



ТЕХНИЧЕСКИ УНИВЕРСИТЕТ – СОФИЯ

**Факултет по Телекомуникации
Катедра „Комуникационни мрежи“**

Маг. инж. Радостина Росенова Петкова

МЕТОДИ И АЛГОРИТМИ ЗА РЕАЛИЗИРАНЕ НА ХОЛОГРАФСКА КОМУНИКАЦИЯ СЪС СВРЪХ-НИСКА ЛАТЕНТНОСТ

А В Т О Р Е Ф Е Р А Т

на дисертация за придобиване на образователна и научна степен
"ДОКТОР"

Област: 5. Технически науки

Професионално направление: 5.3. Комуникационна и компютърна техника

Научна специалност: Комуникационни мрежи и системи

Научен ръководител: Проф. д.т.н. Владимир Пулков

СОФИЯ, 2026 г.

Дисертационният труд е обсъден и насочен за защита от Катедрения съвет на катедра „Комуникационни мрежи“ към Факултет по Телекомуникации на ТУ-София на редовно заседание, проведено на 20.04.2026 г.

Публичната защита на дисертационния труд ще се състои на 08.07.2026 г. (сряда) от 13:00 часа в Конферентната зала на БИЦ на Технически университет – София на открито заседание на научното жури, определено със заповед № ОЖ-5.3-38 / 04.05.2026 г. на Ректора на ТУ-София в състав:

1. Проф. д-р Георги Илиев – председател
2. Доц. д-р Георги Балабанов – научен секретар
5. Доц. д-р Николай Атанасов
3. Проф. д-р Емилия Димитрова
4. Проф. д-р Валентина Маркова

Рецензенти:

1. Доц. д-р Георги Балабанов
2. Проф. д-р Емилия Димитрова

Материалите по защитата са на разположение на интересуващите се в канцеларията на Факултет по Телекомуникации на ТУ-София, блок 1, кабинет 1439-Б.

Дисертантът е докторант към катедра „Комуникационни мрежи“ на Факултет по Телекомуникации. Изследванията по дисертационната разработка са направени от автора, като някои от тях са подкрепени от научноизследователски проекти.

Автор: маг. инж. Радостина Петкова

Заглавие: Методи и алгоритми за реализиране на холографска комуникация със свръх-ниска латентност

Тираж: 20 броя

Отпечатано в ИПК на Технически университет – София

I. ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИОННИЯ ТРУД

Актуалност на проблема

Всяко следващо поколение мобилни мрежи предлага на своите потребители по-високи скорости, по-ниска латентност, по-голяма плътност на свързаност и повишена надеждност. Успоредно с това се създават предпоставки за възникването на качествено нови комуникационни парадигми. След прехода от гласови услуги към видеоразговори настоящият технологичен фокус е насочен към реализирането на пространствено-обогатена комуникация, чиято крайна форма се асоциира с т.нар. холографски тип комуникация (Holographic-Type Communication, НТС).

НТС е термин, представящ идеята за революционна технология, която реализира предаването на холографски данни в реално време през комуникационната мрежа. Поради своята същност да осигурява заснемане, предаване и възпроизвеждане на информация, произтичаща от петте основни човешки сетива, тя обещава да предложи много по-висока степен на ангажираност между потребителите в сравнение с възможните към момента видеоконферентни разговори. Отвъд ролята си на пореден метод за комуникация, множество професионални области, включително образованието, здравеопазването, индустрията и други, заявяват намеренията си да внедрят НТС с цел подобряване на функционалността и достъпността на предлаганите от тях услуги. Освен това се оказва, че НТС е ключов фактор за реализирането и развитието на т.нар. метавселена [1], [2], в която не само хората, но и цели физически системи биват пренесени в единна глобална, дигитална и паралелна действителност.

Цел на дисертационния труд, основни задачи и методи за изследване

Дисертационният труд има за цел да представи особеностите при реализацията на НТС чрез анализ на съществуващите предизвикателства и изследване на възможните подходи за тяхното преодоляване.

Задачите, обхващащи изпълнението на целта на представения труд, са следните:

1. Изследване функционалните компоненти на НТС системите, наричани още системи за телеприсъствие (Holographic Telepresence Systems, НТРС), и идентифициране на основните предизвикателства пред тяхната реализация. Изследване на наличните реализации на подобен тип системи.
2. Представяне на архитектурна рамка за практическа реализация на НТРС, последвано от нейната приложна имплементация. Внедряване на алгоритми за обработка на тримерни (Three Dimensional, 3D) визуални данни в разработената НТРС и оценка на нейната работа.
3. Анализирание ролята на бъдещите поколения мобилни мрежи в реализирането и внедряването на НТС за масова употреба.
4. Разработване на алгоритми за ефективна обработка на 3D визуални данни, съобразени с изискването за свръх-ниска латентност и демонстриращи идеята за интелигентно проектиране на НТРС.

Методологията на изследванията в дисертацията следва комплексен научен подход, съчетаващ аналитични, експериментални и моделиращи методи. Приложен е системен анализ с цел идентифициране на ключови компоненти и предизвикателства пред НТРС. Извършена е експериментална реализация и оценка на НТРС. Оценено е влиянието на комуникационните мрежи чрез анализ на ключовите показатели за ефективност и на потенциалните подходи за решаване на основните НТС предизвикателства. В рамките на

дисертацията са приложени методи за разработка на алгоритми, базирани на невронни мрежи, последвани от експериментална валидация.

Научна новост

В дисертационния труд са идентифицирани основните НТС предизвикателства и са дефинирани потенциалните подходи за тяхното преодоляване. Разработени са алгоритми на изкуствен интелект за регистрация и компресия на 3D визуални данни, чийто резултати дефинират следните открития. Първо, регистрацията на облаци от точки, представящи един обект, заснет от различни изгледи, е постижима в абстрактното латентно пространство на техните компресирани представяния. Второ, ефективността на компресия на облаци от точки, представящи един обект, заснет от различни изгледи, базирана на разпределено кодиране на източника, се повивава с увеличаване на корелацията между изгледите.

Практическа приложимост

Дисертационният труд представя основните функционални компоненти на НTPS и предлага модулен архитектурен подход за последваща практическа реализация, които да послужат като насоки за работа в областта. Също така, имплементираната НTPS има за цел да послужи като платформа за интегриране и изследване на различни подходи и алгоритми за реализация на НТС. Накрая, представените алгоритми за регистрация и компресия на облаци от точки се явяват потенциални решения за някои от изпитаните предизвикателства.

Публикации

Основните постижения и резултати от дисертацията са разпространени чрез 8 научни статии, 3 от които са публикувани в престижни международни научни списания, а останалите 5 са представени на международни научни конференции.

Международните списания са: *MDPI Sensors 2022 & 2023* и *IEEE Access 2024*.

Международните конференции са: IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom) 2022, IEEE International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST) 2025, IEEE International Symposium on Wireless Personal Multimedia Communications (WPMC) 2025, Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering (ECTI DAMT & NCON) 2026.

Структура и обем на дисертационния труд

Дисертационният труд е в обем от **136** страници, като включва въведение, **5** глави за решаване на формулираните задачи, списък на приносите, списък на авторските публикации и използвана литература. Цитирани са общо **176** литературни източника, като **163** са на латиница, а останалите са интернет адреси. Работата включва общо **40** фигури и **15** таблици. Забелязани са **54** цитирания на осемте авторски публикации. Настоящата номерация съответства на тази в дисертационния труд.

В първа глава е предложена блокова диаграма на НTPS, като са дефинирани основните функционални компоненти на системата, и са идентифицирани най-важните предизвикателства пред тяхното реализиране. Представени и сравнени са съществуващите към момента НTPS.

Втора глава предлага подход за практическа реализация на НTPS, основан на модулен архитектурен дизайн. Имплементирана и оценена е реално разработена НTPS, следвайки предложения подход. Представени и оценени са алгоритми за обработка на 3D визуални данни, внедрени в практически реализираната система.

Трета глава разглежда влиянието на петото поколение мобилни мрежи (Fifth Generation Mobile Networks, 5G), както и потенциала на бъдещото шесто поколение мобилни мрежи (Sixth Generation Mobile Networks, 6G) в реализацията и внедряването на НТС и оспорва безпрецедентния им потенциал за сметка на проектиране и интегриране на интелигентни методи и алгоритми за обработка на данните.

В четвърта глава са представени разработени алгоритми за регистрацията на 3D обекти, наложителна при заснемането им от множество изгледи.

Пета глава представя разработени алгоритми за компресия на 3D визуални данни, необходима поради огромния обем данни, генериран при създаването на дигитално 3D съдържание.

II. СЪДЪРЖАНИЕ НА ДИСЕРТАЦИОННИЯ ТРУД

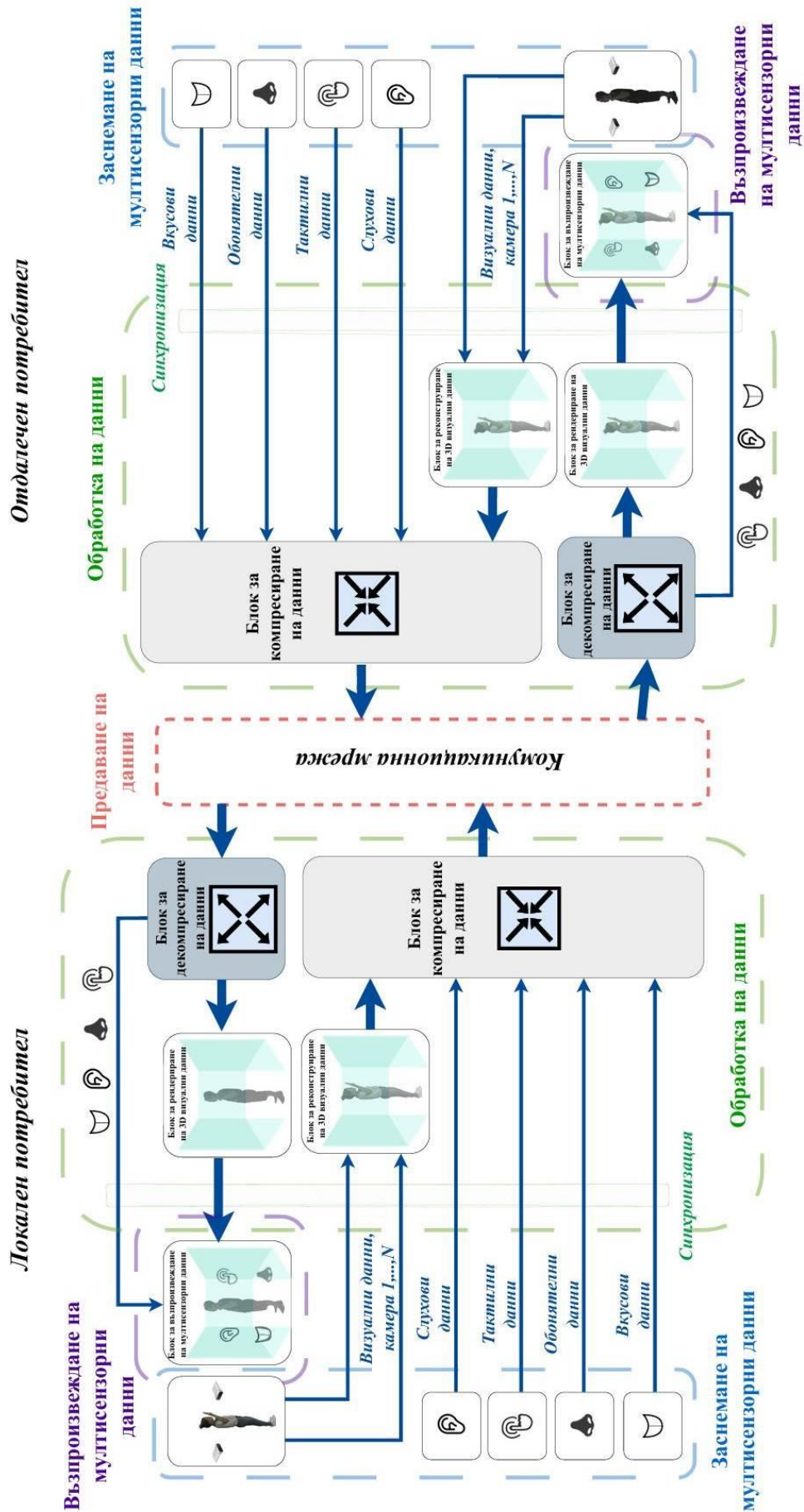
ГЛАВА 1. ОСНОВНИ ПРЕДИЗВИКАТЕЛСТВА ПРЕД РЕАЛИЗАЦИЯТА НА HTTPS

1.1. Функционална диаграма на HTTPS

Всяка система за телеприсъствие трябва да изпълнява множество операции, за да позволи на даден потребител да бъде участник в НТС. Тези операции могат да бъдат групирани в четири основни функционални блока [A1], визуализирани на Фигура 1.1 и оградени с пунктир. Първият функционален блок обхваща заснемането на мултисензорни данни, като в идеалния случай се събира информация, произтичаща от петте основни човешки сетива, в т.ч. визуално, слухово, тактилно, обонятелно и вкусово. Нека се има предвид, че за осигуряване на висока степен на реалистичност при заснемането на визуални данни е наложително използването на множество камери, които да обхванат дадената сцена или обект от различни изгледи. Вторият функционален блок обединява процесите по обработка на данните, включващи най-вече реконструкцията на 3D обекти и сцени и 3D рендерирането. Други по-специфични процеси могат да бъдат компресия, регистрация, сегментация, филтрация на шума и т.н. Синхронизацията на отделните потоци от данни също е част от този блок. Третият блок включва функции, свързани с предаването на данните по комуникационния канал и пренасянето им по мрежата. Накрая, след като потребителят получи и обработи холографските данни, четвъртият функционален блок отговаря за тяхното синхронно възпроизвеждане пред потребителя. Следва да се отбележи, че даден потребител обикновено играе ролята както на предавател, така и на приемник в комуникацията, поради което същият трябва да поддържа изпълнението на всички изброени функции.

1.2. Предизвикателства пред реализирането на HTTPS

Идентифицираните предизвикателствата пред реализацията на НТС са разпределени в три групи, а именно основни технически предизвикателства, предизвикателства, свързани с виртуалното представяне на участниците и други предизвикателства [A1]. Основните технически предизвикателства обединяват предизвикателствата, свързани с входно-изходните (Input/Output, I/O) технологии за НТС, с обработката на холографските данни, с предаването на холографските данни и с мащабируемостта на HTTPS. Предизвикателствата, свързани с виртуалното представяне на участниците включват проблеми, отнасящи се до създаването на автентични 3D виртуални аватари, и проблеми, свързани с поддържането на жестове, поглед и емоции. Останалите предизвикателства се състоят в необходимостта от надеждна оценка на HTTPS, вземаща под внимание както качеството на обслужване (Quality of Service, QoS), така и качеството на възприятие (Quality of Experience, QoE), и в гарантирането на висока сигурност, поверителност и етичност на комуникацията. Предизвикателствата са представени в Таблицы 1.1 и 1.2.



Фигура 1.1: Функционална блокова схема на система за холографска комуникация [A1].

Таблица 1.1: Основни технически предизвикателства

Категория	Предизвикателство
I/O технология	Дълбочинни камери: ниска прецизност, малка резолюция, ограничен обхват, добавяне на шум, тясно зрително поле (Field of View, FOV) [71]
	Множество камери: необходимост от калибрация, ниска прецизност на реконструкция, времемко, трудно инсталиране;
	360° заснемане: ограничени движения на потребителя [9];
	Усложнена обработка и предаване при висока кадрова честота;
	Очила (Head Mounted Display, HMDs): по-ниска резолюция и по-тясно FOV от това на човешката зрителна система [72];
	HMDs: ограничен живот на батерията, неудобни, инвазивни, висока цена, изисквания към изчислителния хардуер [12];
	HMD: конфликт конвергенция-акомодация, причинявайки гадене, замайване, дезориентация [10], [11], [14];
	Социална приемственост на HMD [13], ограничено използване до един потребител;
	Дисплеи със светлинно поле и обемни дисплеи: ограничен обем на панела – ограничена дълбочина, голямо количество данни [12], [14], [15];
	Предоставяне на потапящо преживяване, подкрепено от висока резолюция, голямо FOV, интерактивност [11];
	Необходимост от мултисензорни взаимодействия [18], [20], [24] – [26];
	Необходимост от последователност на видео и аудио (необходимост от пространствено аудио) [18] – [20];
	Интегриране на тактиленост в системите за HTC [21] – [23];
	Ограничено включване на мирис и вкус, висока цена за интегрирането на стай за сензитивна реалност.
Обработка на данни	Обработка на голямо количество данни;
	Създаване на висококачествена реконструкция (при заснемане с няколко камери);
	Прилагане на ефективни техники за компресия/декомпресия;
	Изискване за голяма изчислителна мощност;
	Компромис между латентността от обработката и латентността при предаване по мрежата [14], [28], [32], [36] – [38];
	Интегриране на edge computing [36], [40], [41];
Предаване на данни	Синхронизация между локалните сигнали и тези, пристигащи от отдалечения потребител [37], [39], [40].
	Необходимост от по-високи скорости, следователно по-високи честотни обхвати и ефективни техники за модулация;
	Прилагане на адаптивно предаване на данните [43], [44];
	Необходимост от семантично знание и прогнозиране посоката на погледа [36];
	Предаване с ниска латентност < 15 ms "motion-to-photon" и < от 50 – 100 ms от край до край [9], [45] – [52];
	Разработване на оптимизиран механизъм за предаване на холографски данни [37];
Мащабируемост	Оптимизиране на структурата на мрежата [37], [62] – [64].
	Нарастващи технологични изисквания – по-висока честотна лента и синхронизация между много потребители;
	Необходимост от централизиран контрол [44];
	Мащабируемост относно: различни устройства, различен брой участници, различна степен на виртуалност [65].

Таблица 1.2: Предизвикателства, свързани с виртуалното представяне на участниците в комуникацията, оценяването на HTPS, сигурността и поверителността

Категория	Предизвикателство
Представяне чрез аватар	Осигуряване на усещане за присъствие и реализъм, детайлност на лицата на аватара и изразените емоции;
	Завършеност на телата на аватара и тяхната позиция, вярност на движенията и жестовете;
	Синхронизиране на движенията на аватара с тези на потребителя;
	Чувство за собственост на тялото и разпознаване на собствените движенията и местоположение;
	Сътрудничество в пространството на взаимодействие [2];
	Необходимост от визуална и поведенческа точност на аватара [17], [19], [20], [26], [46], [57], [59], [68], [74] – [77].
Поддържане на жестове	Динамично проследяване на тялото и пресъздаване на жестовете и движенията му чрез виртуалния аватар [19];
	Калибриране и синхронизация между различни потребителски координатни системи.
Поддържане на поглед	Динамично и точно проследяване на погледа;
	Оценка и визуализация на посоката на погледа [78].
Поддържане на емоции	Динамично заснемане и проследяване на лица, синхронизиране движението на устните с говора;
	Закриване на лицата поради носенето на HMDs [79] – [84].
Оценка на HTPS	QoS - недостатъчно за оценка на цялостното представяне на системата, изпитвано от крайните потребители;
	OoE - съвкупност от множество фактори, влияещи върху цялостното представяне на системата [88], [89];
	Оценка на различни HTPS аспекти като реализъм, приложимост, време за изпълнение на поставена задача и т.н;
	Оценяване от много голям брой участници, изпълняващи различни задачи при променливи условия;
	Нови показатели за оценка, съвместими с 3D формата на данни [90];
	Разработване на модели за оценка на OoE чрез машинно обучение (Machine Learning, ML) и изкуствен интелект (Artificial Intelligence, AI) [89], [91].
Защита на данните	Пренебрегване повишаването на сигурността и поверителността в системите на HTC [92];
	Необходимост от защита на 3D данните, представлящи човешки тела, обекти и околна среда;
	Необходимост от защита на чувствителни биометрични данни и данни за поведенчески характеристики [2];
	Необходимост от защита на специфични за приложението данни;
	Подобряване на сигурността на ниво мрежа [95];
	Повишаване на потребителското доверие в различни HTC приложения.

1.3. Литературен обзор на реализирани HTPS

Идентифицираните предизвикателства се потвърждават допълнително от анализа на редица публикации, свързани с практически реализации на HTPS. Системите са сравнени по отношение на използваните I/O технологии за заснемане и възпроизвеждане на данните, мрежовата технология за предаване на данните, мащабируемостта на системата, реализирания тип аватар, поддържането на невербални комуникационни сигнали,

наличието на оценка на системата и възможностите за защита на данните. Резултатите показват, че са интегрирани различни I/O устройства, предназначени основно за снемане и възпроизвеждане на аудио-визуални данни. Най-често се използват 3D камери и очила за добавена и виртуална реалност, осигуряващи най-голямо потапящо изживяване. Относно мрежовата част, не съществува унифициран мрежов протокол, оптимизиран за предаване на холографски данни. Поради тази причина са интегрирани технологии като MPEG-DASH, RTSP, WebRTC, RabbitMQ и PUN, първоначално разработени за различни цели. Макар и засегната в много от изследваните системи, мащабируемостта на HTTPS е ограничена само до няколко потребители. Относно виртуалното представяне на участниците предпочитаните фотореалистични аватари са по-рядко реализирани за сметка на тези, базирани на персонажи. Интегрирането на невербални комуникационни сигнали е осъществено в доста системи, като пресъздаването на жестовете е най-често застъпено. Поддържането на поглед и лицеви изражения остава ограничено. Оценка за работата на HTTPS е почти винаги налична. Тя обаче не зачита потребителския контекст и се основава на ограничен брой потребители. Според предоставените оценки на латентността, терминът „реално време“ е неприложим, тъй като различните НТС сценарии налагат различни изисквания. Следователно оценката трябва да бъде специфична за конкретното приложение и изпълняваните задачи. Спазване на изискването за свръх-ниска латентност като цяло си остава съществен проблем. Накрая, нито една от разгледаните HTTPS не реализира стратегии за гарантиране на сигурността и поверителността на потребителските данни.

1.4. Изводи

Реализацията на същинска НТС все още не е постигната поради наличието на предизвикателства от различен характер. Липсата на усъвършенствани и специализирани технологии за изпълнение на основните функционални компоненти на HTTPS налага внедряването на допълнителни алгоритми за обработка на данните, което от своя страна води до увеличение на латентността от край до край. Практическата имплементация на HTTPS представлява съвкупност от технологични и алгоритмични решения, които трябва да удовлетворяват основните изисквания на НТС.

1.5. Приноси към първа глава

- 1) Дефинирани са функционалните компоненти на HTTPS и са идентифицирани основните предизвикателства пред тяхното реализиране.
- 2) Анализирани и сравнени са съществуващите практически разработки в областта.

ГЛАВА 2. ПРАКТИЧЕСКА РЕАЛИЗАЦИЯ НА HTTPS

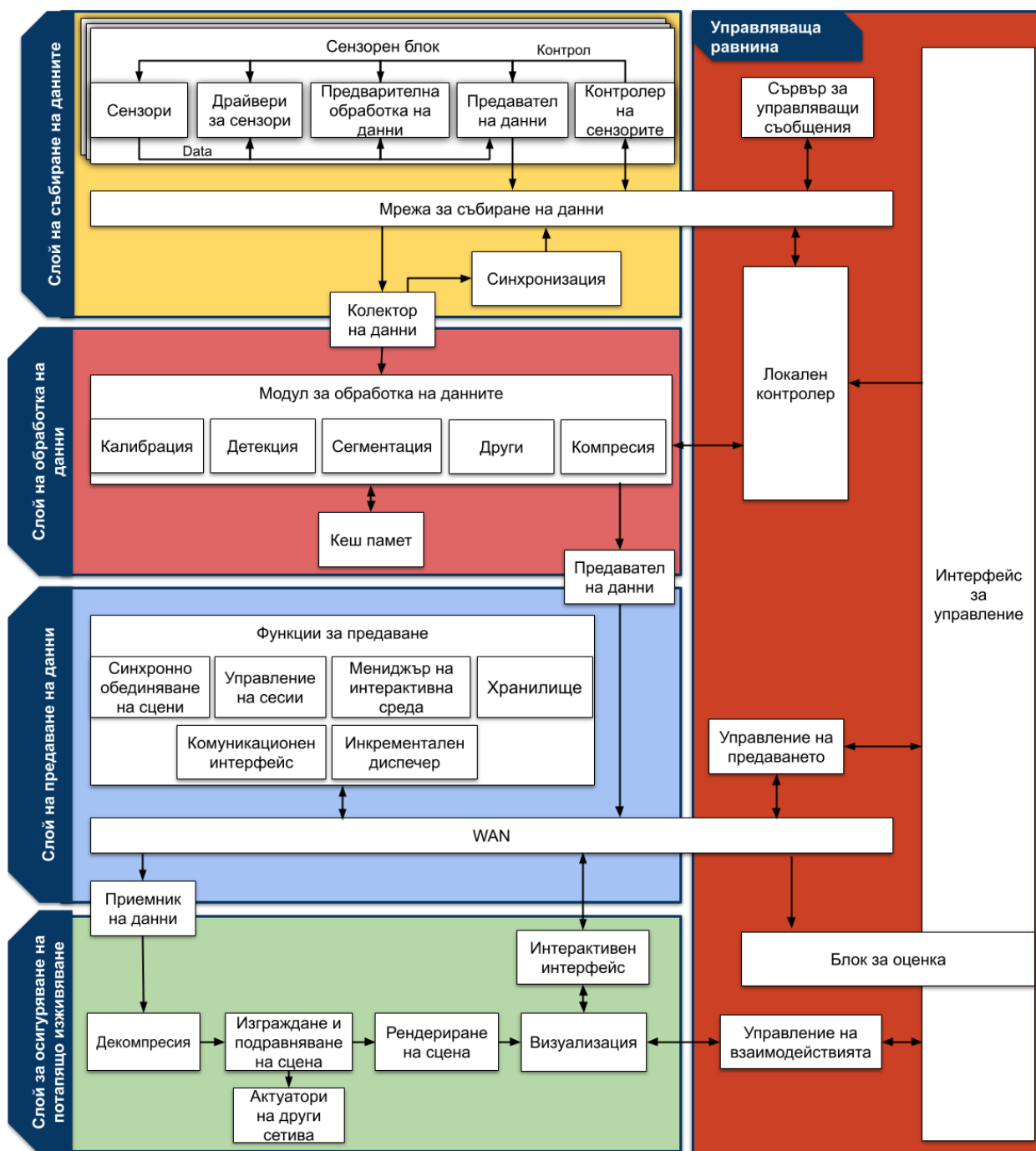
2.1. Модулен архитектурен дизайн на HTTPS

Предложен е архитектурен подход за реализация на HTTPS [A2], фокусиран върху осигуряването на модулност, гъвкавост, адаптивност и лесна интеграция на необходимите НТС технологии. Дефинирани са четири основни слоя, всеки от които е предназначен за изпълнението на конкретна функция:

- Слой за придобиване на данни, отговарящ за първоначалното събиране на сурови сензорни данни и служещ като входна точка към платформата; Тук принадлежи и мрежата за събиране на данни (Data Acquisition Network, DAN);
- Слой за обработка на данни, приемащ входните данни и прилагащ различни алгоритми за тяхната манипулация, подобрене и превръщане във виртуален обект или сцена;
- Слой за предаване на данни, управляващ сигурната и ефективна доставка на обработените данни до всички възли, участващи в сесията;
- Слой за осигуряване на потапящо изживяване, където данните се трансформират в ориентирано към потребителя потапящо изживяване, оформяйки крайния холографски

результат чрез визуално рендериране, ангажиране на останалите сетива и осигуряване на взаимодействие с потребителя.

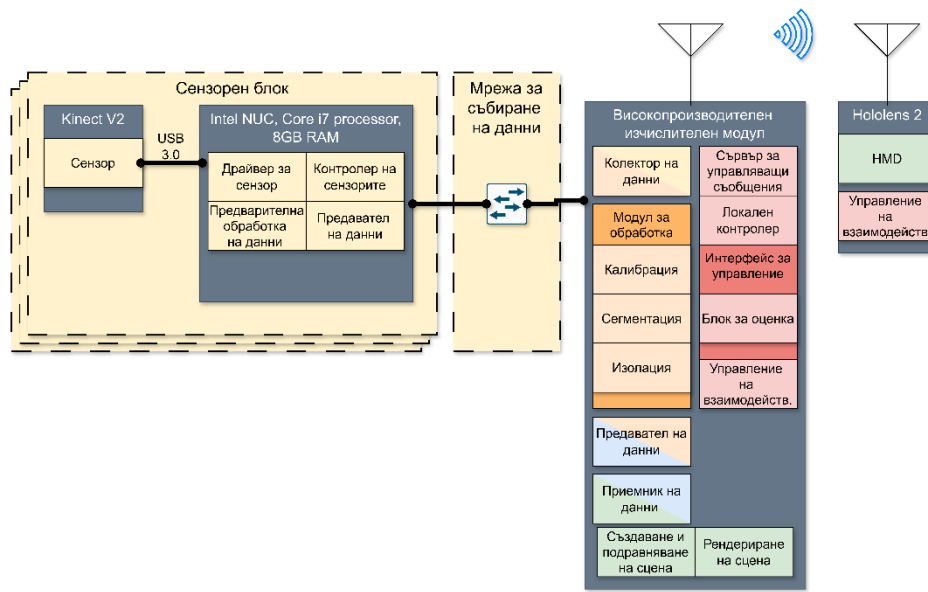
Введена е и управляваща равнина, чието предназначение е да координира и контролира работата на цялата система. Тя дава възможност за управление и оценка от страна на потребителя с цел оптимизиране на изследователските усилия и ускоряване процеса на разработка. Фигура 2.1 илюстрира дизайна на предложената архитектура. От нея се вижда, че всеки слой съдържа множество архитектурни блокове, представляващи основните елементи, необходими за функционирането на HTPS. Всеки от тях може да бъде имплементиран чрез различни избираеми и взаимозаменяеми модули в зависимост от изискванията на конкретния HTC сценарий. По този начин се цели както осигуряването на гъвкавост и адаптивност към новите технологии, така и на възможности за изследване и оценка представянето на системата в зависимост от конкретната ѝ реализация.



Фигура 2.1: Предложена архитектура за HTPS [A2].

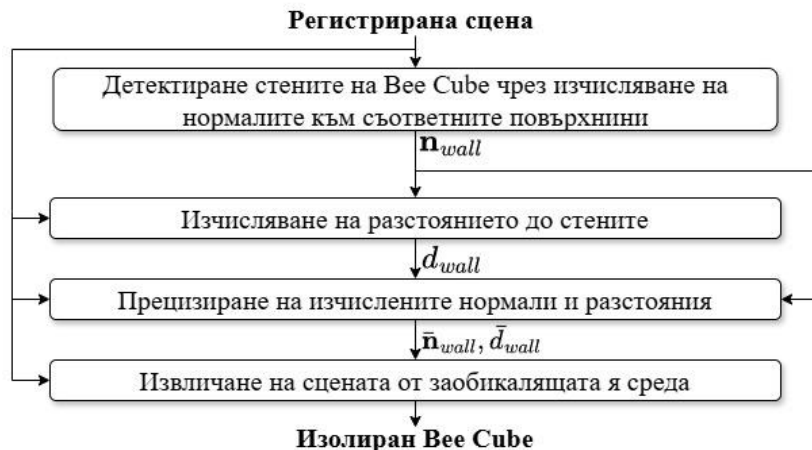
2.2. Имплементация на HTPS и внедрени алгоритми

За практическата реализация на HTPS върху наличната физическа инфраструктура са имплементирани модули, част от слоя на събиране на данни, слоя на обработка на данни, слоя за осигуряване на потапящо изживяване, както и модули от управляващата равнина (Фигура 2.2). Физическа инфраструктура се състои от три сензора Kinect V2, всеки от които е свързан към мини персонален компютър Intel NUC5i7RYH посредством USB 3.0 интерфейс. Трите мини компютъра от друга страна са свързани към високопроизводителна изчислителна единица (High-performance Computing Unit, HPCU) чрез гигабитова комутируема локална мрежа.



Фигура 2.2: Физическо разпределение на архитектурните блокове върху хардуерната инфраструктура на HTPS [A2]

За слоя на обработка на данни специално е проектиран и впоследствие интегриран алгоритъм за изолиране на реконструираната 3D сцена [A3]. HTPS заснема контролирано физическо пространство, ограничено от шест стени, формиращи хексагонална клетка и поради това наречено „Bee Cube“. Заедно с него обаче се заснемат и допълнителни, несъществени данни, налагащи изпълнението на подобен алгоритъм. Блокова диаграма е представена на Фигура 2.6, дефинираща изпълнението на няколко последователни стъпки.



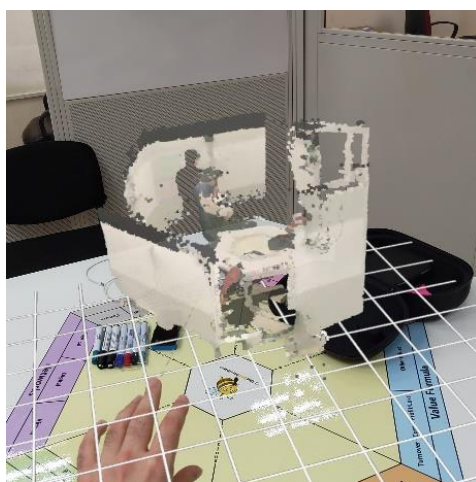
Фигура 2.6: Графично представяне на алгоритъма за извличане на сцена [A3]

2.3. Резултати от изпитване на системата

Въз основа на изчисления полезен товар на транспортния слой за един кадър, възлизаш на 3,256,634В, е установена максималната кадрова честота при едновременна работа на трите Kinect V2 сензора. Резултатите показват, че същата е ограничена до едва 8 fps от възможните 30. Характеристиките на избрания протокол на транспортно ниво не благоприятстват реализирането на подобен тип системи, допълнително утежнявайки натоварването и създавайки предпоставки за десинхронизация между сензорните блокове. Наблюдаваното отклонение от 27 ms при период на кадъра 125 ms, (21%) е в потвърждение на казаното. Изследваната изчислителна ефективност на онлайн обработките е недостатъчна за удовлетворяване на изискванията за минимална латентност. Това се дължи на комбинация от фактори, сред които сложността на алгоритмите, ограничената изчислителна мощност на НРСУ, използването на Python като интерпретиращ език и други. Освен това, наличието на продължителна офлайн фаза за предварително изчисление на определени параметри при фиксирана инсталация на Kinect V2, категорично ограничава реализирането на сценарии, при които се изисква мобилност. Количествени резултати, свързани с оценка на реализираната система са представени в Таблица 2.1. Фигура 2.15 визуализира реконструираната и изолирана след изпълнение на предложения алгоритъм за изолация сцена, наблюдавана през Hololens.

Таблица 2.1. Количествени резултати на различни показатели по отношение представянето на реализираната НTPS

Тип измерване	Стойност
Колектор на данни, Сензор fps=8, Период на девиация (ms)	27
Време за изчисление на офлайн етапа на калибрация (s)	161
Време за изчисление на онлайн етапа на калибрация (ms)	126
Време за изчисление на офлайн етапа на изолация (s)	3419
Време за изчисление на онлайн етапа на изолация (ms)	140
Време за изчисление на сегментацията (ms)	30



Фигура 2.15: Реконструираната сцена, наблюдавана през Hololens2

2.4. Изводи

Предаването на 3D данни е предизвикателство. За да бъде увеличена кадрова честота, съответно QoE, е необходимо да се интегрират алгоритми за компресия още в рамките на сензорния блок или да се използва DAN с по-голям комуникационен капацитет, което обаче е непрактично от потребителска гледна точка. Ефективността на обработката на данни също трябва да се подобри чрез разработването на „леки“ алгоритми, функциониращи във всевъзможни НТС сценарии, независещи от фиксираната инсталация на НTPS.

2.5. Приноси към втора глава

- 1) Прилагане на модулен архитектурен подход за реализиране на HTPS, включително хардуерна имплементация на системата.
- 2) Програмна реализация на архитектурни блокове от слоя на събиране на данни, слоя на обработка на данни и слоя за осигуряване на потапящо изживяване, а именно: колектор на данни, модул за обработка на данни и модул за изграждане и подравняване на сцена.
- 3) Внедряване на алгоритъм за калибрация и на алгоритъм за извличане на 3D сцена в разработения модул за обработка на данни.
- 4) Разработване на алгоритъм за извличане на 3D сцена от заобикалящата среда чрез детектиране на семантични 3D обекти.
- 5) Изследване и оценка на реализираната HTPS.

ГЛАВА 3. ВЛИЯНИЕ НА МОБИЛНИТЕ МРЕЖИ ВЪРХУ РЕАЛИЗАЦИЯТА НА НТС

Трета глава изследва влиянието на 5G и потенциала на 6G върху осъществяването на НТС и внедряването на HTPS сред потребителите. Тя се стреми да дефинира ясна взаимовръзка между основните НТС изисквания, а именно необходимост от голяма пропускателна способност и съответно високи скорости на предаване, минимална латентност, стриктна синхронизация и необходимост от прилагане на сложни и изчислително тежки алгоритми за обработка, и потенциалните подходи на мобилните мрежи, чрез които тези изисквания да бъдат посрещнати.

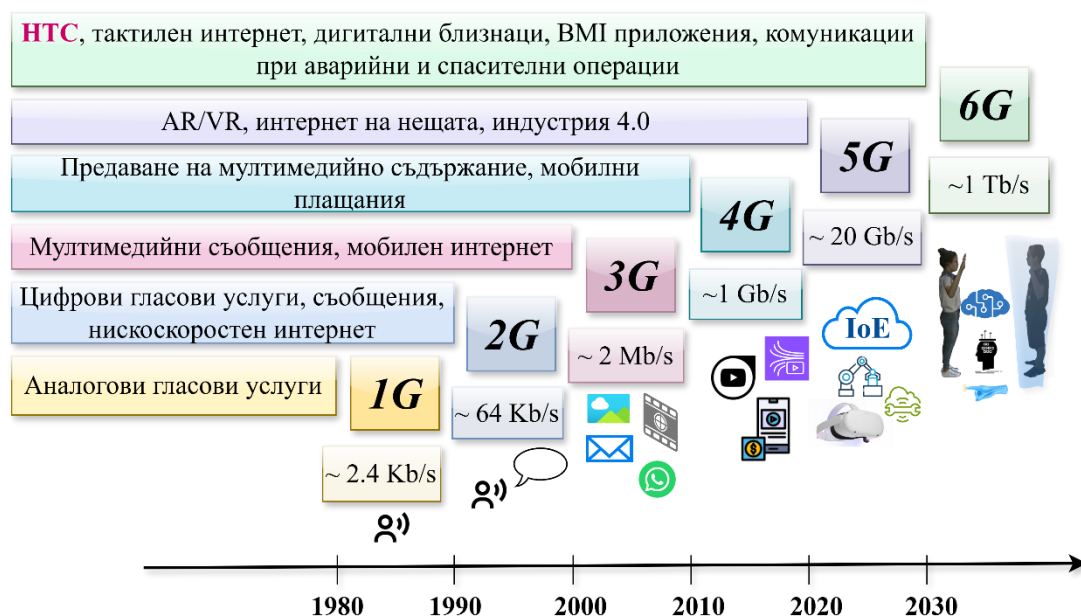
3.1. НТС и новите поколения мобилни мрежи

Удовлетворяването на споменатите технологични изисквания е фундаментално за реализирането на същинска НТС – със свръх-ниска латентност и висока степен на QoE [129], [130]. Оказва се обаче, че преносът на холографско съдържание в 5G мрежите е силно ограничен поради невъзможността на настоящите технологии да посрещнат тези изисквания. Сведения за това са предоставени в Таблица 3.2, която сравнява изискванията на НТС и съответните ключови показатели за работоспособност на мрежата (Key Performance Indicators, KPIs), дефинирани за 5G и 6G. Очевидно е същественото увеличение в скоростите на предаване при 6G, което създава предпоставки за преноса на значително по-голям обем данни спрямо 5G. Наред с това, редуцирането на латентността и джитера, в съчетание с повишената надеждност, осигурява предимство както по отношение на по-добрата синхронизация, така и за успешното реализиране на високоинтерактивни сценарии. Това обосновава схващането, че именно 6G е перспективното решение за нейната същинската реализация на НТС.

След провеждане на аналитичен обзор на редица научни публикации с цел изследване мястото на НТС в бъдещите 6G комуникации, са изведени следните твърдения. Очаква се холографските технологии и комуникации да заемат централна роля в социален и професионален контекст. Почти всички разгледани публикации изрично подчертават НТС като ключово приложение, което предстои да бъде реализирано благодарение на напредъка в 6G технологиите. Нещо повече, декларирана е появата на нов тип 6G услуга, която е едновременно широколентова, с висока надеждност и свръх-ниска латентност. В този контекст Фигура 3.1 илюстрира еволюцията на поколенията мобилни мрежи по отношение на скоростите на предаване на данни и предоставяните услуги.

Таблица 3.2: Сравнение на основните НТС изисквания със съответните KPIs в 5G и 6G мрежи

Параметър	НТС изискване	5G	6G
Скорост на предаване	0.5-2 Gb/s Облак от точки; 0.1-2 Tb/s холограма, пресъздадена от множество последователни изображения [37]	0.02 Tb/s Пикова скорост; 0.1 Gb/s Скорост, изпитана от потребителя	1 Tb/s Пикова скорост; 10 Gb/s Скорост, изпитана от потребителя
Латентност	Латентност от край-до-край: 50-100 ms (буфериране); 20-50 ms (силно интерактивни приложения); ~1 ms (приложения със свърх- ниска латентност) [52]	Латентност при предаване по мрежата: 1 ms	Латентност при предаване по мрежата: 0.1 ms
Синхронизация	Синхронизиране на множество потоци, пристигащи от различни локации в крайната точка [39], [123]	Латентност до 1 ms, джитер – 1 ms	Латентност до 0.1 ms, джитер до 0.001 ms
Изчисления	Изискване за изпълнение на изчислително тежки алгоритми за реконструкция на 3D модели, компресия и др.	Въвеждане на мрежови облачни изчисления	6G, поддържаща както изчисления, така и комуникации



Фигура 3.1: Еволюция на поколенията мобилни мрежи по отношение на скоростите на предаване на данни и предоставяните услуги [A4]

3.2. Потенциални подходи за реализация на НТС

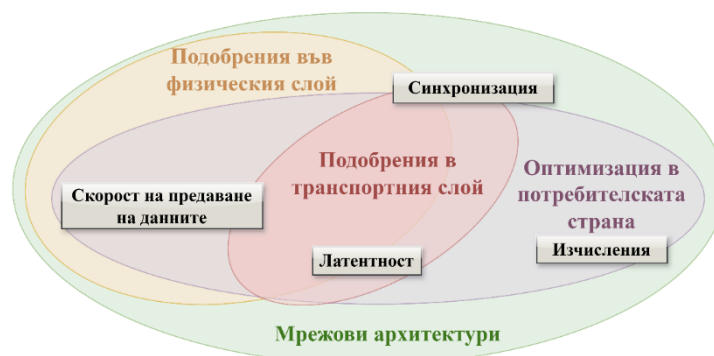
Позовавайки се на схващането, че НТС ще бъде една от основните 6G услуги, са разгледани възможните подобрения, които се очаква да направят това реалност. В тази връзка са представени потенциалните подходи за справяне с технологични изисквания на НТС, класифицирани в две основни групи. Първата група включва подходи, пряко свързани с оптимизацията на комуникационната мрежа и в частност 6G, а втората — подходи,

отнасящи се до оптимизации в потребителската страна. Първата група е допълнително категоризирана в три подгрупи, а именно подобрения във физическия слой, мрежови архитектури и подобрения в транспортния слой. Таблица 3.4 представя тези подходи, класифицирани по описания по-горе начин. Въз основа на влиянието им върху мрежовите KPIs [140], е направена оценка на тяхната приложимост относно различните НТС изисквания [A4], вследствие на което Таблица 3.4 показва и връзката между тях и съответните технологични предизвикателства.

Таблица 3.4: Влияние на потенциалните подходи за подобряване на НТС върху основните технологични предизвикателства при НТС

Група подходи	Подходи	Скорост на предаване	Латентност	Синхронизация	Изчисления
Подобрения на физическия слой	Използване на по-високи честотни ленти над 300 GHz, включително видимата светлина	Да	Да	Не	Не
	Използване на нови техники за модулация	Да	Не	Не	Не
	Усъвършенстване на кодирането на физическия слой	Да	Да	Не	Не
	Прилагане на технологии за многоканални антенно-предавателни системи и beamforming техники заедно с алгоритми за обработка на сигнали	Да	Не	Не	Не
	Разгръщане на реконфигурируеми интелигентни повърхности върху сгради, магистрали и други обекти	Да	Не	Не	Не
	Интегрирано наблюдение и локализация	Да	Да	Не	Не
	Използване на AI и ML във физическия слой	Да	Да	Да	Да
Мрежови архитектури	Разгръщане на гъвкави, адаптивни и мащабируеми мрежови архитектури в 6G	Да	Да	Да	Да
	Интеграция на софтуерно-дефинирани мрежи и мрежово сегментиране (network slicing)	Да	Да	Да	Да
	Интеграция на изчисления в края на мобилната мрежа	Не	Да	Да	Да
	Използване на свързаност, различна от наземната	Да	Да	Да	Не
	Интегриране на AI в 6G	Да	Да	Да	Да
Подобрения на транспортния слой	Нови протоколи за транспортния слой	Не	Да	Да	Не
	Адаптивно предаване на холографско съдържание чрез поддръжка на потоци с различен приоритет	Да	Да	Не	Не
Оптимизация при крайния потребител	Оценка на състоянието на мрежата и на изискванията за различни НТС сценарии	Да	Да	Да	Да
	Иновативни методи за компресия на данни	Да	Да	Не	Да
	Адаптивно предаване според FOV и фокуса на потребителя	Да	Не	Не	Не
	Използване на семантика за реализиране на семантични комуникации	Да	Не	Не	Не

Позовавайки се на Таблица 3.4, Фигура 3.2 илюстрира цялостното въздействие на четирите групи подходи върху технологичните изисквания на НТС. Всяка елипса обозначава отделна група и обхваща изискванията, които има потенциала да адресира. Макар при този начин на представяне да липсва конкретна количествена оценка, той предлага яснота относно фокуса на различните групи подходи върху удовлетворяването на съответните НТС изисквания. Трябва да се подчертае обаче, че въвеждането на различните подобрения не трябва да се разглежда изолирано. Докато някои подходи са насочени към решаване на конкретни НТС предизвикателства, същите могат да окажат негативно влияние върху други. Поради това следва да се търси оптимален компромис между различните решения, съобразен с конкретните НТС сценарии.



Фигура 3.2: Връзка между основните технологични изисквания на НТС и категориите подходи, свързани с тяхното решение [A4]

3.3. AI и НТС

С оглед едновременната оптимизация на комуникационно и потребителско ниво, интеграцията на AI се очертава ключово решение, подкрепено от следните аргументи [A4]. Първо, в потребителската страна AI ще играе важна роля за постигане на адаптивно QoE. По-конкретно, AI-базирани методи могат да анализират поведението на потребителя, факторите на околната среда и наличните мрежови ресурси с цел динамични корекции в обработката и предаването на данни. Второ, AI ще допринесе за високоскоростната обработка на съдържанието чрез използването на предварително обучени модели. Трето, AI ще играе ключова роля и за подобряване на семантичното разбиране на съдържанието и за улесняване на контекстно-осъзната комуникация. Неговото участие ще се изрази основно в прехода от прецизно предаване на данни към точното представяне на тяхната семантика. Четвърто, AI ще спомогне за анализа на големи обеми от данни [152]. Това включва извличане на исторически контекст, диагностични функции за автономно откриване на повреди, възможности за прогнозиране на бъдещото състояние на мрежата, появата на критични събития, потребителско поведение, популярност на съдържанието и др., както и предлагането на възможни решения. Това би улеснило адаптивното планиране на мрежата и управлението в реално време. В резултат бъдещите мрежи имат потенциала да се трансформират в гъвкави и адаптивни системи, способни да поддържат широк спектър от приложения, ориентирани към QoE. Следователно, въпреки значителните ресурси, предвидени за оптималното опериране на 6G мрежата, нейната ефективност ще се реализира в пълна степен, едва когато е интегрирана интелигентност. Накрая нека се има предвид сценарий, при който ресурсите и интелигентността на комуникационната мрежа се комбинират безпроблемно с ресурсите и интелигентността на потребителската страна, като всичко това се оркестрира от координиращ AI модул. Подобен подход създава предпоставки за заобикаляне на зависимостта от наличните ресурси. Следователно, приоритизирането на интелигентен подход, вместо просто увеличаване на ограничени капацитети, е най-ефективната стратегия сред представените в Раздел 3.2.

3.4. Изводи

В сравнение с 5G, 6G се явява ключов фактор за реализирането на НТС. Осигуряването на разширени честотни ресурси, редуцирането на латентността и възможността за изнасяне на изчислителните процеси от потребителските устройства към периферията на мрежата създават съществени предпоставки в подкрепа на тази теза. Въпреки това, имайки предвид историческата тенденция, при която всяка нова технологична парадигма стимулира развитието на следващото поколение комуникационни мрежи, може да се заключи, че 6G не е универсално решение за преодоляване на всички последващи ограничения пред перспективните НТС приложения. Необходимо е прилагането на интелигентни подходи за оптимизация както на ниво комуникационна мрежа, така и в потребителската страна.

3.5. Приноси към трета глава

- 1) Дефиниране и класификация на мрежовите KPIs, свързани с практическата реализация на НТС. Оценка влиянието на 5G и 6G върху реализацията на НТС, съобразно представените KPIs. Аналитичен обзор, изследващ мястото на НТС в бъдещите 6G комуникации.
- 2) Дефиниране и класификация на потенциалните подходи за оптимизация на комуникационната мрежа и на потребителската страна по отношение реализацията на НТС. Оценка на представените подходи спрямо основните НТС изисквания. Анализ на ролята на изкуствения интелект в този контекст.

ГЛАВА 4. АЛГОРИТМИ ЗА РЕГИСТРАЦИЯ НА ОБЛАЦИ ОТ ТОЧКИ, БАЗИРАНИ НА ДЪЛБОКО ОБУЧЕНИЕ

4.1. Облак от точки. Стандартни методи за регистрация

Облакът от точки представлява множество от точки, дефинирани в Евклидовото пространство посредством техните 3D координати (x, y, z) и допълнително описани посредством цветови канали, нормали и други характеристики. Тъй като последните не се използват в разработените алгоритми, всяка точка се представя като вектор, съставен единствено от трите координати. Следователно, даден облак P може да бъде представен като $P \in \mathbb{R}^{N \times 3}$, където N е общият брой точки.

За осъществяването на регистрация на два облака е необходимо намирането (калибрацията) и прилагането на оптималните параметри на трансформация между тях (ротацията $R \in SO(3)$ и транслацията $t \in \mathbb{R}^3$), така че същите да бъдат подравнени и представени в обща координатна система. Когато не е известно съответствието на точките от двата облака, трябва да бъде решена следната оптимизационна задача:

$$\min_{R \in SO(3), t \in \mathbb{R}^3} \sum_{i=1}^N \|Rp_i + t - q_i\|^2 \quad (4.8)$$

Най-използваният подход за нейното решение е чрез прилагане на алгоритъма за итеративно намиране на най-близки точки (Iterative Closest Points, ICP) [155]. При него се извършва итеративно минимизиране на разстоянието между облаците чрез търсене на съответствия посредством алгоритъма за намиране на най-близки съседи и многократно изчисление на Уравнение 4.8, докато грешката бъде минимизирана или бъде достигнат максималният брой итерации.

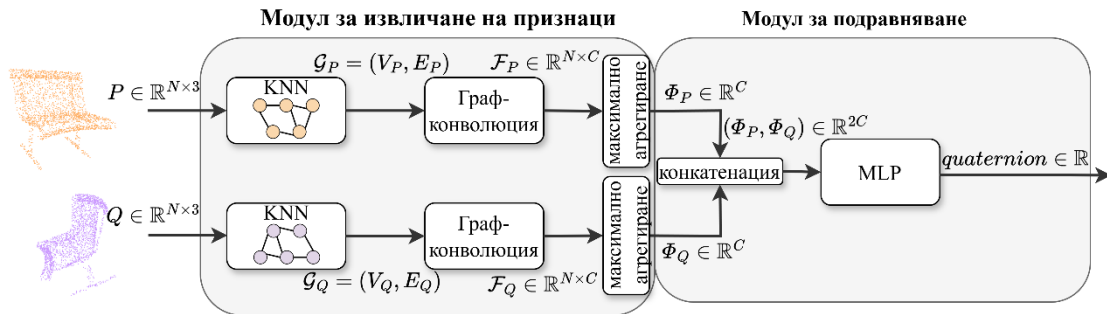
4.2. AI-базирана регистрация на напълно идентични облаци от точки

Предложен е AI-базиран алгоритъм за регистрация на два идентични облака от точки ($P \in \mathbb{R}^{N \times 3}$ и $Q \in \mathbb{R}^{N \times 3}$) чрез регресия на ротационен кватернион [A5]. Нека се има предвид, че алгоритъмът е фокусиран основно върху предсказването на R , бивайки по-сложната задача,

като се пренебрегва трансляцията t . Така оптимизационната задача от Уравнение 4.8 придобива следния вид:

$$\min_{R \in SO(3)} \sum_{i=1}^N \|R\mathbf{p}_i - \mathbf{q}_i\|^2 \quad (4.10)$$

Предложеният подход реализира директна регресия чрез параметризация на Уравнение 4.10 посредством модел на невронна мрежа. Блокова диаграма е илюстрирана на Фигура 4.1. Тя се състои от два основни модула. Първият е модулът за извличане на признаци. Той обработва P и Q паралелно, използвайки споделен единичен граф-конволюционен слой с цел извличане на локални геометрични характеристики, след което чрез максимално агрегиране предоставя и съответните глобални характеристики за всеки от облаците. Вторият модул е т.нар. модул за подравняване. Той обединява глобалните геометрични характеристики и ги прекарва през плитък многослоен перцептрон (Multilayer Perceptron, MLP), за да предскаже ротационния кватернион, $quaternion \in \mathbb{R}^4$, от който директно се изчислява матрицата на ротация.



Фигура 4.1: Блокова диаграма на предложения AI-базиран алгоритъм за регистрация на два идентични облака от точки

4.3. AI-базирана регистрация на частично припокриващи се облаци от точки, заснети от множество изгледи, в латентното пространство на компресираните им представяния

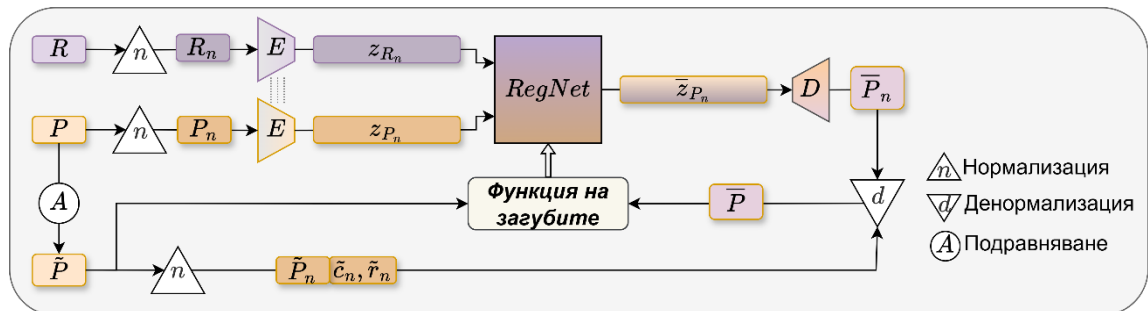
Представен е метод за директна регистрация на частично припокриващи се облаци от точки, представящи един и същ 3D обект, но заснет от различни изгледи [А6]. Алгоритъмът е допълнително надграден да извършва регистрацията след компресиране на отделните облаци, оперирайки изцяло в полученото латентно пространство (т.е. не в Евклидовото). Нека $Q \in \mathbb{R}^{N \times 3}$ бъде облакът от точки, представящ физическия обект от референтен изглед, а $P \in \mathbb{R}^{N \times 3}$ бъде облакът от точки, представящ същия обект, но от допълнителен и различен изглед. Нека между тях е налице частично геометрично припокриване. Стандартният подход за осъществяване на регистрация в този случай се нуждае от набор от съответстващи си точки, т.е. $Q_c = \{q_{i_k}\}_{k=1}^K \subseteq Q$, $P_c = \{p_{j_k}\}_{k=1}^K \subseteq P$ и $C = \{(q_{i_k}, p_{j_k})\}_{k=1}^K$, където Q_c и P_c са подмножествата на Q и P , формиращи т.нар. множество от съответствия C . Подходът, който подравнява P_c към референтното подмножество Q_c , изцяло следва Уравнение 4.8, което в този случай придобива следния вид:

$$\min_{R \in SO(3), t \in \mathbb{R}^3} \sum_{k=1}^K \|R\mathbf{p}_{j_k} + \mathbf{t} - \mathbf{q}_{i_k}\|^2 \quad (4.18)$$

Задачата се параметризира с модел на невронна мрежа, при което горното уравнение бива представено като:

$$\min_{CD} = (\tilde{P}, \bar{P}), \quad (4.19)$$

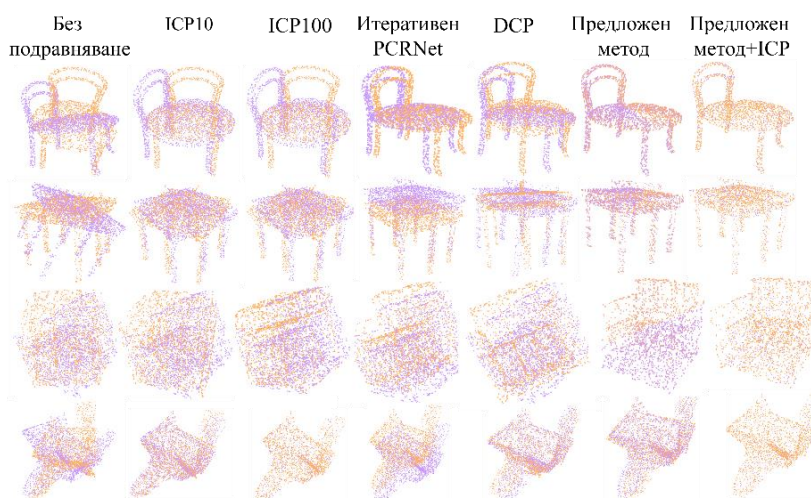
където $\tilde{P} = PR_{gt}^T + t_{gt}$ описва действителното (еталонно) подравняване, съпоставящо P към Q , $\bar{P} = f_{Decoder}(f_{RegNet}(z_Q, z_P))$ е подравняването, осъществено в латентното пространство, $z_Q = f_{Encoder}(Q)$ и $z_P = f_{Encoder}(P)$, като $f_{Encoder}$ и $f_{Decoder}$ са съответно енкoderът и декодерът, част от внедрената автоенкодер архитектура, използвана за компресия и декомпресия на P и Q , а f_{RegNet} е апроксимация на предложения подход за подравняване в латентното пространство. CD обозначава симетричното разстояние по Чамфер (Chamfer distance). Фигура 4.2 представя блокова схема на предложения метод.



Фигура 4.2: Блокова схема на предложения метод за регистрация на облаци от точки, извършена изцяло в латентното пространство

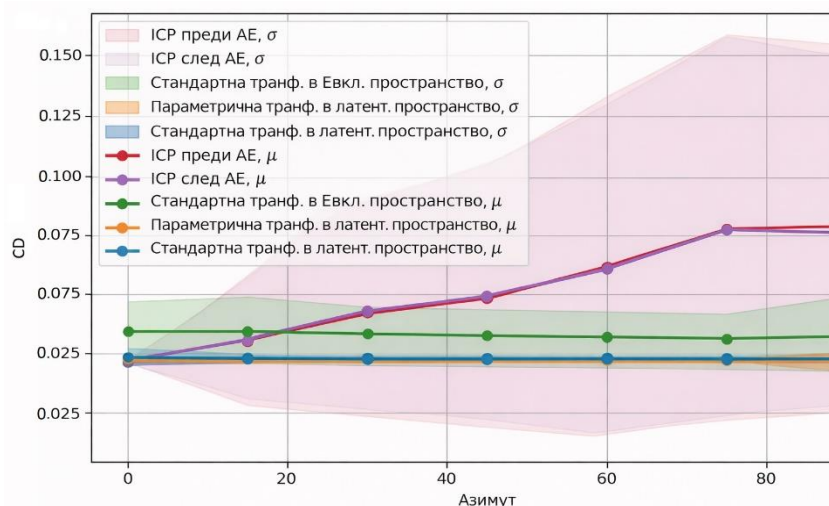
4.4. Имплементация и експериментални резултати

Методът за регистрация на идентични облаци от точки е обучен, използвайки базата данни ShapeNet [158]. Качествено представяне на получените резултати е илюстрирано на Фигура 4.4, където са предоставени и резултатите, постигнати от други подходи за регистрация, базирани на дълбоко обучение, и ICP. Направеното сравнение илюстрира превъзходството на предложия метод, демонстрирайки висока ефективност на регистрацията за всички разгледани 3D модели. Освен това резултатите подчертават важността от подходяща инициализация на съответстващи си точки при ICP. Отнасяйки се до първите два примерни 3D модела, изпълнението на ICP изглежда заседнало в локални минимума и увеличаването на броя итерации има малък или никакъв ефект. За сметка на това последният пример показва, че добра първоначална оценка, осигурена на ICP чрез прилагане на предложия алгоритъм, води до бърза конвергенция и точно подравняване.



Фигура 4.4: Визуално сравнение на резултати от алгоритъма за регистрация на идентични облаци от точки и други подходи

За обучението на модела за регистрация на частично припокриващи се облаци от точки е създадена специална база данни въз основа на ShapeNet [A7]. За целта ShapeNet обектите са заснети от различни изгледи чрез симулация на виртуална камера, която бива последователно позиционирана върху повърхността на сфера около от тях при зададени азимутни ъгли ($\{0,15,30,45,60,75,90\}$). Така за всеки модел биват генерирани по 7 двойки облаци от точки, представени от референтния изглед, генериран при 0 градуса, и съпътстващия го допълнителен изглед, генериран при всеки от посочените азимутни ъгли. Реализирани са два варианта за подравняване в латентното пространство — чрез стандартна трансформация и чрез параметризиране на трансформацията. С цел сравнение е реализиран и трети вариант за регистрация, опериращ в Евклидовото пространство. Той експлицитно регресира трансформационни параметри и ги прилага върху реконструирания от декодера облак от точки. Трите варианта са сравнени с ICP, приложен както преди, така и след автоенкодера, реализиращ компресията и декомпресията. Фигура 4.5 представя резултати от осредненото разстояние по Чамфер спрямо различните азимутни стойности, вследствие на което са направени две основни твърдения. Първо, резултатите потвърждават ефективността на предложените подходи в сравнение с ICP. Второ, грешката при тези методи е относително равномерна за всички азимутни стойности, което показва стабилно и надеждно подравняване дори при ниско ниво на геометрично припокриване между облаци. За разлика от тях, ICP демонстрира известно предимство при по-ниски азимутни стойности и, очаквано, влошава представянето си с увеличаване на азимута.



Фигура 4.5: Статистически резултати от различните методи за регистрация относно разстоянието по Чамфер спрямо различен азимут

4.5. Изводи

Освен прилагането на традиционни методи за регистрация е възможно тя да бъде ефективно реализирана и с помощта на AI-базирани методи. Предимството се състои в липсата на необходимост от идентифициране на съответстващи си точки между облаци, подлежащи на регистрация, както и от използването на калибрационни маркери. Това осигурява възможност за гъвкава промяна на позициите на камерите в един потенциален НТС сценарий. В допълнение предложеният подход за регистрация на частично припокриващи се облаци от точки демонстрира отчетлив успех по отношение на задачата за регистрация, макар и приложен в абстрактното латентно пространство. Нещо повече, той допринася за развитието на възможността за извличане на знание от силно компресирания и абстрактен латентен домейн, с пряка приложимост в 3D визуални системи с ограничена честотна лента, като роботика, добавена и виртуална реалност и, разбира се, НТС.

4.6. Приноси към четвърта глава

- 1) Разработен е алгоритъм за регистрация на идентични облаци от точки, които са трансформирани един спрямо друг в Евклидовото пространство. За целта е разработен модел на невронна мрежа, включващ модули за извличане на локални и глобални признаци, както и модул за регресия на ротационен кватернион.
- 2) Разработен е алгоритъм за регистрация на облаци от точки, представлящи един и същи физически обект от различни изгледи, така че е налице частично геометрично припокриване между тях. Алгоритъмът е надграден да извършва регистрацията изцяло в латентното пространство на компресираните чрез автоенкодер облаци от точки чрез специално разработен за целта RegNet модул.
- 3) Разработена е база данни, съдържаща облаци от точки, представлящи различни изгледи към даден 3D обект.

ГЛАВА 5. АЛГОРИТМИ ЗА КОМПРЕСИЯ НА ОБЛАЦИ ОТ ТОЧКИ, БАЗИРАНИ НА ДЪЛБОКО ОБУЧЕНИЕ

5.1. Традиционни подходи за компресия на облаци от точки

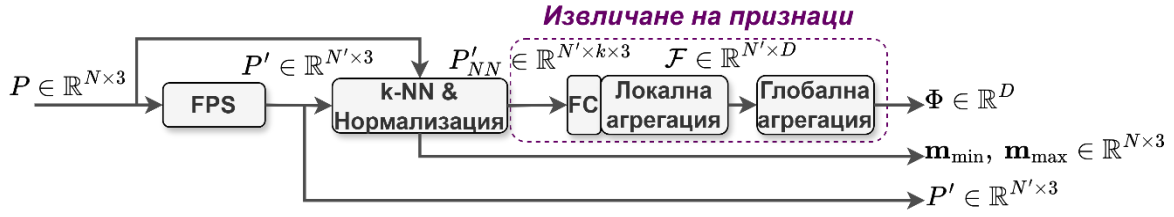
С цел адресиране на нарастващата необходимост от компактно представяне на 3D съдържание MPEG стандартизира два основни кодека, предназначени за компресия на облаци от точки, а именно видео-базиран (Video-based Point Cloud Compression, V-PCC) и геометрично базиран кодек (Geometry-based Point Cloud Compression, G-PCC) [163], [164]. Разработка на Google е библиотеката с отворен код Draco [165], която също предоставя възможности за компресия на облаци от точки и 3D mesh модели. От друга страна, AI-базираните алгоритми за компресия на 3D съдържание придобиват голяма популярност и свидетелстват за широко направление за работа в областта. Нещо повече, според [166] тези подходи постигат съществени подобрения спрямо традиционните методи по отношение на баланса между постигнатото ниво на компресия и полученото изкривяване на декодираните данни (т.нар. Rate-Distortion, RD). Също така се смята, че AI-базираните алгоритми значително повишават ефективността на кодиране, възползвайки се от възможностите за паралелни изчисления, което позволява осъществяването на много по-бърза обработка и съответно благоприятства развитието на услуги като НТС.

5.2. Алгоритъм за AI-базирана компресия на облаци от точки

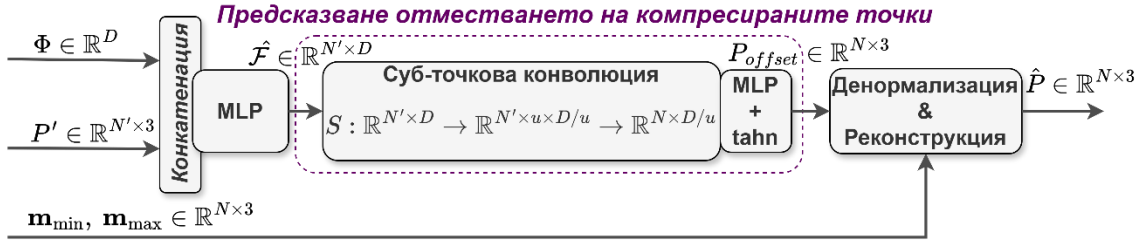
Предложен е алгоритъм за геометрична компресия на облаци от точки, базиран на дълбоко обучение, наречен MilaNet [A8]. Нека $P \in \mathbb{R}^{N \times 3}$ бъде облакът, подложен на разработения метод за компресия, където N е броят точки, а 3 съответства на трите геометрични координати x, y, z , дефинирани в Евклидовото пространство. Нека f_θ апроксимира функцията за компресия на данните, а g_ϕ – функцията за декомпресия:

$$\hat{P} = g_\phi(f_\theta(P)), \quad (5.1)$$

където $\hat{P} \in \mathbb{R}^{N \times 3}$ съответства на декомпресирания облак от точки. В разработения алгоритъм Уравнение 5.1 е параметризирано чрез автоенкодерна архитектура, при която f_θ и g_ϕ представят съответно енкодера и декодера, чиито структури са представени на Фигури 5.1 и 5.2.



Фигура 5.1: Блокова диаграма на енкодер частта на предложениия модел за геометрична компресия на облаци от точки



Фигура 5.2: Блокова диаграма на декодер частта на предложениия модел за геометрична компресия на облаци от точки

5.3. Алгоритъм за AI-базирана компресия на частично припокриващи се облаци от точки, заснети от множество изгледи

Разработен е алгоритъм за компресия на частично припокриващи се облаци от точки, представляващи отделните изгледи на един и същ обект. Алгоритъмът е базиран на дълбоко обучение и принципа за разпределено кодиране на източника (Distributed Source Coding, DSC). Целта е да се изследва възможността за постигане на по-ефективна компресия, възползвайки се от наличието на допълнителна информация (Side information, SI) в декодера. Нека $P \in \mathbb{R}^{N \times 3}$ и $Q \in \mathbb{R}^{N \times 3}$ бъдат два облака от точки, заснети от отделни изгледи, като между тях съществува частично геометрично припокриване. Нека енкодерите \mathcal{E}_P и \mathcal{E}_Q кодират независимо всеки от облациите, прилагайки:

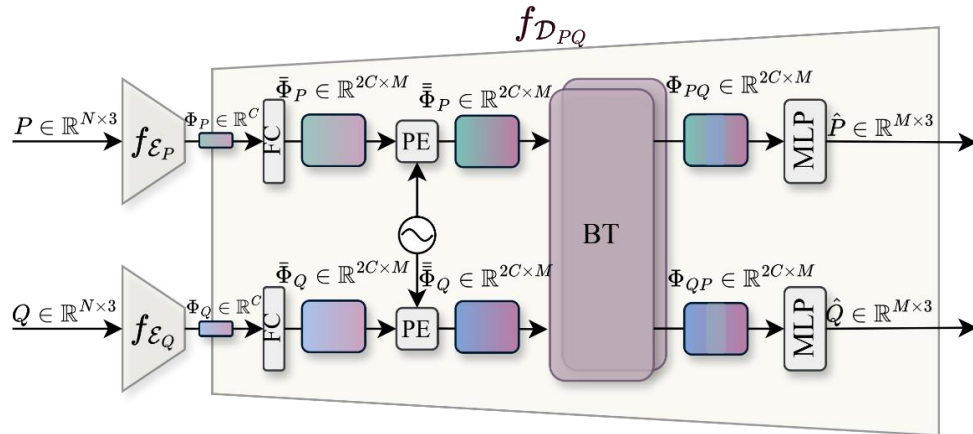
$$\begin{aligned}\Phi_P &= f_{\mathcal{E}_P}(P), \\ \Phi_Q &= f_{\mathcal{E}_Q}(Q),\end{aligned}\tag{5.12}$$

където $f_{\mathcal{E}_P}$ и $f_{\mathcal{E}_Q}$ са функции, описващи действието на \mathcal{E}_P и \mathcal{E}_Q , а $\Phi_P \in \mathbb{R}^C$ и $\Phi_Q \in \mathbb{R}^C$ са глобални признакови вектори, представляващи компресираните латентни представяния на всеки от облациите точки. Нека D_{PQ} е общ декодер, приемащ на входа си Φ_P и Φ_Q и възстановяващ \hat{P} и \hat{Q} с ниво на изкривяване съответно D_x и D_y , следвайки:

$$\{\hat{P}, \hat{Q}\} = f_{D_{PQ}}(\Phi_P, \Phi_Q),\tag{5.13}$$

където $f_{D_{PQ}}$ описва D_{PQ} . Целта на предложениия алгоритъм е P да бъде възстановен, използвайки SI в декодера в лицето на глобалната характеристика Φ_Q . Тъй като в случая не се използва Q , а Φ_Q , замисълът на алгоритъма концептуално следва обобщението на теоремата на Уайнър и Зив за DSC, предложено от Бергер и Тунг (BT). За реализиране на идеята е взимствана готова автоенкодерна архитектура, известна като Transformer Graph Autoencoder (TGAE) [174]. Структурата на енкодера е запазена, докато декодерът е

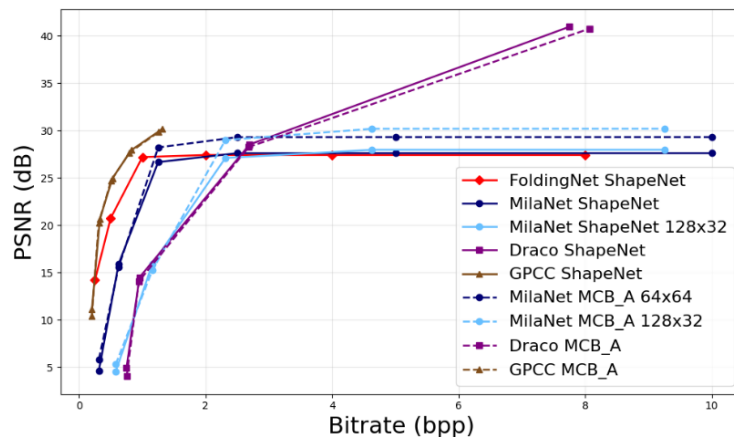
модифициран, за да съответства на становището на Бергер и Тунг. Фигура 5.7 показва блокова диаграма на реализирания модел с акцент върху подобренията в декодера.



Фигура 5.7: Блокова диаграма за компресиране на облаци от точки, заснети от множество изгледи, базирана на DSC

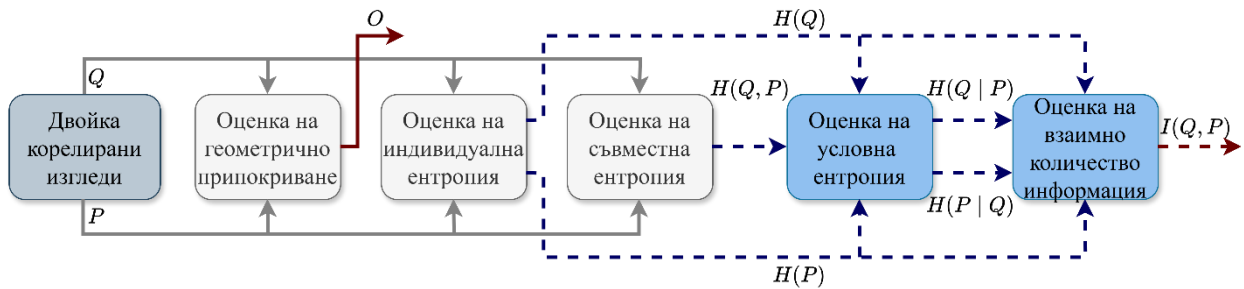
5.4. Имплементация и експериментални резултати

Предложеният MilaNet е обучен върху базите данни ShapeNet и Mechanical Component Benchmark (MCB) [175] и е изследван чрез две различни негови конфигурации, а именно MilaNet 64×64 и MilaNet 132×32. Получените резултати са сравнени с тези, постигнати от проведени експерименти както с традиционните методи за компресия – G-PCC и Google Draco за двата набора от данни, така и с друг невронен автоенкодер – FoldingNet [157], за базата данни ShapeNet. Фигура 5.9 представя RD кривите на приложените методи. По-конкретно, наблюдава се нарастване на пиковото отношение сигнал-шум (Peak signal to noise ratio, PSNR) с увеличаване на броя битове, необходими за кодирането на една точка, като се достига определено ниво на насищане при методите, базирани на дълбоко обучение. В региона с високо ниво на компресия, т.е. при много ниски стойности на bitrate (0 до 2 bpp), предложеният MilaNet 64×64 постига по-висок PSNR спрямо Google Draco и MilaNet 132×32, а след 2 bpp представянето му става съизмеримо с това на FoldingNet. От друга страна MilaNet 132×32 изпреварва FoldingNet при високи стойности на bitrate (над 4 bpp) с приблизително половин dB. Въпреки че на пръв поглед G-PCC се откроява със своята ефективност, тя е до известна степен условна поради някои специфики на кодека. Те се изразяват в значителното редуциране и разреждане на броя изходни точки, индикация за необективно изчисление на постигнатото ниво на компресия.



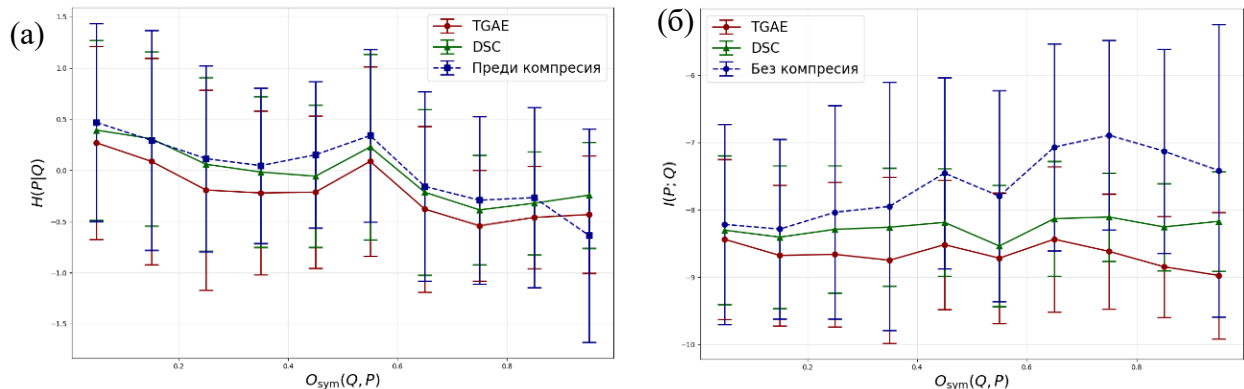
Фигура 5.9: RD криви на резултатите от реализираните методи за компресия, приложени върху ShapeNet и MCB A набори от данни

Алгоритъмът, базиран на DSC, е обучен, използвайки двойките облаци от точки от набора данни, публикуван в [A7]. За оценка на резултатите е предложена рамка за оценка на облаци от точки, заснети от множество изгледи [A7], включваща геометрични и информационно-теоретични метрики. Блокова диаграма на рамката за оценка е представена на Фигура 5.11.



Фигура 5.11: Блокова диаграма на рамката за оценка, илюстрираща работния процес и потоците от данни, използвани за изчисляване на геометричните и информационно-теоретичните метрики [A7]

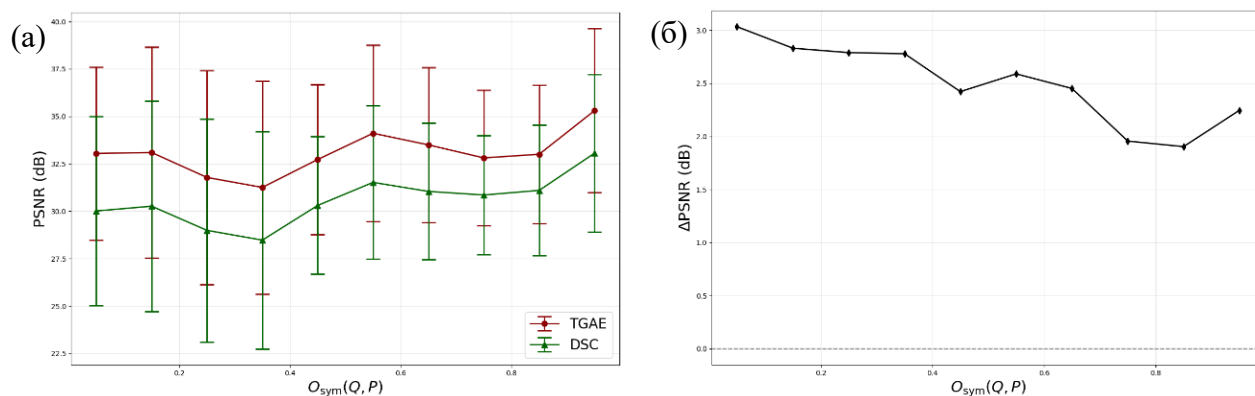
Фигури 5.13а и 5.13б представят условната диференциална ентропия $H(P|Q)$ и взаимното количество информация $I(P; Q)$ спрямо степента на геометрично припокриване $O_{sym}(Q, P)$ за референтния автоенкодер TGAE и предложения модел преди и след компресията. Вижда се, че увеличавайки корелацията между облаци, неопределеността на P спрямо Q намалява, защото посленият все повече го определя, и очаквано взаимното количество информация между тях се увеличава. По-интересно е поведението на графиките след прилагането на компресията, според които $H(P|Q)$ относително повтаря характера на получения преди компресия резултат. По-ниските стойности обаче са свидетелство за това, че след компресия корелацията между двойката изгледи намалява, като по-съществен е ефектът при TGAE. Аналогично заключение може да бъде направено и за $I(P; Q)$. Поради тази причина може да се каже, че предложеният алгоритъм по-добре успява да запази геометричното отношение между двойките изгледи в сравнение с TGAE.



Фигура 5.13: Отношение на условната ентропия на изглед P спрямо референтния изглед Q и на взаимното количество информация, отнесени към степента на геометрично припокриване

От гледна точка на постигнатата ефективност Фигура 5.14а илюстрира зависимостта на PSNR спрямо $O_{sym}(Q, P)$. Вижда се, че предложеният модел не подобрява представянето на TGAE, като резултатите са по-ниски с около 2 dB. Въпреки това Фигура 5.14б, изразяваща тенденцията на разликата $\Delta PSNR$ спрямо $O_{sym}(Q, P)$, засвидетелства намаляване на $\Delta PSNR$ с повишаване на корелацията между двойката изгледи. Макар и кривата да има известно колебание относно своята равномерност, дължаща се основно на характеристиките на

използваната база данни, тя демонстрира потенциала на реализираната идея за DSC и създава перспектива за бъдеща работа в областта.



Фигура 5.14: Отношение на постигнатия при компресия PSNR и $\Delta PSNR$, отнесени към степента на геометрично припокриване

5.5. Изводи

Освен прилагане на традиционни методи за компресия, реализирането на AI-базирани модели е обещаващо направление в областта. Нещо повече, постигнатото време за декодиране и реализирането на „лек“ модел с много малък брой параметри подкрепят идеята за прилагане на подобен тип алгоритми в приемната страна на HTC потребители. От друга страна, резултатите потвърждават и потенциала на AI за изследване на корелацията чрез DSC между латентните представяния на индивидуално компресираните частично припокриващи се облаци от точки. Въпреки че предложеният метод не успява към момента да потвърди становището на Бергер и Тунг, според което съвместното декодиране на изгледите, трябва да подобри ефективността на компресия, същият демонстрира изключително важна тенденция. Зачитането на корелацията между изгледите по време на декодирането повишава качеството на реконструирания облаци от точка с увеличаване степента на геометрично припокриване. Последното обосновава използването на DSC и насърчава бъдеща работа в областта.

5.6. Приноси към пета глава

- 1) Разработен е алгоритъм за компресия на облаци от точки, базиран на дълбоко обучение. За целта е разработен модел на автоенкодер, чийто енкодер редуцира броя точки на входния облак и извлича глобален признаков вектор, описващ компресираната информация, и чийто декодер реконструира обратно входните данни от компресираното латентно представяне.
- 2) Разработен е алгоритъм, базиран на дълбоко обучение, за компресия на облаци от точки, представящ един и същи физически обект, но заснет от различни изгледи, така че е налице частично геометрично припокриване между тях. Алгоритъмът следва принципа на DSC, разгръщайки сценария на Бергер и Тунг за съвместно декодиране, едновременно използвайки компресираните латентни представяния на входните облаци.
- 3) Предложена е рамка за оценка на облаци от точки, заснети от множество изгледи, базирана на геометрични и информационно-теоретични метрики за оценка.

III. НАУЧНО-ПРИЛОЖНИ И ПРИЛОЖНИ ПРИНОСИ

Научно-приложни приноси:

1. Дефиниране на функционалните компоненти на HTPS и идентифициране на основните предизвикателства пред тяхното реализиране.
2. Анализ и сравнение на съществуващите практически разработки в областта.
3. Разработване на алгоритъм за извличане на 3D сцена от заобикалящата среда чрез детектиране на семантични 3D обекти.
4. Изследване и оценка на практически реализирана HTPS.
5. Дефиниране и класификация на мрежовите KPIs, свързани с практическата реализация на НТС. Оценка влиянието на 5G и 6G върху реализацията на НТС, съобразно представените KPIs. Аналитичен обзор, изследващ мястото на НТС в бъдещите 6G комуникации.
6. Дефиниране и класификация на потенциалните подходи за оптимизация на комуникационната мрежа и на потребителската страна по отношение реализацията на НТС. Оценка на представените подходи спрямо основните НТС изисквания. Анализ ролята на изкуствения интелект в този контекст.
7. Разработване на алгоритъм за регистрация на идентични облаци от точки, които са трансформирани един спрямо друг в Евклидовото пространство. Проектиране модел на невронна мрежа, включващ модули за извличане на локални и глобални признаци, както и модул за регресия на ротационен кватернион.
8. Разработване на алгоритъм за регистрация на облаци от точки, представящи един и същи физически обект от различни изгледи, така че е налице частично геометрично припокриване между тях. Усъвършенстване на алгоритъма да извършва регистрацията изцяло в латентното пространство на компресираните чрез автоенкодер облаци от точки чрез специално проектиран за целта модел на невронна мрежа, наречен RegNet модул.
9. Разработване на алгоритъм за компресия на облаци от точки, реализиращ автоенкодерна архитектура. Проектиране на енкодер, редуциращ броя точки на входния облак и извличащ глобален признаков вектор, описващ компресираната информация. Проектиране на декодер, реконструиращ обратно входните данни от компресираното латентно представяне в облак от точки.
10. Разработване на алгоритъм, за компресия на облаци от точки, представящ един и същи физически обект, но заснет от различни изгледи, така че е налице частично геометрично припокриване между тях. Проектиране на декодер, апроксимиращ принципа на DSC, конкретно разгръщащ сценария на Бергер и Тунг за съвместно декодиране.

Приложни приноси:

1. Прилагане на модулен архитектурен подход за реализиране на HTPS, включително хардуерна имплементация на системата.
2. Програмна реализация на архитектурни блокове от слоя на събиране на данни, слоя на обработка на данни и слоя за осигуряване на потапящо изживяване, а именно: колектор на данни, модул за обработка на данни и модул за изграждане и подравняване на сцена.
3. Внедряване на алгоритъм за калибрация и на алгоритъм за извличане на 3D сцена в разработения модул за обработка на данни.
4. Разработване на база данни, съдържаща облаци от точки, представящи различни изгледи към даден 3D обект.
5. Предлагане на рамка за оценка на облаци от точки, заснети от множество изгледи, базирана на геометрични и информационно-теоретични метрики за оценка.

IV. ИЗПОЛЗВАНИ СЪКРАЩЕНИЯ

3D	Three Dimensional	HTPS	Holographic Telepresence System
5G	Fifth Generation Mobile Networks	ICP	Iterative Closest Points
6G	Sixth Generation Mobile Networks	I/O	Input Output
AI	Artificial Intelligence	MCB	Mechanical Component Benchmark
BT	Berger-Tung	ML	Machine Learning
CD	Chamfer Distance	MLP	Multi-layer Perceptron
DAN	Data Acquisition Network	PSNR	Peak signal-to-noise ratio
DSC	Distributed Source Coding	QoE	Quality of Experience
FOV	Field of View	QoS	Quality of Service
G-PCC	Geometry-based Point Cloud Compression	RD	RateDistortion
HMD	Head Mounted Display	SI	Side Information
HPCU	High-performance Computing Unit	TGAE	Transformer Graph Autoencoder
HTC	Holographic-type Communications	V- PCC	Video-based Point Cloud Compression

V. СПИСЪК НА ПУБЛИКАЦИИТЕ ПО ДИСЕРТАЦИОННИЯ ТРУД

[A1] Petkova, R., Poulkov, V., Manolova, A., & Tonchev, K. (2022). Challenges in implementing low-latency holographic-type communication systems. *Sensors*, 22(24), 9617.

[A2] Bozhilov, I., Petkova, R., Tonchev, K., Manolova, A., & Poulkov, V. (2023). HOLOTWIN: A modular and interoperable approach to holographic telepresence system development. *Sensors*, 23(21), 8692.

[A3] Petkova, R., Tonchev, K., Manolova, A., & Poulkov, V. (2022, June). 3D scene extraction using plane detection algorithm. In *2022 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)* (pp. 347-351). IEEE.

[A4] Petkova, R., Bozhilov, I., Manolova, A., Tonchev, K., & Poulkov, V. (2024). On the way to holographic-type communications: Perspectives and enabling technologies. *IEEE Access*, 12, 59236-59259.

[A5] Petkova, R., Bozhilov, I., Tonchev, K., Manolova, A., & Poulkov, V. (2025, June). Learning 3d rotations from point cloud data. In *2025 60th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)* (pp. 1-4). IEEE.

[A6] Petkova, R., Bozhilov, I., Tonchev, K., Manolova, A., & Poulkov, V. (2025, November). Can rigid body transformation be applied in the latent domain of compressed multi-view point clouds?. In *2025 28th International Symposium on Wireless Personal Multimedia Communications (WPMP)* (pp. 1-6). IEEE.

[A7] Petkova, R., Bozhilov, I., Tonchev, K., Manolova, A., & Poulkov, V. (2026, February). Geometric and Information-Theoretic Evaluation of Multi-View Point Cloud Datasets. In *2026 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*. IEEE.

[A8] Aldimirova, M., Petkova, R., Bozhilov, I., Tonchev, K., Manolova, A., & Poulkov, V. (2025, November). Deep Learning-based Geometry Compression of Point Cloud Data. In *2025 28th International Symposium on Wireless Personal Multimedia Communications (WPMP)* (pp. 1-6). IEEE.

VI. ИСПОЛЗВАНА ЛИТЕРАТУРА

- [1] Y. Wang, Z. Su, N. Zhang, etc., “A survey on metaverse: Fundamentals, security, and privacy,” *IEEE Communications Surveys & Tutorials*, 2022.
- [2] L.-H. Lee, T. Braud, P. Zhou, etc., “All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda,” *arXiv preprint arXiv:2110.05352*, 2021.
- [9] T. Rhee, S. Thompson, D. Medeiros, R. Dos Anjos, A. Chalmers, “Augmented virtual teleportation for high-fidelity telecollaboration,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 5, pp. 1923–1933, 2020.
- [10] D. Saredakis, A. Szpak, B. Birchead, H. A. Keage, A. Rizzo, T. Loetscher, “Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis,” *Frontiers in Human Neuroscience*, vol. 14, p. 96, 2020.
- [11] C. Chang, K. Bang, G. Wetzstein, B. Lee, L. Gao, “Toward the next-generation VR/AR optics: a review of holographic near-eye displays from a human-centric perspective,” *Optica*, vol. 7, no. 11, pp. 1563–1578, 2020.
- [12] X. Pan, X. Xu, S. Dev, A. G. Campbell, “3D Displays: Their Evolution, Inherent Challenges and Future Perspectives,” in *Proceedings of the Future Technologies Conference*, Springer, 2021, pp. 397–415.
- [13] J. Gugenheimer, C. Mai, M. McGill, J. Williamson, F. Steinicke, K. Perlin, “Challenges using head-mounted displays in shared and social spaces,” in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–8.
- [14] P.-A. Blanche, “Holography, and the future of 3D display,” *Light: Advanced Manufacturing*, vol. 2, no. 4, pp. 446–459, 2021.
- [15] P. A. Kara, R. R. Tamboli, O. Doronin, etc., “The key performance indicators of projection-based light field visualization,” *Journal of Information Display*, 2019.
- [17] J. Lawrence, D. B. Goldman, S. Achar, etc., “Project Starline: A high-fidelity telepresence system,” 2021.
- [18] J. Orlosky, M. Sra, K. Bektaş, etc., “Telelife: the future of remote living,” *arXiv preprint arXiv:2107.02965*, 2021.
- [19] L. Yoon, D. Yang, C. Chung, S.-H. Lee, “A Full Body Avatar-Based Telepresence System for Dissimilar Spaces,” *arXiv preprint arXiv:2103.04380*, 2021.
- [20] M. Montagud, J. Li, G. Cernigliaro, A. El Ali, S. Fernández, P. Cesar, “Towards socialVR: evaluating a novel technology for watching videos together,” *Virtual Reality*, pp. 1–21, 2022.
- [21] O. Ozioko, R. Dahiya, “Smart tactile gloves for haptic interaction, communication, and rehabilitation,” *Advanced Intelligent Systems*, vol. 4, no. 2, 2022.
- [22] J. Wang, Y. Qi, “A Multi-User Collaborative AR System for Industrial Applications,” *Sensors*, vol. 22, no. 4, p. 1319, 2022.
- [23] Y. Wang, P. Wang, Z. Luo, Y. Yan, “A novel AR remote collaborative platform for sharing 2.5D gestures and gaze,” *The International Journal of Advanced Manufacturing Technology*, vol. 119, no. 9, pp. 6413–6421, 2022.
- [24] H. Regenbrecht, N. Park, S. Duncan, etc., “Atea Presence—Enabling Virtual Storytelling, Presence, and Tele-Co-Presence in an Indigenous Setting,” *IEEE Technology and Society Magazine*, vol. 41, no. 1, pp. 32–42, 2022.

- [25] D. Kim, D. Jo, “Effects on Co-Presence of a Virtual Human: A Comparison of Display and Interaction Types,” *Electronics*, vol. 11, no. 3, p. 367, 2022.
- [26] K. Yu, G. Gorbachev, U. Eck, F. Pankratz, N. Navab, D. Roth, “Avatars for teleconsultation: effects of avatar embodiment techniques on user perception in 3D asymmetric telepresence,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 11, pp. 4129–4139, 2021.
- [27] Sensiks, “Sensory Reality Pods & Platform.” url: <https://www.sensiks.com/>
- [28] S. Schwarz, M. Preda, V. Baroncini, etc., “Emerging MPEG standards for point cloud compression,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 133–148, 2018.
- [32] T. Fujihashi, T. Koike-Akino, S. Chen, T. Watanabe, “Wireless 3D point cloud delivery using deep graph neural networks,” in *ICC 2021 - IEEE International Conference on Communications*, IEEE, 2021, pp. 1–6.
- [36] J. van der Hooft, M. T. Vega, T. Wauters, etc., “From capturing to rendering: Volumetric media delivery with six degrees of freedom,” *IEEE Communications Magazine*, vol. 58, no. 10, pp. 49–55, 2020.
- [37] A. Clemm, M. T. Vega, H. K. Ravuri, T. Wauters, F. De Turck, “Toward truly immersive holographic-type communication: Challenges and solutions,” *IEEE Communications Magazine*, vol. 58, no. 1, pp. 93–99, 2020.
- [38] T. Fujihashi, T. Koike-Akino, T. Watanabe, P. V. Orlik, “HoloCast+: Hybrid digital-analog transmission for graceful point cloud delivery with graph Fourier transform,” *IEEE Transactions on Multimedia*, vol. 24, pp. 2179–2191, 2021.
- [39] S. Anmulwar, N. Wang, A. Pack, V. S. H. Huynh, J. Yang, R. Tafazolli, “Frame Synchronisation for Multi-Source Holographic Teleportation Applications—An Edge Computing Based Approach,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, IEEE, 2021, pp. 1–6.
- [40] I. Selinis, N. Wang, B. Da, D. Yu, R. Tafazolli, “On the Internet-scale streaming of holographic-type content with assured user quality of experience,” in *IFIP Networking Conference*, IEEE, 2020, pp. 136–144.
- [41] P. Qian, V. S. H. Huynh, N. Wang, S. Anmulwar, D. Mi, R. Tafazolli, “Remote Production for Live Holographic Teleportation Applications in 5G Networks,” *IEEE Transactions on Broadcasting*, vol. 68, no. 2, pp. 451–463, 2022.
- [43] W. Zhu, Z. Ma, Y. Xu, L. Li, Z. Li, “View-dependent dynamic point cloud compression,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 2, pp. 765–781, 2020.
- [44] G. Cernigliaro, M. Martos, M. Montagud, A. Ansari, S. Fernandez, “PC-MCU: Point cloud multipoint control unit for multi-user holoconferencing systems,” in *Proceedings of the 30th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, 2020, pp. 47–53.
- [45] C. J. Blackwell, J. Khan, X. Chen, “Holographic 3D Telepresence System with Light Field 3D Displays and Depth Cameras over a LAN,” in *SID Symposium Digest of Technical Papers*, vol. 52, 2021, pp. 761–763.
- [46] D. Roth, K. Yu, F. Pankratz, etc., “Real-time mixed reality teleconsultation for intensive care units in pandemic situations,” in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, IEEE, 2021, pp. 693–694.
- [47] S. F. Langa, M. Montagud, G. Cernigliaro, D. R. Rivera, “Multiparty Holomeetings: Toward a New Era of Low-Cost Volumetric Holographic Meetings in Virtual Reality,” *IEEE Access*, vol. 10, pp. 81856–81876, 2022.

- [48] R. Kachach, P. Perez, A. Villegas, E. Gonzalez-Sosa, “Virtual tour: An immersive low cost telepresence system,” in IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2020, pp. 504–506.
- [49] F. E. Fadzli, A. W. Ismail, “A Robust Real-Time 3D Reconstruction Method for Mixed Reality Telepresence,” *International Journal of Innovative Computing*, vol. 10, no. 2, 2020.
- [50] S. Vellingiri, J. White-Swift, G. Vania, etc., “Experience with a Trans-Pacific Collaborative Mixed Reality Plant Walk,” in IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2020, pp. 238–245.
- [51] R. Li, etc., “Towards a new internet for the year 2030 and beyond,” in Proceedings of the 3rd Annual ITU IMT-2020/5G Workshop Demo Day, 2018, pp. 1–21.
- [52] Qualcomm, “VR and AR Pushing Connectivity Limits,” 2018.
- [57] M. Pakanen, P. Alavesä, N. van Berkel, T. Koskela, T. Ojala, “Nice to see you virtually: Thoughtful design and evaluation of virtual avatar of the other user in AR and VR based telepresence systems,” *Entertainment Computing*, vol. 40, p. 100457, 2022.
- [59] G. Gamelin, A. Chellali, S. Cheikh, A. Ricca, C. Dumas, S. Otmane, “Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments,” *Personal and Ubiquitous Computing*, vol. 25, no. 3, pp. 467–484, 2021.
- [62] Ericsson, “The spectacular rise of holographic communication,” 2022. url: <https://www.ericsson.com/en/about-us/new-world-of-possibilities/imagine-possible-perspectives/holographic-communication>.
- [63] ITU, “Representative use cases and key network requirements for Network 2030,” 2020. url: <https://www.itu.int/pub/T-FG-NET2030-2020-SUB.G1>.
- [64] Huawei, “Communications Network 2030,” 2021. url: <https://www.huawei.com/en/giv/communications-network-2030>.
- [65] V. Memmesheimer, A. Ebert, “Scalable extended reality: A future research agenda,” *Big Data and Cognitive Computing*, vol. 6, no. 1, p. 12, 2022.
- [68] F. Jasche, J. Kirchhübel, T. Ludwig, P. Tolmie, “BeamLite: Diminishing Ecological Fractures of Remote Collaboration through Mixed Reality Environments,” in Proceedings of the 10th International Conference on Communities & Technologies, 2021, pp. 200–211.
- [71] Hackernoon, “3 Common Types of 3D Sensors: Stereo, Structured Light, and ToF.” url: <https://hackernoon.com/3-common-types-of-3d-sensors-stereo-structured-light-and-tof-194033f0>.
- [72] VRcompare, “VRcompare - The Internet's Largest VR & AR Headset Database.” url: <https://vr-compare.com/>
- [74] A. Manolova, K. Tonchev, V. Poulkov, S. Dixir, P. Lindgren, “Context-aware holographic communication based on semantic knowledge extraction,” *Wireless Personal Communications*, vol. 120, no. 3, pp. 2307–2319, 2021.
- [77] H. Bai, P. Sasikumar, J. Yang, M. Billingham, “A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing,” in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–13.
- [78] A. Jing, K. W. May, M. Naem, G. Lee, M. Billingham, “EyemR-Vis: Using bi-directional gaze behavioural cues to improve mixed reality remote collaboration,” in Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 2021, pp. 1–7.
- [79] J. Rekimoto, K. Uragaki, K. Yamada, “Behind-the-mask: A face-through head-mounted display,” in Proceedings of the 2018 International Conference on Advanced Visual Interfaces, 2018, pp. 1–5.

- [80] Y. Zhao, Q. Xu, W. Chen, etc., “Mask-off: Synthesizing face images in the presence of head-mounted displays,” in IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2019.
- [81] G. Song, J. Cai, T.-J. Cham, J. Zheng, J. Zhang, H. Fuchs, “Real-time 3D face-eye performance capture of a person wearing a VR headset,” in Proceedings of the 26th ACM International Conference on Multimedia, 2018, pp. 923–931.
- [82] A. Nijholt, “Capturing obstructed nonverbal cues in augmented reality interactions: A short survey,” in Proceedings of International Conference on Industrial Instrumentation and Control, Springer, 2022, pp. 1–9.
- [83] S.-Y. Chen, Y.-K. Lai, S. Xia, P. Rosin, L. Gao, “3D face reconstruction and gaze tracking in head-mounted displays for virtual interaction,” IEEE Transactions on Multimedia, 2022.
- [84] S. Lombardi, J. Saragih, T. Simon, Y. Sheikh, “Deep appearance models for face rendering,” ACM Transactions on Graphics, vol. 37, no. 4, pp. 1–13, 2018.
- [88] Z. Akhtar, K. Siddique, A. Rattani, S. L. Lutfi, T. H. Falk, “Why is multimedia quality of experience assessment a challenging problem?,” IEEE Access, vol. 7, pp. 117897–117915, 2019.
- [89] G. Kougioumtzidis, V. Poulkov, Z. Zaharis, P. Lazaridis, “A survey on multimedia services QoE assessment and machine learning-based prediction,” IEEE Access, vol. 10, pp. 19507–19538, 2022.
- [90] S. Y. Chin, B. R. Quinton, “Dynamic object comprehension: A framework for evaluating artificial visual perception,” arXiv preprint arXiv:2202.08490, 2022.
- [91] S. Vijayakumar, P. Corcoran, R. Flynn, N. Murray, “AI-derived quality of experience prediction based on physiological signals for immersive multimedia experiences,” in Proceedings of the 13th ACM Multimedia Systems Conference, 2022, pp. 403–407.
- [92] J. A. De Guzman, K. Thilakarathna, A. Seneviratne, “Security and privacy approaches in mixed reality: A literature survey,” ACM Computing Surveys, vol. 52, no. 6, pp. 1–37, 2019.
- [95] W. Mazurczyk, P. Bisson, R. P. Jover, K. Nakao, K. Cabaj, “Challenges and novel solutions for 5G network security, privacy and trust,” IEEE Wireless Communications, vol. 27, no. 4, pp. 6–7, 2020.
- [140] W. Jiang, B. Han, M. A. Habibi, H. D. Schotten, “The road towards 6G: A comprehensive survey,” IEEE Open Journal of the Communications Society, vol. 2, pp. 334–366, 2021.
- [152] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, Y.-J. A. Zhang, “The roadmap to 6G: AI-empowered wireless networks,” IEEE Communications Magazine, vol. 57, no. 8, pp. 84–90, 2019.
- [155] P. J. Besl, N. D. McKay, “A method for registration of 3-D shapes,” in Proceedings of SPIE Sensor Fusion IV: Control Paradigms and Data Structures, vol. 1611, 1992, pp. 586–606.
- [158] A. X. Chang, T. Funkhouser, L. Guibas, etc., “ShapeNet: An information-rich 3D model repository,” arXiv preprint arXiv:1512.03012, 2015.
- [163] ISO/IEC JTC1/SC29/WG7, “V-PCC Codec Description,” 2020.
- [164] ISO/IEC JTC1/SC29/WG7, “G-PCC Codec Description V9,” 2020.
- [165] Google, “Draco: 3D Data Compression.” url: <https://google.github.io/draco/>
- [166] I. Bozhilov, R. Petkova, K. Tonchev, A. Manolova, “A systematic survey into compression algorithms for three-dimensional content,” IEEE Access, vol. 12, pp. 141604–141624, 2024.
- [174] I. Bozhilov, R. Petkova, K. Tonchev, A. Manolova, V. Poulkov, H. V. Poor, “Autoencoder architectures for low-rate sparse point cloud geometry coding,” IEEE Access, vol. 13, pp. 214122–214140, 2025.
- [175] S. Kim, H.-g. Chi, X. Hu, Q. Huang, K. Ramani, “A large-scale annotated mechanical components benchmark for classification and retrieval tasks with deep neural networks,” in Proceedings of the European Conference on Computer Vision (ECCV), 2020.



TECHNICAL UNIVERSITY OF SOFIA
FACULTY OF TELECOMMUNICATIONS
DEPARTMENT “COMMUNICATION NETWORKS”

Radostina Rosenova Petkova, MSc

**METHODS AND ALGORITHMS FOR IMPLEMENTING
HOLOGRAPHIC-TYPE COMMUNICATION WITH ULTRA-LOW
LATENCY**

ABSTRACT of Ph.D THESIS

The continuous evolution of mobile communication networks toward higher throughput, ultra-low latency, greater connection density, and improved reliability is creating the foundation for upcoming immersive communication services. Following the progression from voice communication to video conferencing, the next major paradigm shift is expected to be Holographic-Type Communication (HTC) – a novel technology that enables the real-time transmission and reproduction of spatially rich holographic data, offering a substantially enhanced sense of telepresence and user interaction. However, the practical realization of HTC presents significant technological challenges, particularly in the areas of system architecture development, real-time data processing, and communication efficiency under strict latency constraints. Therefore, this dissertation aims to systematically investigate the key aspects of HTC implementation through a comprehensive analysis of existing challenges and the exploration of effective approaches for their mitigation.

To support this objective, the work provides a systematic study of Holographic Telepresence Systems (HTPS), including an in-depth analysis of their functional components, key implementation challenges, and existing system realizations. Based on these findings, a modular architectural framework for HTPS is proposed, followed by implementation of a real system. Furthermore, the dissertation analyzes the role of next-generation mobile networks as critical enablers for the deployment of HTC, with particular focus on their capability to satisfy stringent system requirements. A central contribution of the thesis is the development of novel algorithms for efficient three-dimensional visual data processing, considering the stringent HTC constraints. These algorithms demonstrate the feasibility of intelligent, task-oriented HTPS design and help bridge the gap between theoretical HTC concepts and their practical realization.



TECHNICAL UNIVERSITY OF SOFIA

**Faculty of Telecommunications
Department “Communication Networks”**

MSc Eng. Radostina Rosenova Petkova

**METHODS AND ALGORITHMS FOR IMPLEMENTING
HOLOGRAPHIC-TYPE COMMUNICATION WITH ULTRA-
LOW LATENCY**

A B S T R A C T

of a dissertation for the acquisition of the educational and scientific degree
“DOCTOR OF PHILOSOPHY”

Field: 5. Technical Sciences

Professional field: 5.3. Communication and Computer Engineering

Scientific specialty: Communication Networks and Systems

Scientific supervisor: Prof. DSc Vladimir Pulkov

SOFIA, 2026 г.

The dissertation has been reviewed and approved for defense by the Department Council of the Department of “Communication Networks” at the Faculty of Telecommunications of the Technical University of Sofia at a regular meeting held on April 20, 2026.

The public defense of the dissertation will take place on July 8, 2026 (Wednesday) at 13:00 in the Conference Hall of the Library Information Center at the Technical University of Sofia, at an open session of the scientific jury appointed by Order No. OJ-5.3-38 / 04.05.2026 of the Rector of the Technical University of Sofia, composed of:

1. Prof. Dr. Georgi Iliev– Chair
2. Assoc. Prof. Dr. Georgi Balabanov – Scientific Secretary
3. Assoc. Prof. Dr. Nikolay Atanasov
4. Prof. Dr. Emilia Dimitrova
5. Prof. Dr. Valentina Markova

Reviewers:

1. Assoc. Prof. Dr. Georgi Balabanov
2. Prof. Dr. Emilia Dimitrova

The materials related to the defense are available to interested parties at the office of the Faculty of Telecommunications at the Technical University of Sofia, Block 1, Room 1439-B.

The PhD candidate is a doctoral student at the Department of “Communication Networks” of the Faculty of Telecommunications. The research in the dissertation has been conducted by the author, with some parts supported by research projects.

Author: MSc Eng. Radostina Petkova

Title: Methods and Algorithms for Implementing Holographic-type Communication with Ultra-Low Latency

Print run: 10 copies

Printed at the Publishing House of the Technical University of Sofia

I. GENERAL CHARACTERISTICS OF THE DISSERTATION

Relevance of the Problem

Each successive generation of mobile networks provides its users with higher data rates, lower latency, greater connection density, and increased reliability. At the same time, conditions are being created for the emergence of fundamentally new communication paradigms. Following the transition from voice services to video calls, the current technological focus is directed toward the realization of spatially enriched communication, whose ultimate form is associated with the so-called Holographic-Type Communication (HTC).

HTC is a term representing the idea of a revolutionary technology that enables the real-time transmission of holographic data over communication networks. By its nature, providing the capture, transmission, and reproduction of information derived from the five primary human senses – it promises a significantly higher level of user engagement compared to currently available video conferencing solutions. Beyond its role as another communication method, numerous professional fields, including education, healthcare, industry, and others, have expressed intentions to adopt HTC in order to improve the functionality and accessibility of their services. Furthermore, HTC is considered a key factor in the realization and development of the so-called metaverse, in which not only people but entire physical systems are transferred into a unified global, digital, and parallel reality.

Objective of the Dissertation, Main Tasks, and Research Methods

This dissertation aims to present the specific features of HTC implementation through an analysis of existing challenges and an exploration of possible approaches for overcoming them.

The tasks associated with achieving the objective of the dissertation are as follows:

1. To investigate the functional components of HTC systems, also referred to as Holographic Telepresence Systems (HTPS), and to identify the main challenges in their implementation. To examine existing implementations of such systems.
2. To propose an architectural framework for the practical implementation of HTPS, followed by its applied realization. To implement algorithms for processing 3D visual data within the developed HTPS and evaluate its performance.
3. To analyze the role of future generations of mobile networks in enabling and deploying HTC for mass use.
4. To develop algorithms for efficient processing of three dimensional (3D) visual data, aligned with the requirement for ultra-low latency and demonstrating the concept of intelligent HTPS design.

The research methodology of the dissertation follows a comprehensive scientific approach, combining analytical, experimental, and modeling methods. A system analysis is applied to identify key components and challenges in HTPS. An experimental implementation and evaluation of HTPS are carried out. The impact of communication networks is assessed through analysis of key performance indicators and potential approaches for addressing the main HTC challenges. Within the dissertation, algorithm development methods based on neural networks are applied, followed by experimental validation.

Scientific Contribution

The dissertation identifies the main challenges associated with HTC and defines potential approaches for overcoming them. Artificial intelligence algorithms for the registration and compression of 3D visual data have been developed, the results of which lead to the following findings. First, the registration of point clouds representing a single object captured from different viewpoints is achievable in the abstract latent space of their compressed representations. Second, the efficiency of point cloud compression – when representing a single object captured from multiple viewpoints and based on distributed source coding – increases with the level of correlation between the views.

Practical Applicability

The dissertation presents the main functional components of HTPS and proposes a modular architectural approach for their subsequent practical implementation, intended to serve as guidelines for work in the field. In addition, the implemented HTPS is designed to serve as a platform for integrating and investigating various approaches and algorithms for HTC realization. Finally, the proposed algorithms for point cloud registration and compression represent potential solutions to some of the identified challenges.

Publications

The main achievements and results of the dissertation have been disseminated through 8 scientific publications, 3 of which are published in prestigious international scientific journals, while the remaining 5 have been presented at international scientific conferences.

The international journals are: *MDPI Sensors 2022 & 2023* and *IEEE Access 2024*.

The international conferences include: IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom) 2022; IEEE International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST) 2025; IEEE International Symposium on Wireless Personal Multimedia Communications (WPMC) 2025; and the Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunication Engineering (ECTI DAMT & NCON) 2026.

Structure and Scope of the Dissertation

The dissertation comprises **136** pages and includes an introduction, **5** chapters addressing the defined tasks, a list of contributions, a list of author publications, and references. A total of **176** sources are cited, of which **163** are in Latin script and the remainder are internet sources. The work contains **40** figures and **15** tables. A total of **54** citations to the eight author publications have been recorded. The numbering corresponds to that used in the dissertation.

Chapter One presents a block diagram of HTPS, defines the main functional components of the system, and identifies the key challenges in their implementation. Existing HTPS solutions are also presented and compared.

Chapter Two proposes an approach for the practical implementation of HTPS based on a modular architectural design. A real HTPS has been implemented and evaluated following the proposed

approach. Algorithms for processing 3D visual data integrated into the implemented system are also presented and evaluated.

Chapter Three examines the impact of fifth-generation mobile networks (5G), as well as the potential of future sixth-generation mobile networks (6G), in the realization and deployment of HTC. It also highlights their unprecedented potential in favor of the design and integration of intelligent methods and algorithms for data processing.

Chapter Four presents developed algorithms for the registration of 3D objects, which is essential when capturing them from multiple viewpoints.

Chapter Five introduces developed algorithms for the compression of 3D visual data, necessitated by the enormous volume of data generated during the creation of digital 3D content.

II. CONTENT OF THE DISSERTATION

CHAPTER 1. MAIN CHALLENGES IN THE IMPLEMENTATION OF HTPS

1.1. Functional diagram of HTPS

Each telepresence system must perform multiple operations to enable a user to participate in HTC. These operations can be grouped into four main functional blocks [A1], illustrated in Figure 1.1 and enclosed by a dashed line. The first functional block involves the capture of multisensory data, ideally collecting information derived from the five primary human senses, including visual, auditory, tactile, olfactory, and gustatory data. It should be noted that achieving a high degree of realism in capturing visual data requires the use of multiple cameras to cover a given scene or an object from different viewpoints. The second functional block integrates data processing operations, primarily including the reconstruction of 3D objects and scenes, as well as 3D rendering. Other more specific processes may include compression, registration, segmentation, noise filtering, and so on. The synchronization of individual data streams is also part of this block. The third block includes functions related to data transmission through the communication channel and across the network. Finally, after the user receives and processes the holographic data, the fourth functional block is responsible for their synchronized rendering to the user. It should be noted that a given user typically acts as both a transmitter and a receiver in the communication process, and therefore must support the execution of all the listed functions.

1.2. Challenges in the implementation of HTPS

The identified challenges in the implementation of HTC are categorized into three groups, namely core technical challenges, challenges related to the virtual representation of participants, and other challenges [A1]. The core technical challenges encompass those related to input/output (I/O) technologies for HTC, holographic data processing, holographic data transmission, and the scalability of HTPS. Challenges related to the virtual representation of participants include issues concerning the creation of authentic 3D virtual avatars, as well as problems associated with supporting gestures, gaze, and emotions. The remaining challenges involve the need for reliable evaluation of HTPS, taking into account both Quality of Service (QoS) and Quality of Experience (QoE), as well as ensuring high levels of security, privacy, and ethical standards in communication. These challenges are presented in Tables 1.1 and 1.2.

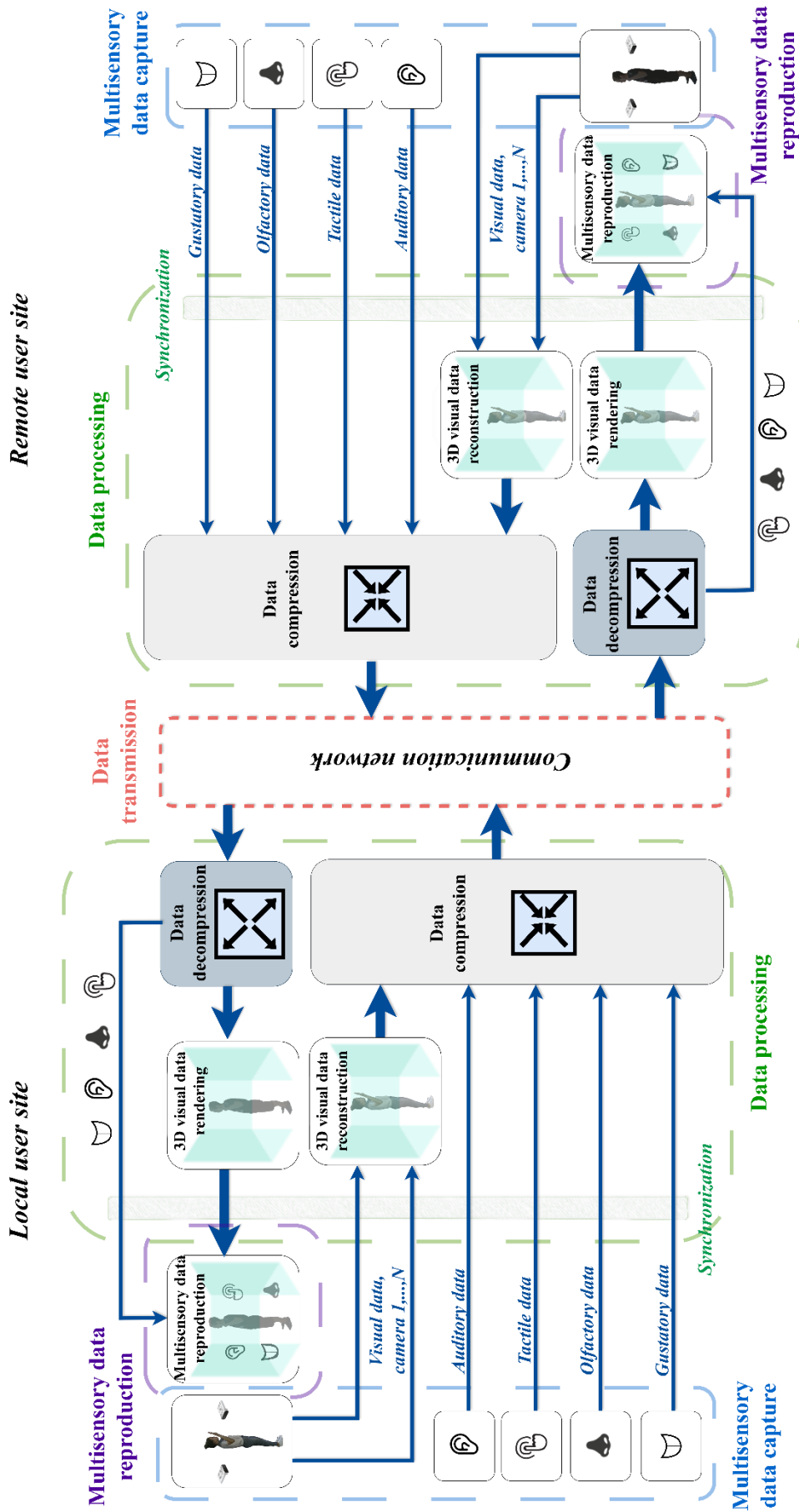


Figure 1.1: Functional block diagram of a holographic-type communication system [A1].

Table 1.1: Main technical challenges

Category	Challenge
I/O Technology	Depth cameras: low accuracy, low resolution, limited range, noise addition, narrow field of view (FOV) [71]
	Multiple cameras: need for calibration, low reconstruction accuracy, time-consuming, difficult installation
	360° capture: limited user movement [9]
	Complex processing and transmission at high frame rates
	Head-Mounted Displays (HMDs): lower resolution and narrower FOV than the human visual system [72]
	HMDs: limited battery life, uncomfortable, invasive, high cost, hardware requirements [12]
	HMDs: convergence–accommodation conflict causing nausea, dizziness, disorientation [10], [11], [14]
	Social acceptance of HMDs [13], limited use to a single user
	Light-field displays and volumetric displays: limited panel volume → limited depth, large data volume [12], [14], [15]
	Providing immersive experience supported by high resolution, large FOV, interactivity [11]
	Need for multisensory interactions [18], [20], [24]–[26]
	Need for video-audio consistency (requirement for spatial audio) [18]–[20]
	Integration of haptics in HTC systems [21]–[23]
	Limited inclusion of smell and taste, high cost for integrating sensory reality rooms
Data Processing	Processing large volumes of data
	Creating high-quality reconstruction (with multi-camera capture)
	Applying efficient compression/decompression techniques
	Requirement for high computational power
	Trade-off between processing latency and network transmission latency [14], [28], [32], [36]–[38]
	Integration of edge computing [36], [40], [41]
	Synchronization between local signals and those received from the remote user [37], [39], [40]
Data Transmission	Need for higher data rates, hence higher frequency bands and efficient modulation techniques
	Application of adaptive data transmission [43], [44]
	Need for semantic awareness and gaze direction prediction [36]
	Low-latency transmission: < 15 ms “motion-to-photon” and < 50–100 ms end-to-end [9], [45]–[52]
	Development of optimized mechanisms for holographic data transmission [37]
	Optimization of network architecture [37], [62]–[64]
Scalability	Increasing technological requirements – higher bandwidth and synchronization among multiple users
	Need for centralized control [44]
	Scalability with respect to: different devices, varying number of participants, different levels of virtuality [65]

Table 1.2: Challenges related to the virtual representation of participants, HTPS evaluation, security, and privacy

Category	Challenge
Avatar-Based Representation	Ensuring a sense of presence and realism, including facial detail of the avatar and expressed emotions
	Completeness of avatar bodies and their positioning, fidelity of movements and gestures
	Synchronization of avatar movements with those of the user
	Sense of body ownership and recognition of one's own movements and location
	Collaboration within the interaction space [2]
	Need for visual and behavioral accuracy of the avatar [17], [19], [20], [26], [46], [57], [59], [68], [74]–[77]
Gesture Support	Dynamic body tracking and reproduction of gestures and movements through the virtual avatar [19]
	Calibration and synchronization between different user coordinate systems
Gaze Support	Dynamic and accurate eye-tracking
	Estimation and visualization of gaze direction [78]
Emotion Support	Dynamic facial capture and tracking, synchronization of lip movements with speech
	Facial occlusion caused by wearing HMDs [79]–[84]
HTPS Evaluation	QoS insufficient for evaluating the overall system performance experienced by end users
	QoE representing a combination of multiple factors influencing the overall system performance [88], [89]
	Need of evaluation of different HTPS aspects such as realism, applicability, task completion time, etc.
	Need for evaluation that involves a very large number of participants performing different tasks under varying conditions
	Need of new evaluation metrics compatible with 3D data representation [90]
	Need for development of evaluation models for QoE using Machine Learning (ML) and Artificial Intelligence (AI) techniques [89], [91]
Data Security and Protection	Neglect of security and privacy enhancement in HTC systems [92]
	Need to protect 3D data representing human bodies, objects, and environments
	Need to protect sensitive biometric data and behavioral characteristic data [2]
	Need to protect application-specific data
	Improving network-level security [95]
	Increasing user trust in various HTC applications

1.3. Literature review of implemented HTPS

The identified challenges are further confirmed through the analysis of a number of publications related to practical implementations of HTPS. The systems are compared in terms of the I/O technologies used for data capture and reproduction, the network technology for data transmission, system scalability, the implemented type of avatar, support for non-verbal communication signals, the presence of system evaluation, and data protection capabilities.

The results show that various I/O devices have been integrated, primarily intended for capturing and rendering audio-visual data. The most commonly used are 3D cameras and HMDs for augmented and virtual reality, which provides the highest level of immersive experience. Regarding the networking part, there is no unified network protocol optimized for holographic data transmission. For this reason, technologies such as MPEG-DASH, RTSP, WebRTC, RabbitMQ, and PUN have been integrated, although they were originally developed for different purposes. Although addressed in many of the investigated systems, HTTPS scalability is limited to only a few users. Regarding the virtual representation of participants, photorealistic avatars are less commonly implemented in comparison to character-based ones. The integration of non-verbal communication signals has been implemented in many systems, with gesture reproduction being the most frequently supported feature. However, gaze tracking and facial expression support remain limited. System evaluation of HTTPS is almost always present. However, it does not consider user context and is based on a limited number of participants. According to the reported latency evaluations, the term “real-time” is not applicable, as different HTC scenarios impose different requirements. Therefore, evaluation should be application-specific and task-dependent. Compliance with the ultra-low latency requirement remains a major challenge overall. Finally, none of the reviewed HTTPS implementations provide strategies for ensuring user data security and privacy.

1.4. Conclusions

The realization of true HTC has not yet been achieved due to challenges of different nature. The lack of advanced and specialized technologies for implementing the core functional components of HTTPS necessitates the introduction of additional data processing algorithms, which in turn leads to increased end-to-end latency. The practical implementation of HTTPS represents a combination of technological and algorithmic solutions that must satisfy the fundamental requirements of HTC.

1.5. Contributions to Chapter 1

- 3) The functional components of HTTPS are defined, and the main challenges in their implementation are identified.
- 4) Existing practical developments in the field are analyzed and compared.

CHAPTER 2. PRACTICAL IMPLEMENTATION OF HTTPS

2.1. Modular architectural design of HTTPS

An architectural approach for the implementation of HTTPS is proposed [A2], focusing on modularity, flexibility, adaptability, and easy integration of the required HTC technologies. Four main layers are defined, each designed to perform a specific function:

- Data Acquisition Layer, responsible for the initial collection of raw sensor data and serving as the entry point to the platform; This also includes the Data Acquisition Network (DAN);
- Data Processing Layer, which receives input data and applies various algorithms for their manipulation, enhancement, and transformation into a virtual object or scene;
- Data Transmission Layer, managing the secure and efficient delivery of processed data to all nodes participating in the session;
- Immersive Experience Layer, where data is transformed into a user-oriented immersive experience, forming the final holographic output through visual rendering, engagement of the remaining senses, and enabling user interaction.

A control plane has also been introduced, whose purpose is to coordinate and manage the operation of the entire system. It enables user-driven control and evaluation with the aim of optimizing research efforts and accelerating the development process. Figure 2.1 illustrates the design of the proposed architecture. It can be seen that each layer contains multiple architectural blocks representing the core elements required for the operation of HTS. Each of them can be implemented through different selectable and interchangeable modules, depending on the requirements of a specific HTC scenario. In this way, the goal is to ensure both flexibility and adaptability to new technologies, as well as to provide opportunities for studying and evaluating system performance depending on its specific implementation.

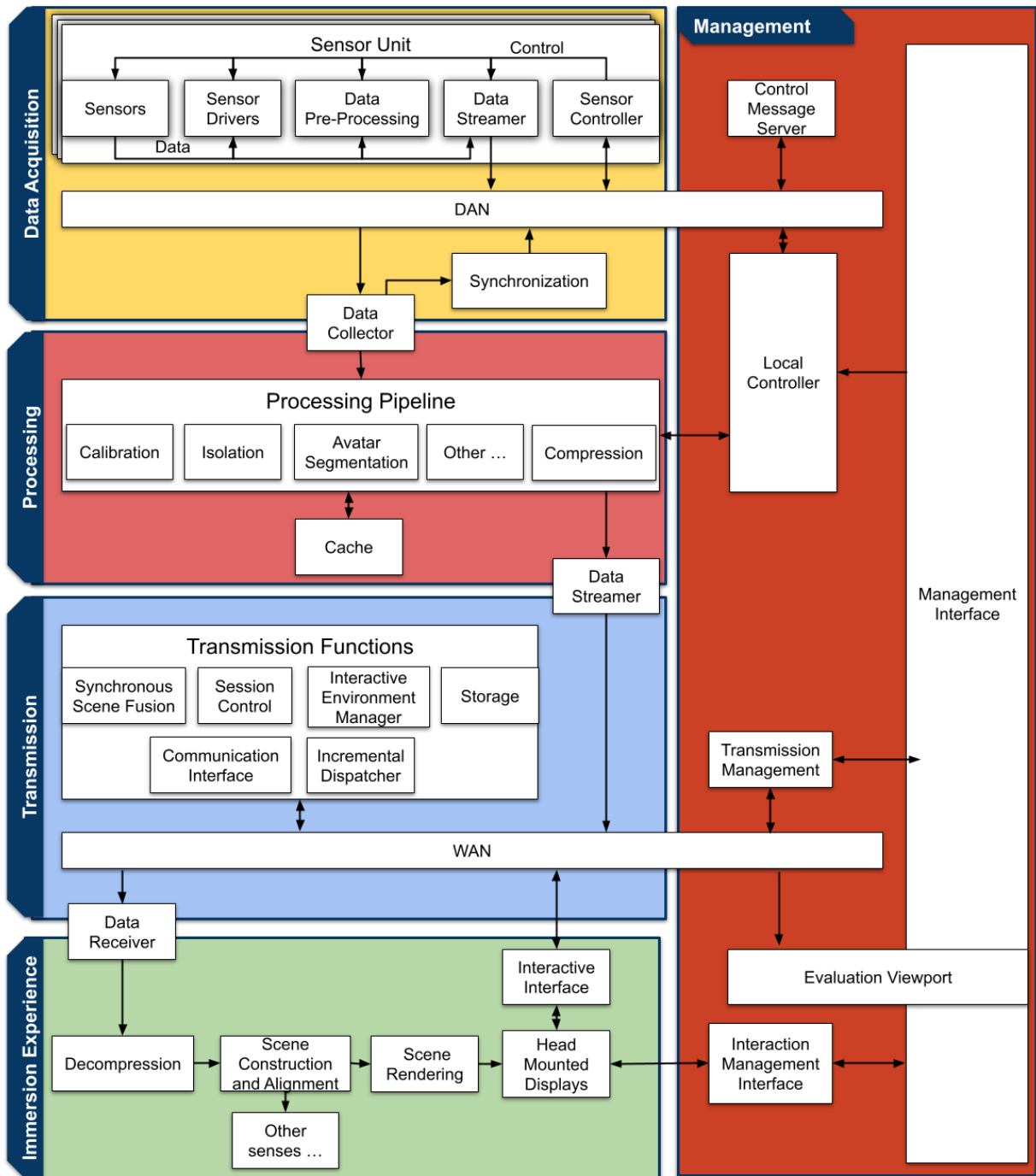


Figure 2.1: Proposed architecture for HTS [A2]

2.6. Implementation of HTPS and integrated algorithms

For the practical implementation of HTPS on the available physical infrastructure modules from the data acquisition layer, data processing layer, immersive experience layer, as well as modules from the control plane have been implemented (Figure 2.2). The physical infrastructure consists of three Kinect V2 sensors, each connected to an Intel NUC5i7RYH mini personal computer (NUC) via a USB 3.0 interface. The three mini computers, in turn, are connected to a high-performance computing unit (HPCU) through a gigabit-switched local area network.

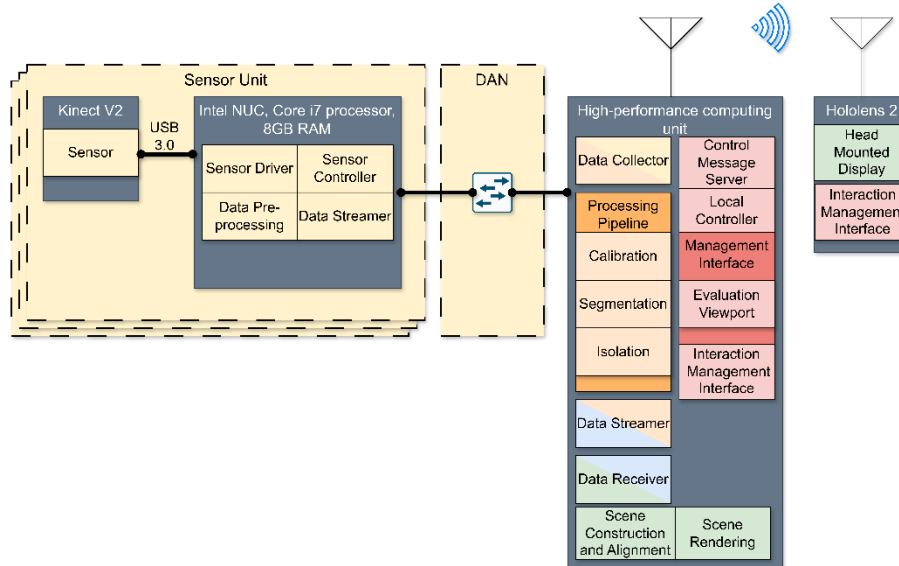


Figure 2.2: Physical distribution of the architectural blocks on the HTPS hardware infrastructure [A2]

For the data processing layer, an algorithm for isolating the reconstructed 3D scene has been specifically designed and subsequently integrated [A3]. HTPS captures a controlled physical space bounded by six walls forming a hexagonal cell, and therefore referred to as the “Bee Cube.” However, additional non-essential data are also captured, which necessitates the execution of such an algorithm. A block diagram is presented in Figure 2.6, defining the execution of several sequential steps.

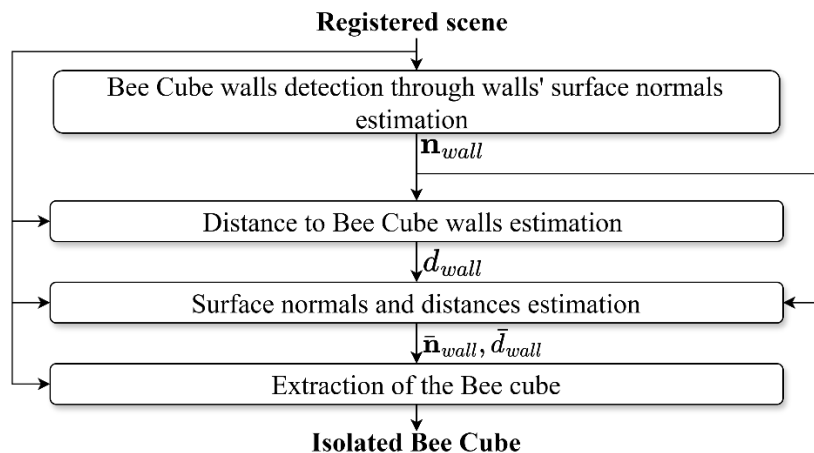


Figure 2.6: Graphical representation of the scene extraction algorithm

2.7. System testing results

Based on the calculated transport-layer payload for a single frame, amounting to 3,256,634 B, the maximum frame rate under simultaneous operation of the three Kinect V2 sensors has been determined. The results show that it is limited to only 8 FPS out of the possible 30. The characteristics of the selected transport-layer protocol are not favorable for the implementation of such systems, further increasing the load and creating conditions for desynchronization between the sensor units. The observed deviation of 27 ms at a frame period of 125 ms (21%) confirms this statement. The evaluated computational efficiency of online processing is insufficient to meet the requirements for minimal latency. This is due to a combination of factors, including algorithmic complexity, the limited computational power of the HPCU, the use of Python as an interpreted language, and others. In addition, the presence of a prolonged offline phase for pre-computing certain parameters in a fixed Kinect V2 installation significantly limits the feasibility of mobility-required scenarios. Quantitative results related to the evaluation of the implemented system are presented in Table 2.1. Figure 2.15 visualizes the reconstructed and isolated scene after applying the proposed scene isolation algorithm, as observed through HoloLens.

Table 2.1. Quantitative results of various performance indicators of the implemented HTPS

Type of measurement	Value
Data collector, sensor FPS = 8, deviation period (ms)	27
Computation time of the offline calibration stage (s)	161
Computation time of the online calibration stage (ms)	126
Computation time of the offline isolation stage (s)	3419
Computation time of the online isolation stage (ms)	140
Computation time of the segmentation stage (ms)	30

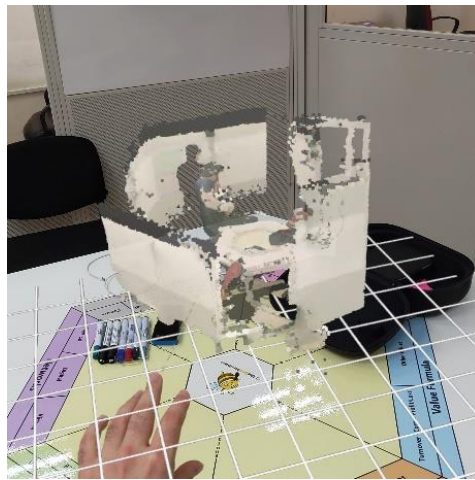


Figure 2.15: Reconstructed scene as observed through HoloLens 2

2.8. Conclusions

The transmission of 3D data is a challenge. In order to increase the frame rate, and consequently the QoE, it is necessary to integrate compression algorithms already within the sensor block or to use a DAN with higher communication capacity, which, however, is impractical from a user perspective. The efficiency of data processing must also be improved through the development of “lightweight” algorithms capable of operating across a wide range of HTC scenarios, independent of a fixed HTPS installation.

2.9. Contributions to Chapter 2

- 6) Application of a modular architectural approach for the implementation of HTPS, including hardware implementation of the system.
- 7) Software implementation of architectural blocks from the data acquisition layer, data processing layer, and immersive experience layer, namely: a data collector, a data processing module, and a scene reconstruction and alignment module.
- 8) Implementation of a calibration algorithm and a 3D scene extraction algorithm within the developed data processing module.
- 9) Development of a 3D scene extraction algorithm for isolating the scene from the surrounding environment through semantic 3D object detection.
- 10) Investigation and evaluation of the implemented HTPS.

CHAPTER 3. IMPACT OF MOBILE NETWORKS ON THE IMPLEMENTATION OF HTC

Chapter Three examines the impact of 5G and the potential of 6G on the realization of HTC and the deployment of HTPS among users. It aims to define a clear relationship between the main HTC requirements, namely the need for high bandwidth and consequently high data transmission rates, minimal latency, strict synchronization, and the need for complex and computationally intensive data processing algorithms, and the potential approaches of mobile networks through which these requirements can be met.

3.6. HTC and the new generations of mobile networks

Addressing the aforementioned technological requirements is fundamental for the realization of true HTC with ultra-low latency and a high level of QoE [129], [130]. However, it turns out that the transmission of holographic content over 5G networks is significantly limited due to the inability of current technologies to meet these requirements. Evidence of this is provided in Table 3.2, which compares HTC requirements with the corresponding network Key Performance Indicators (KPIs) defined for 5G and 6G. A significant increase in data transmission rates in 6G is promised, which would enable the transmission of substantially larger data volumes compared to 5G. In addition, the reduction of latency and jitter, combined with increased reliability, would provide advantages both in terms of improved synchronization and the successful realization of highly interactive scenarios. This supports the view that 6G is the promising solution for the actual realization of HTC.

After conducting an analytical review of numerous scientific publications with the purpose of examining the role of HTC in future 6G communications, the following statements have been derived. Holographic technologies and communications are expected to play a central role in both social and professional contexts. Almost all reviewed publications explicitly highlight HTC as a key application that will be enabled by advances in 6G technologies. Moreover, the emergence of a new type of 6G service is anticipated – simultaneously broadband, highly reliable, and ultra-low latency. In this context, Figure 3.1 illustrates the evolution of mobile network generations in terms of data transmission rates and provided services.

Table 3.2: Comparison of the main HTC requirements with the corresponding KPIs in 5G and 6G networks

Parameter	HTC requirement	5G	6G
Transmission rate	0.5-2 Gb/s Point cloud; 0.1-2 Tb/s light-field view hologram [37]	0.02 Tb/s Peak data rate; 0.1 Gb/s User-experienced data rate	1 Tb/s Peak data rate; 10 Gb/s User-experienced data rate
Latency	End-to-end latency: 50–100 ms (buffering); 20–50 ms (highly interactive applications); ~1 ms (ultra-low latency applications) [52]	Network transmission latency: 1 ms	Network transmission latency: 0.1 ms
Synchronization	Synchronization of multiple data streams arriving from different locations at the endpoint [39], [123]	Latency up to 1 ms, jitter – 1 ms	Latency up to 0.1 ms, jitter – 0.001 ms
Computation	Requirement for executing computationally intensive algorithms for 3D model reconstruction, compression, etc.	Introduction of network-based cloud computing	6G supporting both computation and communication

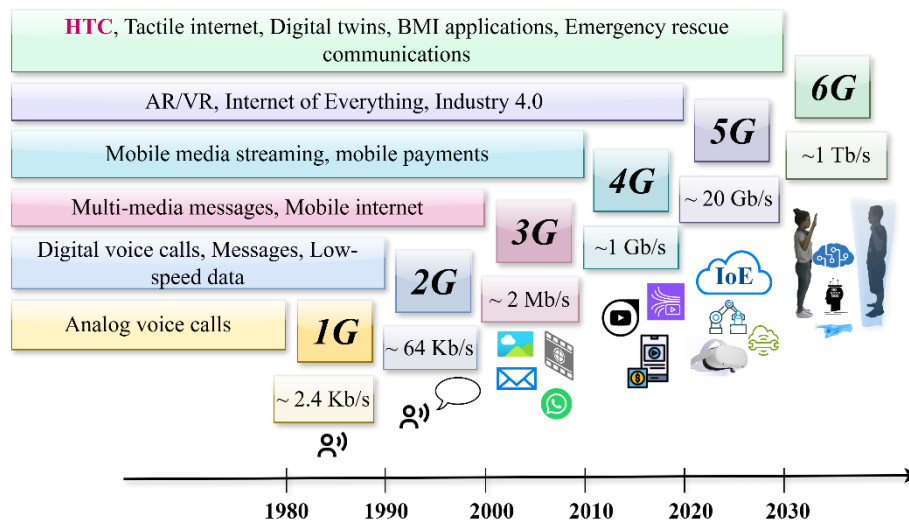


Figure 3.1: Evolution of mobile network generations in terms of data transmission rates and provided services [A4]

3.7. Potential approaches for the realization of HTC

Referring to the notion that HTC will be one of the core 6G services, the possible improvements expected to make this a reality are discussed. In this context, potential approaches for addressing the technological requirements of HTC are presented, classified into two main groups. The first group includes approaches directly related to the optimization of the communication network, in particular 6G, while the second group refers to optimizations on the user side. The first group is further categorized into three subgroups, namely improvements in the physical layer, network

architectures, and transport layer enhancements. Table 3.4 presents these approaches classified according to the above-described structure. Based on their impact on network KPIs [140], an assessment of their applicability to different HTC requirements [A4] has been performed, whereby Table 3.4 also illustrates the relationship between them and the corresponding technological challenges.

Table 3.4: Impact of potential approaches for improving HTC on the main technological challenges in HTC

Group of approaches	Approaches	Transmission rate	Latency	Synchronization	Computation
Physical layer improvements	Use of higher frequency bands above 300 GHz, including visible light	Yes	Yes	No	No
	Use of new modulation techniques	Yes	No	No	No
	Improvement of physical layer coding	Yes	Yes	No	No
	Application of MIMO and beamforming techniques together with signal processing algorithms	Yes	No	No	No
	Deployment of RIS on building surfaces, highways, and other objects	Yes	No	No	No
	Integrated sensing and localization	Yes	Yes	No	No
	Use of AI and ML at the physical layer	Yes	Yes	Yes	Yes
Network architectures	Deployment of flexible, adaptive, and scalable network architectures in 6G	Yes	Yes	Yes	Yes
	Integration of software-defined networking and network slicing	Yes	Yes	Yes	Yes
	Integration of MEC (Multi-access Edge Computing)	No	Yes	Yes	Yes
	Use of non-terrestrial connectivity	Yes	Yes	Yes	No
	Integration of AI in 6G	Yes	Yes	Yes	Yes
Transport layer improvements	New transport layer protocols	No	Yes	Yes	No
	Adaptive transmission of holographic content through support for streams with different priorities	Yes	Yes	No	No
User-side optimization	Assessment of network conditions and requirements for different HTC scenarios	Yes	Yes	Yes	Yes
	Innovative data compression methods	Yes	Yes	No	Yes
	Adaptive transmission based on user field of view (FOV) and focus of attention	Yes	No	No	No
	Use of semantics for enabling semantic communications	Yes	No	No	No

Referring to Table 3.4, Figure 3.2 illustrates the overall impact of the four groups of approaches on the technological requirements of HTC. Each ellipse represents a separate group and covers the requirements that it has the potential to address. Although this type of representation does not provide a quantitative evaluation, it offers clarity regarding the focus of the different groups of approaches in satisfying the corresponding HTC requirements. It should be emphasized, however, that the introduction of the various improvements should not be considered in isolation. While some approaches are targeted at solving specific HTC challenges, they may have a negative impact on others. Therefore, an optimal trade-off between different solutions should be sought, tailored to the specific HTC scenario.

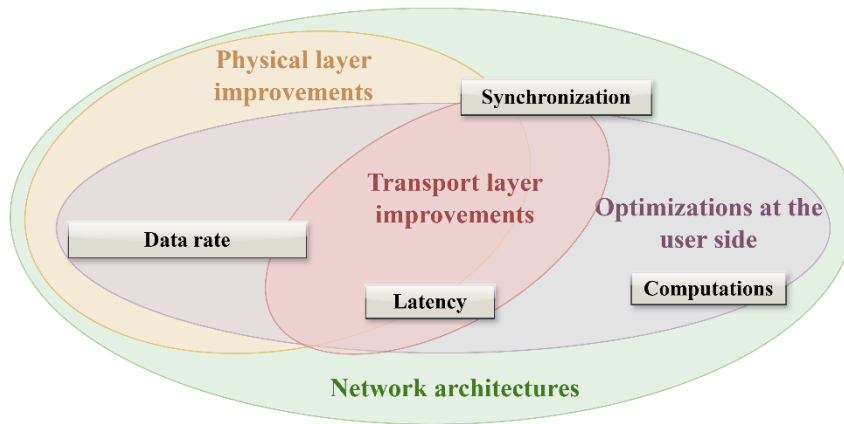


Figure 3.2: Relationship between the main technological requirements of HTC and the categories of approaches associated with their solution [A4]

3.8. AI and HTC

Regarding simultaneous optimization at both the communication network and user side, the integration of AI emerges as a key solution, supported by the following arguments [A4]. First, on the user side, AI will play an important role in achieving adaptive QoE. Specifically, AI-based methods can analyze user behavior, environmental factors, and available network resources in order to enable dynamic adjustments in data processing and transmission. Second, AI will contribute to high-speed content processing through the use of pre-trained models. Third, AI will also play a key role in improving semantic content understanding and enabling context-aware communication. Its involvement will mainly be expressed through the transition from precise data transmission to accurate representation of its underlying semantics. Fourth, AI will facilitate the analysis of large volumes of data [152]. This includes extracting historical context, providing diagnostic functions for autonomous fault detection, enabling prediction of future network states, critical events, user behavior, content popularity, and more, as well as suggesting possible solutions. This would support adaptive network planning and real-time management. As a result, future networks have the potential to evolve into flexible and adaptive systems capable of supporting a wide range of QoE-oriented applications. Therefore, despite the significant resources allocated for the optimal operation of 6G, its full efficiency will only be realized when intelligence is integrated. Finally, consider a scenario in which the resources and intelligence of the communication network are seamlessly combined with the resources and intelligence of the user side, all orchestrated by a coordinating AI module. Such an approach creates conditions for bypassing strict dependence on available resources. Consequently, prioritizing an intelligent approach rather than merely increasing limited capacity is the most effective strategy among those presented in Section 3.2.

3.9. Conclusions

Compared to 5G, 6G emerges as a key enabler for the realization of HTC. The provision of extended frequency resources, latency reduction, and the possibility of offloading computational processes from user devices to the network edge create significant prerequisites supporting this claim. However, considering the historical trend in which each new technological paradigm stimulates the development of the next generation of communication networks, it can be concluded that 6G is not the panacea for overcoming all future limitations of emerging HTC applications. Instead, the application of intelligent approaches for optimization is necessary both at the communication network level and on the user side.

3.10. Contributions to Chapter 3

- 3) Definition and classification of network KPIs related to the practical implementation of HTC. Evaluation of the impact of 5G and 6G on the realization of HTC, in accordance with the presented KPIs. Analytical overview examining the role of HTC in future 6G communications.
- 4) Definition and classification of potential optimization approaches on both the communication network side and the user side with respect to the realization of HTC. Evaluation of the presented approaches against the main HTC requirements. Analysis of the role of artificial intelligence in this context.

CHAPTER 4. POINT CLOUD REGISTRATION ALGORITHMS BASED ON DEEP LEARNING

4.7. Point cloud. Standard registration methods

A point cloud represents a set of points defined in Euclidean space through their 3D coordinates (x, y, z) and optionally described by additional attributes such as color channels, normals, and other features. Since the latter are not used in the developed algorithms, each point is represented as a vector composed solely of the three coordinates. Therefore, a given point cloud P can be represented as $P \in \mathbb{R}^{N \times 3}$, where N is the total number of points.

To perform registration of two point clouds, it is necessary to estimate (calibrate) and apply the optimal transformation parameters between them (rotation $R \in SO(3)$ and translation $\mathbf{t} \in \mathbb{R}^3$), such that they are aligned and represented in a common coordinate system. When the correspondence between the points of the two point clouds is unknown, the following optimization problem must be solved:

$$\min_{R \in SO(3), \mathbf{t} \in \mathbb{R}^3} \sum_{i=1}^N \|R\mathbf{p}_i + \mathbf{t} - \mathbf{q}_i\|^2 \quad (4.8)$$

The most commonly used approach to solve this problem is the Iterative Closest Points (ICP) algorithm [155]. It performs iterative minimization of the distance between the point clouds by establishing correspondences using a nearest-neighbor search algorithm and repeatedly computing Equation 4.8 until the error is minimized or the maximum number of iterations is reached.

4.8. AI-based registration of fully identical point clouds

An AI-based algorithm for the registration of two identical point clouds ($P \in \mathbb{R}^{N \times 3}$ and $Q \in \mathbb{R}^{N \times 3}$) is proposed through regression of a rotational quaternion [A5]. It should be noted that the algorithm

is primarily focused on predicting R , as the more complex component of the transformation, while the translation t is neglected. Thus, the optimization problem in Equation 4.8 takes the following form:

$$\min_{R \in SO(3)} \sum_{i=1}^N \|R\mathbf{p}_i - \mathbf{q}_i\|^2 \quad (4.10)$$

The proposed approach performs direct regression by parameterizing Equation 4.10 through a neural network model. A block diagram is illustrated in Figure 4.1. It consists of two main modules. The first is the feature extraction module. It processes P and Q in parallel using a shared single graph convolutional layer in order to extract local geometric features, and then, through max aggregation, provides the corresponding global features for each of the point clouds. The second module is the so-called point cloud alignment module. It combines the global geometric features and passes them through a shallow multilayer perceptron (Multilayer Perceptron, MLP) to predict the rotational quaternion, $quaternion \in \mathbb{R}^4$, from which the rotation matrix is directly computed.

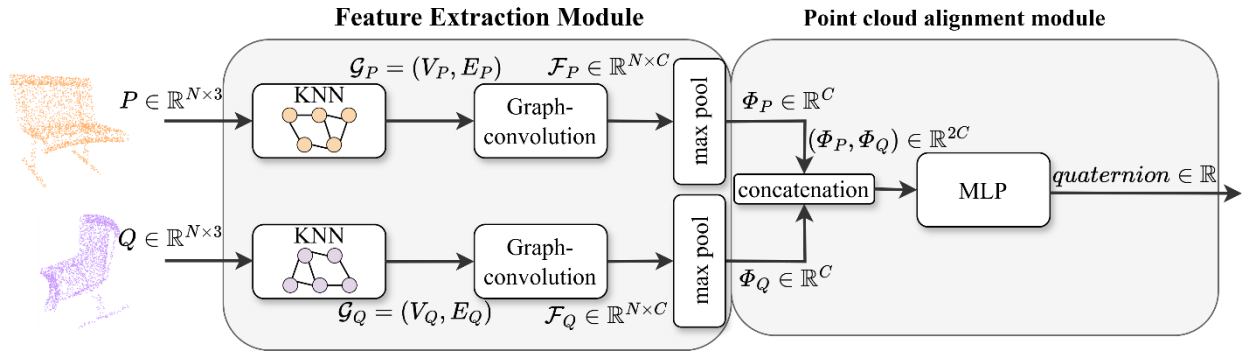


Figure 4.1: Block diagram of the proposed AI-based algorithm for registration of two identical point clouds

4.9. AI-based registration of partially overlapping point clouds in the latent space of their compressed representations

A method for direct registration of partially overlapping point clouds representing the same 3D object but captured from different viewpoints is presented [A6]. The algorithm is further extended to perform registration after compressing the individual point clouds, operating entirely in the resulting latent space (i.e., not in Euclidean space). Let $Q \in \mathbb{R}^{N \times 3}$ be the point cloud representing the physical object from a reference viewpoint, and let $P \in \mathbb{R}^{N \times 3}$ be the point cloud representing the same object but from an additional and different viewpoint. Assume that partial geometric overlap exists between them. The standard approach for performing registration in this case requires a set of corresponding points, i.e. $Q_c = \{q_{i_k}\}_{k=1}^K \subseteq Q$, $P_c = \{p_{j_k}\}_{k=1}^K \subseteq P$ and $C = \{(q_{i_k}, p_{j_k})\}_{k=1}^K$, where Q_c and P_c are subsets of Q and P , forming the so-called correspondence set C . The approach that aligns P_c to the reference subset Q_c , follows Equation 4.8 entirely, which in this case takes the following form:

$$\min_{R \in SO(3), t \in \mathbb{R}^3} \sum_{k=1}^K \|R\mathbf{p}_{j_k} + \mathbf{t} - \mathbf{q}_{i_k}\|^2 \quad (4.18)$$

The problem is parameterized using a neural network model, whereby the above equation is reformulated as:

$$\min_{CD} = (\tilde{P}, \bar{P}), \quad (4.19)$$

where $\tilde{P} = PR_{gt}^T + t_{gt}$ describes the ground-truth alignment mapping P to Q , $\bar{P} = f_{Decoder}(f_{RegNet}(z_Q, z_P))$ represents the alignment performed in the latent space. Here $z_Q = f_{Encoder}(Q)$ and $z_P = f_{Encoder}(P)$, where $f_{Encoder}$ and $f_{Decoder}$ are respectively the encoder and decoder components of the implemented autoencoder architecture used for compression and decompression of P and Q , while f_{RegNet} denotes the approximation of the proposed latent-space alignment network. CD denotes the symmetric Chamfer distance. Figure 4.2 presents a block diagram of the proposed registration method.

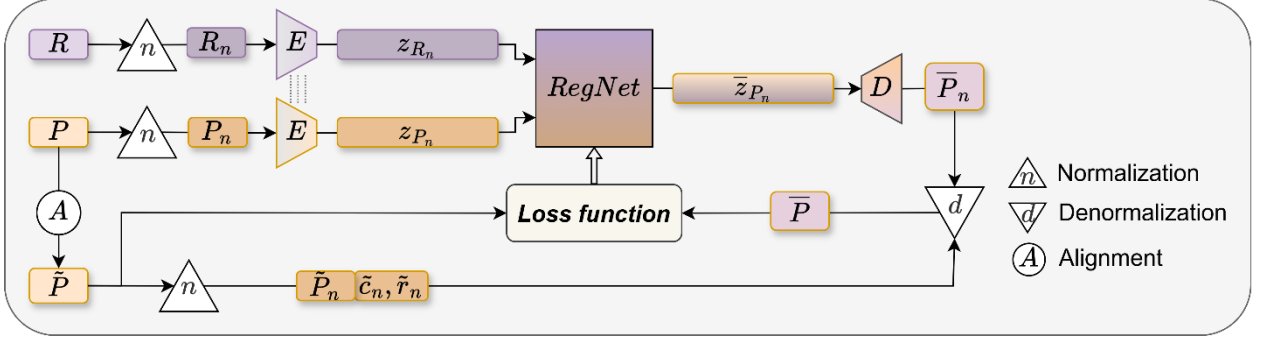


Figure 4.2: Block diagram of the proposed method for point cloud registration performed entirely in the latent space

4.10. Implementation and experimental results

The registration method for identical point clouds is trained using the ShapeNet dataset [158]. A qualitative representation of the obtained results is illustrated in Figure 4.4, where results achieved by other deep learning-based registration approaches and ICP are also provided. The comparison demonstrates the superiority of the proposed method, showing high registration efficiency across all considered 3D models. In addition, the results highlight the importance of proper initialization of point correspondences in ICP. Regarding the first two example 3D models, ICP appears to be trapped in local minima, and increasing the number of iterations has little or no effect. In contrast, the last example shows that a good initial estimate provided to ICP via the proposed algorithm leads to fast convergence and accurate alignment.

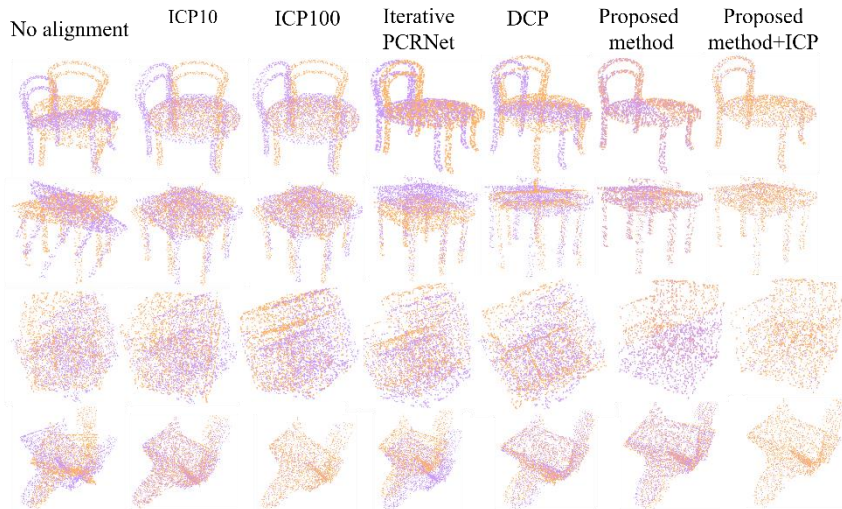


Figure 4.4: Visual comparison of results from the identical point cloud registration algorithm and other approaches

For training the registration model of partially overlapping point clouds, a dedicated dataset was created based on ShapeNet [A7]. For this purpose, ShapeNet objects were captured from different viewpoints by simulating a virtual camera, which was sequentially positioned on the surface of a sphere around them at predefined azimuth angles ($\{0,15,30,45,60,75,90\}$). Thus, for each model, 7 pairs of point clouds were generated, consisting of a reference view captured at 0 degrees and a corresponding additional view captured at each of the specified azimuth angles. Two variants of latent-space alignment were implemented, one using a standard transformation and another based on transformation parameterization. For comparison purposes, a third registration variant operating in Euclidean space was also implemented. This method explicitly regresses transformation parameters and applies them to the point cloud reconstructed by the decoder. All three variants are compared against ICP, applied both before and after the autoencoder responsible for compression and decompression. Figure 4.5 presents the results in terms of average Chamfer distance with respect to different azimuth angles, from which two main conclusions can be drawn. First, the results confirm the effectiveness of the proposed approaches compared to ICP. Second, the error of these methods remains relatively stable across all azimuth angles, indicating consistent and reliable alignment even at low levels of geometric overlap between point clouds. In contrast, ICP shows slightly better performance at lower azimuth angles and, as expected, degrades as the azimuth increases.

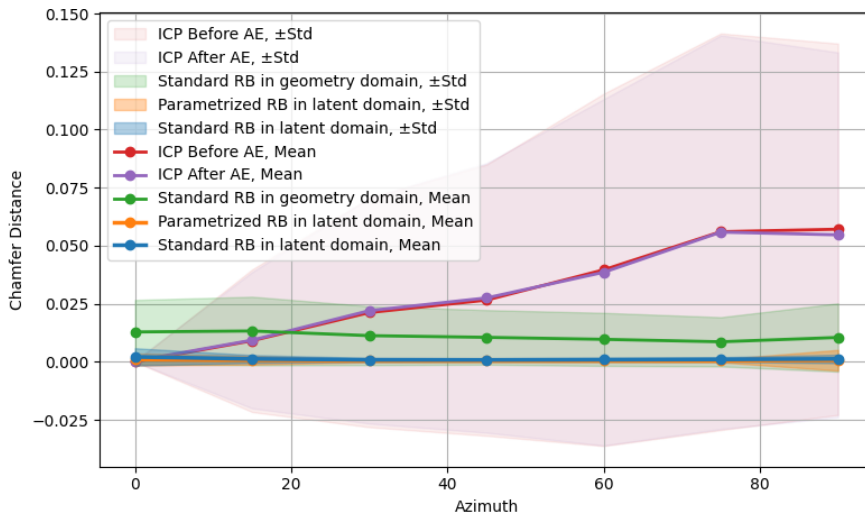


Figure 4.5: Statistical results of different registration methods in terms of Chamfer distance with respect to varying azimuth angles

4.11. Conclusions

In addition to traditional registration methods, it is also possible to effectively achieve registration using AI-based approaches. The advantage lies in the absence of the need to identify corresponding points between the point clouds being registered, as well as the elimination of calibration markers. This enables flexible changes in camera positions within a potential HTC scenario. In addition, the proposed approach for registration of partially overlapping point clouds demonstrates clear success with respect to the registration task, even when applied in an abstract latent space. Moreover, it contributes to the development of knowledge extraction from highly compressed and abstract latent domains, with direct applicability in 3D visual systems with limited bandwidth, such as robotics, augmented and virtual reality, and, of course, HTC.

4.12. Contributions to Chapter 4

- 4) A registration algorithm for identical point clouds transformed with respect to each other in Euclidean space has been developed. For this purpose, a neural network model has been designed, including modules for extracting local and global features, as well as a module for rotational quaternion regression.
- 5) A registration algorithm for point clouds representing the same physical object captured from different viewpoints such as there exists partial geometric overlap between them has been developed. The algorithm has been extended to perform registration entirely in the latent space of the compressed point cloud representations, through a specifically designed RegNet module.
- 6) A special dataset has been developed containing point clouds representing different views of a given 3D object.

CHAPTER 5. POINT CLOUD COMPRESSION ALGORITHMS BASED ON DEEP LEARNING

5.7. Traditional approaches for point cloud compression

In order to address the growing need for compact representation of 3D content, MPEG has standardized two main codecs designed for point cloud compression, namely Video-based Point Cloud Compression (V-PCC) and Geometry-based Point Cloud Compression (G-PCC) [163], [164]. An alternative solution developed by Google is the open-source library Draco [165], which also provides capabilities for compressing point clouds and 3D mesh models. On the other hand, AI-based algorithms for 3D content compression are gaining significant popularity and represent a broad research direction in the field. Moreover, according to [166], these approaches achieve substantial improvements over traditional methods in terms of the trade-off between achieved compression rate and resulting distortion of decoded data (so-called Rate–Distortion, RD). It is also believed that AI-based algorithms significantly increase coding efficiency by leveraging parallel computation capabilities, enabling much faster processing and thereby supporting the development of services such as HTC.

5.8. AI-based point cloud compression algorithm

A deep learning-based geometric point cloud compression algorithm, referred to as MilaNet [A8], is proposed. Let $P \in \mathbb{R}^{N \times 3}$ denote the point cloud subjected to the developed compression method, where N is the number of points and 3 corresponds to the geometric coordinates x, y, z , defined in Euclidean space. Let f_θ approximate the data compression function, and g_ϕ – denote the decompression function:

$$\hat{P} = g_\phi(f_\theta(P)), \quad (5.1)$$

where $\hat{P} \in \mathbb{R}^{N \times 3}$ corresponds to the decompressed point cloud. In the proposed algorithm, Equation 5.1 is parameterized using an autoencoder architecture, where f_θ and g_ϕ represent the encoder and decoder, respectively, whose structures are shown in Figures 5.1 and 5.2.

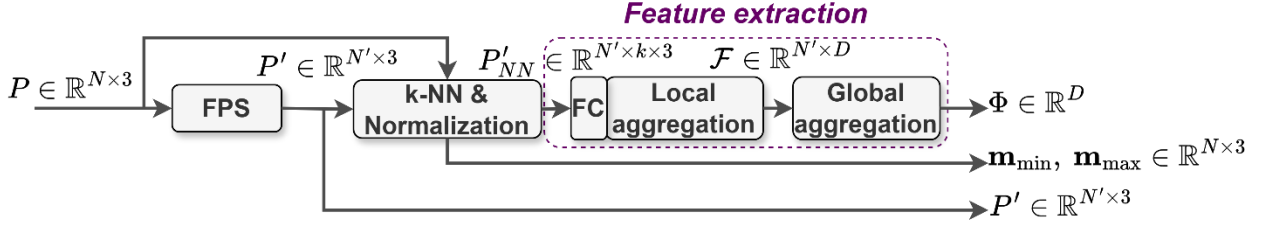


Figure 5.1: Block diagram of the encoder part of the proposed model for geometric point cloud compression

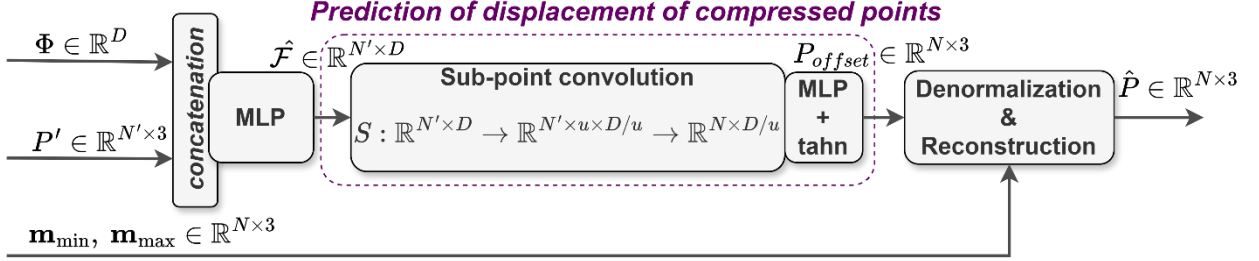


Figure 5.2: Block diagram of the decoder part of the proposed model for geometric point cloud compression

5.9. AI-based compression algorithm for partially overlapping multi-view point clouds

A compression algorithm for partially overlapping point clouds representing different views of the same object has been developed. The algorithm is based on deep learning and the Distributed Source Coding (DSC) principle. The goal is to investigate the possibility of achieving more efficient compression by exploiting the availability of side information (SI) at the decoder. Let $P \in \mathbb{R}^{N \times 3}$ and $Q \in \mathbb{R}^{N \times 3}$ be two point clouds captured from different viewpoints, with partial geometric overlap between them. Let the encoders \mathcal{E}_P and \mathcal{E}_Q independently encode each of the point clouds, applying:

$$\begin{aligned}\Phi_P &= f_{\mathcal{E}_P}(P), \\ \Phi_Q &= f_{\mathcal{E}_Q}(Q),\end{aligned}\tag{5.12}$$

where $f_{\mathcal{E}_P}$ and $f_{\mathcal{E}_Q}$ are functions describing the operation of \mathcal{E}_P and \mathcal{E}_Q , while $\Phi_P \in \mathbb{R}^C$ and $\Phi_Q \in \mathbb{R}^C$ are global feature vectors representing the compressed latent representations of each of the point clouds. Let D_{PQ} be a shared decoder that takes as input Φ_P and Φ_Q and reconstructs \hat{P} and \hat{Q} with corresponding distortion levels D_x and D_y , following:

$$\{\hat{P}, \hat{Q}\} = f_{D_{PQ}}(\Phi_P, \Phi_Q),\tag{5.13}$$

where $f_{D_{PQ}}$ describes D_{PQ} . The goal of the proposed algorithm is to reconstruct P using SI at the decoder in the form of the global feature Φ_Q . Since Q , is not directly used in this case, but rather Φ_Q , the concept of the algorithm follows the generalization of the Wyner–Ziv theorem for DSC, proposed by Berger and Tung (BT). To implement this idea, an existing autoencoder architecture known as the Transformer Graph Autoencoder (TGAE) [174] is adapted. The encoder structure is

preserved, while the decoder is modified to comply with the Berger–Tung formulation. Figure 5.7 shows a block diagram of the implemented model, emphasizing the decoder improvements.

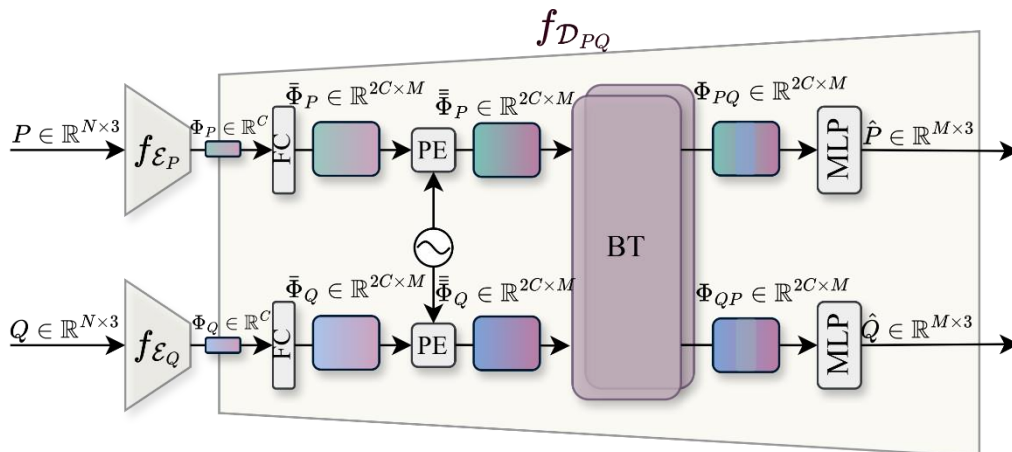


Figure 5.7: Block diagram for compression of multi-view point clouds based on DSC

5.10. Implementation and experimental results

The proposed MilaNet is trained on the ShapeNet and Mechanical Component Benchmark (MCB) datasets [175] and evaluated using two different configurations, namely MilaNet 64×64 and MilaNet 132×32. The obtained results are compared to those achieved by both traditional compression methods – G-PCC and Google Draco for both datasets – and another neural autoencoder, FoldingNet [157], for the ShapeNet dataset. Figure 5.9 presents the RD curves of the applied methods. An increase in the Peak signal to noise ratio (PSNR) is observed with the increase in the number of bits required to encode a single point, reaching a certain saturation level in deep learning-based methods. In the high-compression region, i.e., at very low bitrate values (0 to 2 bpp), the proposed MilaNet 64×64 achieves higher PSNR compared to Google Draco and MilaNet 132×32. After 2 bpp, its performance becomes comparable to that of FoldingNet. On the other hand, MilaNet 132×32 outperforms FoldingNet at high bitrate values (above 4 bpp) by approximately 0.5 dB. Although G-PCC appears to demonstrate strong performance at first glance, this is partially conditional due to specific characteristics of the codec, which significantly reduce and sparsify the number of output points, indicating a non-uniform evaluation of the achieved compression efficiency.

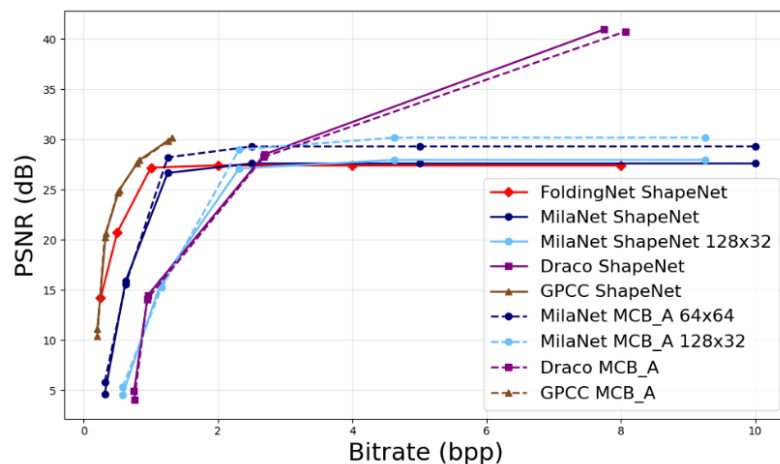


Figure 5.9: RD curves of the results from the implemented compression methods applied to the ShapeNet and MCB-A datasets

The DSC-based algorithm is trained using pairs of point clouds from the dataset published in [A7]. For evaluation purposes, a point cloud evaluation framework for multi-view captured data is proposed [A7], incorporating both geometric and information-theoretic metrics. A block diagram of the evaluation framework is presented in Figure 5.11.

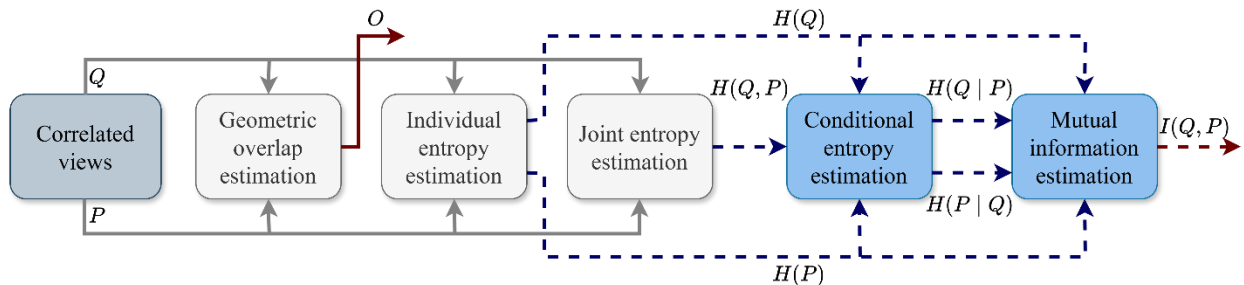


Figure 5.11: Block diagram of the evaluation framework illustrating the workflow and data flows used for computing geometric and information-theoretic metrics [A7]

Figures 5.13a and 5.13b present the conditional differential entropy $H(P|Q)$ and the mutual information $I(P; Q)$ with respect to the degree of geometric overlap $O_{\text{sym}}(Q, P)$ for the reference autoencoder TGAE and the proposed model, both before and after compression. It can be observed that, as the correlation between the point clouds increases, the uncertainty of P given Q decreases, since the latter increasingly determines it, and accordingly the mutual information between them increases. More interesting is the behavior of the curves after applying compression, where $H(P|Q)$ largely preserves the trend observed before compression. However, the lower absolute values indicate that compression reduces the correlation between the view pairs, with a more pronounced effect observed for TGAE. A similar conclusion can be drawn for $I(P; Q)$. For this reason, it can be stated that the proposed algorithm better preserves the geometric relationship between paired views compared to TGAE.

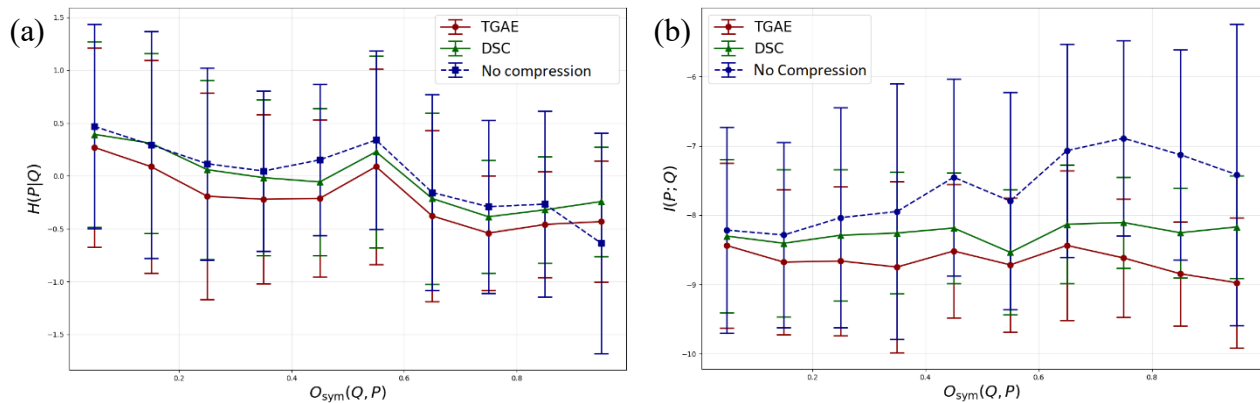


Figure 5.13: Ratio of the conditional entropy of view P with respect to the reference view Q and the mutual information, expressed as a function of the degree of geometric overlap

From a performance perspective, Figure 5.14a illustrates the dependence of PSNR on $O_{\text{sym}}(Q, P)$. It can be observed that the proposed model does not improve the performance of TGAE, with results being approximately 2 dB lower. However, Figure 5.14b, which expresses the trend of the PSNR difference $\Delta PSNR$ with respect to $O_{\text{sym}}(Q, P)$, shows a reduction in $\Delta PSNR$ as the correlation between the view pairs increases. Although the curve exhibits some fluctuations in its smoothness, mainly due to the characteristics of the dataset used, it demonstrates the potential of the implemented DSC-based idea and opens perspectives for future work in this direction.

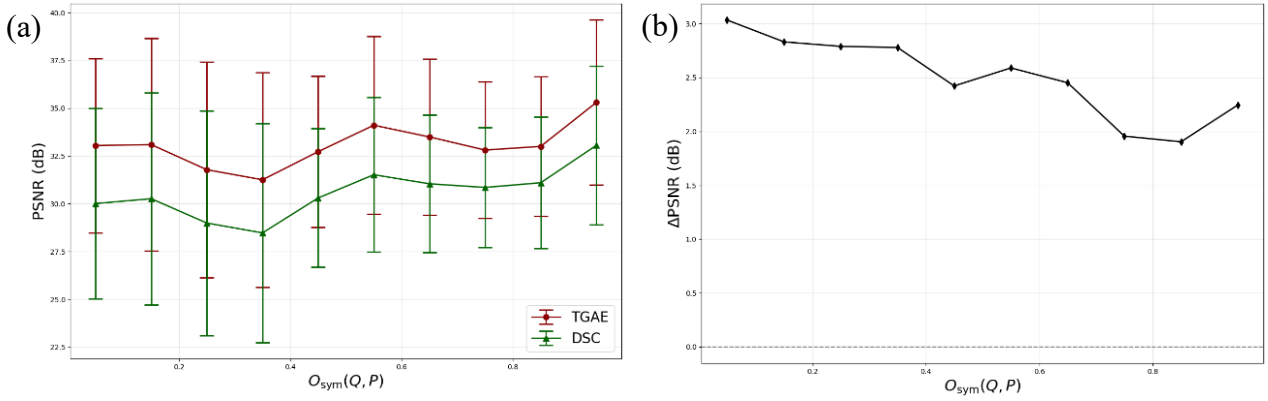


Figure 5.14: Relationship between the achieved compression PSNR and $\Delta PSNR$, as a function of the degree of geometric overlap

5.11. Conclusions

In addition to the application of traditional compression methods, the development of AI-based models represents a promising direction in this field. Moreover, the achieved decoding time and the implementation of a “lightweight” model with a very small number of parameters support the feasibility of deploying such algorithms on the receiving side of HTC users. On the other hand, the results also confirm the potential of AI for studying correlation via DSC between latent representations of individually compressed, partially overlapping point clouds. Although the proposed method does not currently confirm the statement of Berger and Tung, according to which joint decoding of views should improve compression efficiency, it nevertheless demonstrates an important trend. Accounting for inter-view correlation during decoding improves the quality of the reconstructed point clouds as the degree of geometric overlap increases. This finding justifies the use of DSC and encourages further research in this direction.

5.12. Contributions to Chapter 5

- 1) A deep learning-based point cloud compression algorithm has been developed. For this purpose, an autoencoder model has been designed, where the encoder reduces the number of points in the input point cloud and extracts a global feature vector representing the compressed information, while the decoder reconstructs the original input data from the compressed latent representation.
- 2) A deep learning-based algorithm for compressing point clouds representing the same physical object but captured from different viewpoints, such as partial geometric overlap between the views is presented, has been developed. The algorithm follows the DSC principle, extending the Berger and Tung joint decoding scenario by simultaneously exploiting the compressed latent representations of the input point clouds.
- 3) A multi-view point cloud evaluation framework has been proposed, based on geometric and information-theoretic metrics for performance assessment.

III. SCIENTIFIC-APPLIED AND APPLIED CONTRIBUTIONS

Scientific-Applied Contributions:

1. Definition of the functional components of HTPS and identification of the main challenges in their implementation.
2. Analysis and comparison of existing practical developments in the field.
3. Development of an algorithm for extracting a 3D scene from the environment through detection of semantic 3D objects.
4. Investigation and evaluation of a practically implemented HTPS.
5. Definition and classification of network KPIs related to the practical implementation of HTC. Evaluation of the impact of 5G and 6G on HTC implementation based on the defined KPIs. Analytical overview exploring the role of HTC in future 6G communications.
6. Definition and classification of potential approaches for optimizing both the communication network and the user side with respect to HTC implementation. Evaluation of these approaches against key HTC requirements. Analysis of the role of artificial intelligence in this context.
7. Development of an algorithm for registration of identical point clouds transformed relative to each other in Euclidean space. Design of a neural network model including modules for local and global feature extraction, as well as a module for regression of a rotational quaternion.
8. Development of an algorithm for registration of point clouds representing the same physical object from different views, with partial geometric overlap. Enhancement of the algorithm to perform registration entirely in the latent space of autoencoder-compressed point clouds through a specially designed neural network model called the RegNet module.
9. Development of a point cloud compression algorithm based on an autoencoder architecture. Design of an encoder that reduces the number of input points and extracts a global feature vector representing the compressed information. Design of a decoder that reconstructs the input data from the compressed latent representation into a point cloud.
10. Development of an algorithm for compressing point clouds representing the same physical object captured from different views, with partial geometric overlap. Design of a decoder approximating the principles of DSC, specifically unfolding the Berger–Tung scenario for joint decoding.

Applied Contributions:

1. Application of a modular architectural approach for implementing HTPS, including hardware implementation of the system.
2. Software implementation of architectural blocks from the data acquisition layer, data processing layer, and immersive experience layer, namely: a data collector, a data processing module, and a scene construction and alignment module.
3. Integration of a calibration algorithm and a 3D scene extraction algorithm into the developed data processing module.
4. Development of a database containing point clouds representing different views of a given 3D object.
5. Proposal of an evaluation framework for multi-view point clouds based on geometric and information-theoretic metrics.

IV. ABBREVIATIONS

3D	Three Dimensional	HTPS	Holographic Telepresence System
5G	Fifth Generation Mobile Networks	ICP	Iterative Closest Points
6G	Sixth Generation Mobile Networks	I/O	Input Output
AI	Artificial Intelligence	MCB	Mechanical Component Benchmark
BT	Berger-Tung	ML	Machine Learning
CD	Chamfer Distance	MLP	Multi-layer Perceptron
DAN	Data Acquisition Network	PSNR	Peak signal-to-noise ratio
DSC	Distributed Source Coding	QoE	Quality of Experience
FOV	Field of View	QoS	Quality of Service
G-PCC	Geometry-based Point Cloud Compression	RD	RateDistortion
HMD	Head Mounted Display	SI	Side Information
HPCU	High-performance Computing Unit	TGAE	Transformer Graph Autoencoder
HTC	Holographic-type Communications	V- PCC	Video-based Point Cloud Compression

V. LIST OF PUBLICATIONS RELATED TO DISSERTATION WORK

[A1] Petkova, R., Poulkov, V., Manolova, A., & Tonchev, K. (2022). Challenges in implementing low-latency holographic-type communication systems. *Sensors*, 22(24), 9617.

[A2] Bozhilov, I., Petkova, R., Tonchev, K., Manolova, A., & Poulkov, V. (2023). HOLOTWIN: A modular and interoperable approach to holographic telepresence system development. *Sensors*, 23(21), 8692.

[A3] Petkova, R., Tonchev, K., Manolova, A., & Poulkov, V. (2022, June). 3D scene extraction using plane detection algorithm. In *2022 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)* (pp. 347-351). IEEE.

[A4] Petkova, R., Bozhilov, I., Manolova, A., Tonchev, K., & Poulkov, V. (2024). On the way to holographic-type communications: Perspectives and enabling technologies. *IEEE Access*, 12, 59236-59259.

[A5] Petkova, R., Bozhilov, I., Tonchev, K., Manolova, A., & Poulkov, V. (2025, June). Learning 3d rotations from point cloud data. In *2025 60th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)* (pp. 1-4). IEEE.

[A6] Petkova, R., Bozhilov, I., Tonchev, K., Manolova, A., & Poulkov, V. (2025, November). Can rigid body transformation be applied in the latent domain of compressed multi-view point clouds?. In *2025 28th International Symposium on Wireless Personal Multimedia Communications (WPMC)* (pp. 1-6). IEEE.

[A7] Petkova, R., Bozhilov, I., Tonchev, K., Manolova, A., & Poulkov, V. (2026, February). Geometric and Information-Theoretic Evaluation of Multi-View Point Cloud Datasets. In *2026 Joint International Conference on Digital Arts, Media and Technology with ECTI Northern Section Conference on Electrical, Electronics, Computer and Telecommunications Engineering (ECTI DAMT & NCON)*. IEEE.

[A8] Aldimirova, M., Petkova, R., Bozhilov, I., Tonchev, K., Manolova, A., & Poulkov, V. (2025, November). Deep Learning-based Geometry Compression of Point Cloud Data. In *2025 28th International Symposium on Wireless Personal Multimedia Communications (WPMC)* (pp. 1-6). IEEE.

VI. BIBLIOGRAPHY

- [1] Y. Wang, Z. Su, N. Zhang, etc., “A survey on metaverse: Fundamentals, security, and privacy,” *IEEE Communications Surveys & Tutorials*, 2022.
- [2] L.-H. Lee, T. Braud, P. Zhou, etc., “All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda,” *arXiv preprint arXiv:2110.05352*, 2021.
- [9] T. Rhee, S. Thompson, D. Medeiros, R. Dos Anjos, A. Chalmers, “Augmented virtual teleportation for high-fidelity telecollaboration,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 5, pp. 1923–1933, 2020.
- [10] D. Saredakis, A. Szpak, B. Bircckhead, H. A. Keage, A. Rizzo, T. Loetscher, “Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis,” *Frontiers in Human Neuroscience*, vol. 14, p. 96, 2020.
- [11] C. Chang, K. Bang, G. Wetzstein, B. Lee, L. Gao, “Toward the next-generation VR/AR optics: a review of holographic near-eye displays from a human-centric perspective,” *Optica*, vol. 7, no. 11, pp. 1563–1578, 2020.
- [12] X. Pan, X. Xu, S. Dev, A. G. Campbell, “3D Displays: Their Evolution, Inherent Challenges and Future Perspectives,” in *Proceedings of the Future Technologies Conference*, Springer, 2021, pp. 397–415.
- [13] J. Gugenheimer, C. Mai, M. McGill, J. Williamson, F. Steinicke, K. Perlin, “Challenges using head-mounted displays in shared and social spaces,” in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–8.
- [14] P.-A. Blanche, “Holography, and the future of 3D display,” *Light: Advanced Manufacturing*, vol. 2, no. 4, pp. 446–459, 2021.
- [15] P. A. Kara, R. R. Tamboli, O. Doronin, etc., “The key performance indicators of projection-based light field visualization,” *Journal of Information Display*, 2019.
- [17] J. Lawrence, D. B. Goldman, S. Achar, etc., “Project Starline: A high-fidelity telepresence system,” 2021.
- [18] J. Orlosky, M. Sra, K. Bektaş, etc., “Telelife: the future of remote living,” *arXiv preprint arXiv:2107.02965*, 2021.
- [19] L. Yoon, D. Yang, C. Chung, S.-H. Lee, “A Full Body Avatar-Based Telepresence System for Dissimilar Spaces,” *arXiv preprint arXiv:2103.04380*, 2021.
- [20] M. Montagud, J. Li, G. Cernigliaro, A. El Ali, S. Fernández, P. Cesar, “Towards socialVR: evaluating a novel technology for watching videos together,” *Virtual Reality*, pp. 1–21, 2022.
- [21] O. Ozioko, R. Dahiya, “Smart tactile gloves for haptic interaction, communication, and rehabilitation,” *Advanced Intelligent Systems*, vol. 4, no. 2, 2022.
- [22] J. Wang, Y. Qi, “A Multi-User Collaborative AR System for Industrial Applications,” *Sensors*, vol. 22, no. 4, p. 1319, 2022.
- [23] Y. Wang, P. Wang, Z. Luo, Y. Yan, “A novel AR remote collaborative platform for sharing 2.5D gestures and gaze,” *The International Journal of Advanced Manufacturing Technology*, vol. 119, no. 9, pp. 6413–6421, 2022.
- [24] H. Regenbrecht, N. Park, S. Duncan, etc., “Atea Presence—Enabling Virtual Storytelling, Presence, and Tele-Co-Presence in an Indigenous Setting,” *IEEE Technology and Society Magazine*, vol. 41, no. 1, pp. 32–42, 2022.

- [25] D. Kim, D. Jo, “Effects on Co-Presence of a Virtual Human: A Comparison of Display and Interaction Types,” *Electronics*, vol. 11, no. 3, p. 367, 2022.
- [26] K. Yu, G. Gorbachev, U. Eck, F. Pankratz, N. Navab, D. Roth, “Avatars for teleconsultation: effects of avatar embodiment techniques on user perception in 3D asymmetric telepresence,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 11, pp. 4129–4139, 2021.
- [27] Sensiks, “Sensory Reality Pods & Platform.” url: <https://www.sensiks.com/>
- [28] S. Schwarz, M. Preda, V. Baroncini, etc., “Emerging MPEG standards for point cloud compression,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 133–148, 2018.
- [32] T. Fujihashi, T. Koike-Akino, S. Chen, T. Watanabe, “Wireless 3D point cloud delivery using deep graph neural networks,” in *ICC 2021 - IEEE International Conference on Communications*, IEEE, 2021, pp. 1–6.
- [36] J. van der Hooft, M. T. Vega, T. Wauters, etc., “From capturing to rendering: Volumetric media delivery with six degrees of freedom,” *IEEE Communications Magazine*, vol. 58, no. 10, pp. 49–55, 2020.
- [37] A. Clemm, M. T. Vega, H. K. Ravuri, T. Wauters, F. De Turck, “Toward truly immersive holographic-type communication: Challenges and solutions,” *IEEE Communications Magazine*, vol. 58, no. 1, pp. 93–99, 2020.
- [38] T. Fujihashi, T. Koike-Akino, T. Watanabe, P. V. Orlik, “HoloCast+: Hybrid digital-analog transmission for graceful point cloud delivery with graph Fourier transform,” *IEEE Transactions on Multimedia*, vol. 24, pp. 2179–2191, 2021.
- [39] S. Anmulwar, N. Wang, A. Pack, V. S. H. Huynh, J. Yang, R. Tafazolli, “Frame Synchronisation for Multi-Source Holographic Teleportation Applications—An Edge Computing Based Approach,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, IEEE, 2021, pp. 1–6.
- [40] I. Selinis, N. Wang, B. Da, D. Yu, R. Tafazolli, “On the Internet-scale streaming of holographic-type content with assured user quality of experience,” in *IFIP Networking Conference*, IEEE, 2020, pp. 136–144.
- [41] P. Qian, V. S. H. Huynh, N. Wang, S. Anmulwar, D. Mi, R. Tafazolli, “Remote Production for Live Holographic Teleportation Applications in 5G Networks,” *IEEE Transactions on Broadcasting*, vol. 68, no. 2, pp. 451–463, 2022.
- [43] W. Zhu, Z. Ma, Y. Xu, L. Li, Z. Li, “View-dependent dynamic point cloud compression,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 2, pp. 765–781, 2020.
- [44] G. Cernigliaro, M. Martos, M. Montagud, A. Ansari, S. Fernandez, “PC-MCU: Point cloud multipoint control unit for multi-user holoconferencing systems,” in *Proceedings of the 30th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, 2020, pp. 47–53.
- [45] C. J. Blackwell, J. Khan, X. Chen, “Holographic 3D Telepresence System with Light Field 3D Displays and Depth Cameras over a LAN,” in *SID Symposium Digest of Technical Papers*, vol. 52, 2021, pp. 761–763.
- [46] D. Roth, K. Yu, F. Pankratz, etc., “Real-time mixed reality teleconsultation for intensive care units in pandemic situations,” in *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, IEEE, 2021, pp. 693–694.
- [47] S. F. Langa, M. Montagud, G. Cernigliaro, D. R. Rivera, “Multiparty Holomeetings: Toward a New Era of Low-Cost Volumetric Holographic Meetings in Virtual Reality,” *IEEE Access*, vol. 10, pp. 81856–81876, 2022.

- [48] R. Kachach, P. Perez, A. Villegas, E. Gonzalez-Sosa, “Virtual tour: An immersive low cost telepresence system,” in IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2020, pp. 504–506.
- [49] F. E. Fadzli, A. W. Ismail, “A Robust Real-Time 3D Reconstruction Method for Mixed Reality Telepresence,” *International Journal of Innovative Computing*, vol. 10, no. 2, 2020.
- [50] S. Vellingiri, J. White-Swift, G. Vania, etc., “Experience with a Trans-Pacific Collaborative Mixed Reality Plant Walk,” in IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), IEEE, 2020, pp. 238–245.
- [51] R. Li, etc., “Towards a new internet for the year 2030 and beyond,” in Proceedings of the 3rd Annual ITU IMT-2020/5G Workshop Demo Day, 2018, pp. 1–21.
- [52] Qualcomm, “VR and AR Pushing Connectivity Limits,” 2018.
- [57] M. Pakanen, P. Alavesä, N. van Berkel, T. Koskela, T. Ojala, “Nice to see you virtually: Thoughtful design and evaluation of virtual avatar of the other user in AR and VR based telepresence systems,” *Entertainment Computing*, vol. 40, p. 100457, 2022.
- [59] G. Gamelin, A. Chellali, S. Cheikh, A. Ricca, C. Dumas, S. Otmane, “Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments,” *Personal and Ubiquitous Computing*, vol. 25, no. 3, pp. 467–484, 2021.
- [62] Ericsson, “The spectacular rise of holographic communication,” 2022. url: <https://www.ericsson.com/en/about-us/new-world-of-possibilities/imagine-possible-perspectives/holographic-communication>.
- [63] ITU, “Representative use cases and key network requirements for Network 2030,” 2020. url: <https://www.itu.int/pub/T-FG-NET2030-2020-SUB.G1>.
- [64] Huawei, “Communications Network 2030,” 2021. url: <https://www.huawei.com/en/giv/communications-network-2030>.
- [65] V. Memmesheimer, A. Ebert, “Scalable extended reality: A future research agenda,” *Big Data and Cognitive Computing*, vol. 6, no. 1, p. 12, 2022.
- [68] F. Jasche, J. Kirchhübel, T. Ludwig, P. Tolmie, “BeamLite: Diminishing Ecological Fractures of Remote Collaboration through Mixed Reality Environments,” in Proceedings of the 10th International Conference on Communities & Technologies, 2021, pp. 200–211.
- [71] Hackernoon, “3 Common Types of 3D Sensors: Stereo, Structured Light, and ToF.” url: <https://hackernoon.com/3-common-types-of-3d-sensors-stereo-structured-light-and-tof-194033f0>.
- [72] VRcompare, “VRcompare - The Internet's Largest VR & AR Headset Database.” url: <https://vr-compare.com/>
- [74] A. Manolova, K. Tonchev, V. Poulkov, S. Dixir, P. Lindgren, “Context-aware holographic communication based on semantic knowledge extraction,” *Wireless Personal Communications*, vol. 120, no. 3, pp. 2307–2319, 2021.
- [77] H. Bai, P. Sasikumar, J. Yang, M. Billingham, “A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing,” in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp. 1–13.
- [78] A. Jing, K. W. May, M. Naem, G. Lee, M. Billingham, “EyemR-Vis: Using bi-directional gaze behavioural cues to improve mixed reality remote collaboration,” in Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, 2021, pp. 1–7.
- [79] J. Rekimoto, K. Uragaki, K. Yamada, “Behind-the-mask: A face-through head-mounted display,” in Proceedings of the 2018 International Conference on Advanced Visual Interfaces, 2018, pp. 1–5.

- [80] Y. Zhao, Q. Xu, W. Chen, etc., “Mask-off: Synthesizing face images in the presence of head-mounted displays,” in IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, 2019.
- [81] G. Song, J. Cai, T.-J. Cham, J. Zheng, J. Zhang, H. Fuchs, “Real-time 3D face-eye performance capture of a person wearing a VR headset,” in Proceedings of the 26th ACM International Conference on Multimedia, 2018, pp. 923–931.
- [82] A. Nijholt, “Capturing obstructed nonverbal cues in augmented reality interactions: A short survey,” in Proceedings of International Conference on Industrial Instrumentation and Control, Springer, 2022, pp. 1–9.
- [83] S.-Y. Chen, Y.-K. Lai, S. Xia, P. Rosin, L. Gao, “3D face reconstruction and gaze tracking in head-mounted displays for virtual interaction,” IEEE Transactions on Multimedia, 2022.
- [84] S. Lombardi, J. Saragih, T. Simon, Y. Sheikh, “Deep appearance models for face rendering,” ACM Transactions on Graphics, vol. 37, no. 4, pp. 1–13, 2018.
- [88] Z. Akhtar, K. Siddique, A. Rattani, S. L. Lutfi, T. H. Falk, “Why is multimedia quality of experience assessment a challenging problem?,” IEEE Access, vol. 7, pp. 117897–117915, 2019.
- [89] G. Kougioumtzidis, V. Poulkov, Z. Zaharis, P. Lazaridis, “A survey on multimedia services QoE assessment and machine learning-based prediction,” IEEE Access, vol. 10, pp. 19507–19538, 2022.
- [90] S. Y. Chin, B. R. Quinton, “Dynamic object comprehension: A framework for evaluating artificial visual perception,” arXiv preprint arXiv:2202.08490, 2022.
- [91] S. Vijayakumar, P. Corcoran, R. Flynn, N. Murray, “AI-derived quality of experience prediction based on physiological signals for immersive multimedia experiences,” in Proceedings of the 13th ACM Multimedia Systems Conference, 2022, pp. 403–407.
- [92] J. A. De Guzman, K. Thilakarathna, A. Seneviratne, “Security and privacy approaches in mixed reality: A literature survey,” ACM Computing Surveys, vol. 52, no. 6, pp. 1–37, 2019.
- [95] W. Mazurczyk, P. Bisson, R. P. Jover, K. Nakao, K. Cabaj, “Challenges and novel solutions for 5G network security, privacy and trust,” IEEE Wireless Communications, vol. 27, no. 4, pp. 6–7, 2020.
- [140] W. Jiang, B. Han, M. A. Habibi, H. D. Schotten, “The road towards 6G: A comprehensive survey,” IEEE Open Journal of the Communications Society, vol. 2, pp. 334–366, 2021.
- [152] K. B. Letaief, W. Chen, Y. Shi, J. Zhang, Y.-J. A. Zhang, “The roadmap to 6G: AI-empowered wireless networks,” IEEE Communications Magazine, vol. 57, no. 8, pp. 84–90, 2019.
- [155] P. J. Besl, N. D. McKay, “A method for registration of 3-D shapes,” in Proceedings of SPIE Sensor Fusion IV: Control Paradigms and Data Structures, vol. 1611, 1992, pp. 586–606.
- [158] A. X. Chang, T. Funkhouser, L. Guibas, etc., “ShapeNet: An information-rich 3D model repository,” arXiv preprint arXiv:1512.03012, 2015.
- [163] ISO/IEC JTC1/SC29/WG7, “V-PCC Codec Description,” 2020.
- [164] ISO/IEC JTC1/SC29/WG7, “G-PCC Codec Description V9,” 2020.
- [165] Google, “Draco: 3D Data Compression.” url: <https://google.github.io/draco/>
- [166] I. Bozhilov, R. Petkova, K. Tonchev, A. Manolova, “A systematic survey into compression algorithms for three-dimensional content,” IEEE Access, vol. 12, pp. 141604–141624, 2024.
- [174] I. Bozhilov, R. Petkova, K. Tonchev, A. Manolova, V. Poulkov, H. V. Poor, “Autoencoder architectures for low-rate sparse point cloud geometry coding,” IEEE Access, vol. 13, pp. 214122–214140, 2025.
- [175] S. Kim, H.-g. Chi, X. Hu, Q. Huang, K. Ramani, “A large-scale annotated mechanical components benchmark for classification and retrieval tasks with deep neural networks,” in Proceedings of the European Conference on Computer Vision (ECCV), 2020.



TECHNICAL UNIVERSITY OF SOFIA
FACULTY OF TELECOMMUNICATIONS
DEPARTMENT “COMMUNICATION NETWORKS”

Radostina Rosenova Petkova, MSc

**METHODS AND ALGORITHMS FOR IMPLEMENTING
HOLOGRAPHIC-TYPE COMMUNICATION WITH ULTRA-LOW
LATENCY**

ABSTRACT of Ph.D THESIS

The continuous evolution of mobile communication networks toward higher throughput, ultra-low latency, greater connection density, and improved reliability is creating the foundation for upcoming immersive communication services. Following the progression from voice communication to video conferencing, the next major paradigm shift is expected to be Holographic-Type Communication (HTC) – a novel technology that enables the real-time transmission and reproduction of spatially rich holographic data, offering a substantially enhanced sense of telepresence and user interaction. However, the practical realization of HTC presents significant technological challenges, particularly in the areas of system architecture development, real-time data processing, and communication efficiency under strict latency constraints. Therefore, this dissertation aims to systematically investigate the key aspects of HTC implementation through a comprehensive analysis of existing challenges and the exploration of effective approaches for their mitigation.

To support this objective, the work provides a systematic study of Holographic Telepresence Systems (HTPS), including an in-depth analysis of their functional components, key implementation challenges, and existing system realizations. Based on these findings, a modular architectural framework for HTPS is proposed, followed by implementation of a real system. Furthermore, the dissertation analyzes the role of next-generation mobile networks as critical enablers for the deployment of HTC, with particular focus on their capability to satisfy stringent system requirements. A central contribution of the thesis is the development of novel algorithms for efficient three-dimensional visual data processing, considering the stringent HTC constraints. These algorithms demonstrate the feasibility of intelligent, task-oriented HTPS design and help bridge the gap between theoretical HTC concepts and their practical realization.