

Мелешко М.А., канд. техн. наук
 Розорінов Г.М., д-р техн. наук
 Седхесамеддін Дабірсіагі,
 Денисенко С.М.

ЗМЕНШЕННЯ ШВИДКОСТІ ЦИФРОВОГО ПОТОКУ В КОМП'ЮТЕРНИХ СИСТЕМАХ ОБРОБКИ ЗВУКОВИХ СИГНАЛІВ

**Інститут комп'ютерних технологій
 Національного авіаційного університету**

Розглянуті можливості зменшення швидкості цифрового потоку в комп'ютерних системах обробки звукових сигналів шляхом скорочення статистичної та психоакустичної надлишковості з урахуванням психоакустичних закономірностей слухового сприйняття.

Вступ

Важливою проблемою при цифровому представленні звукових сигналів є скорочення наявних у них статистичної та психоакустичної надлишковості [1]. Скорочення статистичної надлишковості ґрунтується на врахуванні властивостей самих звукових сигналів. Вона зумовлена наявністю кореляційного зв'язку між сусідніми відліками звукового сигналу при його дискретизації. Усунення статистичної надлишковості не призводить до значного зменшення швидкості цифрового потоку.

Постановка задачі

Найбільш перспективними, з цієї точки зору, виявились методи скорочення психоакустичної надлишковості звукових сигналів, що враховують властивості слухового сприйняття. Ступінь врахування психоакустичних закономірностей слухового сприйняття визначає якість систем кодування із стисненням цифрових даних. Методами усунення психоакустичної надлишковості можна забезпечити стиснення цифрових аудіоданих в 10...12 раз без суттєвих втрат у якості [2]. Будь-яка слухова система описується частотно-пороговими кривими, які відображають граничні значення інтенсивності (або потужності) звукового тиску, у залежності від частоти звукових коливань. Частотно-порогові криві, отримані для слухової си-

стеми в цілому, отримали назву кривих чутності [3].

Основна частина

Чутливість слуху оцінюється мінімальною інтенсивністю звуку, при якій людина може відрізнити звук від постійно існуючого фону власних шумів. Інтенсивність, при якій звук виявляється з вірогідністю 0,5, називається порогом чутності або абсолютним порогом для даного сигналу [3]. Крива порогу чутності (рис. 1), що характеризує найменшу інтенсивність звуку певної частоти, котрий може бути почутий людиною в тиші, добре апроксимується за допомогою нелінійної функції:

$$T(f) = 3,64(f/1000)^{-0,8} - 6,5e^{-0,6(f/1000-3,3)^2} + 10^{-3}(f/1000)^4, \quad (1)$$

де $T(f)$ – порог чутності, дБ; f – частота звукового сигналу, Гц. Дана крива являє собою геометричне місце точок, котрі відповідають тональним компонентам різних частот, що мають таку ж гучність.

Аналіз усередненого графіку порога чутності показує, що слухова система людини чутлива до частот в діапазоні 0,01...20 кГц. Найбільша чутливість у людини спостерігається на середніх частотах 2...5 кГц. Поріг чутності сигналу на частоті біля 3 кГц складає приблизно 0 дБ. Поріг чутності значно підвищується в області низьких (менше 200 Гц) і високих (більше 10 кГц) частот і може досягати

значення гучності звуку 40 дБ і вище. Відповідно, для слабких сигналів амплітудно-частотна характеристика слуху є нелінійною, і, як наслідок, слабкі низько-частотні та високочастотні складові сигналу можуть бути нечутні. Криві чутності дозволяють оцінити працездатність слухової системи за рядом ознак: діапазону частот, що сприймаються, смузі частот, в межах якої чутливість максимальна, абсолютним значенням звукового тиску та їх змінами в залежності від частоти.

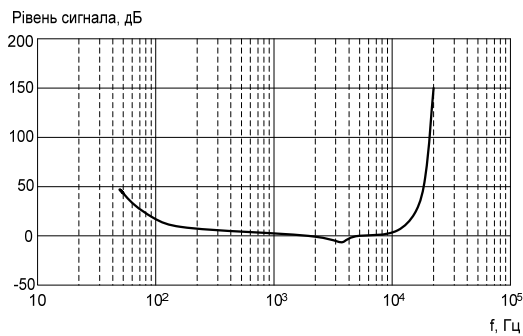


Рис. 1. Крива мінімального порогу чутності слухової системи людини в тиші

Одним із важливих властивостей слухової системи людини є її інтегруюча здатність, тобто групування частотних складових звуку у певні частотні критичні смуги. Слухова система порівнює корисний сигнал і шум, що заважає, по інтенсивності в межах критичних смуг слуху, оцінюючи поріг чутності. Критична смуга слуху є смугою, за межами якої суб'єктивні відчуття звуку сильно змінюються. На рис. 2, 3 схематично показано метод визначення ширини критичної смуги слухової системи.

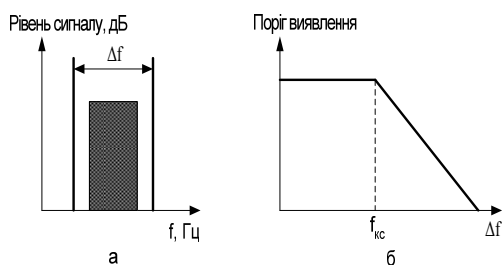


Рис. 2. Визначення ширини критичної смуги при маскуванні вузькосмугового шумового компоненту двома тональними сигналами

Вузькосмуговий шумовий сигнал маскують двома близько розташованими тональними компонентами, відстань Δf між якими поступово збільшують (рис. 2, а). На рис. 2, б представлена залежність зміни порогу виявлення шумової компоненти від відстані між двома тональними компонентами Δf . Даний поріг лишається постійним доти, доки частотна відстань між тональними компонентами лишається в межах критичної смузи $f_{кc}$. Поза смугою поріг швидко зменшується. На рис. 3, а наведено випадок, коли в якості сигналу, що маскується, використовується тональна компонента, а в якості маскуючих сигналів – дві вузькосмугові шумові компоненти. Зміна порогу виявлення для тональної компоненти (рис. 3, б) аналогічна рис. 2, б.

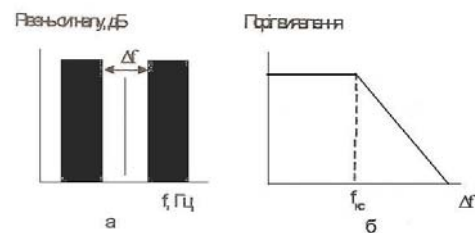


Рис. 3. Визначення ширини критичної смуги при маскуванні тонального компонента двома вузькосмуговими шумовими сигналами

В таблиці наведено розбиття діапазону частот, що чути, на критичні смуги з відповідними центральними частотами і шириною смуг ідеалізованого банку фільтрів (рис. 4) з прямокутними амплітудно-частотними характеристиками, який дозволяє якісно відтворити звукові сигнали.

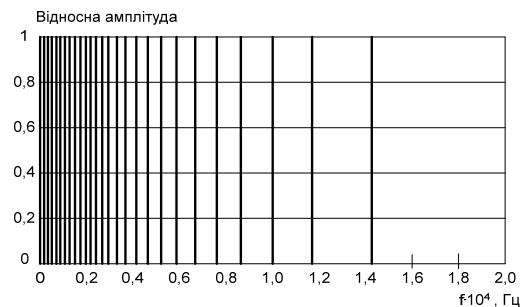


Рис. 4. Ідеалізований банк фільтрів як спрощена модель слухової системи

Таблиця 1. Розбиття частотного діапазону, що чути, на критичні смуги

Номер критичної смуги	Центральна частота, Гц	Нижня і верхня частоти, Гц	Ширина критичної смуги, Гц
1	50	20 - 100	100
2	150	100 - 200	100
3	250	200 - 300	100
4	350	300 - 400	100
5	450	400 - 510	110
6	570	510 - 630	120
7	700	630 - 770	140
8	840	770 - 920	150
9	1000	920 - 1080	160
10	1170	1080 - 1270	190
11	1370	1270 - 1480	210
12	1600	1480 - 1720	240
13	1850	1720 - 2000	280
14	2150	2000 - 2320	320
15	2500	2320 - 2700	380
16	2900	2700 - 3150	450
17	3400	3150 - 3700	550
18	4000	3700 - 4400	700
19	4800	4400 - 5300	900
20	5800	5300 - 6400	1100
21	7000	6400 - 7700	1300
22	8500	7700 - 9500	1800
23	10500	9500 - 12000	2500
24	13500	12000 - 15500	3500
25	19500	15500 - 22500	7000

З таблиці видно, що критична смуга не відповідає діапазону з фіксованими нижньою і верхньою границями. Ширина критичної смуги лишається приблизно постійною (біля 100 Гц) аж до 500 Гц і збільшується приблизно на 20% від центральної частоти при 500 Гц і вище:

$$\Delta f(f) = \begin{cases} 100 \text{ \AA}, & f < 500 \text{ \AA}, \\ 0,2 f \text{ \AA}, & f > 500 \text{ \AA}. \end{cases} \quad (2)$$

З виразу (2) видно, що в частотній області до 500 Гц органи слуху сприймають інтенсивність звуку, розділяючи її на ділянки постійної абсолютної ширини, а вище 500 Гц – на ділянки постійної відносної ширини $\Delta f(f)$.

Для середнього слухача ширина критичної смуги слухових фільтрів (рис. 5) може бути визначена у вигляді

$$\Delta f(f) = 25 + 75 \left[1 + 1,4(f/100)^2 \right]^{0,69}. \quad (3)$$

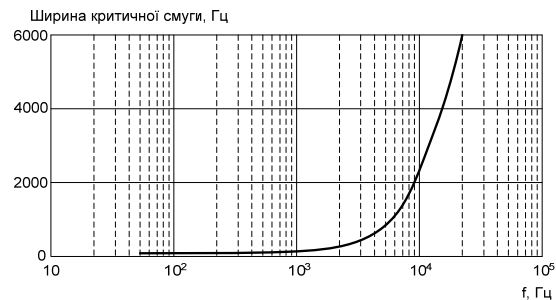


Рис. 5. Ширина критичних смуг як функція центральної частоти

З рис. 5 видно, що по мірі збільшення центральної частоти, критична смуга розширюється. Оскільки має місце однозначна залежність між частотою синусоїдального коливання і висотою тону, то по аналогії з частотною групою слуху, було введено поняття тональної групи слуху. Ширина тональної групи не залежить від висоти тону і складає 100 мел (один мел відповідає висоті звуку, що відчувається, частотою 1000 Гц при рівні інтенсивності 40 дБ). Тому виявилось зручним ввести

одиницю висоти тону в одну тональну групу слуху. В честь німецького фізика Г. Баркгаузена вона носить назву барк (1барк=100 мел), тобто рівна ширині критичної смуги. Перехід від однієї смуги до іншої відповідає зміні висоти в 1 барк. Функція переходу від частоти у герцах до частоти у барках має вигляд (рис. 6):

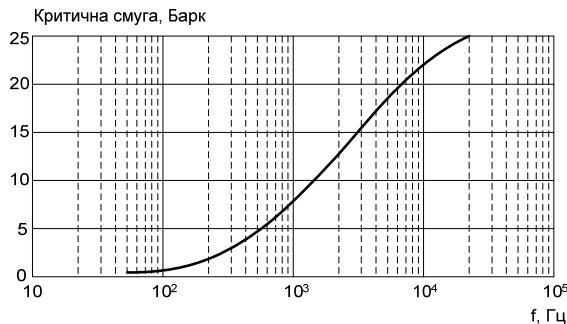


Рис. 6. Відстань між центральними частотами критичних смуг у барках

$$z(f) = 13 \arctg(0,00076 f) + 3,5 \arctg \left[\left(\frac{f}{7500} \right) \right]^2, \quad (4)$$

де $z(f)$ являє собою частоту f у барках.

В основі способів виділення у звуковому сигналі частини інформації, що чується, і видалення із сигналу всіх інших нечутних деталей, лежить ефект маскування. Він пов'язаний з процесом взаємодії сигналів, що призводять до зміни слухової чуттєвості до сигналу, що маскується, у присутності маскуючого. Ефекти слухового маскування залежать від спектральних та часових характеристик сигналу, що маскується, і сигналу маскування, та можуть бути розділені на дві основні групи: частотне (одночасне) і часове (неодночасне) маскування. Маскування за частотою полягає у тому, що, якщо два сигнали одночасно знаходяться в обмеженій частотній області, то більш слабкий сигнал стає нечутним на фоні більш сильного. Маскування за часом ґрунтується на тому ефекті, що більш слабкий сигнал стає нечутним за 5...20 мс до включення сигналу маскування і стає чутним через 50...200 мс після його виключення.

Маскування безпосередньо пов'язане з нелінійністю слуху. Вона ви-

являється в тому, що при дії на барабанну перетинку достатньо гучного синусоїдального звуку з частотою f , у слуховому апараті формуються гармоніки цього звуку з частотами $2f, 3f$ і т.д. Оскільки у первинному діючому тоні цих гармонік немає, вони отримали назву суб'єктивних гармонік.

Аналіз кривих маскування (кривих порогу чутності при наявності маскування) показує, що маскування сигналів, які мають частоту, котра лежить нижче частоти заважаючого тону, проявляється значно слабше, ніж частот, що лежать вище. Це дає основу припустити, що маскуючі дії зумовлені виникненням суб'єктивних гармонік тону, що заважає.

Якщо в якості звуку, що заважає, використовувати вузькосмуговий білий шум, пороги чутності визначаються більш чітко, а самі криві виходять більш симетричними. Це пов'язано з відсутністю гармонічного биття. Великий практичний інтерес являють криві порогу чутності, отримані при масуванні широкосмуговими шумами.

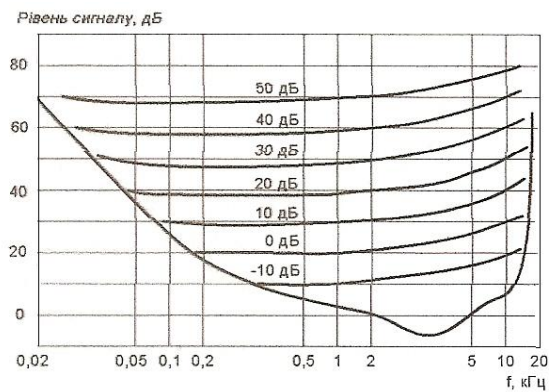


Рис. 7. Криві порогу чутності при масуванні білим шумом

На рис. 7 наведені криві порогів чутності синусоїдальних сигналів, отримані при масуванні білим шумом із різною спектральною щільністю. Величини рівней спектральної щільності маскуючого білого шуму представлені на рис. 7 над відповідними кривими. З рис. 7 видно, що на низьких частотах криві практично не залежать від частоти (приблизно до 500 Гц). При подальшому збільшенні частоти рівень порогу чутності підвищується на 3

дБ при кожному подвоєнні частоти, що приблизно відповідає залежності розширення ширини критичних смуг із збільшенням центральної частоти (формула (2)).

Із цього слідує, що білий шум неоднаково ефективний для маскування різних частот. Властивість широкосмугових сигналів чинити максимальний вплив на маскування сигналу у межах критичних смуг, покладено в основу психоакустичних алгоритмів стиснення сигналів.

Ефект маскування у частотній області пов'язаний з тим, що в присутності великих звукових амплітуд людське вухо нечутливе до малих амплітуд близьких частот (рис. 8).

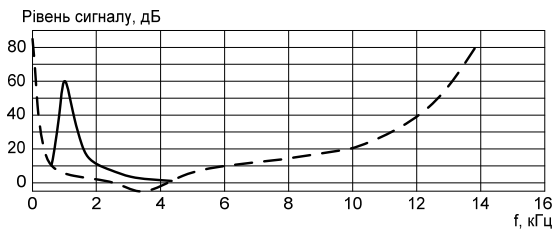


Рис. 8. Зміна порогу чутливості слухової системи при масуванні тональним сигналом

Таким чином, для маскування важливіше не абсолютна гучність звуків, а відношення потужності гучного сигналу до потужності тихого. Крім того, чим ближче частота до частоти маскуючого сигналу, тим сильніший ефект маскування (рис. 8). Ступінь маскування – це різниця у децибелах між рівнем порогу чутності тону, що маскується, у присутності маскуючого сигналу, і його рівнем порогу чутності у тиші. На рис. 8 показані залежність абсолютного порогу чутливості вуха і характер його зміни у присутності тону з частотою 1 кГц та амплітудою 60 дБ. З рис. 8 видно, що поріг чутності змінюється: він підвищується при появі будь-яких звуків й знижується у тиші. Для того, щоб почути більш слабкі звуки на фоні більш сильних, їх частоти повинні значно розрізнятися. Крива порогу чутливості багато в чому подібна функції, що змінюється у часі, короткочасній спектральній щільності потужності та має локальні ма-

ксимуми у відповідності з тональними компонентами високого рівня. З рис. 9 видно, що із збільшенням середньої частоти маскуючого сигналу, діапазон частот, де проявляється маскування, стає ширшим відповідно (2). Криві маскування несиметричні, вони мають крутий спад у сторону низьких частот та пологий спад у сторону високих частот.

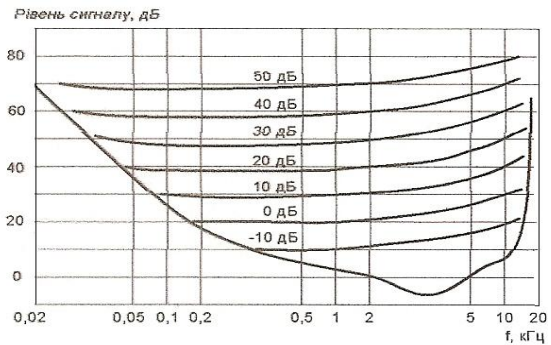


Рис. 9. Зміна порогу чутливості слухової системи при масуванні сигналами із середніми частотами смуг 0,25; 1; 4; 8 кГц та рівнем інтенсивності 60 дБ

Якщо при побудові кривих маскування по осі абсцис відкладати не частоти f , а значення висоти тону z у барках, то криві маскування, при різних значеннях z маскуючого сигналу, будуть схожими по ширині відповідно (4), при одному й тому ж значенні рівня маскуючого сигналу (рис. 10). Їх форма при такому представленні буде залежати тільки від маскуючого сигналу і не буде залежати від величини z . Маскування має місце не тільки усередині критичної смуги, воно розповсюджується й на сусідні смуги. Маскуючий сигнал, центрований у межах критичної смуги, чинить певний вплив на поріг виявлення у сусідніх критичних смугах. Функція протяжності маскування, яка враховує маскування як усередині, так і між смугами, задається аналітичним виразом

$$SF(x) = 15,81 + 7,5(x + 0,474) - 17,5\sqrt{1 + (x + 0,474)^2}, \quad (5)$$

де x і $SF(x)$ представлені відповідно у барках і дБ.

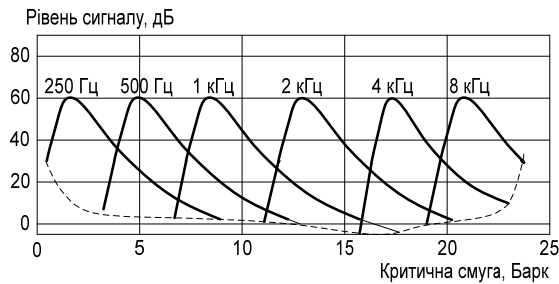


Рис. 10. Частотне маскування із частотною шкалою у барках

Внаслідок розділення спектру звукового сигналу на частотні смуги, слух реагує не на загальну потужність сигналу, а на потужність, зосереджену в окремих частотних групах. При цьому, більш інтенсивні частотні групи, при певних умовах, можуть маскувати менш інтенсивні групи. Щодо стиснення, це значить, що у групах меншої інтенсивності допустимий більший рівень шуму квантування, тобто допустиме зменшення числа розрядів при кодуванні відповідних спектральних складових, що, в свою чергу, зменшує швидкість цифрового потоку, що передається.

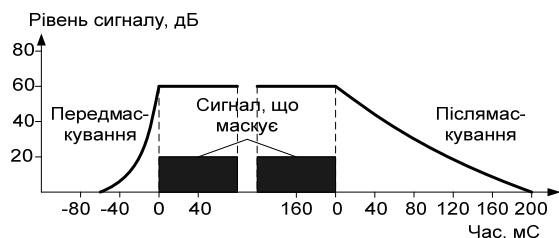


Рис. 11. Залежність зміни порогу чутності сигналу, що маскується, від часового інтервалу між маскуючим сигналом і тим, що маскується

Маскування у часовій області (рис. 11) характеризує динамічні властивості слуху, показуючи зміну у часі відносно порогу чутності (порог чутності одного сигналу у присутності іншого), коли маскуючий сигнал і той, що маскується, звучать не одночасно.

Висновки

Розрізняють передмаскування (сигнал, що маскується, перед маскуючим) та післямаскування (сигнал, що маскується, після маскуючого). Обидва види маскування відбуваються, якщо маскуючий сигнал і той, що маскується, одночасно зна-

ходяться у часовому вікні шириною 100...200 мс. Припускається, що за цей час сигнал цілком обробляється з виділенням усіх його ознак. У цьому випадку слухову систему можна розглядати як процесор для обробки звукових сигналів у часовому вікні шириною приблизно 200 мс. З рис. 11 видно, що тривалість типових проміжків часу, у межах яких діє передмаскування й післямаскування, складає відповідно 5...50 мс і 50...200 мс. Передмаскування проявляється на значно більш короткому часовому інтервалі, причому його тривалість великою мірою залежить від досвіду слухача. Ступінь маскування звукового сигналу залежить від інтервалу між маскуючим сигналом і тим, що маскується, рівня інтенсивності та тривалості впливу маскуючого сигналу. Зближення у часі маскуючого сигналу з тим, що маскується, посилює ефект маскування. Ступінь часового маскування залежить від співвідношення частот маскуючого сигналу і того, що маскується, точно так, як і при частотному маскуванні. Воно проявляється ще сильніше, якщо маскуючий сигнал і той, що маскується, близькі за частотою.

Список літератури

1. Розоринів Г.Н. Кодирование звука в XXI веке // Арсенал XXI века, №1, 2000. – С. 56–60.
2. Теория прикладного кодирования: Учеб. пособие. В 2 т. Т.1 / В.К. Конопелько, В.А. Липницкий, и др.; Под ред. проф. В.К. Конопелько. – Минск: БГУИР, 2004. – 285 с.
3. Справочник по акустике / В.К. Иофе, В.Г.Корольков, М.А. Сапожков: Под ред. М.А. Сапожкова. – М.: Связь, 1979. – 312 с.