

Манов П.А. Доклад MP3Pro

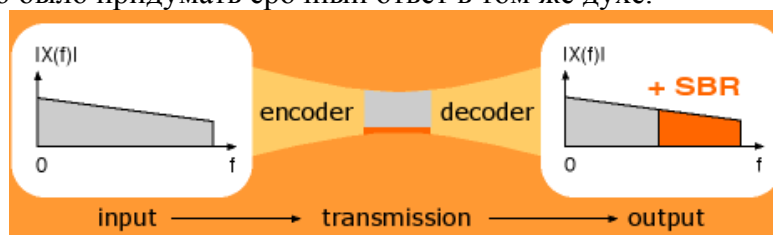
Частная шведская компания Coding Technologies (<http://www.codingtechnologies.com/>) с 1997 года занимается улучшением перцепционных кодеков, таких как mp3 или AAC, – с сохранением обратной совместимости – посредством революционной технологии SBR (Spectral Band Replication – копирование полосы спектра).

14 июня 2001 года ее деятельность вызвала к жизни mp3PRO – новый формат аудиосжатия, обеспечивающий лучшее, чем у mp3, качество при меньшем битрейте.

Не стоит много писать о том, какой стала наша жизнь с появлением mp3. Об этом уже написаны гигабайты патетических текстов на всех языках мира, включая эсперанто и C++. И вряд ли найдется в цивилизованном мире человек, который не знает, что в основе mp3 лежат алгоритм психоакустики, иначе говоря, удаления тех частотных составляющих спектра, которые среднестатистический человек по причине своей глуховатости и невнимательности просто не слышит, и сжатие отфильтрованного психоакустикой звука упрощенным аналогом алгоритма Huffman'a. Во всяком случае, если это и стало для Вас откровением, теперь-то Вы совершенно точно знаете, что такое mp3.

Создатель mp3 немецкая компания Fraunhofer IIS (<http://www.iis.fhg.de/>) утверждает, что человек (среднестатистический, разумеется) не способен отличить mp3 128 kbps (килобит в секунду) от CD-аудио (1411 kbps), а это означает, что без ощутимых потерь качества можно сжать аудио в 12 раз. И казалось бы сейчас в эпоху 40-гигабайтовых винчестеров, всеобщего DSL и грошовых CD-R носителей такого сжатия более чем достаточно, но аудио-компрессия продолжает волновать умы. Во-первых, довольно много людей могут услышать артефакты кодирования и на более высоких битрейтах, а во-вторых, для Интернет вещания или хранение музыки на портативных плеерах хотелось бы еще больше сжать звук – с приличным качеством, разумеется.

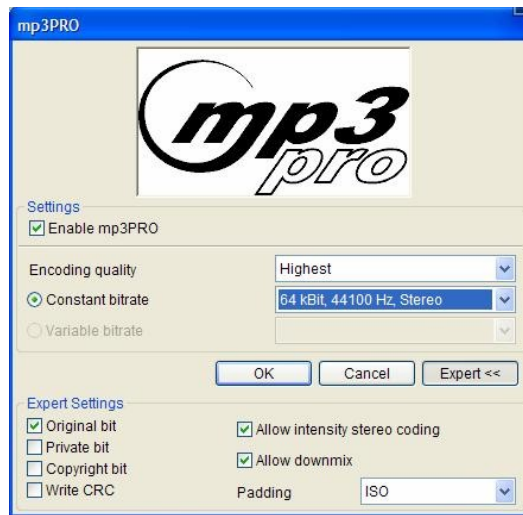
А еще – что, наверное, не менее важно – Fraunhofer IIS и Thomson Multimedia – компании, занимавшиеся разработкой и маркетингом mp3, а теперь инвесторов и стратегических партнеров Coding Technologies – волнует выход в свет средства сжатия аудио в WMA – собственный мелкомягкий формат, обеспечивающий, как утверждает Microsoft, качество mp3 128 kbps на битрейте в 64 kbps. Чтобы не быть вытесненными с рынка аудио-компрессии, нужно было придумать срочный ответ в том же духе.



Вот здесь и пригодилась SBR (Spectral Band Replication – копирование полосы спектра) – «революционная» технология эффективного кодирования высоких частот от Coding Technologies. Вот в чем смысл: кодек должен передавать более узкий диапазон частот, а при декодировании высокие частоты будут восстанавливаться из анализа информации о низшей части спектра. Для обеспечения точного восстановления в закодированном файле сохраняется некоторая дополнительная информация.

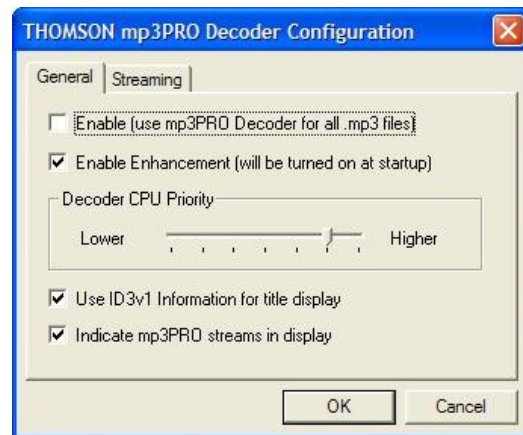
Становится очевидным, что совместно с SBR нужно использовать уже существующий кодек, который будет сожмет нижнюю часть спектра. Для mp3PRO таким «подлежащим» кодеком стал mp3. Кроме того, SBR – это главным образом преобразование звукового потока на стадии декодирования, а значит требования к ресурсам компьютера возрастают.

Уже существует первый и единственный полнофункциональный кодировщик mp3PRO – plugin ко всеми любимой программе записи CD-R(W) Nero – Burning Rom. Trial-версия его, работающая 30 раз, включена в Nero 5.5.4.0 и более поздних.



Можно выбрать любой из поддерживаемых битрейтов (32, 40, 48, 56, 64, 80 и 96 kbps), и между высокой скоростью и высоким качеством, или, я бы сказал, убогим качеством и черепашной скоростью :-). Ну а если серьезно, я особой разницы не заметил.

Если убрать галочку с “Enable mp3PRO” получим стандартный Фраунхоферский кодировщик mp3.



Заодно Coding Technologies выпустила и соответствующий plug-in для Winamp’a, дабы наколенные звуки не пропали даром. Version 0.98 (Build 5) не предвещает ничего хорошего в плане стабильности; этот plug-in берет на себя функции проигрывания обычных mp3 (поскольку файлы что mp3, что mp3PRO имеют одинаковое расширение *.mp3) и редактирования ID3 тегов, причем ID3v2 не поддерживаются.

Есть еще Thomson mp3PRO Audio Player, но он заслуживает разве что упоминания. Все равно все пользуются Winamp’ом. Здесь (<http://www.mp3prozone.com/download.htm>) можно поглядеть на скриншот, почитать, что он умеет, и загрузить.

Методика тестирования

Отрывки формата PCM 44Hz, 16 bit, Stereo, взятые с аудио CD, кодировались в mp3 и mp3PRO plug-in’ом для Nero – Burning Rom на следующих битрейтах:

- 1) mp3 128 kbps vs. mp3PRO 64 kbps
- 2) mp3 192 kbps vs. mp3PRO 96 kbps

Во всех случаях качество кодирования устанавливалось в Highest.

Может быть Вам интересно, почему собственно я не стал кодировать mp3 LAME’ом? Потому что формат mp3 не предусматривает никаких ограничений на используемую кодеком частотную фильтрацию и психоакустику, алгоритмы LAME’а и plug-in’а для Nero могут различаться как день и ночь. А поскольку кодер mp3PRO включает в себя алгоритм mp3, интересно его сравнить с mp3, полученным с использованием того же алгоритма. Кроме того, используемый алгоритм от IIS Fraunhofer (включенные в l3enc, mp3 Producer, и многие коммерческие продукты для работы со звуком, например, Steinberg WaveLab), идеален для низких битрейтов.

В качестве декодировщика в WAV (PCM) использовался Winamp 2.77 с plug-in'ом in_mp3.dll версии 2.22 от Fraunhofer IIS и in_mp3pro.dll версии 0.98 (Build 5) от Coding Technologies.

К сожалению, не существует никаких консольных, вроде LAME'а, кодеков mp3PRO.

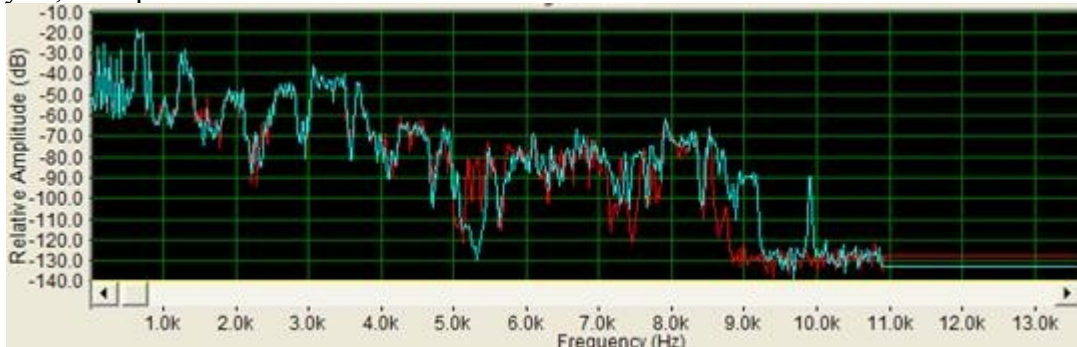
Для анализ результатов применялись Cool Edit Pro 1.2a и SpectraLAB 4.32.17.

Музыка, подвергшая столь нещадному издательству: G. Caccini. "Ave Maria". Высокие регистры органа Рижского Домского Собора и оперное сопрано. Великолепное звучание, мягкие и прозрачные «верха». Именно их кодеры должны будут сохранить, а передать их сложнее всего.

"White Stones" инструментальная музыка норвежской группы Secret Garden. Множество колоритных инструментов: скрипка, ирландский whistle, испанская гитара etc. Чтобы достоверно передать насыщенную накладываемыми друг на друга тембрами живых инструментов оркестровку, кодерам придется потрудиться.

mp3 128 kbps vs. mp3PRO 64 kbps

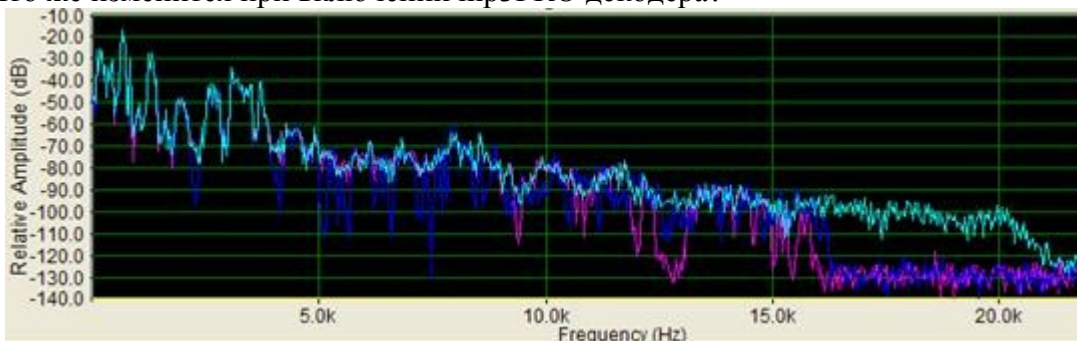
Как я уже говорил, основной кодека для mp3PRO – mp3, а значит mp3PRO можно разжать в PCM как самый обыкновенный mp3. Иначе говоря, мы можем услышать и увидеть на АЧХ, что же представляет собой mp3PRO как таковой. Для этого отключаем in_mp3pro.dll, декодируем, смотрим...



mp3PRO 64 mp3 64

Так вот что интересно. Обещанное уменьшение битрейта в два раза получено не столько срезанием высоких, сколько даунсемплингом! Частота дискретизации полученного файла – 22 KHz, а это грозит появлением звонов, хрипов, нечеткости звучания. Фраунхоферский кодек не позволяет получить в mp3 64 kbps дискретизацию 44 KHz и стерео одновременно, хотя это противоречит маркетинговым заявлениям, что битрейт экономится исключительно за счет срезания высоких частот. Частоты все-таки срезаются на 8-9 KHz, но это скорее заслуга Фраунхоферского алгоритма. И еще можно заметить: АЧХ mp3PRO выглядит хуже, нужно оставить место для некой таинственной информации, которая поможет декодеру достоверно передать высшую составляющую спектра.

Что же изменится при включении mp3PRO декодера?



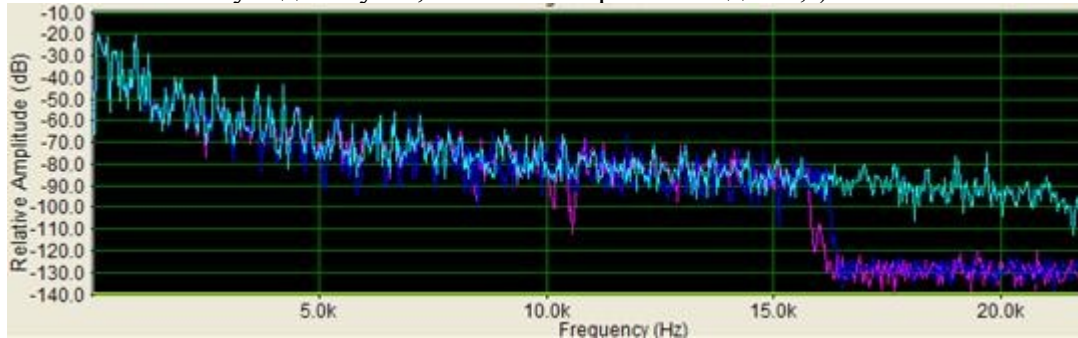
mp3PRO 64 WAV mp3 128

Стоит сравнить две части АЧХ: если до 8.7 KHz mp3 практически идентичен оригинальному PCM в отличие от «скачущего» mp3PRO, то затем mp3PRO оказывается гораздо детализованней и ровнее mp3, но некоторые детали просто непонятно откуда взяты.

Впрочем, ничего удивительного. Поведение mp3PRO в низшей части спектра (0-8 KHz) объясняется ограничениями битрейта, а затем вступает в действие SBR, который синтезирует высшую часть (8-16KHz), и, если внимательно изучить АЧХ, можно заметить, что эти две части повторяют друг друга, но форма верхней изменена достаточно, чтобы походить на оригинал. Не зря SBR означает копирование полосы спектра.

А теперь о моих субъективных впечатлениях. В целом mp3 и mp3PRO звучат примерно одинаково. Звук органа сильно размазан, особенно органной педали, которая звучит как один непрерывный гул, ни о какой прозрачности верхов речи не идет, хотя вокал звучит довольно прилично. Да что там говорить! Ну невозможно втиснуть хорошую музыку в такой битрейт!

Взглянем теперь на отрывок из Secret Garden, когда мы имеем дело в несколькими инструментами, звук которых локализован с разных частях спектра. На АЧХ этого не видно, но на то человеку и даны уши, чтобы тестировать кодеки ;-)



mp3PRO 64 WAV mp3 128

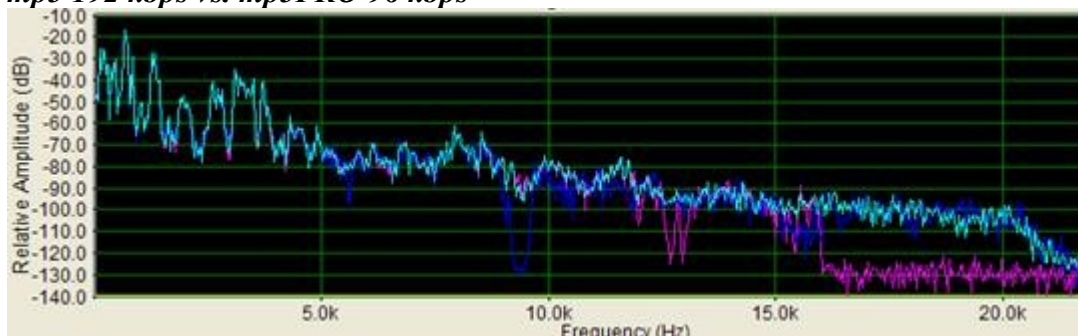
В общем ничего нового. АЧХ mp3, как всегда, наиболее близка к оригиналу, а mp3PRO синтезирует такие высокие частоты, о необходимости которых авторы даже и не догадывались. ;-)

Полифония прощупывается с трудом. Это общий грех психоакустики – превращать звук несолирующих инструментов в однородный смутный гул, во всяком случае на таких битрейтах.

Звучание mp3 оказалось глуховатым на верхах. А mp3PRO, напротив, ярким... черезчур ярким. Во-первых, появилось нечто вроде хай-хетов – транспонированием более низких ударных. Во-вторых куда-то исчезла скрипка. Нет, не совсем исчезла – кое-что вместо нее есть, но мне звук этого инструмента напоминает... губную гармошку. Тут следует вспомнить о ирландском whistle и станет ясным, откуда SBR извлек такую оригинальную интерпретации скрипичного тембра.

Я бы назвал полученный mp3PRO файл не столько сжатым оригиналом, сколько отдельной версией, новой оркестровкой или, если хотите, ремиксом ;-)))

mp3 192 kbps vs. mp3PRO 96 kbps



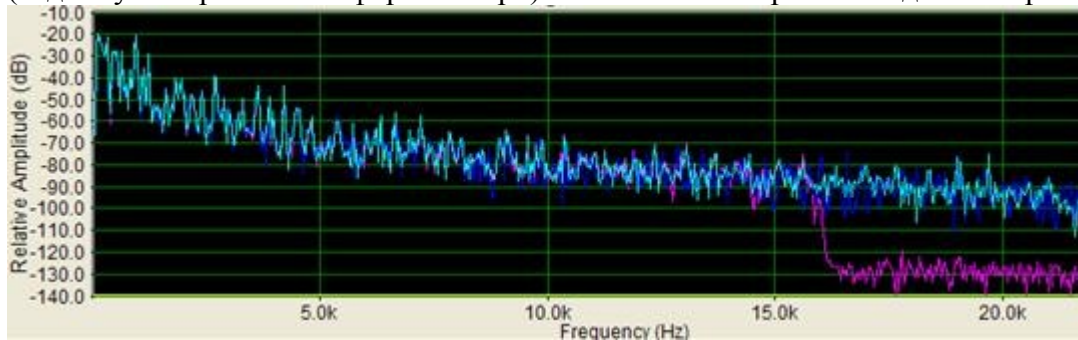
mp3PRO 96 WAV mp3 192

Как и следовало ожидать, чем выше битрейт, тем выше и качество. Только вот представления о качестве у каждого кодека по-прежнему свои. Например, mp3 не стал покрывать большего частотного диапазона, зато детализованность и приближенность в оригиналу средних и низких частот возрасла, например исчез провал на 12-13 KHz. Алгоритм mp3PRO, как Вы поняли, озабочен не достоверностью, а скорее похожестью. Во всяком случае, заметно, что эту свою задачу он выполнил вполне. mp3 96 kbps 22 KHz передает

частоты до 11 KHz, а значит SBR может покрыть весь диапазон слышимости (0-22 KHz) – это как раз и есть предел «похожести».

Большого битрейта из mp3PRO выжать нельзя еще и потому, что 96 – максимальный из возможных битрейтов с дискретизацией 22 KHz.

mp3 отличился наибольшей реалистичностью несмотря на ощутимый недостаток верхов. mp3PRO же наоборот размазал органную педаль и давил на верхи. mp3 избавился от многих своих недостатков и не зря потерял битрейт, а mp3PRO по четкости и полноте звучания был ближе скорее к mp3 128 kbps. Тем не менее от такого низкого для mp3 битрейта (ведь звук сохраняется в формате mp3) такого качества раньше ждать не приходилось.



mp3PRO 96 WAV mp3 192

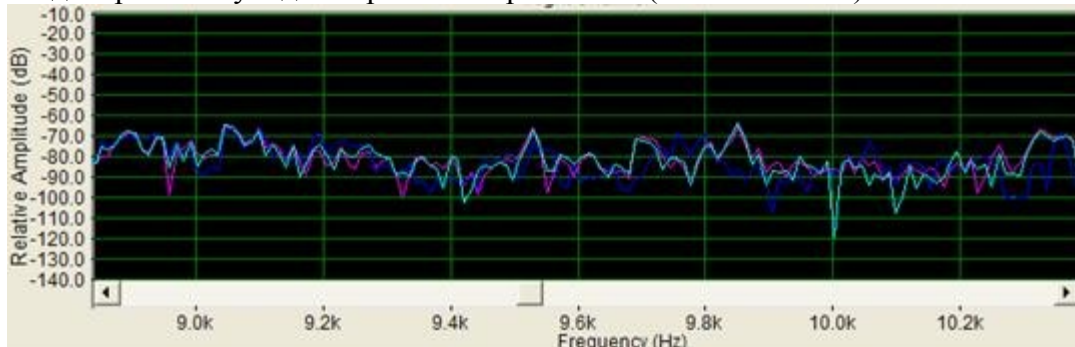
Обратите внимание: как ни старается АЧХ mp3PRO прижаться к оригинальному, все равно ненужные пики выдают ту истину, что верхняя часть спектра – всего лишь синтетическое подобие.

Поэтому ничего удивительного нет в том, что тембры реальных инструментов искажены и неестественны даже на таком сумасшедшем :-) битрейте.

Несчастной скрипке снова досталось – хотя она уже гораздо больше похожа на скрипку, чем на губную гармошку. Теперь чувствуется какие-то звенящие призвуки. Ну и черт с ними, в конце концов, я и не собираюсь хранить музыку в формате mp3PRO.

Разоблаченный mp3PRO

На следующем графике можно рассмотреть в подробностях волнующий нас момент, когда mp3PRO «уходит отрыв» от mp3 и PCM (на отметке 9.6)



mp3PRO 96 WAV mp3 192

Теперь мы располагаем достаточными сведениями, чтобы сделать предположения о процессе кодирования / декодирования mp3PRO. Конечно, тайну алгоритма хранят в Coding Technologies как зеницу ока, но если Вы – сотрудник этой компании, а я в чем-то ошибся – пишите :-))

Кодирование

Оригинальный звук в формате PCM с частотой дискретизации 44 KHz при помощи специальной версии Фраунхоферовского кодирующего алгоритма сжимается в mp3 заказанного битрейта и частотой дискретизации 22 KHz; именно специальной версии, чуть более строгой, позволяющей сохранить несколько kbps для последующей записи информации о высоких частотах. Затем определяется, какой частотный диапазон удалось сжать, и следующий за ним диапазон того же размера разделяется на поддиапазоны и для каждого

вычисляется амплитуда. Пары поддиапазон – амплитуда помещают в mp3 поток на место сэкономленных при сжатии килобитов.

Декодирование

Из потока mp3 извлекаются пары поддиапазон – амплитуда и восстанавливается приблизительный вид высшего диапазона АЧХ. mp3 поток декодируется, частота дискретизации поднимается до 44 КHz; полученный диапазон АЧХ копируется в конец себя самого и изменяется в соответствии с восстановленной информацией о амплитудах его поддиапазонов.

Выводы

Принцип mp3 прост: зачем хранить или передавать гигантские объемы звуковой информации, если человек все равно не способен услышать большУю ее часть? Идеальным вариантом стало бы составление индивидуальной психоакустической модели и, соответственно, оптимизированная сжатия под конкретные уши. В этом случае человек, прослушав исходный и сжатый звук, не заметил никакой разницы между ними. Но заточка психоакустики под каждого отдельного человека невысказана, поэтому мы довольствуемся некой «средней» психоакустической моделью. Высокое качество кодирования означает идентичность в исходном и сжатом аудио именно тех участков АЧХ, которые слышит «средний» человек.

mp3PRO – тот же mp3 с даунсемплингом на стадии кодирования, что позволяет уменьшить битрейт в два раза, и обработкой звука апсемплингом и питчем на стадии декодирования, что позволяет иногда приблизить качество к первоначальному. Такая обработка может привести к обратному: mp3PRO живет своей собственной жизнью, связанной с оригиналом лишь геометрическим сходством АЧХ.

Короче говоря, это довесок к mp3, способный улучшить его звучание на низких битрейтах. Ничего революционного. Никаких чудес.

Чем же, спросите Вы меня, отличается от SBR от enhancer'a, вроде DFX for Winamp (<http://www.fxsound.com/>), который не только добавляет верхи и требует за полную работоспособную версию себя денег ;-), но и усиливает басы, расширяет стереобазу и создает объем. Тем, что DFX прибавляет верхи вообще невпопад. Теперь уже mp3PRO оказывается достовернее, ибо запоминает хотя бы вид верхней части АЧХ. А это чтонибудь, да значит.