Инновации, представленные на IBC 2026

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.

Проходящая ежегодно в середине сентября в выставочном центре RAI Amsterdam, Международная вещательная конвенция (IBC) – это старейшая в Европе выставка оборудования, программного обеспечения и услуг для медиаиндустрии.

В этом году посещаемость выставки немного снизилась, а число экспонентов свидетельствует о заметной ротации участников. Тем не менее мероприятие в целом сохранило свою направленность и привлекательность. В приводимом ниже обзоре представлены избранные доклады технической конференции IBC, посвященные вопросам, которые сегодня находятся в центре внимания вещательных компаний. Не рассчитывайте на легкое чтение: технологии развиваются стремительно.

Унифицированное управление вещанием

Георг Фюрст (Georg Fürst) из Австрийской телерадиокомпании ORF рассказал о преимуществах унифицированного управления вещательными процессами в докладе «От разрозненных систем записи и воспроизведения к единому решению для аппаратной управления эфиром в ORF» (From heterogeneous record and playout systems to a unified control room solution at ORF): *«В современных вещательных комплексах эффективные и адаптивные системы управления аппаратными играют ключевую роль. Традиционные архитектуры, объединяющие множество различных систем, усложняют эксплуатацию и повышают требования к обслуживанию. Одновременно сокращение штата требует решений, позволяющих одному оператору эффективно контролировать несколько процессов воспроизведения и записи одновременно.*

SFERAVIDEO

Проектирование, поставка и инсталляция программно-аппаратных комплексов оборудования для обработки, хранения и кодирования медиаданных для студий производства и пост-производства цифрового кино, систем онлайнового и «холодного» хранения медиаданных на жёстких дисках с возможностью реставрации и восстановления контента.

В ORF внедрили новое решение, которое объединяет управление файлами, запись и воспроизведение материалов для подачи в эфир в рамках одного приложения. Его задача – упростить взаимодействие с системой и повысить эффективность рабочих процессов.

Выбранный подход предусматривает модернизацию и консолидацию инфраструктуры видеосерверов, а также переход к единому типу устройств. Это упрощает техническое обслуживание, снижает потребность в поддержке и способствует развитию компетенций, применимых в разных подразделениях. В результате новая архитектура становится стандартом технической инфраструктуры ORF, повышая надежность и масштабируемость телевизионного производства. Она создает условия для устойчивого выпуска высококачественного контента и повышает организационную устойчивость в условиях меняющихся требований.

Успешное внедрение решения и интерес со стороны других вещательных компаний свидетельствуют о его потенциале для более широкого применения. В перспективе решение может стать образцом для построения аппаратных управления эфиром нового поколения.

С самого начала команда проекта руководствовалась принципом минимизации количества элементов управления, отображаемых во время прямого эфира, при сохранении доступа ко всем необходимым функциям. Поэтому на раннем этапе было решено предоставить расширенные настройки через сворачиваемые вкладки и контекстные меню, вызываемые правой кнопкой мыши. Благодаря этому во время производства оператор видит только основные элементы управления, а на этапах подготовки и обработки материала может получить доступ ко всему набору функций в рамках одного приложения.

Модуль Player характеризуется удобным интерфейсом с вкладками Player Window, Playlist (Cliplist), Recorder и Marker, обеспечивающими комплексное управление воспроизведением видео и аудио. К основным операциям относятся запуск воспроизведения, управление списком клипов, связанным с каналом, запись, синхронное воспроизведение и настройка клипов, в том числе создание субклипов.

Окно Channel содержит множество функций, которые оператор может вызывать по мере необходимости. Это позволяет отображать во время производства только основные элементы управления,

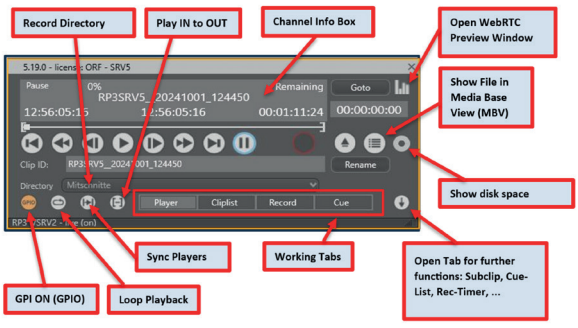


Рис. 1. Элементы окна Channel Window

сохраняя при этом полный доступ ко всем возможностям системы.

WebRTC Player обеспечивает предварительный просмотр загруженного клипа и отображает выходной сигнал канала с сервера. Среди других функций – открытие представления Media Base View с каталогом файлов для доступа к ним и их организации, а также отображение оставшегося свободного места на сервере.

Новая архитектура задает ориентир для технической инфраструктуры аппаратных ORF, повышая надежность и масштабируемость повседневного телевизионного производства».

AoIP-аудиовещание

Т. Ониси и М. Окано (T. Onishi, M. Okano, Японская вещательная корпорация NHK) рассмотрели тему «Возможности и стратегии внедрения облачных аудиоконсолей на базе AoIP» (Feasibility and deployment strategies for cloud-based AoIP audio consoles).

В докладе говорится: *«Спрос на гибкие и ресурсоэффективные решения для телевизионного и радиовещания продолжает расти, особенно там, где технические ресурсы ограничены. IP-технологии и распределенные производственные процессы широко обсуждаются как альтернатива традиционным архитектурам, сосредоточенным в пределах одной инфраструктуры. Облачные модели производства также рассматриваются как потенциальное решение для отдельных рабочих процессов. Однако перенос профессиональных вещательных аудиосистем в облачную среду сопряжен с серьезными техническими трудностями.*

Публичные облачные платформы обладают особенностями, которые осложняют прямое применение традиционных архитектур AoIP. Маршруты сетевого трафика динамически управляются для балансировки нагрузки и обеспечения отказоустойчивости, поэтому их нельзя считать статичными или привязанными к определенной топологии. Вариации задержки пакетов зависят не только от номинальной пропускной способности, но и от конкуренции за ресурсы между несколькими пользователями. Кроме того, исходная маршрутизация IP multicast, на которую опираются профессиональные AoIP-системы, изначально не поддерживается в публичных облачных сетях. Для ее реализации обычно требуются дополнительные сервисные компоненты или изменения архитектуры, внедрение которых, как правило, обходится дорого.

Для синхронизации времени между географически разнесенными площадками можно рассмотреть два основных подхода. Первый заключается в установке на каждой площадке собственного генератора PTP Grandmaster (GM) и использовании независимых доменов PTP. Либо в расширении единого домена PTP путем передачи сигналов PTP с одной площадки на другую с помощью единого ведущего генератора GM.

Второй подход может быть эффективным в выделенных сетевых средах, однако передача PTP через сети с нестабильными характеристиками джиттера, например сети 5G или публичные сети без гарантированного качества обслуживания, остается сложной

задачей. Отсутствие поддержки multicast в публичных облачных средах дополнительно ограничивает практическую применимость этого подхода даже при наличии выделенного подключения. Прямое использование стандартов AoIP в таких условиях может приводить к непредсказуемому поведению системы синхронизации, нестабильной работе буферизации и потенциальной потере фазовой согласованности.

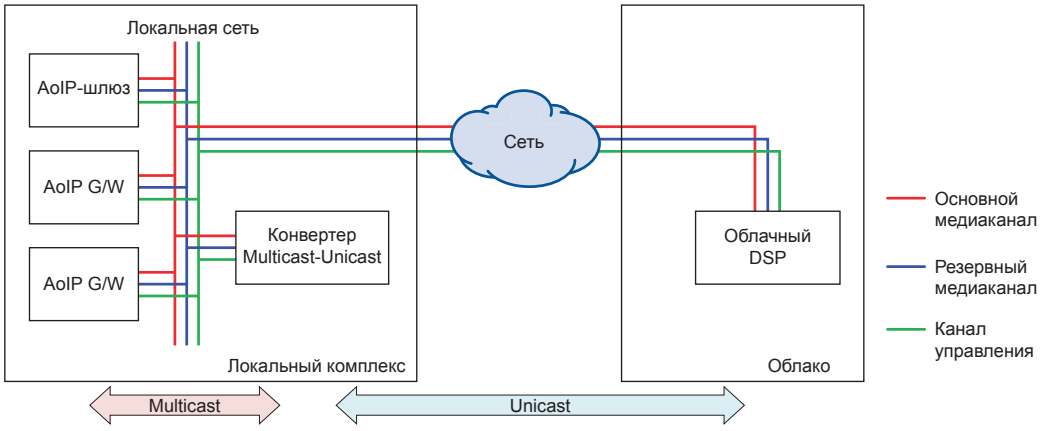


Рис. 2. Архитектура облачной AoIP-системы с раздельными доменами транспорта и синхронизации

Вещательные системы PTP ориентируются на международное атомное время (TAI), тогда как системное время в публичных облачных средах основано на всемирном координированном времени (UTC). Поскольку облачные платформы опираются на аппаратные часы, привязанные к UTC, а не к PTP, ведомом от TAI, для компенсации временного дрейфа между доменами необходимо напрямую управлять буферизацией. При этом для операций микширования, не привязанных к внешнему тактовому сигналу, выравнивание с точностью до отдельного аудиосэмпла не требуется».

На рис. 2 представлена общая архитектура облачной AoIP-системы, в которой транспортный и временной домены разделены. Для проверки практической реализуемости концепции была создана экспериментальная система, объединяющая передачу AoIP на основе unicast и гибридную синхронизацию времени. В ходе испытаний использовался один несжатый 32-канальный аудиопоток PCM, сформированный локальным AoIP-оборудованием ввода-вывода. Внутри локальной сети звук передавался с использованием multicast-транспорта на базе AES67/SMPTE ST 2110-30. На границе облачной среды поток преобразовывался в unicast и направлялся на аудиокон-

соль, размещенную в облаке на виртуализированном сервере. Внутренняя задержка обработки оцененной аудиоконсоли составляла 128 сэмплов.

«Результаты экспериментов показали, что наиболее эффективный способ внедрения облачных AoIP-систем заключается в использовании их для постепенного расширения функциональности существующей вещательной звуковой инфраструктуры, а не для ее полной замены», – делают вывод авторы доклада.

Экономически эффективное виртуальное производство

Х. Реймонд-Хейлинг (H. Raymond-Hayling) и его коллеги из исследовательского подразделения BBC представили Osmia VP – недорогую мобильную систему виртуального производства, предназначенную для создания визуальных эффектов непосредственно в кадре (in-camera VFX):

«Наша цель – снизить порог входа в виртуальное производство, которое требует специализированных аппаратно-программных средств и поэтому обходится слишком дорого для проектов со средним и низким бюджетом. При этом такие проекты составляют значительную часть из 20 тыс. часов контента, ежегодно производимого BBC в рамках ее общественной миссии.

Мы демонстрируем систему и ее возможности интеграции захвата сцены с отслеживанием камеры в режиме реального времени и проекцией – все на базе стандартного коммерческого оборудования. Система разработана с учетом потребностей производственных команд BBC и позволяет разделить процессы съемки локации и записи актерской игры. Благодаря этому реалистичные виртуальные фоны, созданные с помощью технологии Gaussian Splatting, можно объединять с видеоматериалами, снятыми в студии.

Техническая архитектура системы построена на основе решений с открытым исходным кодом. В Osmia VP (рис. 3) используются Gaussian Splatting для создания виртуальной среды, модифицированный подход SLAM (на базе AsLAM – асинхронной локализации и картографирования (Asynchronous Location and Mapping)) с печатными маркерами для отслеживания положения камеры в режиме реального времени, а также модифицированное ядро визуализации Gaussian Splatting для проекции с корректной перспективой.

Такой рабочий процесс позволяет существенно сократить расходы и уменьшить воздействие на окружающую среду как при съемках на натуре, так и при телевизионном производстве в целом.

Съемку сцены может выполнить один оператор с обычной камерой или мобильным телефоном. Затем на основе фотографий или видео создается Gaussian Splat – трехмерное представление сцены, сформированное из множества небольших элементов, содержащих информацию о цвете и пространственном расположении. Для обучения сцены используются облачные вычислительные ресурсы и программная реализация Gsplat. В ходе испытаний исследователи применяли графический ускоритель NVIDIA A100: обучение сцены по 400 изображениям занимало около 2 ч.

Точное отслеживание положения камеры необходимо для виртуального производства, поскольку так обеспечивается правильное совмещение виртуального фона с изображением, снимаемым камерой в режиме реального времени. Традиционные системы такого типа чрезвычайно дороги, требуют специально оборудованных помещений и высококлассных систем оптического трекинга. Решение Osmia VP не нуждается в специализированном оборудовании и может быть развернуто по мере необходимости практически в любом студийном пространстве, в том числе в помещениях, изначально не предназначенных для производства видео, например в офисах.

Главное ограничение системы связано с масштабом съемки: использование одного потребительского экрана ограничивает размер кадра и количество исполнителей, которых можно одновременно снимать. Одним из возможных решений без существенного увеличения затрат может стать объединение изображения с нескольких экранов меньшего размера. Для более широкого внедрения технологии также потребуется постоянный обмен опытом, чтобы адаптировать методы производства и планирования к возможностям виртуального производства.

Поскольку этот рабочий процесс предназначен в том числе для ситуаций, когда систему необходимо развернуть в кратчайшие сроки, желательно автоматизировать как можно больше этапов. В частности, полезно было бы разработать инструмент автоматической цветокоррекции виртуального фона с учетом изображения на переднем плане. Условия съемки могут существенно ограничивать возможности управления освещением реальных объектов, тогда как цвет виртуального фона изменить гораздо проще.

Дополнительные движущиеся элементы на заднем плане могли бы повысить реалистичность изображения и расширить творческие возможности. Быстрое создание полноценной четырехмерной сцены на основе Gaussian Splatting может оказаться сложной задачей. Однако естественное движение можно добавить и другим способом –

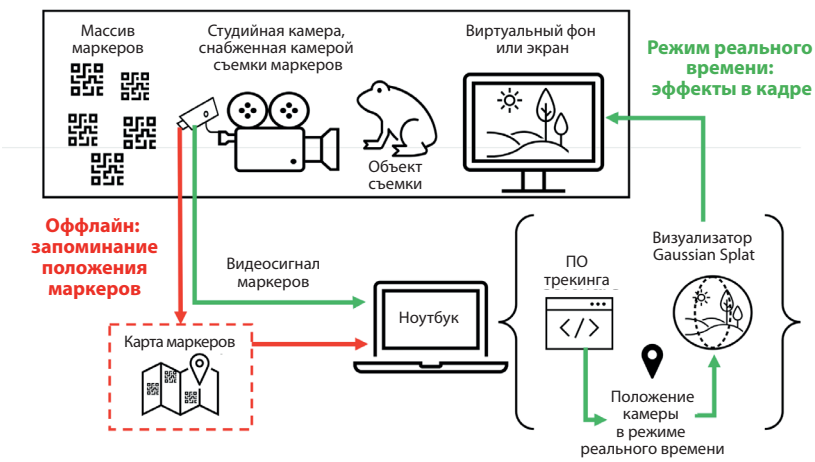


Рис. 3. Структурная схема Osmia VP, демонстрирующая оффлайновый процесс картографирования маркеров и процесс съемки в режиме реального времени с применением виртуального фона

разместив несколько двумерных элементов видео внутри более крупной трехмерной сцены.

Еще одной перспективной возможностью является использование Gaussian Splatting на переднем плане. В зависимости от содержания сцены отдельные объекты, например, листья в джунглях, можно динамически накладывать поверх видеосигнала основной камеры.

Мы считаем, Osmia VP демонстрирует, что для получения высококачественного изображения по технологии виртуального производства необязательно требуются специализированная инфраструктура и дорогостоящее оборудование. Разработанный подход дает возможность снизить как финансовые затраты, так и экологическую нагрузку, связанную с телевизионным производством».

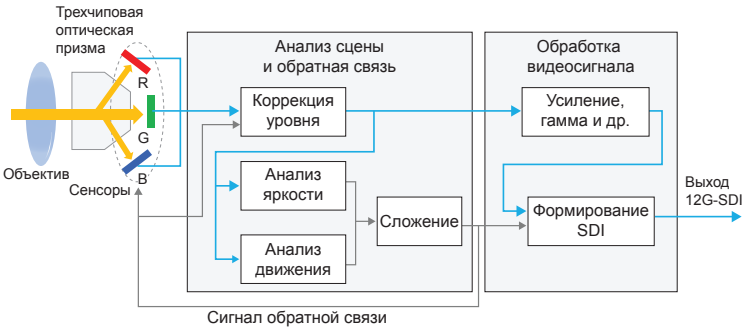


Рис. 4. Блок-схема камерной системы с адаптивными к сцене настройками

Камеры с адаптивными настройками в зависимости от сцены

К. Томиока (K. Tomioka) и его коллеги из японской вещательной корпорации NHK совместно с С. Кавахито (S. Kawahito) из Университета Сидзуоки представили доклад «Камера с адаптацией к сцене: динамическое управление параметрами изображения для отдельных областей съемки» (Scene-adaptive camera: dynamic imaging parameter control for localised shooting areas):

«Одновременное достижение высокого пространственного разрешения, высокой частоты кадров и широкого динамического диапазона по-прежнему остается сложной задачей для систем обработки изображений. Это обусловлено неизбежными компромиссами между скоростью считывания данных, энергопотреблением и ограничениями пропускной способности.

Чтобы преодолеть эти ограничения, мы предлагаем камеру с адаптацией к сцене. Вместо применения единых настроек ко всему кадру в ней оптимизируются параметры съемки для отдельных областей изображения. Такой подход позволяет локально управлять пространственным разрешением, частотой кадров и временем экспозиции с точностью до блоков размером 4×4 пикселя.

Для демонстрации работоспособности концепции мы разработали камерную систему на основе нового CMOS-сенсора разрешением 4K, способного независимо назначать один из четырех режимов съемки каждому блоку. Система также содержит контур обратной связи в режиме реального времени, который анализирует локальную яркость и движение объектов в сцене и синхронно с работой сенсора обновляет параметры изображения».

На рис. 4 показан разработанный CMOS-сенсор. Его эффективная матрица содержит 3904 пикселя по горизонтали и 2224 по вертикали, что соответствует приблизительно 8,7 Мпк. Матрица разделена на блоки размером 4×4 пикселя – 976 блоков по горизонтали и 556 по вертикали.

«В зависимости от характеристик снимаемого объекта каждому блоку можно независимо назначить один из четырех режимов: Normal, Bright, Low-Light и Fast. Разработанный CMOS-сенсор разрешением 8,7 Мпк обеспечивает исключительно точное управление на уровне блоков 4×4 пикселя благодаря новой архитектуре, в которой схемы проходных транзисторов интегрированы непосредственно в схемотехнику столбцов матрицы.

Эта архитектура позволяет точно адаптировать параметры съемки к контурам объектов и сложному распределению яркости, независимо управляя пространственным разрешением, частотой кадров и временем экспозиции в пределах одного кадра.

Сенсор был интегрирован с трехматричной оптической системой RGB и системой обратной связи, работающей в режиме реального времени. В результате удалось создать практическую камерную платформу,

которая одновременно обеспечивает широкий динамический диапазон и частоту съемки 240 кадр/с, сокращая смаз движущихся объектов и сохраняя качество видео 4K.

Перенос базовых технологий в полноценную систему 4K, пригодную для вещательного производства, стал важным этапом данного исследования.

Полученные результаты подтверждают эффективность локального распределения ресурсов обработки изображения в зависимости от характеристик сцены. Предлагаемая технология создает основу для систем визуализации следующего поколения, способных максимально повышать качество изображения при ограничении требований к скорости передачи данных. Среди потенциальных областей применения – иммерсивные медиа.

Дальнейшие исследования будут направлены на совершенствование алгоритмов анализа сцены. Кроме того, благодаря уникальным возможностям предлагаемого подхода разработчики намерены внедрять эту технологию в самых разных областях – от вещания до более широких систем общественной инфраструктуры.

Эксперименты показали, что предлагаемая камера существенно повышает качество изображения: в пределах одного кадра она одновременно подавляет пересветы и недоэкспонированные участки, а также уменьшает смаз при движении».

Рис. 5. Модуль AI-обработки, потребляющий и производящий синхронизированные медиа через MXL-интерфейсы, разделяя при этом предобработку, инференс, визуализацию и контроль

Конвергенция IP-систем и искусственного интеллекта

Гийом Артюис (Guillaume Arthuis) и его коллеги из французской компании BBright совместно с С. Мубайедом (S. Moubayed) и его коллегами из французской Hexalia представили доклад о применении MXL (Media eXchange Layer) в качестве уровня взаимодействия между системами искусственного интеллекта и вещательной инфраструктурой:

«В данной работе представлен архитектурный подход на основе Media eXchange Layer – открытого стандарта, продвигаемого Европейским вещательным союзом (EBU). Он устанавливает четкую границу взаимодействия между вещательной инфраструктурой и системами обработки данных с помощью искусственного интеллекта. Основная гипотеза заключается в том, что проблемы интеграции AI эффективнее решать посредством стандартизированного абстрактного уровня, который отделяет транспортировку медиаданных от их интеллектуальной обработки».

На рис. 5 показана структура модуля AI-обработки, предназначенного для получения и передачи синхронизированных медиаданных через интерфейсы MXL. При этом предварительная обработка, инференс, визуализация результатов и контроль выполняются отдельными компонентами.

«Такое разделение функций позволяет AI-ядрам работать независимо от базовой инфраструктуры, сокращая объем интеграционных работ, необходимых для конкретных рабочих процессов. Модуль AI не ограничивается инференсом: он обеспечивает четыре функциональных этапа – предварительную обработку, инференс, интерпретацию результатов и визуализацию либо формирование метаданных.

На этапе предварительной обработки вещательный сигнал преобразуется в формат, необходимый модели. Инференс формирует результаты работы модели, обычно в виде метаданных обнаружения объектов, включая зоны, метки и показатели уверенности. Затем на этапе интерпретации результатов выходные данные модели преобразуются в действия, предусмотренные рабочим процессом. На заключительном этапе эти действия применяются к выходному медиапотоку либо передаются другим системам в виде метаданных.

Это разграничение принципиально важно, поскольку многие AI-модели выдают информацию, которую нельзя напрямую использовать в вещательной цепочке.

Первые результаты внедрения свидетельствуют о снижении сложности интеграции, повышении масштабируемости и улучшении совместимости по сравнению с традиционными схемами точечного подключения AI-систем.

Главным препятствием для внедрения искусственного интеллекта в вещательную отрасль является отсутствие не столько подходящих алгоритмов, сколько стабильной границы между инфраструктурой передачи медиаданных в режиме реального времени и AI-системами обработки.

Вещательные процессы требуют детерминированной синхронизации, целостности сигнала и непрерывности работы. Системы AI, напротив, нуждаются в гибких средах выполнения, возможности обновления моделей и ускоренных вычислениях. Их непосредственная интеграция приводит к созданию индивидуальных решений, которые сложно поддерживать и которые могут оказаться ненадежными.

В работе предлагается архитектура на основе MXL, где Media eXchange Layer выполняет роль такого связующего уровня. Вещательная инфраструктура отвечает за обработку медиапотоков по SMPTE ST 2110, синхронизацию и непрерывность вещания, тогда как модули AI работают как сервисы в контейнерах, взаимодействуя с системой через четко определенные плоскости данных и управления».

Авторы приходят к выводу, что промышленное внедрение искусственного интеллекта в вещании зависит в

большей степени от стандартизации архитектуры, чем от дальнейших алгоритмических инноваций. Создавая стабильную границу между системами AI и вещательной инфраструктурой, MXL позволяет объединить надежность вещательного уровня с быстрым развитием технологий искусственного интеллекта.

Виртуальное производство на базе светодиодных экранов

К. Боровски (C. Borowski) из немецкой общественной телерадиокомпании Südwestrundfunk (SWR) рассказал о преимуществах подхода с чередованием кадров при частоте 100 Гц для переключения между камерами в прямом эфире при использовании светодиодных экранов в виртуальном производстве.

SWR – общественный вещатель юго-западной Германии – готовится к обновлению студийной инфраструктуры в Баден-Бадене. В рамках подготовки компания осуществила экспериментальный проект по виртуальному производству, чтобы проверить в реальных производственных условиях, а не в закрытой лаборатории, насколько технология на базе LED-экранов подходит для форматов, которые планируется выпускать в будущих студиях.

«Проверка концепции проходила в течение трех месяцев – с сентября по ноябрь 2025 года. Большая часть технического оборудования была арендована на этот период.

У SWR есть основные площадки в Баден-Бадене, Штутгарте и Майнце, каждая со своим профилем производства. В Штутгарте и Майнце выпускают большую часть новостных программ и студийного контента с высокой частотой обновления. Основные студии ежедневно обслуживают несколько программ, а смена декораций занимает около 15 мин. Баден-Баден специализируется на развлекательных, культурных и информационно-документальных программах. Срок эксплуатации студийной инфраструктуры этой площадки приближается к концу, поэтому ожидается ее масштабное обновление.

В рамках экспериментального проекта рассматривались три ключевых вопроса. Во-первых, соответствие производственным требованиям. Можно ли использовать виртуальное производство на базе LED-экранов в условиях прямого эфира с несколькими камерами, учитывая, что опубликованный опыт применения этой технологии в основном связан с кинопроизводством, где используются одна камера и не ведется прямой эфир?

Во-вторых, выбор площадки. На какой площадке SWR инвестиции в эту технологию будут наиболее оправданными с учетом различий в загрузке студий?

И в-третьих, соответствие формату. В каких случаях виртуальное производство действительно необходимо для конкретного формата, а когда оно служит лишь визуальным усовершенствованием?

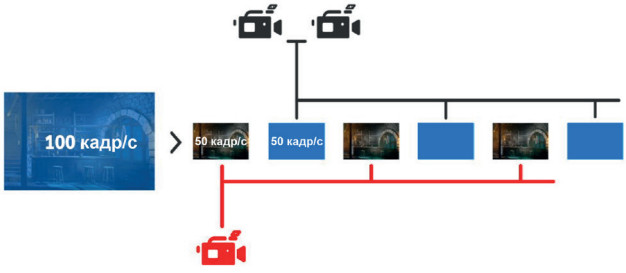
Основным испытанием стал восьмичасовой прямой эфир развлекательного формата с использованием бумаги и ручки, предназначенный для канала ARD на Twitch. Производство велось с помощью трех отслеживаемых камер и с коммутацией изображения в режиме реального времени. В качестве фона использовалась LED-стена размером 10×4 м с частотой обновления 100 Гц.

Главный технический результат – метод переключения между камерами в прямом эфире на основе чередования кадров (рис. 6). Нечетные кадры на LED-стене содержат перспективно-корректное изображение, рассчитанное для активной камеры, а четные — равномерный фон для рирпроекции. Для каждой камеры предусмотрен отдельный узел рендеринга, который работает непрерывно. Таким образом, при переключении между камерами меняется фаза отображения на LED-стене, а не сам процесс рендеринга. Предварительный просмотр с неактивных камер формируется с помощью рирпроекции и параллельных процессов визуализации в сниженном качестве.

Результаты реализации проекта позволили с высокой степенью уверенности ответить на все три вопроса, которые прежде оставались открытыми.

Соответствие производственным требованиям: виртуальное производство на базе LED-экранов можно использовать в условиях прямого эфира с несколькими камерами, если техническое решение изначально учитывает ограничения многокамерной съемки. Предложенный метод чередования кадров успешно работал на протяжении восьмичасового прямого эфира.

Выбор площадки: основным ограничением является не повседневная эксплуатация, а разработка форматов, для которой требуется непрерывное время в



Подход многокамерной съемки с чередованием кадров

студии с реальным оборудованием для рендеринга и LED-экраном. Для SWR это означает, что важнее определить не столько площадку для проведения съемок, сколько место размещения ресурсов для разработки. Баден-Баден, где преобладают программы более продолжительного формата и планируется обновление студий, остается подходящим кандидатом для совмещения разработки и эксплуатации технологии.

Соответствие формату: здесь были выявлены два организационных фактора, а не технических ограничения. Они актуальны для любой общественной телерадиокомпании, рассматривающей аналогичные инвестиции.

Первый фактор – это необходимость создания специализированной команды для работы с графикой. Эффективность виртуального производства зависит от специалистов, которые создают и поддерживают контент. Экономия на традиционных декорациях становится реальной исключительно при наличии такой команды, причем ее необходимо подключать уже на этапе разработки концепции, а не только непосредственно перед съемками.

А второй фактор заключается в важности постоянного внутреннего взаимодействия. Редакционные команды не могут предлагать идеи для форматов,

основанные на технологии, которую они никогда не видели в действии. Поэтому будущие инвестиционные проекты должны с самого начала содержать план информирования сотрудников и ознакомления редакционных команд с новыми возможностями, а не рассматривать эту работу как второстепенную задачу».

Монтаж с помощью искусственного интеллекта

Доклад «Суверенный монтаж с помощью AI: MCP против межагентной оркестрации при работе с конфиденциальным аудиовизуальным контентом» (Sovereign AI-assisted editing: MCP versus agent-to-agent orchestration for confidential audiovisual content) представили А. Руэксель (A. Rouxel) и П. Фуше (P. Fouché) – оба из Европейского вещательного союза (Швейцария), А. Мессина (A. Messina, итальянская телерадиокомпания Radiotelevisione Italiana) и Р. Саломоне (R. Salomone, SRG SSR – Швейцарская вещательная корпорация):

«Вещательные компании все чаще стремятся внедрять AI в монтаж, сохраняя при этом неопубликованные видео- и аудиоматериалы внутри собственной контролируемой инфраструктуры. В докладе представлена архитектура, ориентированная на метаданные и предусматривающая локальную обработку исходных медиаматериалов. Редакционные AI-агенты получают только метаданные, синхронизированные по времени: расшифровки речи, границы планов, кластеры лиц, описания сцен и акустические события.

Архитектура состоит из трех уровней: определения намерений и планирования, специализированных агентов для редакционных решений и сервисов анализа медиаматериалов на базе Model Context Protocol (MCP). Дополнительно может использоваться локальный шлюз обезличивания, который удаляет или заменяет заданные идентификаторы перед передачей выбранных метаданных внешней модели рассуждения».

Авторы сравнивают два способа координации одних и тех же агентов и сервисов: централизованный рабочий процесс на базе MCP и совместную координацию по модели Agent-to-Agent (A2A). Сравнение носит качественный характер и учитывает управляемость, воспроизводимость, гибкость, возможность аудита, конфиденциальность и эксплуатационные затраты. Предлагается гибридный подход: централизованная оркестрация для фиксированных и повторяемых задач, а взаимодействие A2A — для решений, зависящих от недостаточно конкретного редакционного задания.

«Редакции сталкиваются с противоречием между стремительным внедрением генеративного и мультимодального AI и обязательствами по обеспечению конфиденциальности, управлению данными и соблюдению прав на вещательный контент. В данном исследовании суверенитет означает, что вещательная компания сохраняет контроль над тем, где исходные материалы хранятся, обрабатываются и становятся доступными. Неопубликованные съемки могут содержать информацию под эмбарго, сведения о защищенных источниках, данные, позволяющие идентифицировать людей, а также архивные и спортивные материалы с ограниченными правами использования. Передача таких видео- или аудиоматериалов сторонним AI-сервисам может раскрыть чувствительную редакционную информацию и создать риски, связанные с защитой данных, правами собственности и соблюдением нормативных требований.

Тем не менее, AI может быть полезен в редакционных и архивных процессах. Предыдущие исследования распознавания лиц в телевизионном контенте показали, как синхронизированные по времени метаданные, ориентированные на потребности пользователей, помогают специалистам по работе с медиаматериалами. MCP и A2A представляют собой взаимодополняющие стандарты: первый обеспечивает подключение AI-систем к инструментам, а второй позволяет агентам взаимодействовать друг с другом. В сочетании эти разработки открывают возможности для частично автоматизированных рабочих процессов, в которых контент сегментируется, ранжируется и собирается в монтаж под руководством журналиста, задающего инструкции на высоком уровне».

В докладе предлагается архитектура, в которой исходные аудиовизуальные материалы остаются внутри контролируемой среды вещательной компании, а AI анализирует производные метаданные. Архитектура организована сверху вниз, начиная с пользовательского запроса:

- уровень 1 интерпретирует редакционное задание и устанавливает ограничения;
- уровень 2 содержит специализированных агентов, которые принимают монтажные решения на основе метаданных;
- уровень 3 состоит из локальных сервисов анализа медиаматериалов, формирующих описания исходного контента с привязкой ко времени.

«В дальнейших исследованиях планируется сравнить три варианта развертывания. Полностью локальная модель предполагает, что метаданные и языковые модели находятся внутри защищенного периметра вещательной компании. Гибридная модель предусматривает локальное создание метаданных и передачу внешней языковой модели обезличенного представления данных. Полностью внешняя модель заключается в доступе сторонней мультимодальной модели к исходным материалам. Третий вариант послужит базовым сценарием для оценки компромисса между производительностью и суверенитетом данных при передаче обработки внешним сервисам.

Еще одно направление дальнейшей работы – интеграция с BBC Time-addressable Media Store (TAMS). Это позволит использовать tcp-data для обращения к существующему хранилищу медиаматериалов с временной адресацией, не меняя типизированный слой метаданных».

IBC 2027

Настоящие – живые – выставки продолжают пользоваться популярностью даже в современном мире, где все взаимосвязано через Интернет. Международная вещательная конвенция IBC вновь пройдет в выставочном центре RAI Amsterdam с пятницы, 10 сентября, по понедельник, 13 сентября 2027 года.