

Инновации, представленные на IBC 2025

Дэвид Кёрк, лондонский корреспондент MediaVision

Лондонский корреспондент журнала MediaVision Дэвид Кёрк является генеральным директором базирующейся в Соединенном королевстве компании Stylus Media Communications, которая специализируется на аналитическо-публицистической деятельности в сфере вещательной индустрии. Прежде работавший редактором журнала International Broadcast Engineer, Дэвид – ветеран индустрии, посетивший все выставки и конференции IBC, начиная с 1968 года, а также 35 выставок и конференций NAB подряд.



ние первого и второго способов. И, наконец, в-четвертых, посещение технической конференции.

Четвертый способ полезен для выяснения перспектив, поскольку технические доклады конференции IBC постоянно показывают свою состоятельность как индикаторы направления развития отрасли в ближайшие десятилетия. Конференция также дает возможность задать вопросы докладчикам в процессе их выступления и поговорить с глазу на глаз с отраслевыми коллегами.

То, что приводится ниже, это выдержки из девяти докладов, сделанных на конференции IBC 2025, которые суммарно отражают текущее состояние развития индустрии. Машинное обучение продолжает оставаться острой темой спустя семь лет после того, как было применено компанией NHK Япония для преобразования монохромного киноизображения в цветной видеоконтент.

Есть четыре основных способа выяснить, что нового и важного демонстрируется на выставке вещательных технологий, такой как Международная вещательная конвенция – IBC (International Broadcasting Convention), состоявшаяся в середине сентября в выставочном центре RAI в Амстердаме.

Во-первых, это максимально тщательный и скрупулезный обход выставочных павильонов и открытых площадок. Во-вторых, изучение интернет-сайтов различных производителей с помощью каталога на сайте IBC либо по алфавитному списку компаний-участниц, на что потребуется месяц или около того. В-третьих, это сочета-

Создание исходного контента

Винтер Клинкемайли (W. Clincemaijle) с коллегами из Университета Гента (Бельгия) обратили внимание на контекстное применение дополненной реальности в дикой природе для улучшения трансляций велосипедных гонок (Contextual AR in the wild: enhancing live cycling broadcasts): «Недавние исследования показывают, что есть потребность в инновациях применительно к спортивным трансляциям. Последние разработки в сфере обогащения видео могут внести свой вклад. Пока дополненная реальность успешно улучшает вовлечение болельщиков

Подключайтесь спокойно

- Кабель для инсталляций
- Tактический кабель
- Кабельные сборки
- Надёжно

NETWORK

АО "Ом Нетворк" 195196, Санкт-Петербург, Таллинская, 7
Тел: +7 (812) 612-81-33 +7(812) 309-22-44 www.omnetwork.ru



Рис. 1. Применение элементов дополненной реальности в процессе трансляции велогонки

в таких управляемых средах, как футбольные стадионы и баскетбольные площадки. Ее применение для спортивных дисциплин, состязания по которым проходят под открытым небом, таких как велоспорт, практически отсутствует. В течение прошедших двух лет мы работали над тем, чтобы устранить этот пробел, разрабатывая полностью автоматизированный рабочий процесс для дополнения велогонщиков и других ключевых объектов виртуальными компонентами в процессе прямой трансляции гонки».

На рис. 1 показаны примеры вывода плашек с именами поверх видео.

«На первом этапе рабочего процесса, – продолжают авторы, – выполняется автоматизированная идентификация гонщика и команды, снятых с разных ракурсов. Второй этап предусматривает рассмотрение использования 3D-предсказания позы велогонщиков для создания элементов дополненной реальности. Наконец, мы также уделили внимание разработке автоматического обнаружения интересных объектов в кадре, их распознавания и визуализации.

Мы разработали модульный конвейер для распознавания велосипедистов и точного размещения элементов дополненной реальности. В нем сочетаются обнаружение, отслеживание, классификация команды и 3D-предсказание позы. Все эти данные поступают в систему, которую можно применять как в прямом эфире, так и к записанному материалу. Кроме того, мы представили подход, предусматривающий распознавание и трекинг интересующих объектов в режиме реального времени, что снижает ручную нагрузку и позволяет проводить трансляцию более динамично. Эффективность метода оптимизирована для прямых трансляций.

Весь процесс распознавания занимает примерно 250 мс на каждую итерацию. Хотя процесс можно ускорить за счет применения более мощных аппаратных средств, того, что есть, уже достаточно, поскольку система выполняет эти итерации только один раз в секунду. Для работы в режиме реального времени, что необходимо для соответствия требованиям прямых трансляций, требуется оптический датчик движения. Это достижимо в зависимости от характеристики оборудования и выбранного алго-

ритма трекинга. Например, датчик MedianFlow действует со скоростью порядка 10 мс/кадр, что обеспечивает точный трекинг в режиме реального времени в процессе прямой трансляции».

Вай Джей Ли с коллегами из Korean Broadcasting System представили доклад, описывающий конструкцию и практическое применение системы поддержки вещания. Название доклада – «Разработка системы поддержки на базе искусственного интеллекта для прямых трансляций Олимпиады в Париже: плавание и фехтование» (Developing an AI-based assistant system for Paris Olympics live broadcasts: swimming and fencing): «Система предназначена для устранения сложностей, связанных с прямыми спортивными трансляциями, включая распознавание мельчайших визуальных сигналов, устранение неясностей в ракурсах от широкоугольных камер и интеграцию насыщенных контентом элементов, таких как национальные флаги. Примененная система существенно снизила нагрузку на оператора, улучшила точность синхронизации и сделала контент более привлекательным для аудитории. Забегая вперед, можно сказать, что эта работа закладывает фундамент для более широкого применения, включая другие виды спорта, использование аудио и данных синхронизации, а также движение в направлении вещательных рабочих процессов, в которых автоматизация сочетается с редакторским контролем для обеспечения точной и интересной прямой трансляции».

Авторы доклада также описали демонстрацию AI-производства любительского видео с помощью 8K-камер и монтажа средствами Vvertigo: «KBS и NHK совместно оценили съемочный рабочий процесс с применением AI, который извлекает многочисленные HD-сигналы любительского видео из общего 8K-источника, заменяя многокамерные комплексы. Разработанная в KBS Media технология генерации пространственного изображения Vvertigo распознает исполнителей и связывает их с алгоритмом кластеризации на основе графов. Система была протестирована в процессе съемки вышедшей на канале NHK вечерней новогодней передачи Kohaku Uta Gassen, где три монтажера, прошедшие 4-часовое обучение, обеспечили формирование 8...10-минутных видеофрагментов с исполнителем

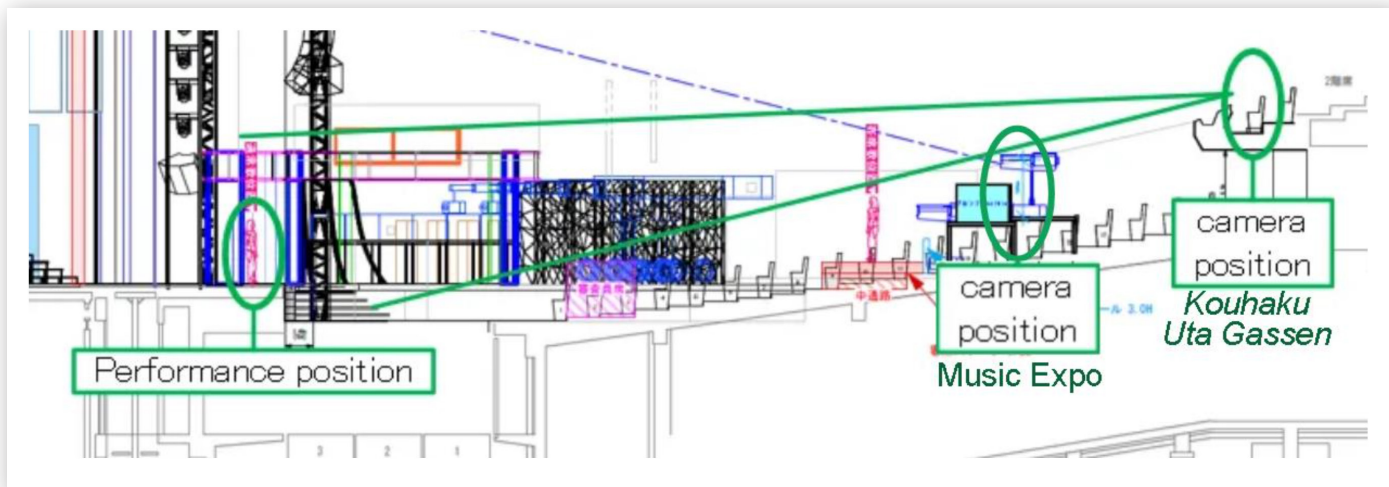


Рис. 2. Камерная позиция для съемки передачи и аналогичное расположение камеры на Music Expo



Рис. 3. Монтажеры KBS и NHK работают с Vvertigo

адаптируется к количеству лиц в каждом кадре и остается надежным даже при обработке плотно заполненных зрителями сцен.

Совместная проверка, проведенная KBS и NHK, демонстрирует, что кадрирование с помощью AI может перевести 8K-съемку из экспериментального формата на уровень практического инструмента для крупных многоракурсных трансляций. Съемка материала одной камерой 8K, из которого на стадии обработки извлекаются «зрительские» ракурсы исполнителя, избавляет от необходимости применения нескольких синхронизированных камер, позволяет убрать загромождающие студию кресла и производить широкий спектр ракурсов, пригодных для нелинейных сервисов. Эти мгновенно доступные ракурсы соответствуют привычкам потребления более молодой аудитории и в то же время позволяют вернуться к основному вещанию, позиционируя 8K и как творческий, и как эксплуатационный катализатор».

в кадре по окончании каждой части передачи. На рис. 2 показана камерная позиция для съемки передачи и аналогичное расположение камеры на Music Expo. Рис. 3 демонстрирует работу монтажеров KBS и NHK с Vvertigo в процессе съемки Kouhaku Uta Gassen».

Исследование показывает, что опирающаяся на данные AI и стандартная съемка в 8K позволяют уложиться в расписание прайм-тайм с параллельным сокращением численности выездных съемочных групп и высвобождением зрительских мест в студии.

«Персонализированный просмотр стал важнейшим способом потребления аудиторией музыкального контента, а «зрительские» кадры с исполнителем часто сильнее привлекают аудиторию, чем официальное видео. Традиционное производство «зрительского» видео требует установки специальных съемочных систем и трудоемкого ручного монтажа. Vvertigo позволяет решить эту проблему за счет применения искусственного интеллекта для генерирования нескольких отдельных «зрительских» ракурсов на базе одного исходного 8K-видео путем автоматизированного кадрирования. В Vvertigo применяется алгоритм кластеризации на основе графов, который

Алексис Аллеманн (Alexis Allemann) из Европейского вещательного союза (EBU, Switzerland) с Себастьяном Нуаром (Sebastien Noir) и Андреем Попеску-Белисом (Andrei Popescu-Belis) – профессором Высшей школы инженерии и менеджмента Кантона Во (HEIG-VD) были признаны авторами «Лучшего технического доклада на IBC 2025» за их презентацию «EBU Neo – развитый многоязыковой чат-бот для надежного поиска новостей».

«Neo был разработан, – отмечают авторы, – в тесном сотрудничестве с членами EBU, особенно со Шведским радио, в ответ на растущую тревогу, связанную с качеством, предвзятостью и прозрачностью коммерческих новостных помощников на базе AI. Наш коллектив уделял внимание оптимизации поисково-расширенного подхода к генерации данных – RAG (Retrieval-Augmented Generation), который улучшает большие языковые модели за счет обучения с применением надежных проверенных источников. Результатом стал многоязыковой чат-бот, опирающийся на обширную базу материалов, ежедневно пополняемую примерно 3 тыс. новых сюжетов. Ожидается, что этот ресурс позволит профессиональным журналистам и общественности получать

информацию о текущих событиях с большей уверенностью в редакторских стандартах и с повышенной прозрачностью.

На основе первых отзывов были запланированы некоторые улучшения в сотрудничестве со Шведским радио, призванные улучшить новостной и общественный аспекты работы Neo. Сюда входит упрощение пользовательского интерфейса при сохранении прозрачности, чтобы пользователи могли свободно продолжать общение, а редакторы имели возможность модифицировать или автоматически генерировать вопросы, относящиеся к злободневным темам».

На рис. 4. Приведен снимок экрана мобильного приложения Swedish Radio с чат-ботом Neo.

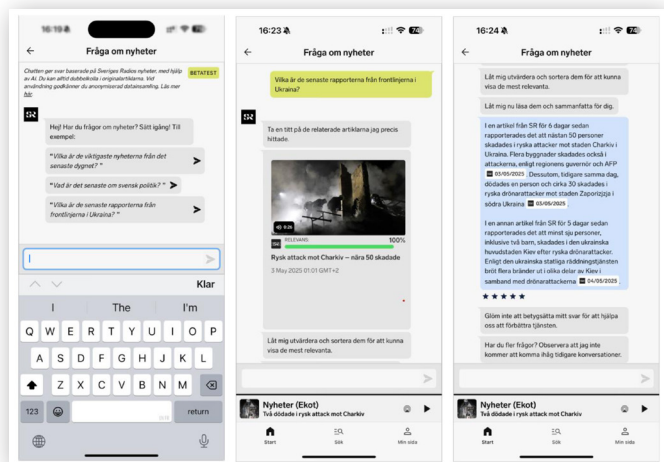


Рис. 4. Снимок экрана мобильного приложения Swedish Radio с чат-ботом Neo.

«Чтобы снять ограничения, выявленные в процессе проверки, – продолжают создатели Neo, – мы изучим гибридные методы поиска, сочетающие семантический и по ключевым словам, чтобы повысить точность отбора материала, в частности, для коротких и очень специфических запросов. Одновременно будет рассмотрено применение модуля более детальной сортировки материала с многокритериальной системой оценки, чтобы лучше отражать релевантность найденного материала. Продолжающееся совершенствование описания также направлено на повышение эффективности работы с неоднозначными или пограничными случаями, чтобы снизить ненужные отказы и повысить точность ответа, а также сделать взаимодействие пользователя с ботом комфортнее.

Еще одно важное наблюдение, сделанное в ходе оценки, заключается в том, что использование только фактических конкретных вопросов не в полной мере отражает действительность и удовлетворяет пользователей. На практике чат-бот используется в основном для срочных новостных запросов, где важен временной контекст. Фактическая оценка тяготеет к акценту на тривиальных вопросах с одним ответом, что полезно для точности сравнительного анализа, но не отражает истинного потенциала конвейера генерации данных на базе расширенно-дополненного поиска. Ведь одно из ключевых достоинств Neo заключается в его стабильности при синтезировании дополнительной информации на основе разных материалов, чтобы дать расширенный обзор и более

подробные ответы. Предстоящие тесты будут направлены на улучшенное отражение этого за счет добавления открытых и привязанных ко времени запросов, а также путем сбора отзывов напрямую от пользователей.

Возможности Neo будут расширены, он получит функции индексирования и поиска по нетекстовым медиаданным, таким как подкасты, что расширит ассортимент доступного контента и сделает работу бота более эффективной для пользователей».

Обработка контента

«Автоматическая нарезка видео в SWR – поиск новостных клипов» (Automatic video cutting at SWR - findings for the 'news clips') – такой практический пример был объектом рассмотрения в докладе В. Ламма и его коллег из Südwestrundfunk (Германия): «Появление AI создало новые сложности для SWR, поставив правовые, технические и относящиеся к контенту вопросы. Чтобы ответить на них, коллектив специалистов решил инициировать тестовый этап для оценки и проверки технологий на базе AI. Внимание уделялось анализу зрелости SWR применительно к использованию AI, потенциальной ценности AI и его важности для перспективных процессов и методов работы. Со всей SWR было собрано более 120 примеров применения. Затем они были оценены в соответствии с предварительно заданными критериями, включая потенциальное воздействие на SWR, вероятность успеха, возможности, риски и устойчивость развития. С ноября 2023 года до конца 2024 года были обработаны 13 из 18 примеров. В их состав была включена и автоматическая нарезка новостных клипов.

Пример применения фокусируется на создании иллюстрированных файлами формата NiF отчетов, в основе каждого из которых лежат текст и исходные видеоматериалы – собственные или полученные от агентства. Файлы NiF (NetImmersive Format) содержат геометрические 3D-объекты, текстуры, анимацию, данные о взаимодействии объектов и другую информацию, применяемую в компьютерных играх. Создание NiF – это сложная и длительная процедура, так что мы хотели выяснить, можно ли применять AI как функционального помощника для автоматизации процессов производства NiF».

Рис. 5 иллюстрирует процесс фильтрации некачественных сегментов в соответствии со стандартами качества SWR.

«Тесты, проведенные совместно с Television.AI с использованием примера NiF, – говорится в докладе, – позволили SWR получить ценные данные о применимости AI для автоматизации рабочих процессов работы с видео. С монтажной точки зрения результаты тестов пока не очень впечатляющие. В отношении качества работы были выявлены недостатки, в частности, касающиеся выбора изображений, последовательностей и склеек.

Тесты применительно к монтажу показали, что создание NiF-материалов с помощью AI заметно проигрывает по времени созданию такого же материала исключительно вручную. Используя собственные устройства, видеомонтажеры, работающие локально или в онлайн-режиме, создавали NiF-материал в среднем вдвое быстрее.

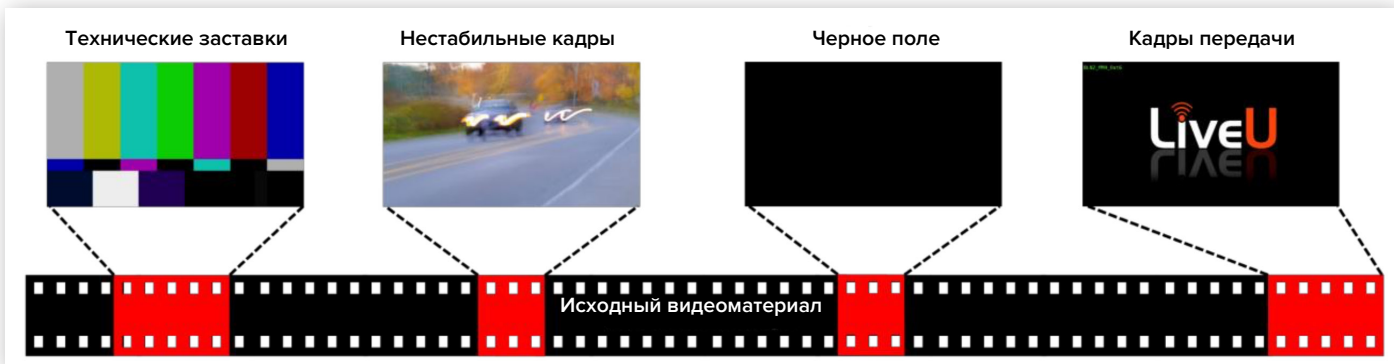


Рис. 5. Процесс отсева сегментов, не соответствующих стандартам качества SWR

В завершение нужно сказать, что испытание привело к нескольким ключевым выводам относительно использования AI в сфере видеопроизводства. Автоматическая NiF-нарезка показала свой существенный потенциал, так что SWR продолжит работать в этом направлении».

Контроль качества

Дж. Реннис и его коллеги из германского Фраунгоферовского института цифровых медиатехнологий (Fraunhofer Institute for Digital Media Technology) обратились к теме безэталонного измерения разборчивости диалога в процессе создания контента, его обработки, анализа и прямых трансляций (Reference-free measurement of dialog intelligibility in (post)-production, content analyses and live broadcast applications): «Последние достижения, опирающиеся на машинное обучение, привели к созданию нового поколения моделей разборчивости речи, которые снимают многие ограничения, присущие традиционным подходам без необходимости тренировок по специальным сценариям прослушивания. Эти модели способны в принципе предсказать пагубное воздействие факторов, нарушающих разборчивость речи, таких как бормотание, нечеткое или невнятное произношение, реверберация, искусственная модификация голоса или дефекты обработки.

Примером автоматического распознавания речи служит LEAP – «Прогнозирование усилий при прослушивании на основе акустических параметров» (Listening

Effort Prediction from Acoustic Parameters). Главная идея состоит в том, что машина становится все более неуверенной в том, какая фонема имеет место в тот или иной момент времени по мере роста искажений. Это выражается в возрастающем размывании вероятностного распределения, которое может быть количественно оценено и наложено на психоакустические шкалы, например, для разборчивости речи или усилий при восприятии, с использованием соответствующих функций отображения. Система автоматического распознавания речи (рис. 6) создает транскрипцию на основе входного аудиосигнала. Зная правильную транскрипцию, можно определить разборчивость речи».

Р.Г. Олдфилд и В.Дж. Ведгвуд (Salsa Sound, UK) с коллегами из Университета Солфорда обсуждали похожую тему в своем докладе «Автоматический контроль качества вещательного аудио» (Automatic quality control of broadcast audio): «Приглушенная, тихая или неразборчивая речь служит основным источником жалоб в процессе многих трансляций. Авторы настоящей работы приняли участие в проекте по созданию набора процессов и алгоритмов, которые позволят проводить комплексную проверку качества восприятия для аудиопотоков в режиме реального времени. Машинное обучение используется для быстрой диагностики и информирования об атрибутах восприятия аудио, таких как шум дождя и ветра или сбоя в потоках, которые иначе было бы сложно выявить. Также

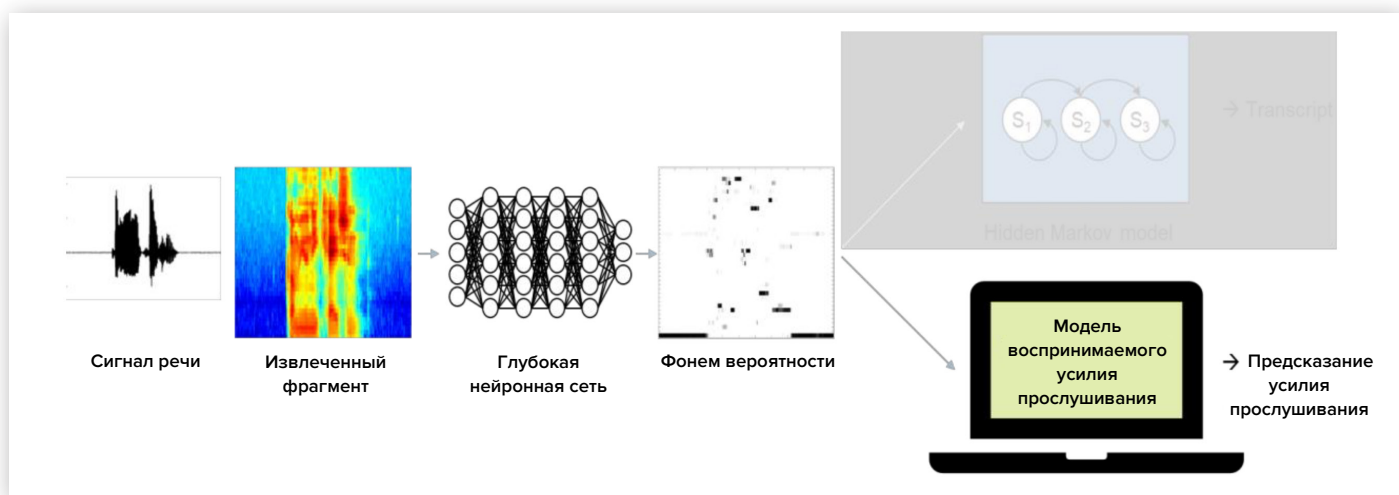


Рис. 6 Типовая структура системы автоматического распознавания речи, которая создает транскрипцию на основе входного аудиосигнала



Рис. 7. Двухэтапный рабочий процесс неинвазивного отслеживания разборчивости речи

обсуждалась разработка речевого процессора, способного идентифицировать язык и ключевые слова для выявления ненормативной лексики и определения темы.

На рис. 7. Показан предложенный авторами двухэтапный рабочий процесс неинвазивного отслеживания разборчивости речи.

Передача контента

Бернардо Азивидо (Bernardo Azevedo) и его коллеги из Globo (Бразилия) обобщили опыт внедрения телевидения следующего поколения в своей стране, сделав доклад «Революция TV 3.0 в Бразилии: маяк для будущего вещания» (The TV 3.0 revolution in Brazil: A lighthouse for the future of broadcasting):

«Бесплатное эфирное телевидение играет в Бразилии ключевую роль, поскольку охватывает более 100 млн зрителей ежедневно и является одним из основных средств потребления аудиовизуального контента, превосходя совокупный рынок платного линейного ТВ и сервисов стриминга на ТВ-приставки. Появление DTV+ знаменует революционное изменение вещательной технологии, обеспечивая улучшение качества для аудитории, повышение эффективности использования частотного ресурса и инфраструктуры доставки контента. Новая технология также объединяет эфирное вещание с широкополосной OTT-доставкой контента.

Бразильский стандарт DTV+ (ранее называвшийся TV 3.0), коммерческое внедрение которого намечено на 2026 год, является инновационным, поскольку в нем сочетаются такие функции физического уровня ATSC 3.0, как «множество входов и множество выходов», передача с двойной поляризацией, идентификация передачи и мультиплексирование с многоуровневым разделением. Успешное выполнение проекта «Маяк» (Lighthouse Project) в апреле 2025 года ознаменовало важную веху во внедрении DTV+ в Бразилии. Продемонстрировав стабильную передачу, эффективный прием и расширенные сервисные возможности в условиях реальной эксплуатации, инициатива доказала техническую состоятельность новой системы. Она также обеспечила важную информацию о наилучших эксплуатационных методах, интеграции инфраструктуры и совместном с предыдущими системами использовании частотного спектра. Как результат, проект заложили надежный фундамент для предстоящего коммерческого внедрения сервисов DTV+ в масштабах всей страны».

Испытания ТВ-вещания по технологии 5G, проводившиеся в крупных городах Италии, нашли отражение в докладе Ф.М. Пандольфи (F.M. Pandolfi) и его коллег из Radiotelevisione Italiana: «С конца 2024 года Rai в сотрудничестве с оператором итальянской вещательной сети Rai Way выдавала в эфир первый крупномасштабный тестовый вещательный 5G-сервис, чтобы проверить качество приема живого контента зрителями в крупнейших городах Италии. Вещание, изначально рассчитанное на Рим и Турин, затем было расширено на Милан, Неаполь и Палермо, охватив потенциально 16 млн жителей (около 27% населения Италии).

В настоящем докладе обобщаются итоги полевых испытаний 5G-вещания в Риме и Турине, проводится сравнение результатов с эмуляцией и предлагается новаторский подход, изучаемый в рамках проекта ESA 5G-EMERGE, направленного на подключение вещательных передатчиков 5G через спутник.

Интенсивные полевые испытания, проведенные Rai в связи с запуском тестовой услуги вещания 5G в крупных городах Италии, позволили оценить достижимую производительность в зависимости от различных условий и возможностей приемного устройства потребителя. Предварительные результаты, несмотря на то что они далеки от финальных коммерческих, уже положительны и показывают, что оптимальная эффективность может быть при минимальной доработке. На момент написания работы, ряд других вещателей и телекоммуникационных компаний уже экспериментировали с различными вариантами этой технологии. Сделан экономический анализ, разработаны планы внедрения. Есть надежда, что результаты этих испытательных кампаний и соображения относительно перспектив внедрения, включая оптимизацию конструкции устройств, послужат основой для запуска первых коммерческих вещательных 5G-сервисов уже в ближайшие годы».

Долгосрочный прогноз

Если попытаться заглянуть в далекое будущее, то нынешний темп совершенствования компьютеров дает основания полагать, что за машинным обучением и искусственным интеллектом последуют потенциально бессмертные сознательные машины, имеющие доступ ко всем знаниям, которые накопило человечество. Пять лет? Десять лет? Кто знает? Мы живем в интересные времена! ▶