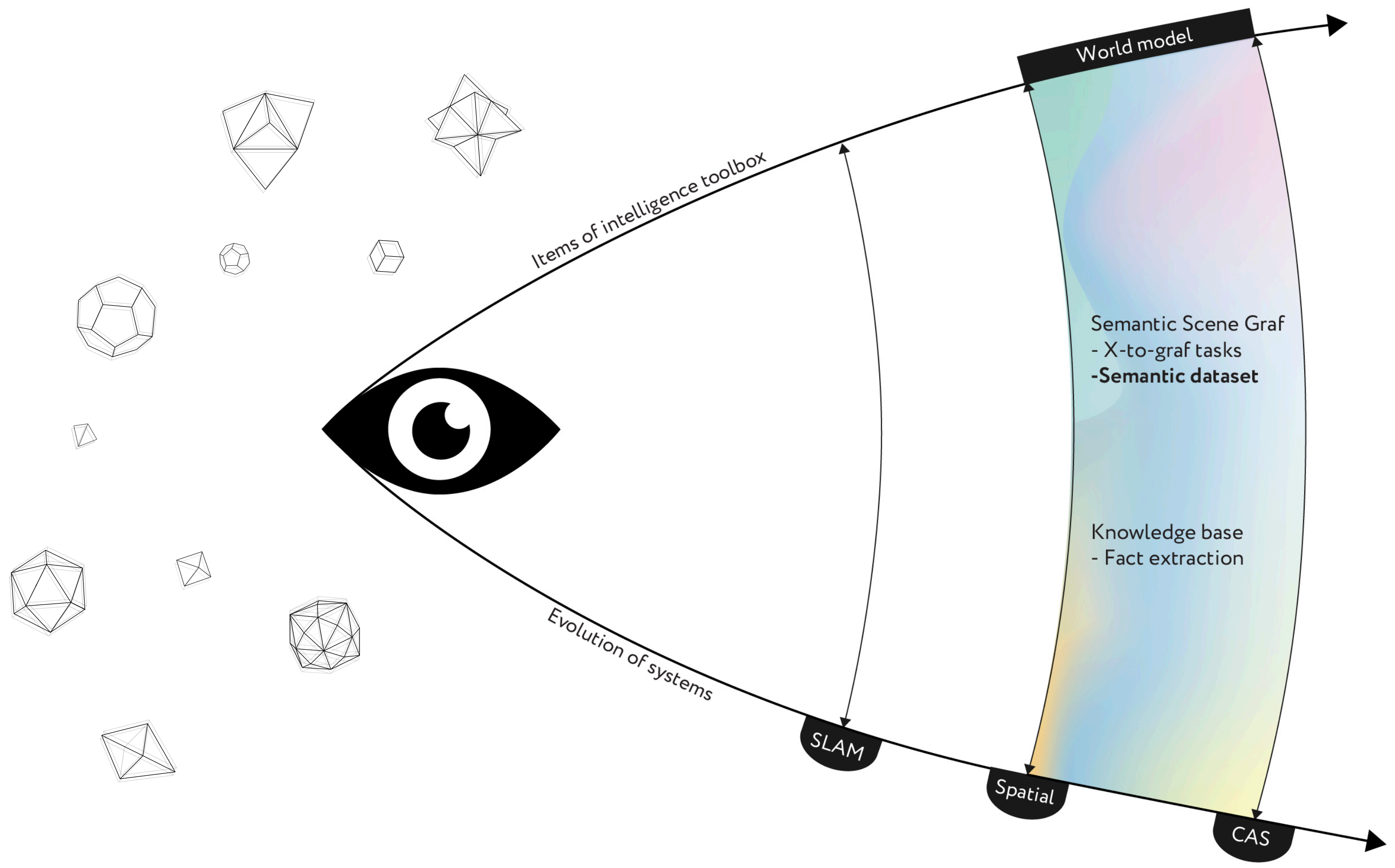
Таким образом, скрытый слой сегментируется на карты функций, где каждый элемент карты функций ищет одну и ту же функцию, но в разных положениях входного изображения.

Карта признаков, по сути визуальная память машины и признаки выучиваются самостоятельно, человек в этом практически не участвует.



Изображение 63 Визуализация признаков: (а) выходные карты признаков и (б) ядра свертки рассматриваемых. Выделение и комбинация глубоких признаков для классификации целей РЛС с синтезированной апертурой (Мусса Амрани, Фэн Цзян, 2017).

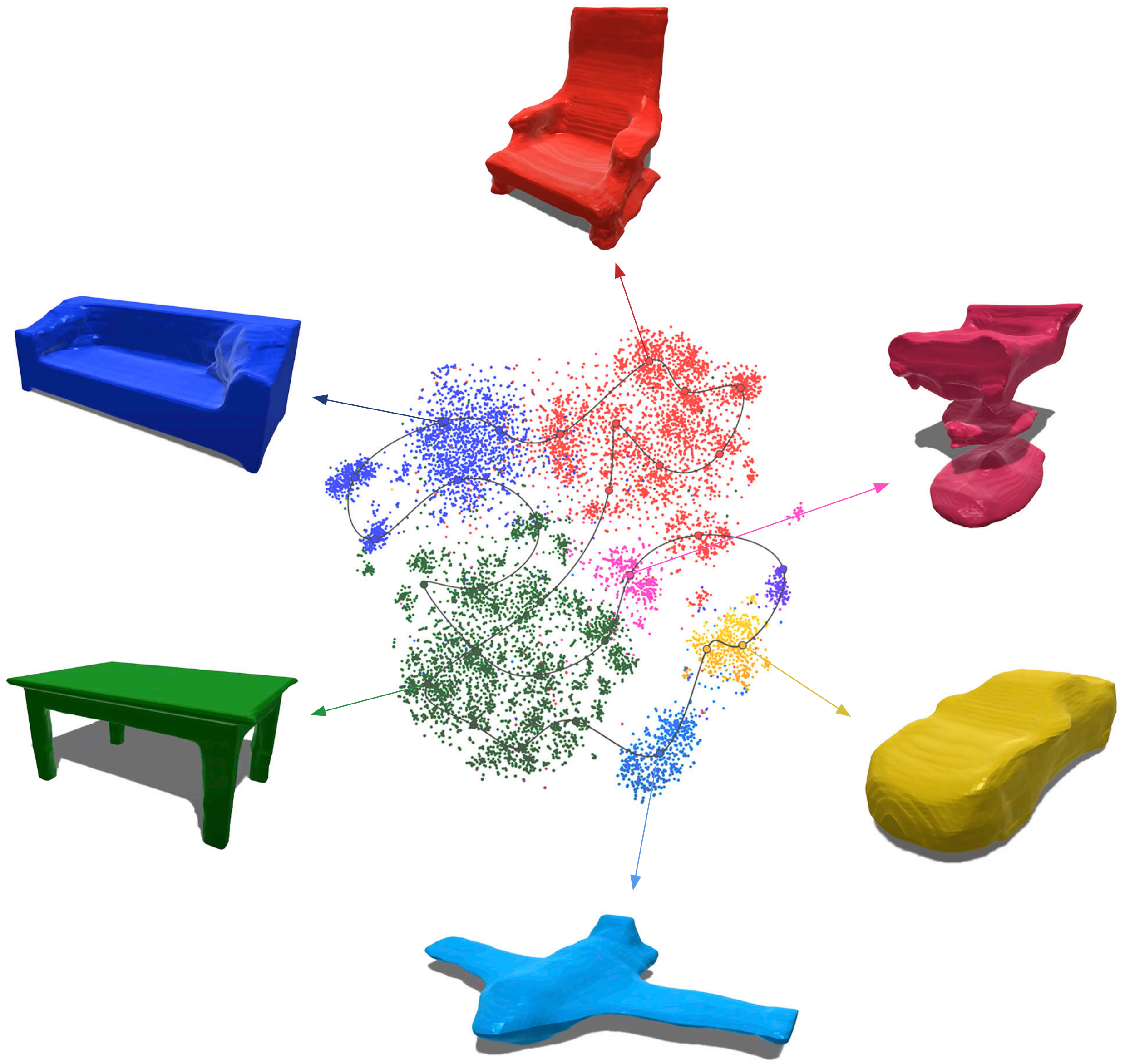
МОДЕЛЬ МИРА



World model
Items of intelligence toolbox
Evolution of systems
Semantic Scene Graf
- X-to-graf tasks
-Semantic dataset
Knowledge base
- Fact extraction
SLAM
Spatial
CAS

SEMANTIC SPACE VISUALIZATION

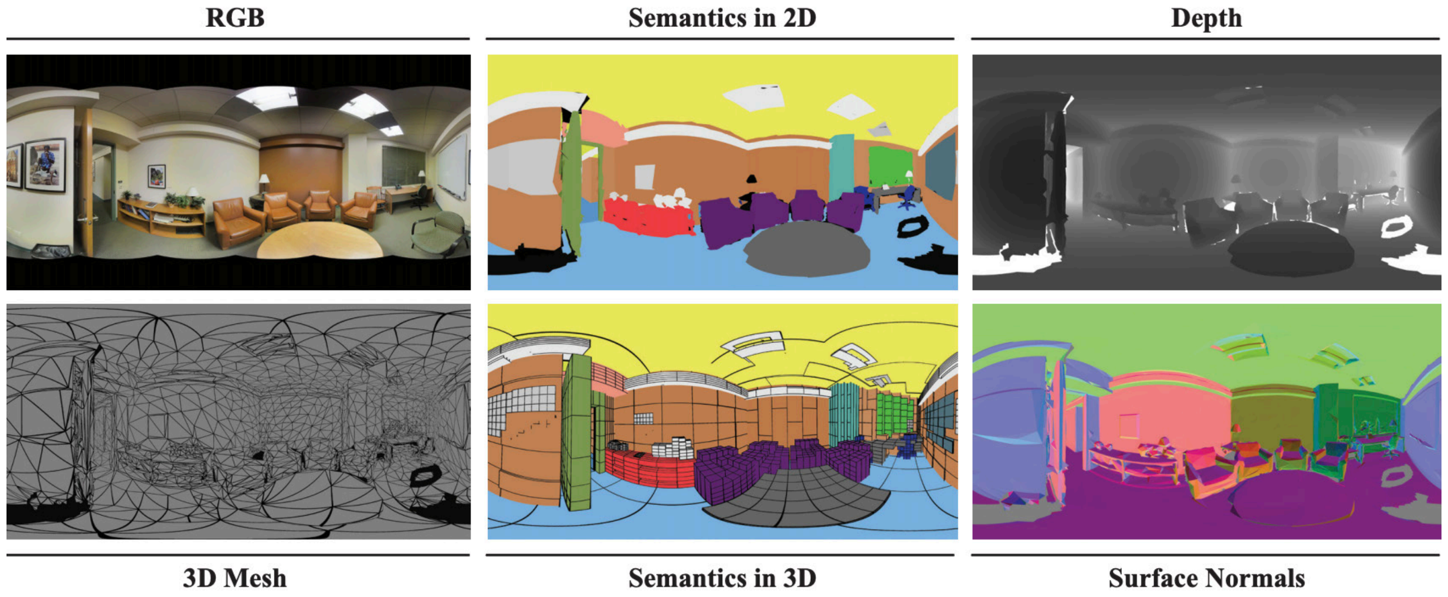
Визуализация семантического пространства - визуализация результатов моделирования должна выполняться в соответствии с восприятием бенефициаров и пониманием профессиональной области. Это означает, что прежде необходимо определить правильные данные. Семантические технологии предоставляют новые способы доступа к данным и получения знаний.



Изображение 64 Визуализация скрытого пространства DeepSDF с использованием t-SNE

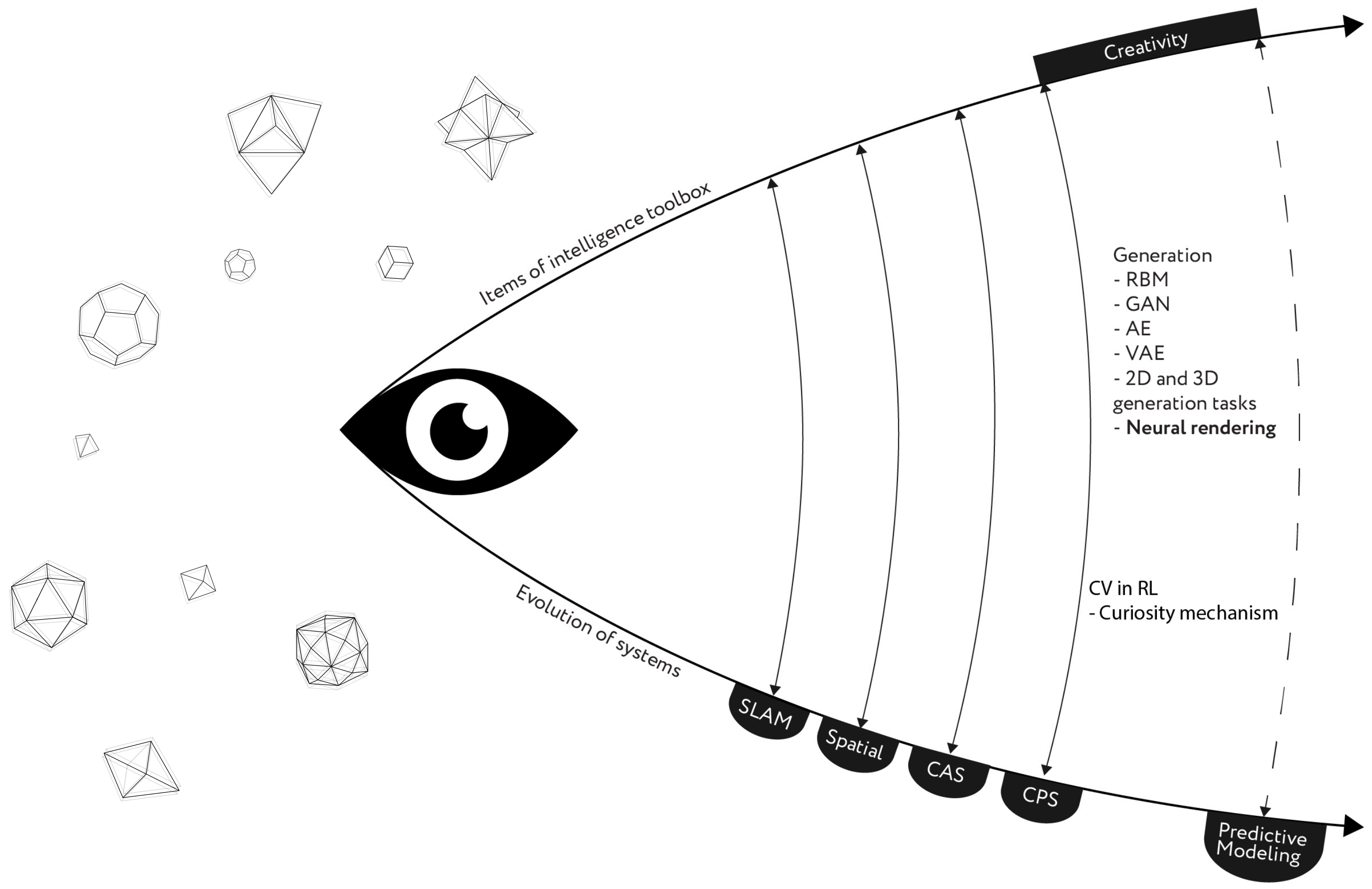
2D-3D СЕМАНТИЧЕСКИЙ НАБОР ДАННЫХ

2D-3D семантический набор данных - предоставляет множество взаимно зарегистрированных модальностей из 2D, 2.5D и 3D доменов с семантическими и геометрическими аннотациями на уровне экземпляра. Он покрывает более 6000 м2 и содержит более 70000 изображений RGB, а также соответствующие глубины, нормали поверхности, семантические аннотации, глобальные изображения XYZ (все в форме как обычных, так и равнопрямоугольных изображений на 360 °), а также информацию о камере. Он также включает зарегистрированные необработанные и семантически аннотированные 3D-сетки и облака точек. Кроме того, семантический набор данных содержит необработанные изображения RGB и глубины, а также соответствующую информацию о камерах для каждого места сканирования. Семантический набор данных позволяет разрабатывать совместные и кросс-модальные модели обучения и потенциально неконтролируемые подходы с использованием закономерностей, присущих крупномасштабным внутренним пространствам.



Изображение 65 Совместные 2D-3D-семантические данные для понимания сцен в помещении (Iro Armen, S.A. Alexander Saks, Amir R. Zamir, Silvio Savarese, 2017).

КРЕАТИВНОСТЬ

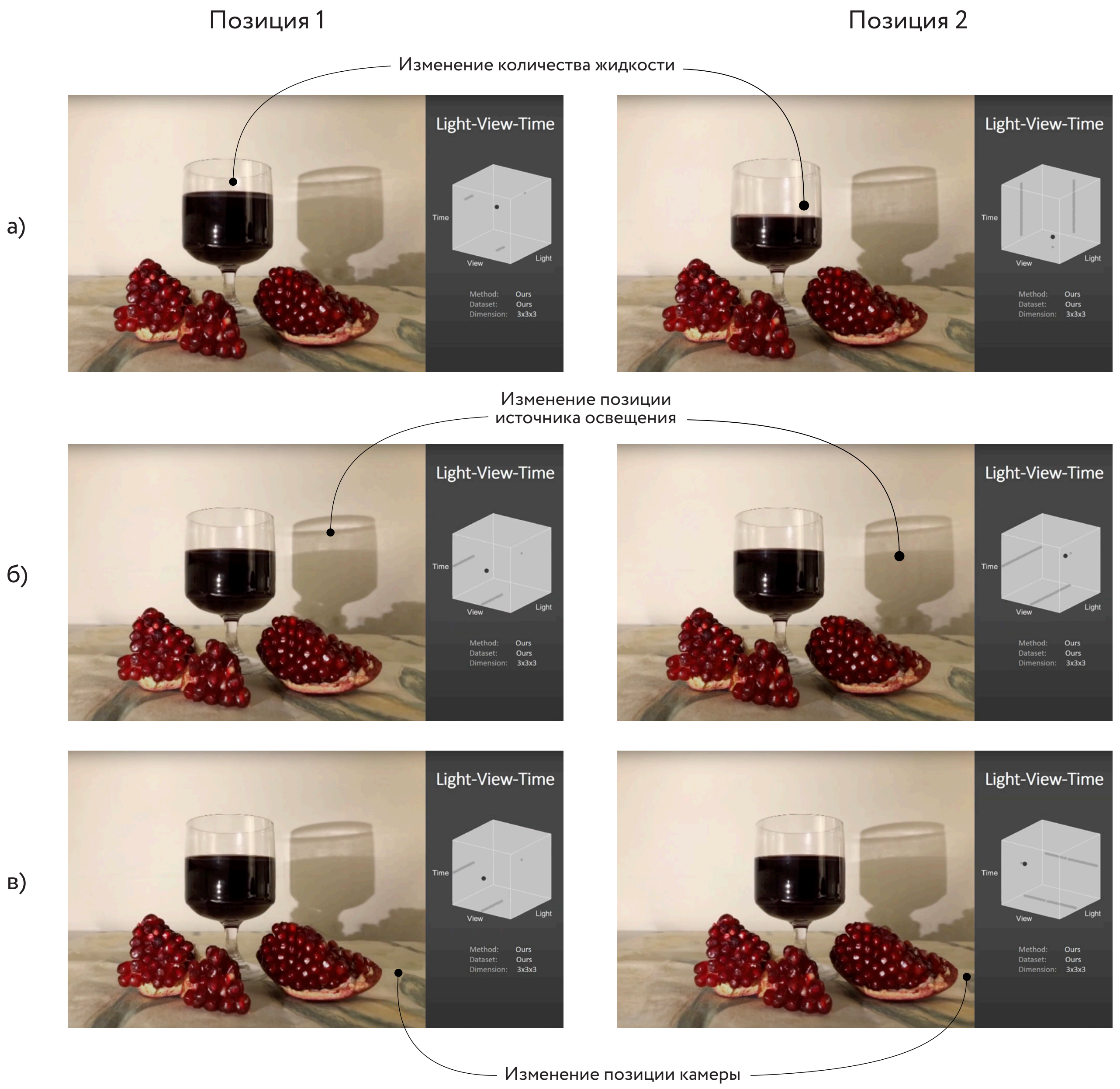


НЕЙРОННЫЙ РЕНДЕРИНГ

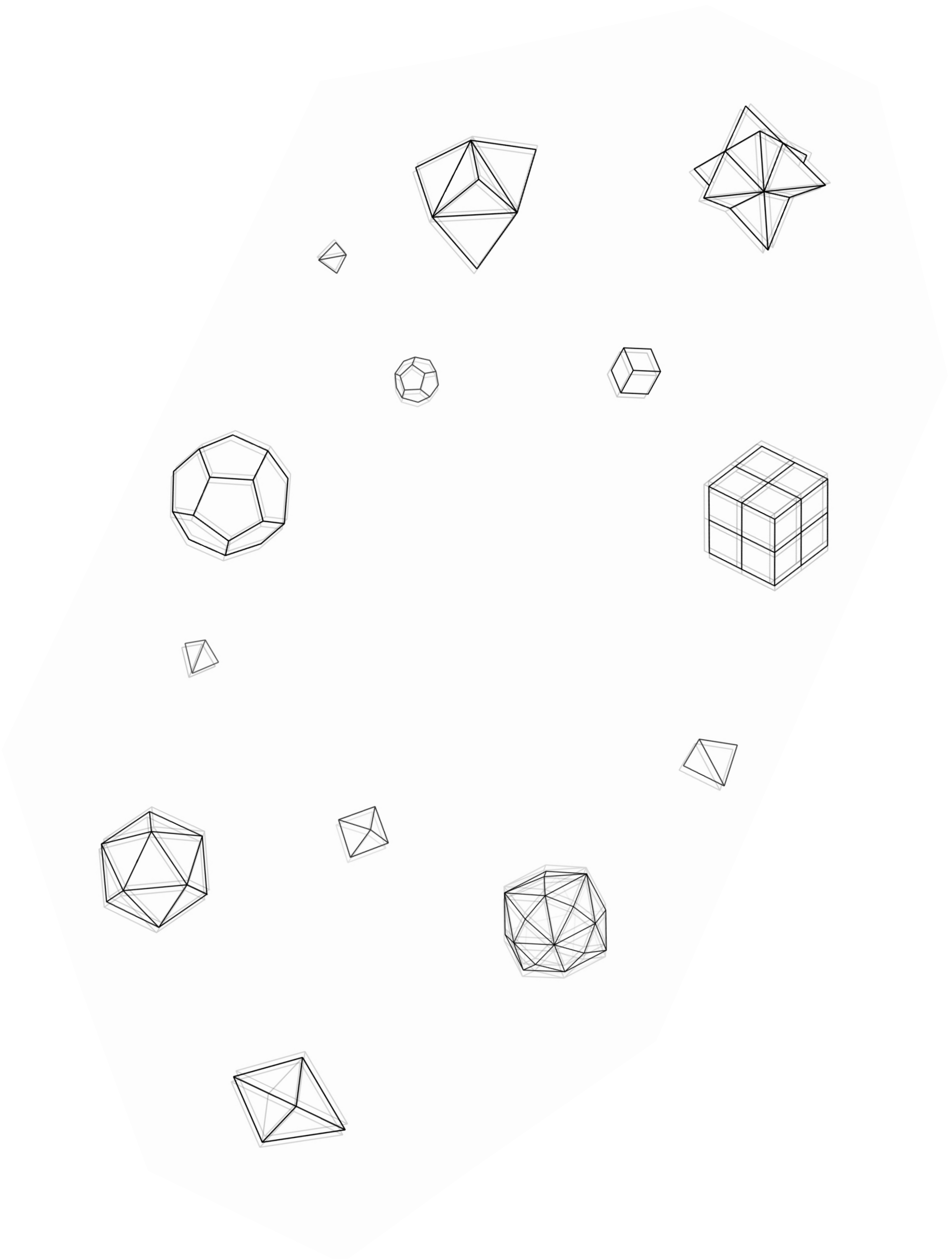
Нейронный рендеринг - это новая и быстро развивающаяся область, сочетающая методы генеративного машинного обучения с физическими знаниями из компьютерной графики, например, путем интеграции дифференцируемого рендеринга в обучение нейронной сети.

В то время как классическая компьютерная графика начинается с точки зрения физики, моделируя, например, геометрию, свойства поверхности и камеры, машинное обучение исходит из статистической точки зрения, то есть обучения на реальных примерах для создания новых изображений. С этой целью качество изображений, созданных с помощью компьютерной графики, зависит от физической корректности используемых моделей, в то время как качество подходов к машинному обучению в основном зависит от тщательно разработанных моделей машинного обучения и качества используемых обучающих данных. Явная реконструкция свойств сцены сложна и подвержена ошибкам и приводит к артефактам. С этой целью методы рендеринга на основе нейросетевого рендеринга пытаются преодолеть эти проблемы. Но в сложных сценах эти методы показывают артефакты, такие как швы или “лики”. Нейронный рендеринг обещает решить как реконструкцию, так и рендеринг с помощью глубокого обучения для изучения отображения изображений в новый домен. Нейронный рендеринг сочетает в себе физические знания, например математические модели проекции, которые получены в ходе обучения моделей, чтобы создать новые и мощные алгоритмы

Нейронный рендеринг еще не получил четкого определения в литературе, но его **определяют как** - подходы на основе глубокого обучения к генерации изображений или видео, которые позволяют явно или неявно управлять свойствами сцены, такими как освещение, параметры камеры, размещение объектов, геометрия, внешний вид и семантическая структура.



Изображение 6 X-Field набор 2D-изображений, снятых для разных ракурсов, времени или условий освещения, например, видео, светового поля, полей отражения или их комбинаций - путем обучения нейронной сети (NN) отображать их вид, координаты времени или света для 2D-изображений.



PHYGITALISM

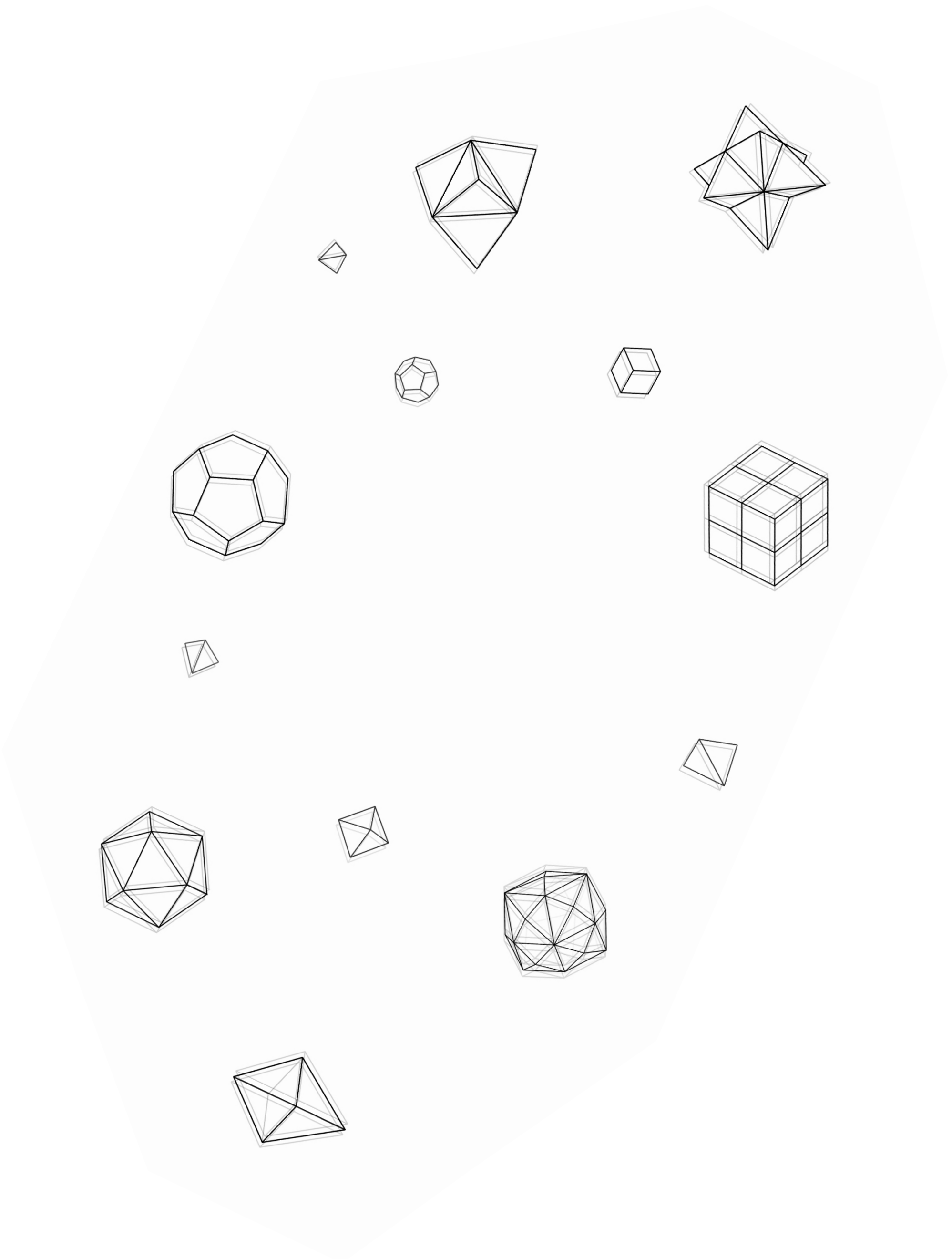
PHYGITALISM – компания разработчик визуальных решений на основе искусственного интеллекта, дополненной и виртуальной реальности.

Наша деятельность лежит в основе термина **Phygital = Physical + Digital**. Это означает взаимопроникновение цифрового и физического миров в единое пространство, в котором мы уже живем. Мы разрабатываем phygital-продукты для бизнеса, науки и искусства. Внутри компании организована исследовательская лаборатория, мы вкладываем силы и время в разработки на стыке 3D и ML технологий.

Название компании PHYGITALISM звучит как направление в искусстве. Мы верим, что phygital вдохновит многих художников на воплощение в жизнь своих захватывающих идей через существующий физический мир и технологии.

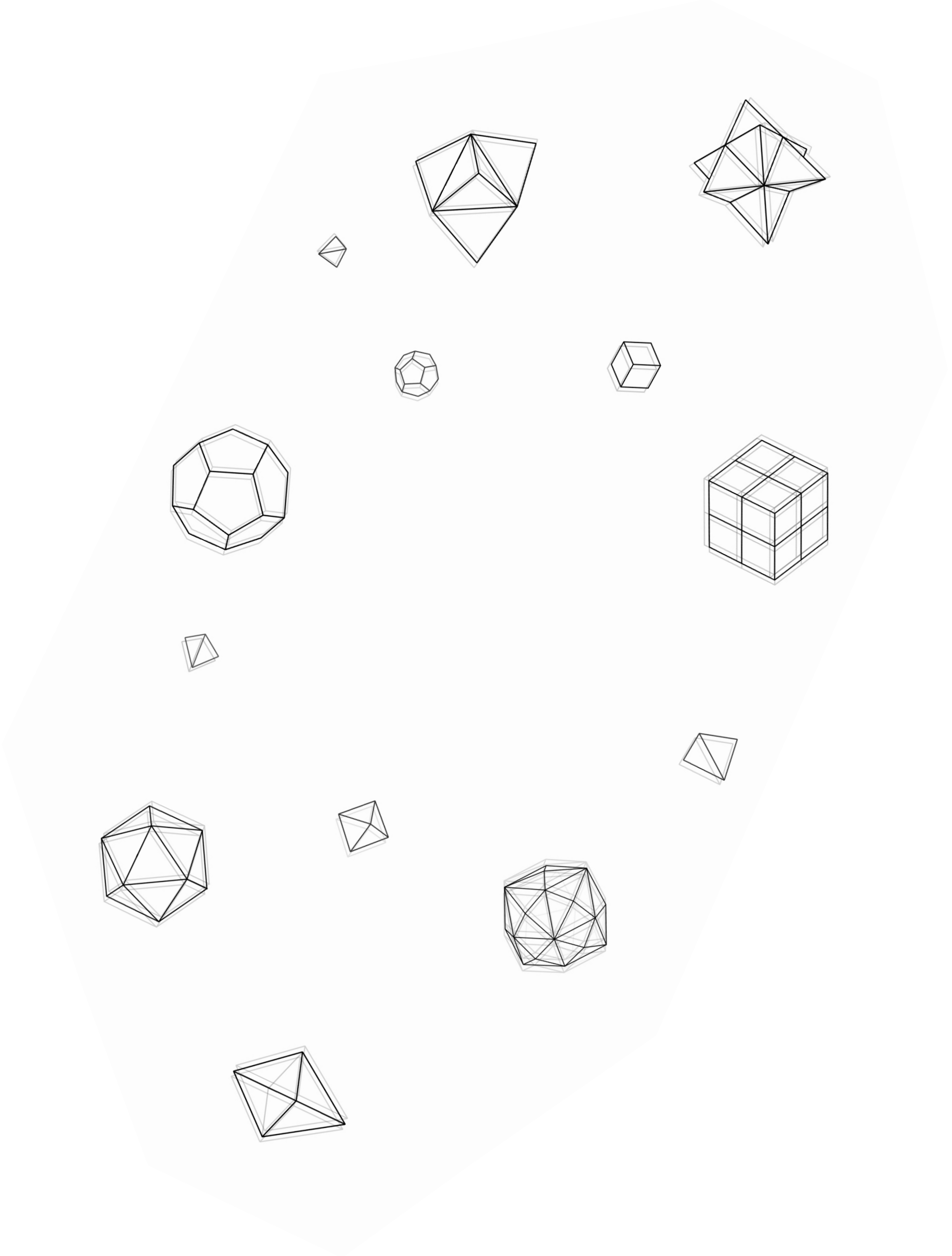
Узнайте больше на **www.phygitalism.com.**

Данное исследование было выполнено исследовательской командой - Камила Юсупова, Вадим Кондаратцев, Олег Юсупов.



АКРОНИМЫ

2D (Two-Dimension) - Двухмерный;
3D (Three-Dimension) - Трехмерный;
CAS (Context Aware System) - Контекстно-зависимая система;
CNN (Convolutional Neural Network) - Свёрточная нейронная сеть;
CPS (Cyber-Physical Systems) - Кибер-физические системы;
CV (Computer Vision) - Компьютерное зрение;
FOV (Field of View) - Поле зрения;
GPU (Graphics Processing Unit) - Графический процессор;
IMU (Inertial Measurement Unit) - Инерциальная единица измерения;
IR (Infra-Red) - Инфракрасный
LiDAR (Light Detection And Ranging) - Система световой идентификации, обнаружения и определения дальности;
PCL (Point Cloud Library) - Библиотека облаков точек;
RADAR (Radio Detection and Ranging) - Радиолокационная станция;
SDK (Software Development Kit) - Комплект для разработки программного обеспечения
SfM (Structure-from-Motion) - Структура по движению;
SI (International System) - Международная система;
SLAM (Simultaneous Localization and Mapping) - Одновременная локализация и отображение;
SNR (Signal-to-Noise Ratio) - Соотношение сигнал-шум;
TOF (Time-of-Flight) - Время полета;
VIO (Visual Inertial Odometry) - Визуальная инерционная одометрия.



ПОЛЕЗНЫЕ ССЫЛКИ

Будьте в курсе Phygital - подписывайтесь на наши телеграм каналы:

PHYGITAL ART

Современное искусство и технологии. Здесь следят за необычными идеями и их воплощением.

PHYGITAL GRAPHICS

Компьютерная графика и кто ее делает. Здесь рассказывают про особенность работы с фиджитал проектами и показывают внутреннюю кухню фиджитал студии в Москве.

PHYGITAL MARKETING

Коммуникации и технологии. Здесь следят за необычными способами общения брендов со своей аудиторией.