

# Диалог в телевизоре

Артем Калинин

**Т**елезритель не склонен напрягаться – речь должна восприниматься легко. А всегда ли так получается?

Конечно, основное в ТВ – это картинка, но если не слышно, то мы возвращаемся на сто лет назад, в эпоху немого кино. Правильно ли это? Как сделать, чтобы было слышно? Решается задача комплексно: используются технические инструменты и организационные меры.

## Организационные меры

В числе организационных мер – правильная расстановка приоритетов. Для начала надо решить, что важнее: показать корреспондента на месте события или донести до зрителя информацию об этом событии? Решение должен принимать не корреспондент и не звукорежиссер – им предлагается только исполнение.

Рассмотрим типичные ситуации. В кадре ведущий в студии, слышен его голос: «Итак, «Формула 1». Наш корреспондент с места события сейчас нам все расскажет». В кадре корреспондент с микрофоном. Слышен его голос на фоне звука автомобильных двигателей, визга шин и криков болельщиков.

Нетрудно при желании ослабить интершум до состояния, когда он не мешает речи (рис. 1): взять направленный микрофон и поднести его поближе к рту. Останется что-то слышно с места



Рис.1. Варианты сведения голос/интершум: подавлен интершум

события, но содержание этих звуков едва угадывается. Речь корреспондента слышна разборчиво, но вместо эффекта присутствия мы получаем звуковой мусор. Не самый лучший результат.

Несложно поднять интершум, чтобы было слышно, что же именно

происходит в кадре (рис. 2): можно взять микрофон с менее выраженной направленностью или отнести его подальше от рта. Но в этом случае будет зашумлена речь корреспондента, и результат тоже едва ли можно назвать хорошим.



Рис. 2. Варианты сведения голос/интершум: избыток интершума

## Обновленная линейка систем VECTORBOX

Подробности и техническая информация о системах на [WWW.VECTORBOX.RU](http://WWW.VECTORBOX.RU)

Спрашивайте у ваших поставщиков профессионального оборудования

Представитель и технический центр VECTORBOX по СНГ – компания «АРТОС». Тел.: 8 (495) 223-9202. E-mail: info@artos.ru

### А ТАКЖЕ

Интеллектуальные медиасерверы для монтажных комплексов на базе локальных сетей

## WWW.EDITSHARE.RU

Одновременная работа до 32 NLE разных производителей: Avid Adrenaline (Mac and Windows), Avid Xpress Studio (Complete and Essentials), Avid Xpress Pro HD (Windows), Avid Xpress Pro (Mac + Windows), Avid Xpress DV (Mac + Windows), Avid Liquid 7, Pinnacle Liquid Chrome HD, Pinnacle Liquid Edition Pro, Apple Final Cut Pro, Adobe Premiere Pro, Sony Vegas, Canonius Edius



Рис. 3. Варианты сведения голос/интервю: временное разделение

Существует ли оптимальное отношение уровней речи и интершума? Скорее нет, чем да. Дело в том, что обычному зрителю трудно одновременно следить за тремя процессами: изображением, речью корреспондента и интершумом.

Что делать? Можно сократить количество потоков информации, действующих одновременно: останется только одно изображение и один звук (рис. 3).

1. Сначала идет анонс из студии (голос + человек в кадре).

2. Корреспондент в кадре рассказывает о событии, а интершум подавлен до уровня, при котором он едва угадывается и не отвлекает: снова только две переменных.

3. Камера показывает место события и идет только интершум. Телезрители имеют возможность увидеть и услышать все без комментариев: звук и видео с места события.

4. Видео продолжает транслироваться с места события, а комментирует происходящее ведущий в студии, интершум ослаблен и не ухудшает разборчивость.

5. Корреспондент в кадре завершает репортаж и прощается, преобладает его голос, которому ничего не мешает.

В каждый момент времени к телезрителю идет изображение и «один звук», но по завершении репортажа остается впечатление, что слышен был и голос корреспондента, и звук с места события, и голос комментатора в студии. Собственно, так все и было, и для достижения этого положительного результата не требовалось передавать все звуки одновременно.

Устанавливая при монтаже баланс речь/интершум, следует помнить о свойствах слуха. Когда вы слушаете на сведении знакомый текст, то вопрос разборчивости не возникает, и на

первый план выходит передача ощущения атмосферы места события, чувства соучастия, которые опираются именно на звуковой фон. Телезритель, слышащий этот текст впервые, пытается разобрать в шуме, что же там сказали. Если речь непонятна, то контакта со зрителем не будет, и все труды пропадут втуне. При сведении сначала надо выставить на слух баланс речь/фон, после чего убавить шумы на 6 дБ – получится именно то, что нужно. Имейте в виду еще и компрессию при трансляции, она «спрессует» звук.

Качество надо обеспечивать на съемках, не откладывая на потом. Телевидение – это поточное производство и «шлифование деталей» плохо вписывается в темп работы. Хотелось бы иметь технологию, при которой не требуется никакая дополнительная работа со звуком. Идеал недостижим, но можно так выстроить процесс, что не понадобится дополнительная чистка звука. В основе такой технологии – деление. Если нужен диктор, то пишем его речь, если нужен интершум – пишем только его. Если нужен комментарий из студии – убираем почти в ноль остальные звуки. В идеале звук при съемках «в поле» хорошо бы писать на несколько каналов, разделив хотя бы ведущего и интершум.

Итак, самый эффективный инструмент обработки звуковых сигналов – аранжировка, то есть организация взаимодействия звуковых сигналов от разных источников. Иными словами – организационные мероприятия, которые ни в коей мере не отменяют технической работы с сигналами и оборудованием.

## На площадке

Слух человека умеет разделять сигналы, приходящие с разных направлений, а микрофон их объединяет. Так что при работе с речью не следует увлекаться микрофонами с широкой диаграммой направленности, а сам микрофон надо держать поближе ко рту, то есть в зоне преобладания прямых сигналов. Не забывая при этом про эффект близости, то есть про подъем АЧХ на низких частотах, возникающий при приближении микрофона к источнику звука. Слух человека адаптируется к стационарным искажениям сигнала, мы уверенно замечаем, например, подъем АЧХ на 3 дБ на частоте 200 Гц, но раздражения такие искажения не вызывают. При изменении положения головы относительно микрофона изменяется и АЧХ, то есть появляется модуляция в такт с движением головы. Если она и не выведет слушателя из состояния равновесия, то отрицательно скажется на качестве звука – это факт. Чтобы уменьшить искажения такого рода, следует правильно выбрать микрофон и научить дикторов им пользоваться.

Практически отсутствует модуляция при работе с гарнитурой, поскольку фиксировано ее положение относительно рта. А вот клипса или ручной микрофон требуют пристального внимания. В крайнем случае, раздражающее действие модуляции АЧХ на низких частотах можно уменьшить,



Двухканальный влестудийный предусилитель, который можно применять в сочетании с рекордером



Гарнитура



Микрофон-пушка

отрезав фильтром с крутизной 12 или даже 18 дБ на октаву проблемную часть спектра речи – ниже 300 Гц.

Мы без напряжения понимаем речь стоящего впереди человека на фоне шума проезжающего сзади транспорта, но в записи она может оказаться не столь разборчивой, если не принять специальных мер: можно использовать «пушку», можно снабдить человека гарнитурой или просто дать ему в руки направленный микрофон.

Если нельзя неприятность предотвратить, то можно компенсировать ее влияние – эта идея тоже широко применяется в борьбе с шумами. Принцип простой: из выходного сигнала микрофона, расположенного рядом с источником полезного сигнала, вычитается сигнал второго микрофона, который расположен чуть поодаль. Практически удается на 20 дБ ослабить влияние источников шума, удаленных более чем на полметра. Выпускается целое семейство микросхем (LMV10XX National Semiconductor), предназначенных для применения в мобильных телефонах, гарнитурах и других ре-

чевых устройствах. Возможно, взывательный вкус звукорежиссера, воспитанного на высоком искусстве, и не будет удовлетворен качеством звука, но не спешите с оценками и приговором, послушайте звуковой пример ([http://www.national.com/analog/audio/noise\\_reduction](http://www.national.com/analog/audio/noise_reduction)).

### After Party, или после бала

Что делать, если запись все же получилась невнятная, с невысокой разборчивостью или зашумленная?

Вариант 1. Дать в эфир как есть. Поточное телевизионное производство не располагает к «полировке до блеска». Кроме того, улучшая одно, можно ухудшить другое.

Вариант 2. Попытаться почистить фонограмму. В этом случае хотелось бы избежать научных изысканий, а опираться на шаблонные решения с предсказуемыми результатами.

### Статическая фильтрация

Теоретически, строго говоря, все реальные сигналы конечной длительности, в том числе и речевой,

спектр имеют бесконечный. Практически же на краях диапазона спектральная плотность довольно быстро падает, и можно говорить о его границах.

Если проанализировать спектр речевого сигнала, то обнаружим, что накрыта практически вся полоса 20 Гц...20 кГц. На низких частотах мы видим последствия взрывных согласных, которые есть практически во всех языках мира: [p], [t], [k]. Вслушайтесь, как звучит слово «пушка», спектр сигнала фонемы [p] начинается почти с нуля. По мере удаления ото рта мы видим ослабление инфранизкочастотных составляющих спектра, а на расстоянии 2...3 м их уже практически невозможно обнаружить, но микрофон расположен обычно гораздо ближе.

Слово «свисали», например, богато высокочастотными спектральными составляющими. Можно привести еще «высокочастотное» имя Василий, приятное кошачьему слуху, диапазон которого простирается далеко за 20 кГц.

<b>ФОРВАРД Т</b> Комплексная автоматизация телевизионного вещания	<b>СТРИМИНГ</b> Врезка рекламы и наложение титров в цифровом ТВ (MPTS)	<b>ГОЛКИПЕР</b> Система для многоканальной записи и замедленных повторов	<b>ФОКУС</b> Виртуальные студии и трехмерная графика реального времени
--	---	---	---

**Нелинейный монтаж – Многослойные титры – Мультиформатность – Автоматизация вещания – Видеосервер – Перепланирование ретрансляции – Распределенная сетевая архитектура – Многоканальный ввод**

11:37:07.00 11:42:45.50 11:42:15.50 11:42:20.51 11:45:00.00 11:45:00.00 11:45:17.82 11:46:11.00	00:34.44 00:01.00 -0:00:35.34 +0:01:11.60 00:17.82 00:03.36 00:15.40	0:10 0:10 0:10 5:30 0:10 0:10 0:10	Daredoon.avi pogoda_16_3.avi mc_24_11_02_Letted.avi vspornaya_16_3.avi
--	--	--	---

ФорвардТ - система нелинейного монтажа и подготовки видео, аудио и графических материалов  
 ФорвардТА - комплексная система автоматизации эфира  
 ФорвардТВ/ТП2 - сервер задержки ретранслируемого сигнала  
 ФорвардТТ - наложение многослойных титров на проходящее видео  
 ФорвардТК - система высококачественной рипроекции, 2D виртуальная студия

**IBC2010**  
 RAI Amsterdam  
 Stand 7.A05 Hall 7

Контактная информация: (383) 333 1067 330 0220 факс: (383) 333 2173 www.softlab.nsk.com/ru/forward/index.html forward@softlab.nsk.com

Можно определить частотный диапазон не только по свойствам сигнала, но и по критерию слуховой заметности ограничения спектра.

В системах телефонной связи используется канал, пропускающий сигналы с частотами 300...3400 Гц, и этого вполне достаточно, чтобы без напряжения узнать голос и понять сказанное.

В свое время был предусмотрен для информационного вещания канал с полосой 100...6300 Гц. Следует согласиться с мнением тех, кто считает эту полосу вполне достаточной для речи.

Итак, спектр сигнала на выходе расположенного вблизи рта микрофона весьма широк, не ограничен пределами 20 Гц...20 кГц. Даже тренированное ухо едва ли способно заметить ограничение спектра при полосе пропускания 30 Гц...15 кГц. Сужение спектра до полосы 50...10000 Гц замечают многие. Наличие фильтра с полосой 100...6300 Гц будет замечено подавляющим большинством слушателей – это факт. И объективным фактом является улучшение субъективного качества речевого сигнала при ограничении его спектра.

Низкочастотные помехи (ниже 100 Гц) представляются не очень важными, поскольку в этой части спектра спадает чувствительность слуха, и акустические системы большинства телевизоров их не воспроизводят. Звуковое давление оказывается ничтожным, хотя диффузор и движется с большой амплитудой. Звуча нет, но есть нелинейные искажения, обусловленные этим движением. И величина этих искажений может быть существенной. Проверить утверждение несложно: добавьте к звуковому сигналу низкочастотную помеху (синусоидальный сигнал с частотой, например, 20 Гц) и послушайте, как воспроизводят эту смесь акустические системы бюджетного уровня. Только осторожно, поскольку можно даже повредить головку низкочастотным сигналом.

Итак, фильтр с крутизной спада АЧХ 12 или даже 18 дБ на октаву, частота среза 100...200 Гц и даже 300 Гц, в зависимости от зашумленности сигнала. Чем больше помех, тем выше следует выбирать частоту среза ФВЧ и крутизну спада АЧХ.

Высокие частоты – если речь идет об информационно-новостном сегменте, то разумным компромиссом будет включение ФНЧ 6...7 кГц с крутизной спада АЧХ 12 или 18 дБ на октаву. В этом случае выигрыш от подавления шумов и помех вполне окупает незначительные потери речевого сигнала.

Работа систем уплотнения в телефонной сети способна привести к появлению высокочастотных помех в слышимом диапазоне частот, хотя и за пределами полосы телефонного канала, то есть выше 3400 Гц. В телефонных гибридах высокого качества имеется специальный фильтр с частотой среза 3400 Гц, и затухание на частоте 5 кГц может превышать 30 дБ, то есть все высокочастотные помехи оказываются подавлены. В аппаратах попроще фильтрация может оказаться не столь высокой и помехи будут слышны как высокочастотный «зуд». Такой звук, увы, не редкость и у «живых» звонков в эфир, и при трансляции предварительно записанных телефонных интервью. Остается без ответа вопрос, есть ли вообще смысл передавать изображение телефонного аппарата, но о качестве звука следует позаботиться, по крайней мере – подавить помехи за пределами телефонного канала 300...3400 Гц.

#### Динамическая фильтрация

Желание уменьшить шумы фонограммы – давнее, практически ровесник самой фонограммы. В основу большого семейства шумоподавителей положена вполне разумная идея: считать шумами слабые составляющие сигнала и давить их. Именно так устроено, например, одно из известных решений: динамический шумоподавитель DNR (Dynamic Noise Reduction).

И на магнитофонах можно было увидеть буквы DNR, и в системе телекоммуникаций использовалось динамическое подавление шума. И до сих пор выпускается специализированная интегральная схема LM1894 National Semiconductor.

Шумоподавитель DNR представляет собой управляемый фильтр низких частот. При номинальном значении уровня сигнала на входе (0 дБ) шумоподавитель имеет прак-

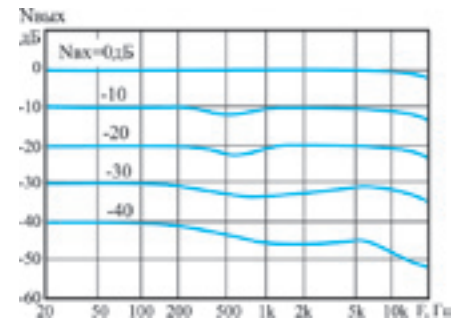


Рис. 4. Семейство АЧХ шумоподавителя DNR

тически горизонтальную АЧХ. При отсутствии сигнала на входе частота среза ФНЧ составляет около 500 Гц.

При уровне входного сигнала -40 дБ относительно номинального искажения АЧХ достигают 6 дБ, что заметно на слух. При этом подавление шумов составляет около 10 дБ, что тоже заметно. Параметры системы DNR – компромисс, при котором выигрыш по шумам еще оправдывает заметность обработки сигналов и потери их высокочастотных составляющих.

Уместно вспомнить и HUSH Stereo Noise Reduction System, для которой тоже выпускалась специализированная микросхема SSM2000 (Analog Devices).

Широкие возможности цифровой обработки сигналов используются и для подавления шумов. Большинство этих программ представляет собой фильтр, параметры которого зависят от свойств сигнала. В какой-то степени шум и помехи удается ослабить, но остается неприятное последствие обработки: модуляция сигнала, которая отрицательно сказывается именно на разборчивости. Так что не переоценивайте возможности вычислительной техники: шум не исчезает, он заменяется искажениями, и пример DNR это вполне убедительно демонстрирует.

#### От теории к практике

Как было уже отмечено, телевидение – это конвейер, а не ателье индпошива. В поточном производстве неуместны поиски, нужны типовые решения. Что для этого потребуется? Немного: комплект микрофонов, двухканальный рекордер, голова и уши. Главное: убираем все лишнее. Телевидение – искусство простого и ясного звука. Неразборчивая речь – это не CNN в Рио-де-Жанейро, это гораздо хуже.