

# Звук – новейшие тенденции по итогам CES, NAB и AES 2018

Екатерина Петухова

История звукозаписи насчитывает чуть больше лет, чем история записи движущихся изображений. Но с тех самых пор как звук стал неотъемлемой частью кинематографа, его запись, хранение и воспроизведение из года в год улучшаются и совершенствуются, и так же, как к изображению, к звуку предъявляются все новые требования. Телевизионное производство тоже немало поспособствовало тому, чтобы звук стал равноправной частью контента, а видеозаписи опер и симфонических концертов уже стали особым видом искусства. С недавних пор в моду вошла практика мировых премьер из ведущих театров и с филармонических площадок, которые транслируются в кинозалы по всему миру, так же как и показы на большом экране мировых спортивных событий. Значительный вклад в развитие технологий звука внесли такие компании, как, например, Dolby Labs, которые и сегодня остаются на передовом рубеже звукового «фронта». Однако же в данной статье речь пойдет не о них. В связи с кардинальными изменениями всей медиаиндустрии, о которых журнал уже писал в предыдущих аналитических обзорах аналогичного типа, ниже делается попытка выделить ключевые моменты развития аудиотехнологий, на которые необходимо обратить внимание. Три глобальных форума этого года – CES и NAB 2018 в Лас-Вегасе и 144-я AES в Милане – дают для этого богатый материал.

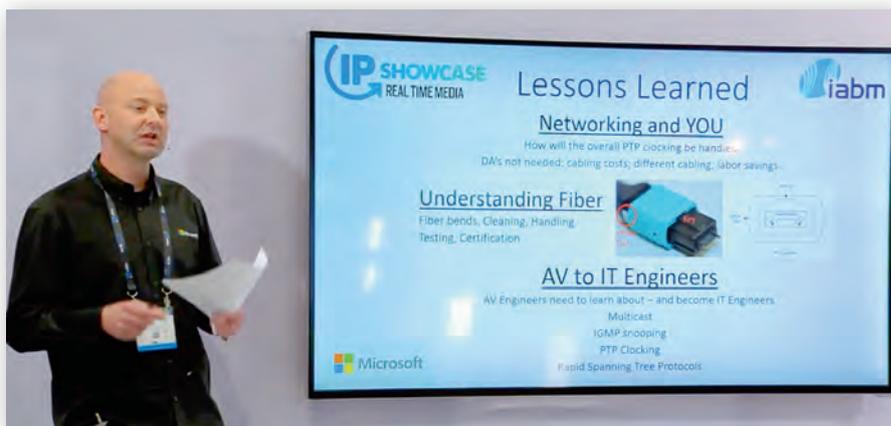
## Аудио по IP

Нет смысла в сотый раз обсуждать в данной статье преимущества перехода на IP-технологии, поскольку они доказаны самой жизнью – множеством уже эксплуатируемых вещательных и производственных комплексов. И даже стандарты IP-аудио уже достаточно подробно обсуждались в ноябрьском номере Mediavision за 2016 г. В дополнение к сказанному там на тему стандартов надо добавить, что AES67 был принят составной частью главного стандарта SMPTE 2110 в сентябре прошлого года, а в этом году получил дальнейшее развитие – в апреле формально принят AES67-2018, в который был добавлен PICS (protocol implementation conformance statement). По этому документу можно проверять совместимость всех устройств в сети, сколь бы большой она ни была, а также параметры самой сети. Аудио профессионального качества – 16 бит/48 кГц или выше – требует высокоскоростной сети, которая обеспечит минимально возможную задержку. Это, как

правило, не является проблемой в студийных комплексах, но когда речь идет о дистанционной трансляции или использовании публичного Интернета, могут возникать проблемы. Тем не менее стандарт AES67 позволяет организовать профессиональную передачу звука в IP-сети независимо от ее размеров и типа. Интересной возможностью IP-аудиомашрутизатора является широковещательный (multicast) режим. Если классический коммутатор просто переключает тот или иной вход на тот или иной выход, широковещательный режим обеспечивает доставку данных сразу на все заданные IP-адреса. И это чрезвычайно важно в мире, где направление взято на TV everywhere. При этом синхронизация аудио и видео осуществляется без проблем с помощью протокола PTP. На выставке NAB 2018 на площадке демонстрации действующих IP-решений теме звука была посвящена отдельная сессия, на которой подробно рассматривались части 30 и 31 стандарта SMPTE ST 2110, имеющие непосредственное отношение к аудио, а также то, как этот стандарт соотносится с AES67. Но кроме теории, на этой сессии прозвучали и вполне практические доклады, например, системного инженера Microsoft Production Studios Джона Болла (John Ball) на тему «Практические уроки: организация перехода к SMPTE ST 2110» (Practical Lessons: Managing Migration Towards SMPTE ST 2110). В своей презентации Джон показал, как трансформировать студию, используя оборудование Cisco, рассказал об уроках этой трансформации – экономии ресурсов и трудозатрат после ее внедрения. Но самое сложное, как он отметил, завершая свое выступление, «... это превратить аудиоинженера в IT-инженера». Вернее, даже не превратить, а научить понимать IT и работать в новой технологической среде. Но у Microsoft получилось, значит, и у нас тоже получится.

## Объемный реалистичный звук и AR/VR

Следующая новая важная тенденция в аудиотехнологиях – это концепция объемного звука. По мере того как изображение получает более широкую цветовую гамму (HDR, WCG) и становится круговым (V360°), звук также должен становиться более глубоким и реалистичным. Впервые парадигма «объемности» была предложена в 2011 году в виде формата Auro-3D для публичных и домашних кинотеатров, а следом, в 2012 году, появилась технология Dolby Atmos. Она получила широкое распространение и в Dolby Cinema, и в сфере домашних кинотеатров. Однако инженерная мысль не стоит на месте, и вот в 2015 году было предложено решение DTS-X, которое не так уж превосходило Atmos по реалистичности звучания, но обладало лучшей «обратной совместимостью» при внедрении. Тем не менее прошло еще два года, прежде чем DTS-X вышел на полноценное коммерческое использование – по данным CES этот формат стандартно поддерживает большинство производителей аудиоборудования. А в январе 2018 года на выставке CES компания DTS анонсировала еще несколько новых звуковых решений, включая DTS:X Premium 1.0, предназначенное для смартфонов и планшетов и поддерживающее до 7.1 каналов, а также звуковой мастер-декодер DTS-HD с четырьмя конфигурируемыми комплектами настроек для контента. DTS Headphone:X 2.0 предназначен для видеоигр, а DTS:X Ultra 1.0 представляет лучший объемный звук для игр, VR и AR. Потому что объемный звук – одно из главных требований контент-производителей VR. Так, поворот головы в очках VR должен вместе со сменой ракурса изображения вызвать изменения вектора звука, и синхронизация этих двух параметров призвана усиливать глубину и реалистичность эффекта присутствия на на-



Джон Болл рассказывает о переходе на инфраструктуру стандарта SMPTE ST 2110

работай с **ПЕРСПЕКТИВОЙ**



реклама



129515, РФ, Москва, ул. Академика  
Королева, д. 13, стр. 1, офис 701



+7 495 290 42 75



INFO@TV-PROSPECT.RU



WWW.TV-PROSPECT.RU

блюдаемом событии. На конгрессе NAB 2018 в секции конференции по вещательному оборудованию и информационным технологиям (Broadcast Engineering and Information Technology Conference) прозвучал совершенно необычный для аудитории инженеров доклад «Мониторинг объемного звука в сериалах, играх и виртуальной реальности» (Immersive Audio Monitoring in Drama, Games and VR). Он был посвящен не столько технологии, сколько оценке физиологии человеческого восприятия звука. Презентацию сделал представитель финской Genelec Томас Лунд (Thomas Lund), а модератором выступил директор и главный звукоинженер NBC Universal Джим Старжински (Jim Starzinsky). Исследователи проводили сравнение восприятия звукового 3D-сопровождения различными испытуемыми в оборудованном колонками помещении и в наушниках для различных видов контента – фильмов, телетрансляций и видеоигр. Субъективные параметры восприятия затем сравнивали с объективными показателями – уровнем, спектральной характеристикой и т.д. То есть был переброшен мост между субъективным и объективным. И аналогичная же – психофизиологическая – научная работа была представлена на 144-й AES, только здесь испанские исследователи Франциско Куадро (Francisco Cuadrado), Изабель Лопес-Кобо (Isabel López-Cobo), Таня Матеос (Tania Mateos) и Беатрис Вальверде (Beatriz Valverde) из университета Севильи сравнивали восприятие и запоминание учебного контента со звуковым сопровождением стерео или 3D. И результат получился достоверно убедительный – 3D рулит! А это означает, что работы в сфере объемного звука будут продолжаться очень интенсивно. В том числе и для мобильных устройств. Кроме упомянутых выше технологий DTS, на арене объемного звука для мобильных устройств мощно выступила шведская компания Dirac (dirac.com). Первая версия их решения Dirac VR – звуковой 3D-платформы – вышла в 2017 году, а в нынешнем появилась уже улучшенная вторая версия, в которой интенсивность и направление (вектор) звука привязываются к положению головы, на которую надеты очки виртуальной реальности, благодаря чему создается полное ощущение присутствия в естественной звуковой среде. Еще одно название, которое важно запомнить, – Noveto (noveto.biz). Эта относительно молодая фирма (основана в 2011 году) из Израиля предлагает «умную систему персонального звука» под брендом Sowlo – Smart Personal Sound (SPS). При помощи специальных 3D-сенсоров, которые отслеживают изменение положения тела, достигается эффект прицельного попадания звуковой волны в ушную раковину, так что можно даже слушать музыку без помех для окружающих.

Из европейских исследователей надо, вне всякого сомнения, упомянуть ученых Fraunhofer

IIS Audio and Media Technologies, которые занимаются объемным звуком уже на протяжении нескольких лет. На IBC 2016 они показывали свое готовое VR-решение Cingo для мобильного просмотра. Интересно будет посмотреть, что они приготовили в этом году – узнаем об этом в сентябре на IBC 2018.

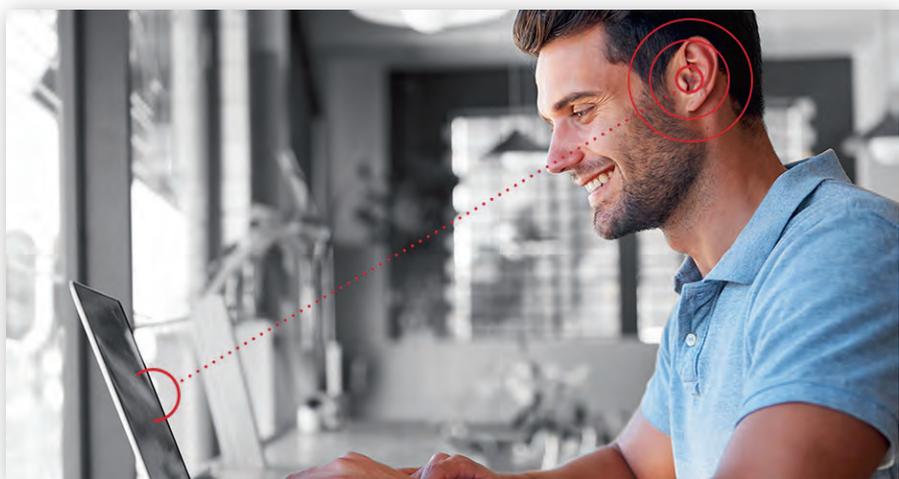
Одна из наиболее фундаментальных работ на 144-й AES была представлена специалистами из Японии. Ее название – «Создание 64-канального сферического микрофонного массива и 122-канального массива звуковоспроизведения для исследования технологии ViReal Dome для записи и воспроизведения звукового 3D-поля» (Development of a 64-Channel Spherical Microphone Array and a 122-Channel Loudspeaker Array System for 3D Sound Field Capturing and Reproduction Technology Research). Это совместное исследование специалистов из фирмы Yamaha и университета Киото. Для исследования была построена специальная лаборатория в форме полусферического купола ViReal Dome с массивом из 6 микрофонов, расположенных по спирали Фибоначчи, и 122 акустических системы. Коллеги действительно подошли к вопросу со всей серьезностью, и вообще складывается такое впечатление, что наиболее интересные исследования ведутся сегодня именно в области звука.

#### Глобализация контента и искусственный интеллект для аудио

И, наконец, третья тенденция, которую необходимо обсудить в связи с переориентацией телевидения на реальную глобальность и модель Direct to Consumer. На самом деле, применение искусственного интеллекта в обработке звука становится все шире. Это и «интеллектуальное шумоподавление», и «умный мастеринг» для записи сложных музыкальных произведений, исполняемых оркестром, который особенно востребован при прямых трансляциях, где на ручной мастеринг просто нет времени. Искус-

ственный интеллект будет также широко востребован для пространственной реконструкции звука в производстве VR-контента, ибо далеко не всегда будет возможность записать звук с использованием 64 микрофонов. Но наиболее актуальное на сегодня приложение – это обработка речи. Со времен падения мифической Вавилонской башни люди разговаривают на разных языках. И если изображение универсально, как, впрочем, и музыка, то вот речь – это реальная проблема. По крайней мере, так было до сих пор. Но на помощь производителю контента и вещателю уверенно приходит искусственный интеллект – AI. Преобразование звукозаписи в титры и конвертирование титров в живую человеческую речь – вовсе не такая простая задача, как может показаться на первый взгляд. Даже в рамках одного языка есть диалекты и акценты. Да что там акценты – интонация, которую человеческое ухо (а вернее сказать, – мозг) легко улавливает и распознает, для машинного переводчика представляет непреодолимое препятствие. На прошедшей NAB в программе технической конференции даже была отдельная сессия – «Искусственный интеллект применительно к субтитрованию» (Artificial Intelligence application for closed captions). На ней выступили представители компаний IBM Watson Media, Ericsson, Telestream. Провести сессию пригласили одного из известнейших и уважаемых в мире специалистов – Криса Хомера (Chris Homer, <http://chrishomerllc.com>), и он отлично с этим справился. Автоматический перевод аудиотреков любого контента – задача очень насущная, и решение ее не за горами.

Какие же выводы можно сделать, глядя на эти тенденции и достижения? Пришла эпоха звука! Об этом надо помнить и обращать пристальное внимание на новейшие звуковые решения. Как ни парадоксально, правильный звук может оказаться самым главным конкурентным преимуществом в теле- и кинопроизводстве и доставке контента. ▶



*Sowlo позволяет передавать звук без классических наушников точно в заданное положение, где находится ухо пользователя*

### PROFNEXT

НОВЫЕ МОДУЛИ

**Модульная система до 16 Гбит/с**

**PN-TDM-066** – 12-канальный оптический трансивер HD-SDI с электрическим временным уплотнением (TDM):

- 6-канальный передатчик (Tx) + приёмник (Rx) HD-SDI
- до 6 сигналов HD/SD-SDI, ASI на одной длине волны
- двунаправленная передача по двум или по одному волокну
- расстояние передачи до 80 км

**Серия PN-TRP-200** – оптические преобразователи 3G/HD/SD-SDI, ASI с автоконфигурированием в зависимости от типа SFP: двухканальный передатчик; двухканальный приёмник; передатчик + приёмник

**Серия PN-CRP-201** – оптические/электрические резерваторы SDI, ASI с автоконфигурированием в зависимости от SFP:

- с оптическими входами, электрическим и оптическим выходами
- с электрическими входами и оптическим выходом
- с оптическим и резервным электрическим входами и электрическим выходом

**Формирователи мультиэкрана (Multiscreen)**

8 - 32 источника сигнала 3G/HD/SD-SDI в составе:

- PN-MSC-030 - процессор мультиэкрана
- PN-MEX-031-1/8, 9/16, 17/24, 25/32 - 4 вида входных блоков
- Для максимального количества источников (32) необходим процессор плюс все четыре входных блока
- ПО для произвольной раскладки окон любого размера
- Соединение процессора с входным блоком 1/8 и входных блоков между собой - внешним кабелем



Процессор PN-MSC-030 с входным блоком PN-MEX-031-1|8

### ProBox – автономные модули

**PBX-ENP-200** – конвертер двунаправленный TsoIP↔ASI

- автономный шлюз TsoIP→ASI и ASI→TsoIP
- алгоритм устранения джиттера на IP-входе
- скорость потока ASI до 216 Мбит/с
- входы/выходы – ASI, IP, GPIO
- выделенный порт Ethernet для настройки
- WEB-интерфейс, поддержка SNMP

**PBX-CC-300** – преобразователь стандартов разложения

- 3G/HD/SD-SDI/HDMI, Up- и Down-преобразование
- синхронизация выходного сигнала с опорным
- поддержка 16 каналов (4 группы) вложенного звука
- пропуск телетекста с входа на выход.
- преобразование формата кадра и масштабирование
- управление с лицевой панели или по Web-интерфейсу

**PBX-STR-500** – H.264-сервер потокового вещания

- трансляция контента с одновременной записью его в формате H.264, скорость потока до 20 Мбит/с
- входы: SDI и HDMI, бесподрывное переключение
- встроенный Chroma Key для наложения графики
- наложение на сигнал аудиопрограммы внешнего источника
- звука и/или сигнала микрофона
- управление по Web-интерфейсу
- подключаемая к USB память (объемом до 1 ТБ)



НОВЫЕ  
УСТРОЙСТВА

### PROFLEX

НОВЫЕ МОДУЛИ

**Модульная система до 3 Гбит/с**

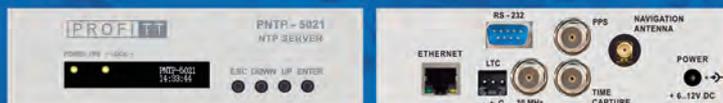
**Серия PDDA-7106-x** – усилители-распределители аудио AES/EBU

**PCOS-7356** – Smart-резерватор синхронных сигналов HD/SD-SDI, бесподрывный аварийный коммутатор; расхождение по времени сигналов основного и резервного каналов – до 15 кадров по видео, до 600 мс по аудио

**PCOA-7105** – резерватор аудио AES/EBU с детектором “тишины”; автоматическое, ручное местное и ДУ-переключение; программируемые критерии перехода; регулировка параметров определения “тишины” – порога и длительности паузы



### PNTP-5021 – сервер точного времени

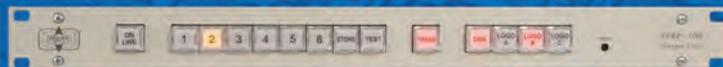


- Синхронизация от GPS/ГЛОНАСС
- Формирование сигналов 10 МГц, 1 PPS и LTC по стандарту EBU/SMPTЕ309M
- Вывод навигационной информации через RS-232 по протоколу NMEA0183
- Кратковременная нестабильность (девиация Аллана) за 1 с – 1×10-11
- Дистанционное управление по протоколу SNMP и web-интерфейсу
- Питание устройства – внешний адаптер 6...15В.

### PDMX-2106(SM)



Шестивходовый  
эфирный микшер HD-SDI



**Конфигурации:**

- PDMX-2106 – системный блок PDMX-2106F (1U) и пульт дистанционного управления PDMX-2106P со встроенным сенсорным экраном
- PDMX-2106SM – системный блок PDMX-2106F (1U) и пульт дистанционного управления PFRP-4106 (1U)