

*Алимов У.Б.  
старший преподаватель  
ТУИТ, Каршинский филиал*

## **ЗВУКОВЫЕ СИГНАЛЫ С ЦИФРОВЫМ КОДИРОВАНИЕМ И КОМПРЕССИОННОЙ МАСКОЙ**

*Аннотация: В некоторых случаях один звук может быть скрыт за другим звуком. Например, разговаривать рядом с железнодорожными путями может быть совершенно невозможно, если мимо проезжает поезд. Такой эффект называется маскировкой. Говорят, что слабый звук маскируется, если он становится неразличимым в присутствии более громкого звука.*

*Ключевые слова: цифровое кодирование, маскирование, компресс, сжатие звука, звуковое сопровождение, сжатие аудиоданных, частотное маскирование, временное маскирование, декодер форматирует звуки*

*Alimov U.B.  
Senior lecturer  
TUIT, Karshi branch*

## **AUDIO SIGNALS WITH DIGITAL ENCODING AND COMPRESSION MASK**

*Abstract: In some cases, one sound may be hidden behind another sound. For example, talking next to railway tracks may be completely impossible if a train is passing by. This effect is called masking. It is said that a weak sound is masked if it becomes indistinguishable in the presence of a louder sound.*

*Keywords: digital encoding, masking, compression, sound compression, soundtrack, audio data compression, frequency masking, time masking, decoder formats sounds*

Одновременная маскировка - любые два звука при одновременном прослушивании влияют на восприятие относительной громкости между ними. Более громкий звук снижает восприятие более слабого, вплоть до исчезновения его слышимости. Чем ближе частота маскируемого звука к частоте маскирующего устройства, тем больше он будет скрываться. Эффект маскирования будет иным, если частота маскируемого звука будет смещена ниже или выше по сравнению с частотой маскирующего устройства. Низкочастотный звук маскирует высокочастотный. Важно отметить, что высокочастотные звуки не могут маскировать низкочастотные.

Временная маскировка - это явление аналогично частотной маскировке, но здесь происходит маскировка во времени. Когда маскирующий звук прекращается, замаскированный звук продолжает оставаться неслышимым в течение некоторого времени. В обычных условиях эффект временной маскировки длится гораздо меньше. Время маскирования зависит от частоты и амплитуды сигнала и может достигать 100 мс.

В случае, когда маскирующий сигнал появляется позже, чем замаскированный, эффект называется постмаскированием. Когда маскирующий тон появляется раньше, чем замаскированный (такой случай также возможен), эффект называется предварительной маскировкой.

Постстимульная усталость - часто после воздействия громких звуков высокой интенсивности слуховая чувствительность человека резко снижается. Восстановление до нормальных пороговых значений может длиться до 16 часов. Этот процесс называется "временным изменением

порога слуховой чувствительности” или “постстимульной усталостью”. Сдвиг порога начинает проявляться, когда уровень звукового давления превышает 75 дБ, и соответственно увеличивается с повышением уровня сигнала. Кроме того, наибольшее влияние на сдвиг порога чувствительности оказывают высокочастотные составляющие сигнала.

Сжатие аудиоданных с потерями основано на несовершенстве человеческого слуха в восприятии звуковой информации. Неспособность человека в определенных случаях различать тихие звуки при наличии более громких, называемая эффектом маскировки, была использована в алгоритмах для уменьшения психоакустической избыточности. Эффекты слуховой маскировки зависят от спектральных и временных характеристик маскируемого и маскировочных сигналов и могут быть разделены на две основные группы:

Частотная (одновременная) маскировка

Временная (неодновременная) маскировка

Эффект маскировки в частотной области обусловлен тем, что при наличии больших

амплитуд звука человеческое ухо нечувствительно к малым амплитудам близких частот. То есть, когда два сигнала одновременно находятся в ограниченной частотной области, более слабый сигнал становится неслышимым на фоне более сильного.

Маскировка во временной области характеризует динамические свойства слуха, показывая изменение во времени относительного порога слышимости (порога слышимости одного сигнала при наличии другого), когда замаскированный и маскировочный маскируемый сигналы не звучат одновременно. В этом случае необходимо различать явления постмаскировки (изменение порога слышимости после сигнала высокого уровня) и предварительной маскировки (изменение порога слышимости перед поступлением сигнала максимального уровня). Более слабый сигнал

становится неслышимым за 5-20 мс до включения маскирующего сигнала и становится слышимым через 50-200 мс после его включения.

Лучшим методом кодирования звука, учитывающим эффект маскировки, является полосовое кодирование. Суть его заключается в следующем. Группа отсчетов входного аудиосигнала, называемая кадром, подается на блок фильтрации, который делит сигнал на частотные поддиапазоны. Выходной сигнал каждого фильтра - это та часть входного сигнала, которая попадает в полосу пропускания этого фильтра. Далее, в каждом диапазоне, используя психоакустическую модель, анализируется спектральный состав сигнала и оценивается, какая часть сигнала должна передаваться без сокращений, а какая лежит ниже порога маскирования и может быть квантована на меньшее количество бит. Чтобы уменьшить максимальный динамический диапазон, определяется максимальная выборка в кадре и вычисляется коэффициент масштабирования, который приводит эту выборку к верхнему уровню квантования. Эта операция аналогична компандированию в аналоговом вещании. Все остальные выборки умножаются на тот же коэффициент.

Коэффициент масштабирования передается в декодер вместе с закодированными данными для корректировки коэффициента усиления последних. После масштабирования оценивается порог маскирования и общее количество битов перераспределяется между всеми диапазонами.

Очевидно, что после устранения психоакустической избыточности звуковых сигналов их точное восстановление при декодировании уже невозможно. Методы устранения психофизической избыточности могут обеспечить сжатие цифровых аудиоданных в 10-12 раз без существенной потери качества.

Многие другие приемы могут служить способом уменьшения объема аудиоинформационных данных. Даже простое сужение полосы частот сигнала вместе с уменьшением динамического диапазона уже можно

назвать сжатием аудиоданных. Например, стандарт сжатия звука сотовой связи использует и то, и другое. Стремясь устранить избыточность звука, кодек с низким качеством сигнала становится избирательным по отношению к определенным словам, упрямо проглатывая их.

Общепринятые методы сжатия данных, такие как RLE, статистические и словарные методы, могут использоваться для сжатия звуковых файлов без потерь, но результат сильно зависит от конкретных аудиоданных. Некоторые звуки хорошо сжимаются с помощью RLE, но плохо - с помощью статистических алгоритмов. Другие звуки больше подходят для статистического сжатия, но при использовании словарного подхода, наоборот, может произойти расширение. Вот краткое описание эффективности этих трех методов сжатия аудиофайлов.

RLE хорошо работает со звуками, которые содержат длинную серию повторяющихся звуковых фрагментов - сэмплов. При 8-битной дискретизации это может происходить довольно часто. Напомним, что разница в электрическом напряжении между двумя 8-битными сэмплами составляет около 4 мВ. Несколько секунд однородной музыки, в течение которых звуковая волна изменится менее чем на 4 мВ, сгенерируют последовательность из тысяч идентичных сэмплов. Очевидно, что при 16-битной дискретизации длинные повторы встречаются реже, и, следовательно, алгоритм RLE будет менее эффективным.

Статистические методы присваивают звуковым сэмплам коды переменной длины в соответствии с их частотой. При 8-битной дискретизации имеется всего 256 различных сэмплов, поэтому в большом звуковом файле сэмплированный звук может быть распределен равномерно. Такой файл не может быть хорошо сжат методом Хаффмана. При 16-битной дискретизации допускается более 65 000 аудиофрагментов. В этом случае возможно, что некоторые выборки будут встречаться чаще,

а другие реже. При сильной асимметрии вероятностей хороших результатов можно добиться с помощью арифметического кодирования.

Методы, основанные на словарном запасе, предполагают, что определенные фразы будут часто встречаться во всем файле. Это происходит в текстовом файле, в котором отдельные слова или их последовательности повторяются несколько раз. Однако звук является аналоговым сигналом, и значения конкретных генерируемых выборок сильно зависят от работы АЦП. Например, при 8-битной выборке волна 8 мВ становится числовой выборкой, равной 2, но волна, близкая к ней, скажем, на 7,6 мВ или 8,5 мВ, может стать другим числом. По этой причине фрагменты речи, содержащие совпадающие фразы и звучащие для нас одинаково, могут немного отличаться при оцифровке. Тогда они попадут в словарь в виде разных фраз, что не даст ожидаемого сжатия. Таким образом, словарные методы не очень подходят для сжатия звука.

#### **Использованные источники:**

1. Ковалгин Ю.А., Вологдин Е.И. Цифровое кодирование аудиосигналов. - Санкт-Петербург: КРАУН-принт, 2004.
2. В.П. Дьяконов. Вейвлеты. От теории к практике. - М.: СОЛОН - Р, 2002. - 448 с.
3. Зайнидинов Х.Н., Зайнутдинова М.Б., Назирова Е.Sh., Юсупов И. Двумерные вейвлет-базисы с компактными несущими в задаче отсчетов сигналов - функций двух переменных. МАТЕРИАЛЫ Научно- практической и духовно-просветительской конференции, посвященной 1235-летию Мухаммада ай - Хорезми, Международной конференции "Значение информационно-коммуникационных технологий в инновационном развитии отраслей экономики", Ташкент, Узбекистан, 5 - 6 апреля 2018 г. стр.834 – 836.

4. Цифровая обработка двумерных сигналов на основе вейвлетов Хаара - Компьютерные системы и технологии. 19-я международная конференция CompSys Tech`18 (индексируется SCOPUS), Русе, Болгария, 13-14 сентября 2018 г., Материалы, стр. 130-133 - Х. Зайнидинов, М. Зайнутдинова, Э. Назирова. Неправильно в списке литературы ставить только URL ресурс, необходимо добавить название статьи или заголовок страницы – это требование ГОСТа.