

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації
(повна назва)

Кафедра медіаінженерії та інформаційних радіоелектронних систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти другий (магістерський)

Методи вирівнювання гучності у телевізійних трансляціях.
(тема)

Виконав:
студент 2 курсу, групи МІМ-22-2
Рибас К.С.
(прізвище, ініціали)

Спеціальність 172 Телекомунікації та
радіотехніка
(код і повна назва спеціальності)

Тип програми освітньо-професійна

Освітня програма Медіаінженерія
(повна назва освітньої програми)

Керівник доц. Шаповалов С.В.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри Володимир КАРТАШОВ
(підпис)

2023 р.

Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації

Кафедра медіаінженерії та інформаційних радіоелектронних систем

Рівень вищої освіти другий (магістерський)

Спеціальність 172 Телекомунікації та радіотехніка
(код і повна назва)

Тип програми освітньо-професійна

Освітня програма Медіаінженерія
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

«_____» _____ 20__ р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту Рибасу Костянтину Сергійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Методи вирівнювання гучності у телевізійних трансляціях.

затверджена наказом по університету від " 20 " 10 2023 р. № 1224 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 10.01.2024 р.

3. Вихідні дані до роботи Розглянути відмінність концепцій нормалізації піків і нормалізації гучності. Оцінити переваги і недоліки нормалізації на приймальній і передавальній стороні. Скласти алгоритм нормалізації контенту із різних джерел. Запропонувати варіанти переходу працівників медіа виробництва на новий спосіб нормалізації по гучності. Скласти алгоритм процесу нормалізації гучності у медіавиробництві. Розглянути найбільш вірогідні сценарії невідповідності нормам трьох ключових параметрів: гучності, діапазону та піків. Скласти алгоритм дій звукооператора для кожного із зазначених сценаріїв. Підсумувати принципи нормалізації гучності для програм різних жанрів.

4. Перелік питань, що потрібно опрацювати в роботі _____

Вступ

1 Аналіз задачі вимірювання і нормалізації гучності в телевізійній і радіотрансляції.

2 Методи нормалізації гучності звуку телевізійних і радіотрансляцій.

3 Практичні рекомендації з вирівнювання гучності телевізійного контенту.

Висновки

Перелік посилань

Додатки

5. Перелік графічного матеріалу із зазначенням обов'язкових креслеників, схем, плакатів, комп'ютерних ілюстрацій

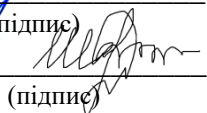
1. Проблемна ситуація (1 аркуш А4).
2. Постановка задачі (1 аркуш А4).
3. Класичні індикатори рівня (1 аркуш А4).
4. Динамічні процесори (1 аркуш А4).
5. Нормалізація гучності (1 аркуш А4).
6. Одиниці гучності LUFS (1 аркуш А4).
7. Способи нормалізації гучності (1 аркуш А4).
8. Нормалізація контенту з різних джерел (1 аркуш А4).
9. Перехід звукорежисера на LUFS (1 аркуш А4).
10. Експеримент з нормалізацією (1 аркуш А4).
11. Діапазон гучності (Loudness Range) (1 аркуш А4).
12. Алгоритм нормалізації (1 аркуш А4).
13. Сценарії невідповідності нормам (1 аркуш А4).
14. Нормалізація музичних програм (1 аркуш А4).
15. Нормалізація реклами (1 аркуш А4).
16. Нормалізація шоу та спорту (1 аркуш А4).
17. Нормалізація кінофільмів (1 аркуш А4).
18. Висновки (1 аркуш А4).

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналітичний огляд літератури	01.09.23–27.09.23	
2	Теоретичний аналіз методів шумоподавлення	28.09.23–11.10.23	
3	Підготовка аудіофрагментів	12.10.23–10.11.23	
4	Експериментальна частина	11.11.23–03.12.23	
5	Обробка результатів	04.12.23–17.12.23	
6	Графічна частина роботи	18.12.23–17.12.23	
7	Перевірка керівником	18.12.23–30.12.23	
8	Перевірка на академічний плагіат	02.01.24–05.01.24	
9	Перевірка завідувачем кафедри, рецензування	06.01.24–09.01.24	

Дата видачі завдання _____ 20.10.2023 р. _____

Студент _____  _____ Костянтин РИБАС _____

Керівник роботи _____  _____ Сергій ШАПОВАЛОВ _____

РЕФЕРАТ

Пояснювальна записка до кваліфікаційної роботи: 68 сторінок, 39 рисунків, 30 джерел.

ГУЧНІСТЬ, ДИНАМІЧНА ОБРОБКА, ДІАПАЗОН ГУЧНОСТЕЙ, ЗВУКОВИЙ ТИСК, КОМПРЕССОР, ЛІМІТТЕР, ПІДСИЛЕННЯ, РІВЕНЬ СИГНАЛУ

Метою кваліфікаційної роботи є аналіз методів вимірювання і нормалізації гучності телевізійного або звукового мовлення з метою формулювання зрозумілих рекомендацій для робітників медіаіндустрії, студентів та випускників відповідних навчальних закладів.

Показано відмінність концепцій нормалізації піків і нормалізації гучності. Розглянуті переваги і недоліки нормалізації на приймальній і передавальній стороні. Складено алгоритм нормалізації контенту із різних джерел. Запропоновано варіанти переходу працівників медіа виробництва на новий спосіб нормалізації по гучності. Складено алгоритм процесу нормалізації гучності у медіавиробництві. Розглянуті найбільш вірогідні сценарії невідповідності нормам трьох ключових параметрів: гучності, діапазону та піків. Складено алгоритм дій звукооператора для кожного із зазначених сценаріїв. Підсумовані принципи нормалізації гучності для програм різних жанрів.

ABSTRACT

Explanatory note to the qualification work: 68 pages, 39 figures, 30 sources.

VOLUME, DYNAMIC PROCESSING, VOLUME RANGE, SOUND PRESSURE, COMPRESSOR, LIMITER, GAIN, SIGNAL LEVEL

The purpose of the qualification work is to analyze the methods of measuring and normalizing the volume of television or sound broadcasting in order to formulate clear recommendations for workers in the media industry, students and graduates of relevant educational institutions.

The difference between the concepts of peak normalization and loudness normalization is shown. The advantages and disadvantages of normalization on the receiving and transmitting side are considered. An algorithm for content normalization from various sources has been developed. Variants of transition of media production workers to a new method of volume normalization are offered. An algorithm for the volume normalization process in media production has been developed. The most likely scenarios of non-compliance with the norms of three key parameters are considered: volume, range and peaks. The sound operator's action algorithm for each of the specified scenarios has been compiled. Summarized principles of volume normalization for programs of various genres.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів.....	8
Вступ.....	9
1 АНАЛІЗ ЗАДАЧІ ВИМІРЮВАННЯ І НОРМАЛІЗАЦІЇ ГУЧНОСТІ В ТЕЛЕВІЗІЙНІЙ І РАДІОТРАНСЛЯЦІЇ.....	11
1.1 Поняття рівня звуку і гучності.....	11
1.2 Методи вимірювання рівня звукових сигналів.....	14
1.3 Методи регулювання рівня звукового сигналу.....	19
1.3.1 Нормалізація звуку.....	19
1.3.2 Динамічна компресія.....	20
1.3.3 Обмежувач (ліміттер).....	25
1.4 Висновки по розділу 1.....	25
2 МЕТОДИ НОРМАЛІЗАЦІЇ ГУЧНОСТІ ЗВУКУ ТЕЛЕВІЗІЙНИХ І РАДІОТРАНСЛЯЦІЙ.....	28
2.1 Загальна концепція нормалізації гучності.....	28
2.1.1 Відмінність нормалізації гучності від пікової нормалізації.....	28
2.1.2 Вимірювання гучності програми в LUFS.....	29
2.1.3 Порівняння нормалізації по сигналу і нормалізації по метаданим.....	31
2.1.4 Цільовий рівень LUFS.....	31
2.1.5 Особливості використання процесорів гучності.....	34
2.2 Стратегії встановлення рівня гучності.....	36
2.2.1 Базовий підхід до зведення.....	36
2.2.2 Вимірювання гучності для виробництва та пост виробництва.....	38
2.2.3 Діапазон гучності (Loudness Range).....	40
2.3 Висновки по розділу 2.....	42

3 ПРАКТИЧНІ РЕКОМЕНДАЦІЇ З ВИРІВНЮВАННЯ ГУЧНОСТІ ТЕЛЕВІЗІЙНОГО КОНТЕНТУ.....	45
3.1 Мовлення на базі файлів.....	45
3.2 Обробка програм різних жанрів.....	50
3.2.1 Рекламні вставки (кліпи) та анонси.....	52
3.2.2 Музика.....	53
3.2.3 Спортивні програми.....	54
3.2.4 Шоу.....	55
3.2.5 Кінофільми.....	56
3.2.6 Використання анкерного сигналу.....	57
3.3 Висновок по розділу 3.....	59
Висновки.....	61
Перелік джерел посилання.....	66
ДОДАТКИ.....	69
Додаток А. Графічний матеріал.....	70
Додаток Б. Відомість кваліфікаційної роботи.....	88

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,
СКОРОЧЕНЬ І ТЕРМІНІВ

- АЦП – аналого-цифрове перетворення;
- РЗТ – рівень звукового тиску;
- DAW – цифрова звукова робоча станція (Digital Audioworkstation);
- DRC – динамічна компресія (Dynamic range compression);
- EBU – європейський мовний союз (European Broadcasting Union);
- FS – повна шкала АЦП (Full Scale);
- ISO – міжнародна організація зі стандартизації (International Organization for Standardization);
- LRA – Loudness Range (діапазон гучності);
- LU – Loudness Units (одиниці вимірювання гучності);
- LUFS – Loudness Units Full Scale (одиниці вимірювання гучності);
- MTPL – максимальний рівень реальних піків (Maximum True Peak Level);
- MSTLL – максимальний рівень короткочасної гучності (Maximum Short-term Loudness Level);
- PEAQ – Perceptual Evaluation of Audio Quality (перцептуальна оцінка якості аудіо);
- PML – Permitted Maximum Level (максимально допустимий рівень);
- PPM – Peak Program Meter (вимірювач пікових значень);
- PWL – рівень потужності (Power Level);
- QPPM – квазіпіковий вимірювач;
- RMS – Root Mean Square (середньоквадратичне);
- SPL – рівень звукового тиску (Sound Pressure Level);
- SPPM – вибіркового пікового вимірювач;
- VU – Volume Unit (одиниця рівня).

ВСТУП

В даний час телевізійне мовлення є одним із найпопулярніших та найефективніших видів ЗМІ.

Зрозуміло, що зображення на екрані телевізора, само собою, передає телеглядачеві дуже великий обсяг інформації. Але загальне враження про телевізійну передачу формується не тільки на основі оцінки якості самої "картинки", поряд з цим величезну роль відіграє якість звукового супроводу. Найдосконаліше зображення залишить глядача байдужим і буде абсолютно знецінено, якщо воно супроводжується поганим та нерозбірливим звуком.

Стрибкоподібна зміна рівня звукового супроводу телевізійних програм створює суттєві психологічні навантаження на кінцевого споживача продукту – тобто на того, заради кого всі ці програми й створювали. Така ситуація виникає як для різних програм одного каналу (особливо анонсів та окремих рекламних вставок), так і при перемиканні з одного каналу на інший. Незважаючи на те, що проблема існує вже не один десяток років, лише останнім часом було запропоновано шляхи для її можливого вирішення.

Проблема нормалізації гучності проявляє себе двома різними ефектами. Перший, найбільш "дратівливий", це періодична зміна "гучності" звучання різних програм, що приймаються по одному каналу. Другий, теж неприємний – зміна рівня гучності звукового супроводу при перемиканні ефірного, супутникового або Інтернет ТВ-приймача з каналу на канал, що змушує щоразу підлаштовувати рівень гучності. Телеглядач при цьому опиняється у комічній у ролі "ді-джея", що постійно маніпулює ТВ-пультом.

Взагалі, термін "гучність звучання" є не зовсім коректним – правильніше було б говорити про "рівні звукового тиску", саме про те, що можна виміряти шумоміром.

Тут з'являється перша суперечність – споживач продукту (звукового супроводу телевізійного чи радіоканалу) – і той, хто його створює

(наприклад, ТВ-компанія) використовують різні методики для оцінки його якості (в даному випадку рівня звучання). Йдеться про суто психоакустичні аспекти сприйняття кінцевим користувачем.

В одному випадку при регулюванні рівня сигналу звукового супроводу "мовник" орієнтується на показання вимірювальних приладів, і керується стандартами, інструкціями та іншими нормативними документами, а в іншому – своїми суто фізичними відчуттями.

Виникає цілком природна проблема – наскільки показання вимірювальних приладів співвідноситься із суто фізичним сприйняттям гучності звучання.

В даний момент в сфері медіавиробництва не існує єдиного рішення для вимірювання і нормалізації гучності телевізійного або звукового мовлення. Є ряд розрізнених рекомендацій в різних сферах медіавиробництва, що при переході з одного способу передачі програми на інший можуть не відповідати між собою і давати спотворені результати.

Метою кваліфікаційної роботи є аналіз методів вимірювання і нормалізації гучності телевізійного або звукового мовлення з метою формулювання зрозумілих рекомендацій для робітників медіаіндустрії, студентів та випускників відповідних навчальних закладів.

1 АНАЛІЗ ЗАДАЧІ ВИМІРЮВАННЯ І НОРМАЛІЗАЦІЇ ГУЧНОСТІ В ТЕЛЕВІЗІЙНІЙ І РАДІОТРАНСЛЯЦІЇ

1.1 Поняття рівня звуку і гучності

Гучність звуку – це суб'єктивне відчуття звукового тиску (інтенсивності звуку), яке дозволяє розташовувати всі звуки за шкалою від тихих до гучних. Гучність звуку переважно залежить від інтенсивності звуку, але і від розподілу енергії звукових коливань за шкалою частот [1]. Також на гучність звуку впливають його локалізація у просторі, тривалість впливу, що маскує дію інших звуків та інші фактори [2, 3].

Гучність звуку складним чином залежить від звукового тиску, частоти та форми коливань. При постійній частоті та формі коливань гучність звуку зростає зі збільшенням звукового тиску. При однаковому звуковому тиску гучність синусоїдальних звуків різної частоти різна – однакову гучність різних частотах можуть мати звуки різної інтенсивності [4].

Для кількісної оцінки абсолютної гучності американським психологом Стенлі Стівенсом було запропоновано спеціальну сон. Гучність 1 сон – це гучність чистого тону (синусоїдального звуку) з частотою 1000 Гц і рівнем звукового тиску (РЗТ) 40 дБ [2] відносно опорного значення 20 мкПа [1].

Відносний рівень гучності прийнято оцінювати у логарифмічних одиницях – фонах. Рівень гучності чистого тону з частотою 1000 Гц у фонах чисельно дорівнює РЗТ у дБ [1].

На рис.1.1 зображено сімейство кривих рівної гучності, які називаються також ізофонами. Вони є стандартизованою (ISO 226:1987) залежністю РЗТ від частоти при заданому рівні гучності. Ізофона "0 фон", позначена пунктиром, характеризує поріг чутності звуків різної частоти для нормального слуху. Кожна крива поєднує чисті тони на різних частотах, однакові за гучністю для слухачів віком 18-20 років [1].

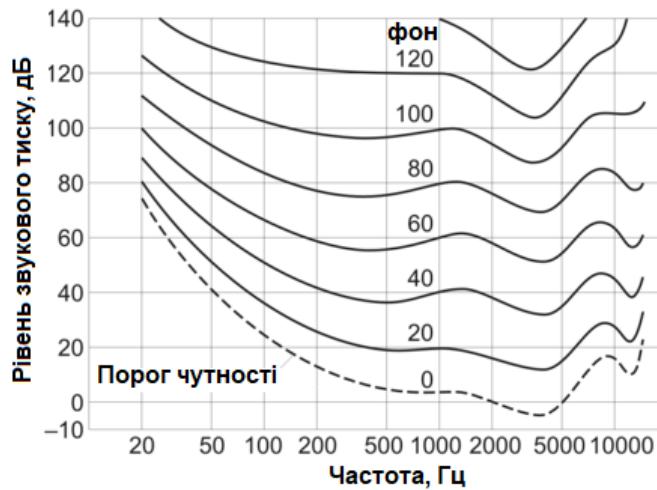


Рисунок 1.1 – Залежність рівня гучності від звукового тиску та частоти.
Криві рівної гучності Робінсона-Дадсона. ISO 226:1987

Уточнені стандартні криві рівної гучності у графічному та табличному вигляді представлені в (ISO 226:2003 Acoustics - Normal equal-loudness-level contours, IDT). Вони отримані за результатами 12 незалежних досліджень, проведених у Данії, Німеччині та Японії у 1983–2002 роках. Криві рівної гучності побудовані за усередненими відчуттями людей із нормальним слухом віком від 18 до 25 років включно (рис.1.2) [5].

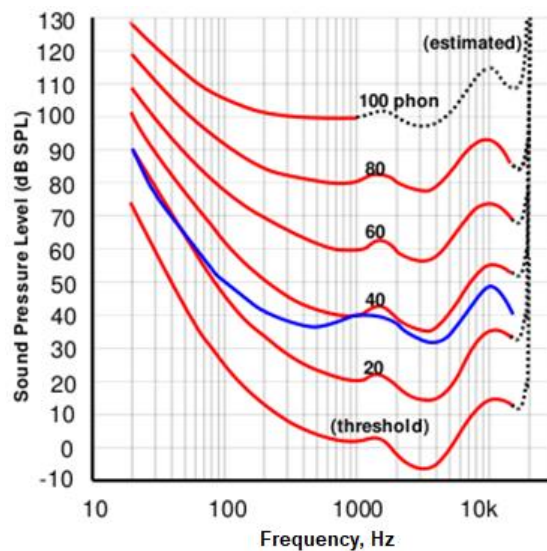


Рисунок 1.2 – Криві рівної гучності ISO 226:2003. Для порівняння показано синім кольором крива 40 фон ISO 226:1987

При малому РЗТ суб'єктивна оцінка рівня гучності залежить від частоти – слух менш чутливий до низьких і високих частот. При великому РЗТ низькі, середні та високі звуки оцінюються за рівнем гучності рівномірніше [2]. Характер кривих рівної гучності показує, що з нормальним слухом найбільш чутливі до звуків в діапазоні частот 2500–4000 Гц [5].

Також звернемо увагу на так званий діапазон комфортного прослуховування (рис.1.3) [5].

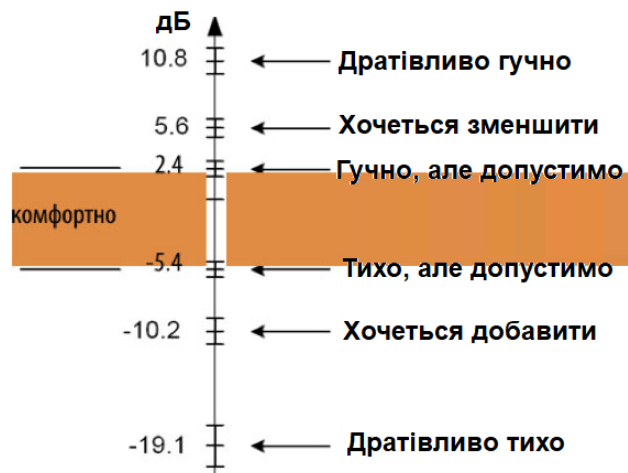


Рисунок 1.3 – Діапазон комфортного прослуховування

Помаранчева зона комфорту складає всього 8 дБ. Це той проміжок, в якому слухачу не хочеться кожні півхвилини хапатися за пульт для збільшення або зменшення гучності телевізора. У той же час значна частина необробленого звукового контенту має істотно більший динамічний діапазон, що часто перевищує 20 дБ. Ще вище розкид середнього рівня гучності різного звукового матеріалу.

Проблема приведення гучності в комфортний для прослуховування діапазон полягає в тому, що традиційне лімітування або компресування, які часто використовуються як інструменти регулювання гучності, досить примітивні і не враховують психоакустичних моделей сприйняття.

На практиці мовний сигнал формується з різних джерел звуку, що чергуються, які можуть бути дуже різними по динамічних діапазонах і середньому рівню гучності. Особливо це помітно при рекламних вставках,

гучність яких часто навмисно максимально завищена для кращого залучення телеглядачів уваги до змісту реклами. В результаті виходить, що після тихої сцени фільму під -24 дБ раптом вилітає реклама з рівнем нехай -10 дБ, наведена по піках, і слухач насилу встигає дотягтися до пульта.

1.2 Методи вимірювання рівня звукових сигналів

Як слідує із властивостей людського слуху, відчуття гучності є логарифмічною величиною. Тому вимірювання рівня звукових сигналів має виконуватися теж в логарифмічних одиницях. Найчастіше використовують децибели (дБ).

Вимірювання рівня в децибелах означає порівняння даного вимірюваного рівня з певним опорним «нульовим» рівнем, позначеним як 0 дБ. Таким чином, позначення «0 дБ» – це так званий «відносний нульовий» рівень, що вказує лише на те, що рівень даного сигналу точно дорівнює певному рівню, умовно прийнятому для цієї точки тракту як опорний, номінальний.

Якщо опорний рівень прийняти рівним напрузі $U_0 = 1$ В, то отримаємо одиниці дБВ (dBV) [6]

$$N = 20 \lg \frac{U}{U_0} = 20 \lg \frac{U}{1V}, [\text{дБВ}]. \quad (1.1)$$

Рівень, що перевищує опорний, позначається у дБ зі знаком «плюс» (наприклад, +3 дБ), а менший за опорний – у дБ зі знаком «мінус» (наприклад, -6 дБ).

У студіях мовлення та звукозапису прийнято використовувати величину дБи. У цьому випадку напруга, що вимірюється, порівнюється з так званим абсолютним нульовим рівнем. Ця напруга приймається рівною 0,775 В [6]:

$$N = 20 \lg \frac{U}{U_0} = 20 \lg \frac{U}{0,775 \text{ В}}, [\text{дБв}]. \quad (1.2)$$

Ця трохи незручна для запам'ятовування цифра $U_0 = 0,775 \text{ В}$ прийшла в студійну практику з тих областей звукотехніки, де для оцінки рівня сигналу важливіше вимірювати його напругу, а електричну потужність. І тоді можна було б оцінювати рівень в одиницях дБм, порівнюючи дану потужність з опорною, за яку приймається $P_0 = 1 \text{ мВт}$ на опорі $R = 600 \text{ Ом}$. Такий опір характерний, наприклад, для мідної телефонної лінії зв'язку. Тоді [6]

$$U_0 = \sqrt{P_0 R} = \sqrt{10^{-3} 600} = 0,775 \text{ В}. \quad (1.3)$$

У сфері акустичних вимірювань користуються тими самими децибелами, але вже позначення рівня звукового тиску (SPL). В цьому випадку за опорну величину 0 дБ SPL приймається звуковий тиск $p_0 = 2 \cdot 10^{-5} \text{ Па}$, тобто звук, що відповідає порогу слухового відчуття [6]:

$$N = 20 \lg \frac{P}{p_0} = 20 \lg \frac{P}{2 \cdot 10^{-5} \text{ Па}}, [\text{дБ SPL}]. \quad (1.4)$$

А рівень потужності акустичного сигналу, що відповідає цьому порогу, вимірюють в дБ PWL по відношенню до опорної величини, що дорівнює $I_0 = 10^{-12} \text{ Вт/м}^2$ [6]:

$$N = 10 \lg \frac{I}{I_0} = 10 \lg \frac{I}{10^{-12} \text{ Вт/м}^2}, [\text{дБ SPL}]. \quad (1.5)$$

Виділяють три типи вимірювачів рівня звуку.

Вимірювач середніх значень (VU-meter, або «волюметр» [7]).

Спочатку це був аналоговий стрілочний прилад, динамічні характеристики якого визначаються інерційними параметрами стрілочного індикатора (рис.1.3).

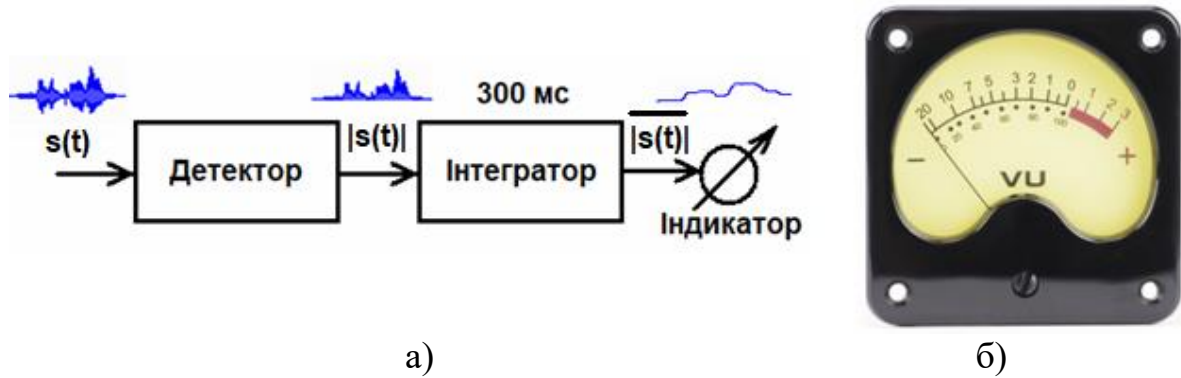


Рисунок 1.3 – Вимірювач середніх значень (VU-meter):

а – структурна схема, б – зовнішній вигляд

Був розроблений у 1939 році Bell Labs, CBS та NBC для вимірювання та стандартизації рівнів у телефонних лініях. Останнім часом такі вимірювачі часто робляться не зі стрілочними індикаторами, а зі світлодіодними чи іншими світловими показчиками [7]. Постійна часу вимірювальної схеми для цього типу вимірювачів рівня становить $\tau=300$ мс, що найбільш приблизно відображає суб'єктивно сприймається людиною гучність.

Вимірювач діючих значень (RMS, середньоквадратичний).

Показує величину напруги, пропорційну реальній довготривалій потужності сигналу $s(t)$, його «тепловий еквівалент». Кращі RMS-вимірювачі побудовані з використанням термоперетворювачів – напруга, що досліджується, нагріває термоелемент, за температурою якого і судять про величину напруги. У зв'язку із зайвою інерційністю застосовують для вимірювання рівня шумів.

У цифровому вигляді вимірювач RMS реалізується за виразом [8]

$$RMS = \frac{1}{N} \sqrt{\sum_{n=1}^N s^2(n)}, \quad (1.6)$$

де n – номер відліку цифрового сигналу;

N – довжина вікна усереднення.

Тривалість вікна усереднення в сучасних цифрових додатках становить одиниці секунд.

Піковий вимірювач (PPM).

Істинний піковий вимірювач (True PPM) – відображає точні пікові значення рівня незалежно від тривалості звукового сигналу [9].

Квазіпіковий вимірювач (QPPM) – показує пікові значення рівнів сигналу, що перевищують задану тривалість часу інтеграції. Значення меншої тривалості, ніж час інтеграції, будуть відображатися з меншим рівнем, ніж при вимірюванні True PPM. Квазіпіковий вимірювач повинен мати час інтеграції 5 мс.

Вибірковий піковий вимірювач (SPPM) – вимірювач для цифрового звукозапису, який показує значення вибірки цифрового сигналу. Може мати одночасно характеристики точного та квазіпікового вимірювача.

Час інтеграції – це величина, що характеризує швидкодію вимірювача. Визначається тривалістю такої одиночної тональної послідовності, при якій показчик індикатора сягає позначки -2 дБ (рис.1.4) [10].

Час повернення – це час, за який показчик індикатора після відключення від його входу сигналу номінального рівня опускається до позначки -20 дБ (рис.1.4).

На відміну від квазіпікових, у VU-вимірювачів немає двох різних часів інтеграції та повернення, а є тільки одне, однакове для обох напрямків переміщення показчика, воно називається постійним часом. У механічних (стрілочних) приладах цей час визначається конструктивними особливостями їхньої рухомої системи.

З появою світлодіодних індикаторів з'явилася можливість поєднувати VU або RMS та пікові вимірювачі на одній шкалі. Також застосування світлодіодів дозволило утримувати індикацію максимального значення

"точкою", яка називається Peak Hold [11]. Через досить великий час її зависання (1-3 с) немає необхідності постійно стежити за індикатором.

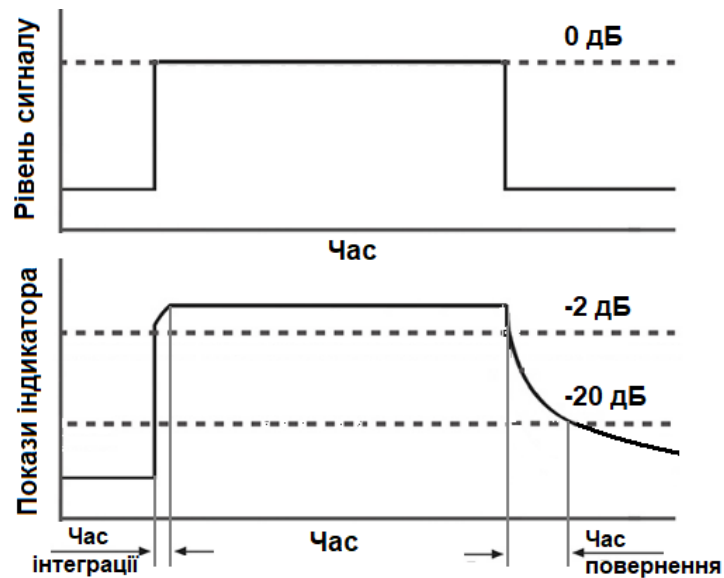


Рисунок 1.4 – Час інтеграції і час повернення показів

Вимірювальний пристрій Dorrrough Loudness Meter, розроблений інженером Майком Дорро в 1980-х роках [12], відображає як середні, так і пікові рівні в одній ітерації з надзвичайно швидким відгуком («балістичним»), і досить точно для цього часу відображає транзєнти (рис.1.5).

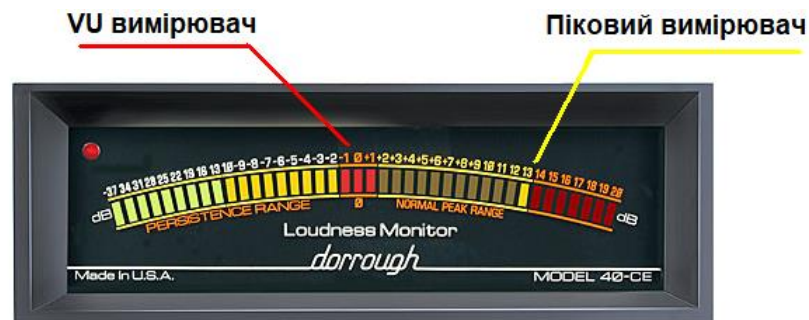


Рисунок 1.5 – Вимірювальний пристрій Dorrrough Loudness Meter

Отже, важлива концепція сучасного вимірювання гучності – це час вимірювання. У законах психоакустики говориться, що ми сприймаємо гучність від середнього рівня, а не миттєві піки – іншими словами, дуже короткий перехідний вибух (навіть на високому рівні) не сильно впливатиме

на наше сприйняття гучності, але подібний, який триває 0,5 с або більше, буде мати великий вплив, настільки, що це може викликати дискомфорт.

За останні кілька років було встановлено та стандартизовано кілька тимчасових шкал, найпоширеніші з них [13]:

- Momentary – миттєві вимірювання 400 мс сигналу, достатні для найкоротшого вимірювання загальної гучності;

- Short Term – короткострокова середня гучність у період від 1 с до кількох секунд, з деяким перекриттям між часовими вікнами;

- Long Term (іноді званий Integrated Loudness) показує середній рівень гучності всього треку;

- True Peak – у цифрових колах навіть миттєві, несуттєві на слух піки рівня передачі стають причиною найсильніших спотворень, що виражаються в клацаннях та трісках. Тому квазіпіковий прилад, що не показує піки коротше 5 мс, цілком задовольняв за точністю вимогам аналогового мовлення, для контролю рівня цифрового сигналу виявився непридатним.

Цифровий сигнал необхідно вимірювати практично безінерційним вимірювачем, що реагує на імпульси тривалістю менше 0,1 мс, тобто. що дозволяє виміряти миттєві значення сигналу (True peak) чи SPPM (Sample peak program meter), тобто, "вимірювачами піків програмного рівня з точністю до семпла".

Вимірювачі миттєвих значень мають шкалу, що відградує в dBFS, тобто у дБ щодо повної цифрової шкали рівня сигналу. Цей рівень, позначений на шкалі приладу 0 dBFS, є межею, перевищення якої призводить до кліпінгу.

1.3 Методи регулювання рівня звукового сигналу

1.3.1 Нормалізація звуку

Нормалізація звуку – вирівнювання максимальної гучності звукової доріжки до заданого рівня [12].

Нормалізація цифрового звуку буває двох видів:

– пікова – коли за точку відліку береться гучність найвищого піку на звуковій доріжці і гучність всієї доріжки збільшується так, що цей пік виявляється рівним заданому максимуму (рис.1.6).

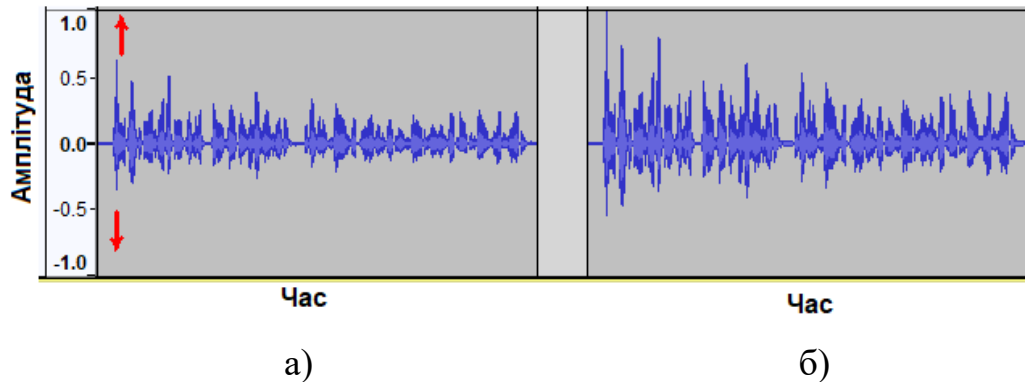


Рисунок 1.6 – Сигнал до (а) і після (б) пікової нормалізації звуку

Даний спосіб не дуже підходить, якщо в записі присутні сильно виділяються піки, оскільки гучність інших частин може виявитися занадто низькою.

RMS нормалізація – тут за основу береться середньоквадратичний рівень гучності для всього файлу і вимірювання ведеться вже в дБ. Цей спосіб більш природний для людського вуха, але за великої гучності можливі спотворення рахунок обрізання (кліпінгу) занадто високих піків.

1.3.2 Динамічна компресія

Динамічна компресія (Dynamic range compression, DRC) – звуження динамічного діапазону фонограми. Динамічний діапазон, це різниця між тихим і найгучнішим звуком. Іноді найтишнішим у фонограмі буде звук трохи гучніший за рівень шуму, а іноді трохи тихіше за найгучніший. Апаратні пристрої та програми, що здійснюють динамічну компресію, називають компресорами [13].

Понижувальна компресія (Downward compression) зменшує гучність звуку, коли вона починає перевищувати певне граничне значення,

залишаючи тихіші звуки в незмінному вигляді. Екстремальним варіантом понижувальної компресії є лімітер.

Підвищуюча компресія (Upward compression), навпаки, збільшує гучність звуку, якщо вона нижча від порогового значення, не торкаючись більш гучних звуків. При цьому обидва види компресії звужують динамічний діапазон аудіосигналу.

На рис. 1.7 показані амплітудні характеристики понижувального (а) і підвищуючого (б) компресора [13].

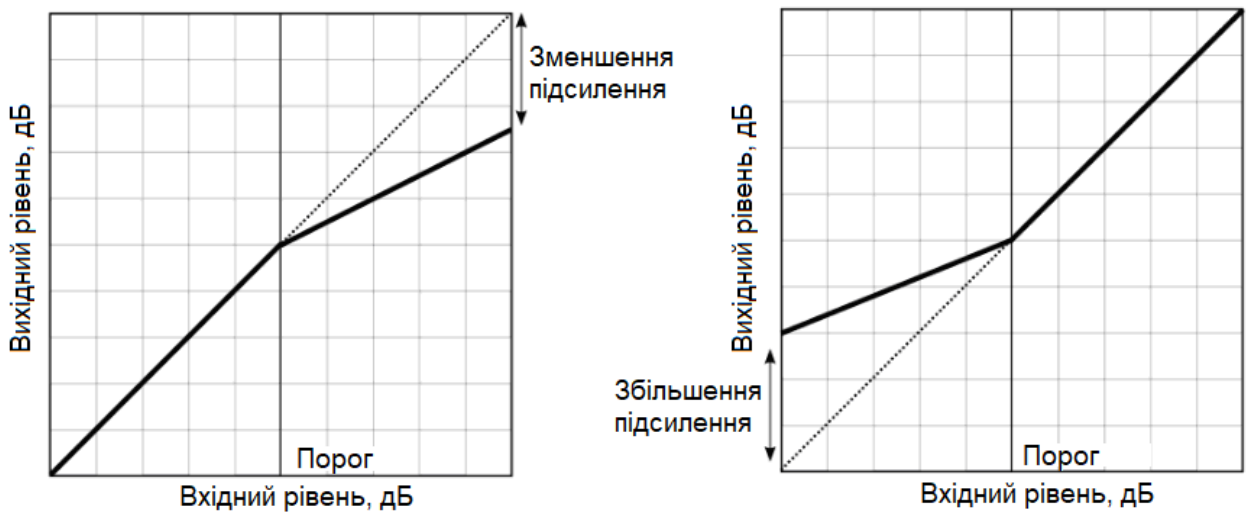


Рисунок 1.7 – Амплітудні характеристики понижувального (а) і підвищуючого (б) компресора

Структурні схеми компресорів показані на рис.1.8 [13].

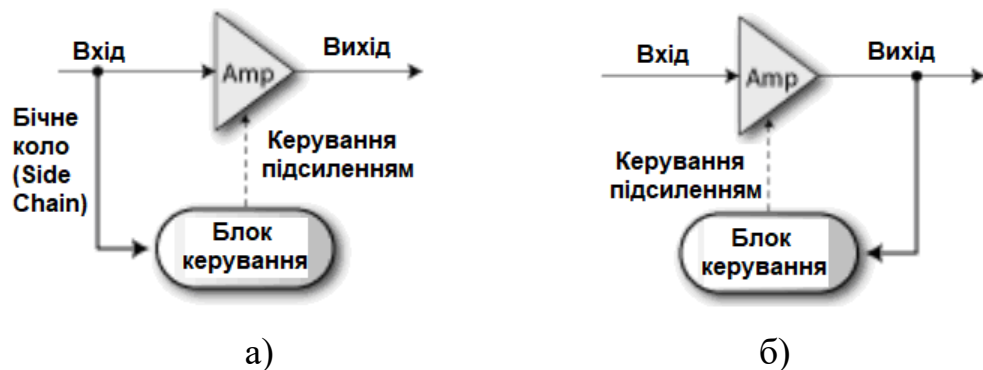


Рисунок 1.8 – Структурні схеми компресорів з прямим (а) і зворотним (б) керуванням

Принцип роботи компресора полягає в наступному. Сигнал, який потрапляє в компресор, поділяється на дві копії. Одна копія спрямовується на підсилювач, у якому ступінь посилення управляється зовнішнім сигналом, друга копія формує цей сигнал. Вона потрапляє в пристрій, званий side-chain (бічне коло), де сигнал вимірюється, і на основі цих даних створюється обвідна, що описує зміну його гучності.

Так влаштована більшість сучасних компресорів, це так званий тип feed-forward (рис.1.8, а). У старіших пристроях (тип feedback) рівень сигналу вимірюється після підсилювача (рис.1.8, б).

Існують різні аналогові технології керованого посилення (variable-gain amplification), кожна зі своїми перевагами та недоліками: лампові, оптичні з використанням фоторезисторів та транзисторні. При роботі з цифровим звуком (у звуковому редакторі або DAW) можуть використовуватись власні математичні алгоритми або емулювати роботу аналогових технологій.

Основні амплітудні параметри компресорів [13]:

- Threshold – компресор зменшує рівень аудіосигналу, якщо його амплітуда перевищує певне граничне значення (threshold). Воно зазвичай вказується в децибелах, при цьому нижчий threshold (наприклад -60 dB) означає, що буде оброблено більше звуку, ніж при більш високому порозі (наприклад, -5 dB);

- Ratio – ступінь зменшення рівня визначається параметром Ratio (відношення): Ratio 4:1 означає, що якщо вхідний рівень на 4 дБ перевищує поріг, рівень вихідного сигналу буде вищим за поріг на 1 дБ.

Компресія з максимальним значенням Ratio ∞ :1 називається обмеженням, обмеженням (limiting). Це означає, що будь-який сигнал вище за пороговий рівень пригнічується до порогового рівня (за винятком короткого періоду після різкого збільшення вхідної гучності).

На рис. 1.9 показані амплітудні характеристики компресора для різних Ratio [13].

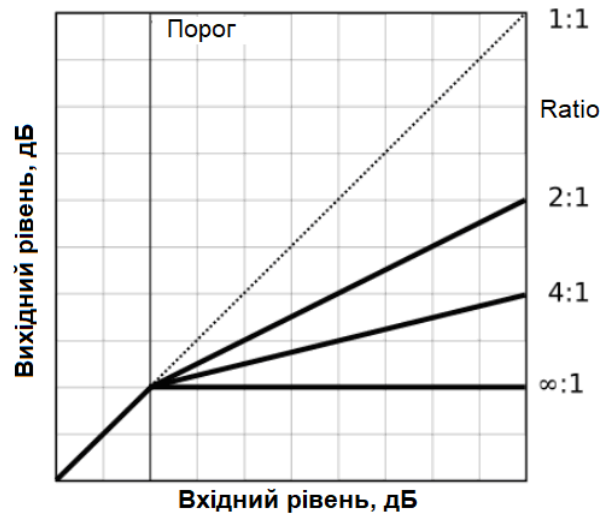


Рисунок 1.9 – Амплітудні характеристики компресора для різних Ratio

Компресор надає певний контроль над тим, як швидко він реагує зміну динаміки сигналу. Основні часові параметри компресорів (Attack та Release):

- Attack визначає час, протягом якого компресор зменшує коефіцієнт посилення рівня, що визначається параметром Ratio;
- Release визначає час, протягом якого компресор, навпаки, збільшує коефіцієнт посилення, або повертає до нормального, якщо рівень вхідного сигналу падає нижче за порогове значення.

На рис.1.10 показані часові діаграми роботи компресора [13].

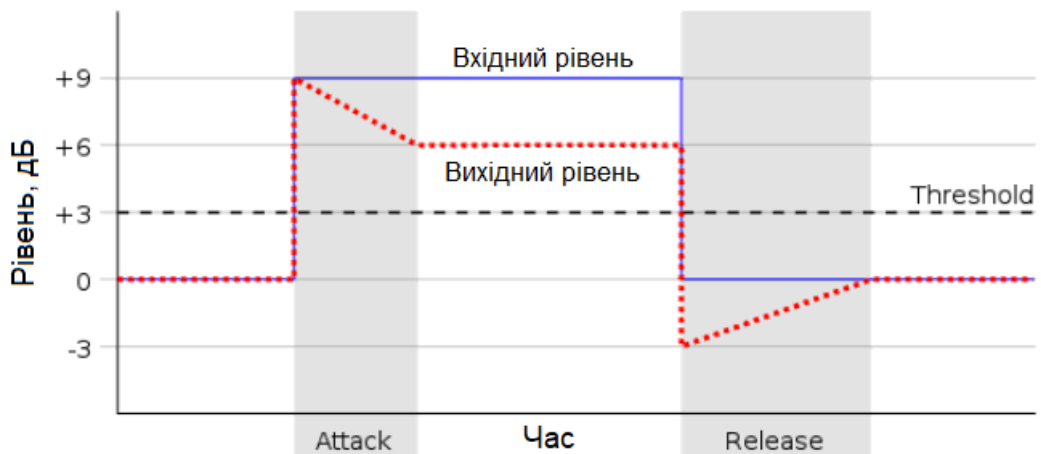


Рисунок 1.10 – Часові діаграми роботи компресора

Ці параметри вказують час (зазвичай у мс), який знадобиться для зміни посилення на певну кількість дБ, зазвичай це 10 дБ. Наприклад, у разі, якщо

Attack встановлено на 1 мс, зменшення посилення на 10 дБ потрібно 1 мс, але в 20 дБ - 2 мс.

В багатьох компресорах параметри Attack і Release можуть налаштовуватися, але у деяких вони задані спочатку і не регулюються. Іноді позначаються як «automatic» чи «program dependent», тобто, змінюються залежно від вхідного сигналу.

Кнее (коліно). Ще один параметр компресора: hard/soft Knee (рис.1.11). Він визначає, чи буде початок застосування компресії різким (hard) чи поступовим (soft) [13].

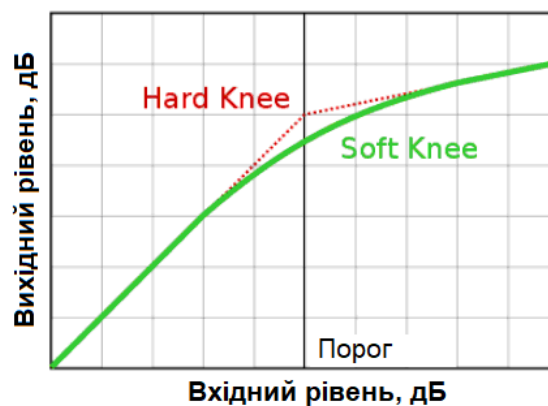


Рисунок 1.11 – Амплітудні характеристики компресора при hard/soft Knee

Soft knee зменшує помітність переходу від необробленого сигналу до сигналу, підданого компресії, особливо при високих значеннях Ratio та різких збільшення гучності.

Компресор може реагувати на пікові (короточасні максимальні) значення або усереднений рівень вхідного сигналу. Використання пікових значень може призводити до різких коливань ступеня компресії, і навіть спотворень. Тому компресори застосовують функцію усереднення (зазвичай RMS) вхідного сигналу при порівнянні його з пороговим значенням. Це дає більш комфортне стиснення, наближене до людського сприйняття гучності.

Головна сфера застосування динамічної компресії – музичне виробництво та мовлення. Компресія надає звуку "щільності".

1.3.3 Обмежувач (ліміттер)

Ліміттер по своїй суті є компресором максимальним значенням $\text{Ratio} = \infty:1$. Така обробка називається обмеженням (limiting). Це означає, що будь-який сигнал вище за пороговий рівень пригнічується до порогового рівня (за винятком Атаки – короткого періоду після різкого збільшення вхідної гучності).

Порівняння амплітудних характеристик компресора і ліміттера показано на рис.1.12 [12].

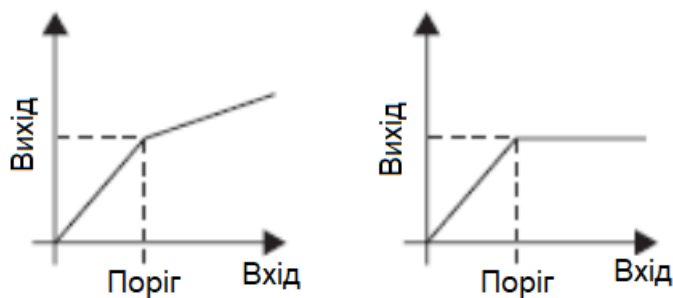


Рисунок 1.12 – Амплітудні характеристики компресора (а) і ліміттера (б)

Часи Attack і Release у ліміттера, як правило, на порядок (в 10 разів) менше, ніж для компресора. Це дозволяє ліміттеру обмежувати короткі піки сигналу, майже не впливаючи на його середній рівень.

Після обмеження піків сигналу з'являється набагато більше можливостей по нормалізації рівня без ризику перевантаження тракту (кліпінгу).

1.4 Висновки по розділу 1

1. Звукова доріжка є не менш важливою складовою телевізійної та відеотрансляції, ніж відеоряд. Найбільше роздратування у глядачів викликає необхідність регулювати рівень при переключенні каналів на ТВ або в Інтернеті, а також стрибки гучності всередині однієї програми. Це здатне спотворити враження від будь-якого досконалого відеоряду.

2. Гучність звуку є результатом суб'єктивного сприйняття. Крім рівня звукового тиску гучність залежить від спектрального складу сигналу та часових характеристик його обвідної. Тому класичне вимірювання рівня часто не відповідає реальній картини зі сприйняттям гучності.

3. Важливими характеристиками класичних індикаторів рівня є час інтегрування і час повернення. За цими показниками існує широкий спектр індикаторів – від істинно-пікових (True Pick) до середньоквадратичних RMS з часом інтегрування одиниці секунд. Найпоширенішими є VU індикатори з постійною часу інтегрування 300...400 мс.

Широка номенклатура індикаторів рівня призводить до появи різного роду комбінованих індикаторів.

4. Пристрої динамічного регулювання рівня, аналогічно звукорежисеру, використовують інформацію про виміряний рівень сигналу для адаптивної зміни коефіцієнту передачі тракту. При цьому результат їхньої роботи, природно, буде залежати від часових параметрів (Attack і Release) детектора рівня.

5. Пристрої динамічного регулювання рівня – нормалізатори, компресори, ліміттери – давно використовуються в радіомовленні. Їх задача – уникнути перевантаження і перемодуляції передавача. Однак часто їх дія призводить до суттєвого зменшення динамічного діапазону трансляції. Якщо для радіопередач, де основний контент – це музика – таке звуження підходить, то для ТВ, де контент набагато різноманітніший, дане рішення не задовольнить.

6. В даний момент в сфері медіавиробництва не існує єдиного рішення для вимірювання і нормалізації гучності телевізійного або звукового мовлення. Є ряд розрізнених рекомендацій в різних сферах медіавиробництва, що при переході з одного способу передачі програми на інший можуть не відповідати між собою і давати спотворені результати.

Метою кваліфікаційної роботи є аналіз методів вимірювання і нормалізації гучності телевізійного або звукового мовлення з метою

формулювання зрозумілих рекомендацій для робітників медіаіндустрії, студентів та випускників відповідних навчальних закладів.

7. Кваліфікаційна робота виконується на кафедрі МІРЕС ХНУРЕ. На кафедрі проводяться дослідження в таких наукових областях, як виявлення та розпізнавання БПЛА за результатами акустичного спостереження [14-17], створення систем зондування атмосфери за допомогою акустичних хвиль [18-21]. Цілий ряд студентських доповідей [22-25] і атестаційних робіт магістрів минулих років [26-28] присвячені дослідженню систем звукозапису. Отже, дослідження в даній роботі пов'язані і ґрунтуються на традиційному напрямку робіт колективу і студентів кафедри МІРЕС.

2 МЕТОДИ НОРМАЛІЗАЦІЇ ГУЧНОСТІ ЗВУКУ ТЕЛЕВІЗІЙНИХ І РАДІОТРАНСЛЯЦІЙ

2.1 Загальна концепція нормалізації гучності

2.1.1 Відмінність нормалізації гучності від пікової нормалізації

Концепція встановлення рівнів звуку з нормалізацією піків щодо максимально допустимого рівня (Permitted Maximum Level – PML, наприклад, -9 dBFS) призвела до однорідних пікових рівнів програм, але дуже різних рівнів гучності. Фактичне коливання залежить від рівня динамічної компресії сигналу.

На відміну від цього, нормалізація гучності дозволяє досягти однакової середньої гучності програм.

Зі зміною піків залежно від контенту, а також художніх та технічних вимог (рис. 2.1), слухач може насолоджуватися однорідним середнім рівнем гучності у всіх програмах, не користуючись пультом для частого регулювання гучності.

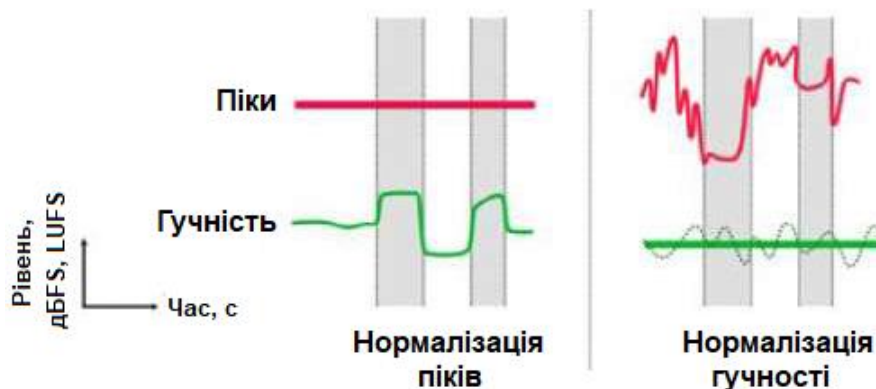


Рисунок 2.1 – Порівняння нормалізації пікового рівня з нормалізацією рівня гучності у серії програм

Сказане не означає, що всередині програми рівень гучності має бути постійним. Це також не означає, що всі окремі компоненти програми (наприклад, премікси або стем-мікси, версія з музикою та ефектами або

ізолювана мовна доріжка) повинні бути на одному рівні гучності. Варіація гучності – художній інструмент, а концепція нормалізації гучності стимулює динамічніше зведення. Це середня, інтегрована гучність цілої програми, яка нормалізована.

2.1.2 Вимірювання гучності програми в LUFS

LUFS – це одиниці вимірювання гучності звуку. Аббревіатура LUFS розшифровується як Loudness Units Full Scale – одиниця гучності щодо повної шкали. Сьогодні це стандартний спосіб вимірювання звуку, який об'єднує гучність, що сприймається людським слухом, і справжню інтенсивність аудіосигналу.

LUFS дозволяє фахівцям галузі встановлювати середні стандарти гучності для застосування у кіно, телебаченні, радіо та потокових сервісах. Два музичні або аудіо твори з однаковим рівнем LUFS повинні мати однакову гучність, що сприймається.

На сьогодні LUFS є найбільш точним способом вимірювання гучності звуку, оскільки він враховує людське сприйняття, на відміну від інших одиниць вимірювання рівня, таких як RMS.

Дуже важливо розуміти різницю між LUFS та RMS. На відміну від True Peak вимірювань, RMS, як і LUFS, вимірює середній рівень сигналу.

Але LUFS здатні більш точно описувати гучність, використовуючи криві Флетчера-Мунсона (рис.1.2). Ці криві також називаються кривими рівної гучності і описують зв'язок між фізичними характеристиками звуку і нашим сприйняттям.

Наприклад, одна з відмінностей між LUFS та піковими або середньоквадратичними показниками полягає в тому, що людське вухо краще сприймає середні частоти, ніж низькі. Тому можна отримати мікс із високим середньоквадратичним або середнім показником гучності та низьким рівнем LUFS.

У треку може бути багато басу, який збільшує середню гучність, але оскільки людському вуху складніше сприймати ці відмінності, ніж, скажімо, середньочастотний інструмент, такий як гітара, рівень LUFS виявляється нижчим.

На LUFS вимірювачі може бути кілька різних значень (рис.2.2).

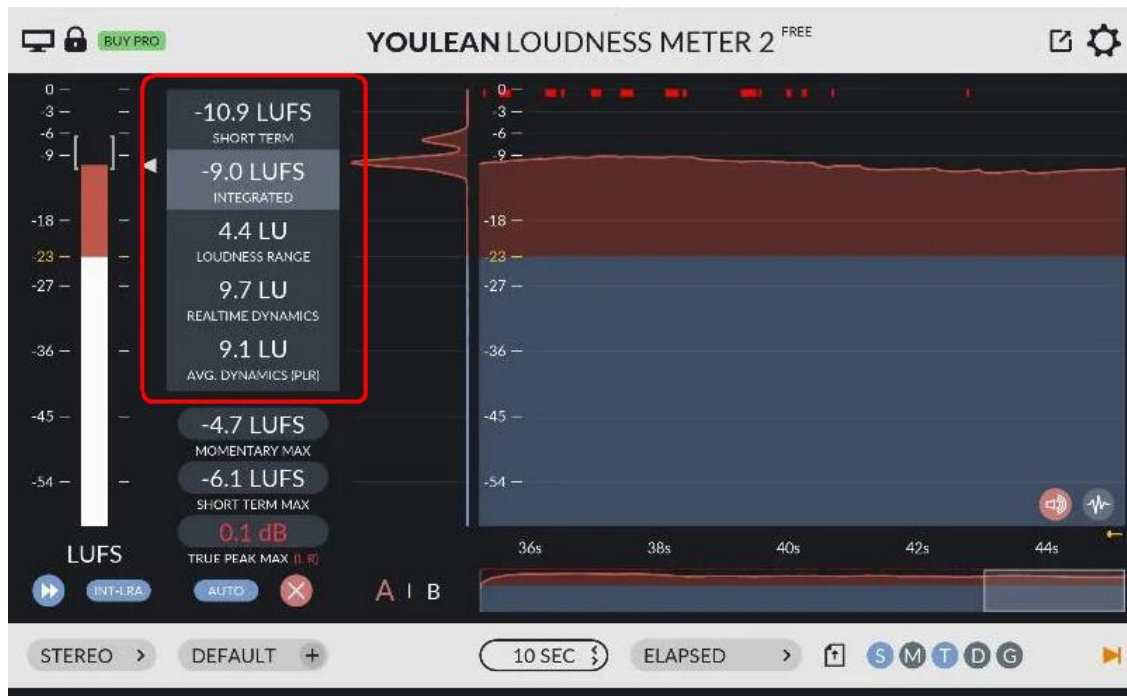


Рисунок 2.2 – Різні варіанти шкали LUFS (обведено червоним)

Нижче наведено основні відомості про різні типи LUFS:

– інтегрований LUFS – призначений для аналізу певного відрізка аудіо та вимірює середнє значення LUFS;

– короткостроковий LUFS аналізує більш коротку ділянку аудіофайлу на відміну від інтегрованого LUFS. Наприклад, ми можемо очікувати, що короткострокові значення LUFS будуть відрізнятися між куплетом і приспівом, оскільки між цими двома розділами зазвичай існує широкий динамічний діапазон;

– миттєвий LUFS – є найбільш короткостроковим виміром LUFS. Ці вимірювання можуть допомогти отримати точні покази LUFS для окремих моментів у треку, але не є репрезентативними для загального показу LUFS треку.

2.1.3 Порівняння нормалізації по сигналу і нормалізації по метаданим

Теоретично рівна гучність може бути отримана шляхом нормалізації аудіосигналу або за допомогою метаданих гучності.

По суті є два способи досягнення нормалізації гучності для споживача:

- перший – сама нормалізація аудіосигналу на передавальній стороні, щоб програми спеціально були в середньому однаково гучними (рис.2.3, а),
- другий метод – використання метаданих гучності, що описують, наскільки гучна програма. В останньому випадку фактичні середні рівні гучності програми можуть змінюватися на нормалізоване значення і, як і раніше, бути дуже різними між програмами (рис.2.3, б).

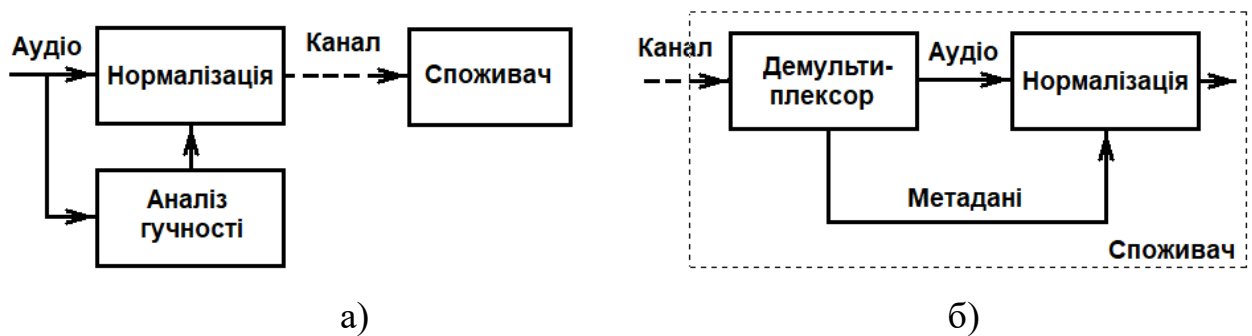


Рисунок 2.3 – Нормалізація звуку по сигналу (а) і по метаданим (б)

При наявності сучасного обладнання нормалізація може проводитися у споживача за допомогою індивідуальних значень метаданих гучності для встановлення діапазону посилення програм на однаковий рівень відтворення.

Нормалізація гучності аудіосигналу у медіавиробництві рекомендується через простоту і потенційне підвищення якості.

2.1.4 Цільовий рівень LUFS

Рекомендація EBU R 128 визначає новий опорний рівень гучності (так званий «цільовий рівень») як:

$$N = -23.0 \text{ LUFS } (\pm 0.5 \text{ LU}). \quad (2.1)$$

Таке значення є компромісним для звукового супроводу різних жанрів. Відомо, що звуковий сигнал описується середньоквадратичним $N_{сер}$ та піковим $N_{п}$ значеннями рівнів. Різницю цих рівнів називають пік-фактором сигналу:

$$k = N_{п} - N_{сер}. \quad (2.2)$$

На рис. 2.4, а показано результат нормалізації звуку по пікам для різних жанрів. По причині, що звуковий супровід різних жанрів має різний пік-фактор, середня гучність різного контенту буде відмінною.

Для вирівнювання суб'єктивної гучності звуку різних жанрів була запропонована компромісна величина гучності -23.0 LUFS, яка враховує пік-фактори різного контенту (рис.2.4, б).

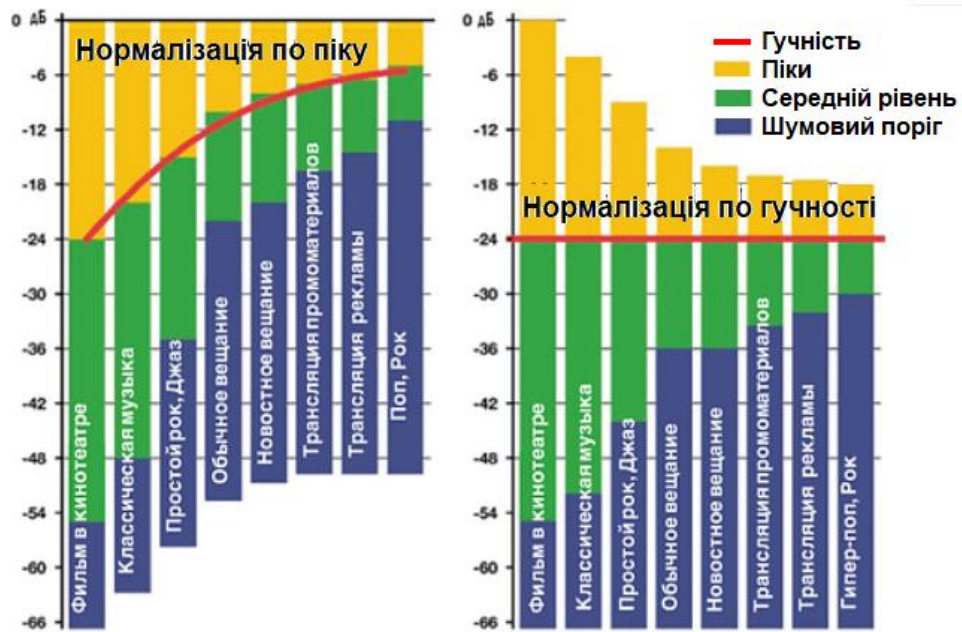


Рисунок 2.4 – Нормалізація звукового контенту різних жанрів:

а – по пікам, б – по гучності (-23.0 LUFS)

Єдина величина має велике значення в поширенні контенту рівної гучності. Оскільки нормалізація гучності джерела по LUFS у певному сенсі

«карає» надстиснуті сигнали, на кшталт тих, що є в рекламі, і таким чином автоматично змушує виробничий персонал подумати про інші, більш динамічні та креативні методи створення враження від своєї програми.

Отже, реальна технічна зміна рівня аудіосигналу шляхом активної нормалізації до -23 LUFS надає прямий позитивний вплив на художній процес. Виробнича частина позбавляється участі у «війні гучності» – невдалий і широко поширений результат парадигми нормалізації піків.

Робота над загальним рівнем гучності означає нову концепцію установки рівнів і всієї роботи зі звуком. Якщо обмежувач піків встановлений на максимально допустимий рівень (зазвичай -9 dBFS, виміряний QPPM), за наявності «захисної стелі», де незалежно від того, з якою силою ти в нього упираєшся, він завжди гарантує «коректний» максимальний рівень (рис.2.5, а).

Парадигма встановлення рівня гучності в порівнянні з цим більше нагадує «політ просто неба» (рис.2.5, б).

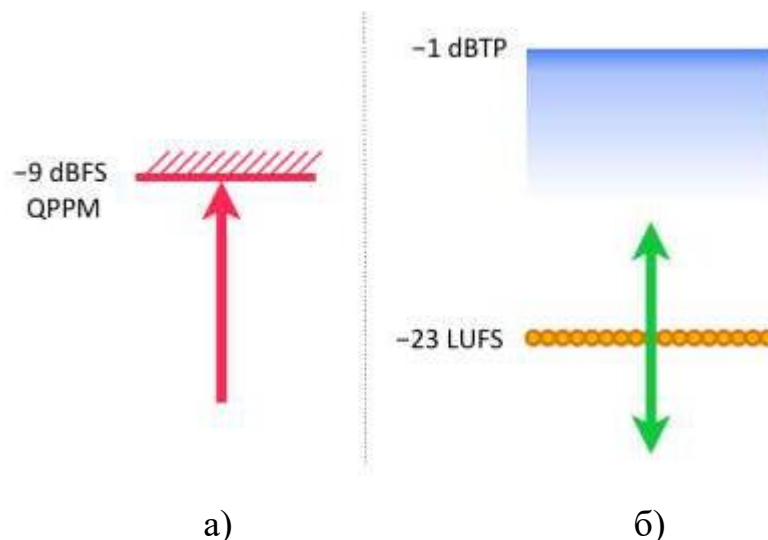


Рисунок 2.5 – Квазіпікова нормалізація рівня («стеля безпеки», а)
і нормалізація рівня гучності (б)

При нормалізації та вимірюванні гучності стеля безпеки зникає. Здається, це може злякати. Оскільки в певному сенсі було «зручно», що не треба так ретельно стежити за рівнями – лімітер наприкінці тракту

гарантував завжди неперевантажений вихід. Але побічний ефект був у тому, що рівні гучності зростали, почалося зловживання парадигмою нормалізації піків та конкуренція у гучності, стимульована складнішими динамічними процесорами.

Регулювання рівнів гучності, з іншого боку, стимулює використання кращого на даний момент вимірювального приладу: вуха. Це сприяє якості звуку. Боротьба «хто голосніше» мають припинитися, загальні рівні знизитися, і у поєднанні з підвищенням максимального рівня реальних піків (-1 dBTP для лінійного звуку) це веде до потенційно більш динамічного зведення з більшою узгодженістю гучності всередині програми.

Тобто динамічна компресія знову стає художнім інструментом, а не зброєю гучності, в результаті якість звуку зростає.

2.1.5 Особливості використання процесорів гучності

Далі у виробництві мовник стикається з необхідністю нормалізації різного контенту із різних джерел. Йому може потрапляти багато звукових програм без нормалізації гучності.

Необхідно розробити стратегію для таких програм, наприклад, автоматичну нормалізацію відразу після завантаження в ефірний сервер або встановлення захисного пристрою для регулювання гучності (процесора гучності) на виході центральної апаратної обробки, наприклад, прямих сигналів, що не встановлені у виробництві на цільовий рівень -23 LUFS.

Однак використання процесора гучності – є питання делікатне. Якщо процесор суворо працює з більш гучними частинами програми, то в результаті може вийти менший рівень гучності програми, «денормалізуючи контент. Отже, система мовлення повинна сигналізувати процесору гучності, коли відтворюється сумісний за гучністю контент (передбачається, що мікс підходить для мовлення). Тоді процесор повинен перейти в режим Bypass або

в попереднє налаштування, де застосовується тільки безпечне обмеження реальних пікових значень (рис.2.6).

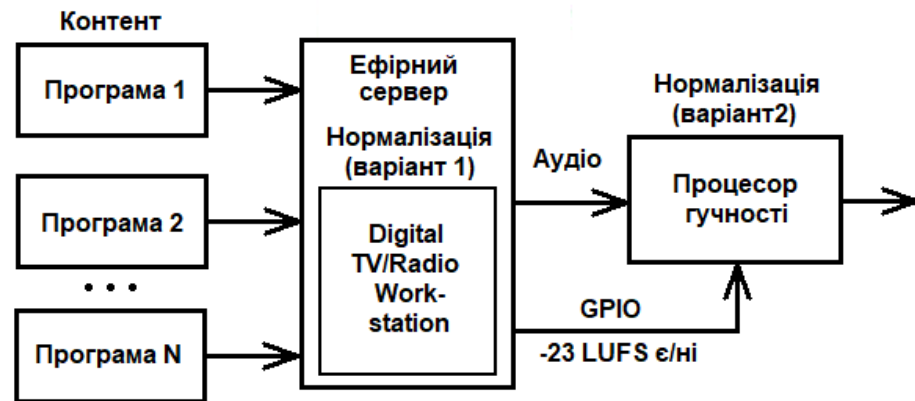


Рисунок 2.6 – Використання нормалізації в програмному забезпеченні (варіант 1) або в процесорі гучності (варіант 2)

Така сигналізація може здійснюватись через інтерфейс GPIO (general-purpose input/output) або мережеві системи контрольних даних.

Процесор гучності також може використовуватися для прямого ефіру, для «гармонізації джерел». З відповідними установками такий процесор може допомагати звукооператору пригнічувати непередбачувано гучні частини прямої трансляції. Взагалі, потрібно стежити за тим, щоб не створювати «сосиску гучності» (рис.2.7) шляхом надмірно агресивної обробки, руйнуючи початковий задум контрастності та динаміки.

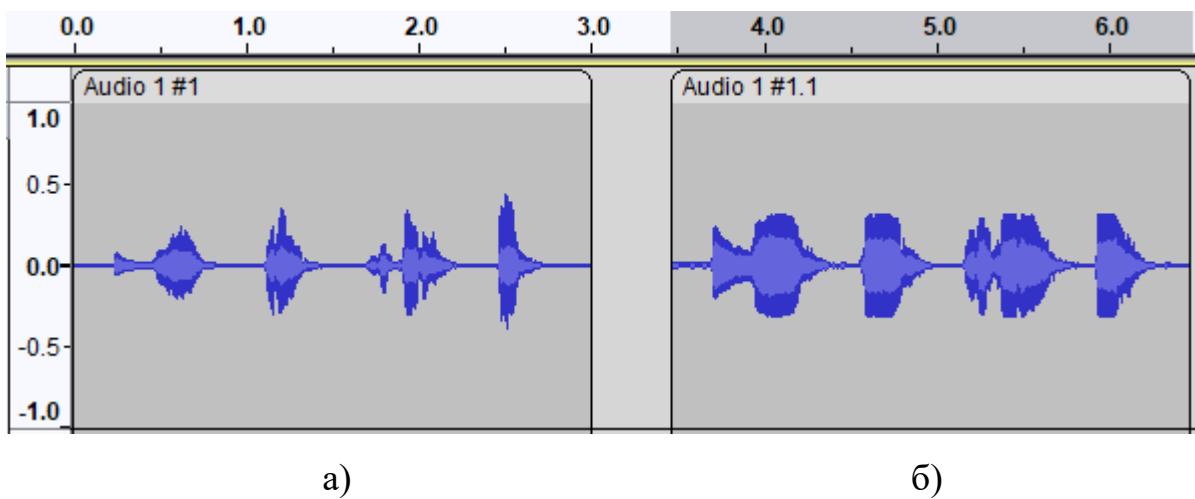


Рисунок 2.7 – Сигнал без перекомпресії (а) і з перекомпресією (б, т. зв. «сосиска» або «кирпич» гучності)

2.2 Стратегії встановлення рівня гучності

2.2.1 Базовий підхід до зведення

Принцип встановлення рівня гучності у медіавиробництві дає два варіанти (рис. 2.8):

- миттєва зміна звичок на зведення та нормалізацію гучності з подальшою необхідністю невеликого зсуву рівнів або без нього (якщо звукооператор не потрапив точно на цільовий рівень),
- або зберегти поточну практику встановлення (пікових) рівнів з необхідністю наступного зсуву.



Рисунок 2.8 – Два базові методи роботи для досягнення однорідної гучності у медіавиробництві та пост виробництві

Перше рішення установки рівнів (перехід на зведення гучності та безпосереднє вимірювання LUFS) можна рекомендувати для звукооператорів і звукорежисерів, що працюють в студії офлайн. Після початкового періоду вимірювань і випробувань за допомогою LUFS-метра минулих програм або треків того ж жанру можна отримати уявлення про те, де знаходяться рівні і яку різниця рівнів LUFS і дБ можна очікувати (рис.2.9).

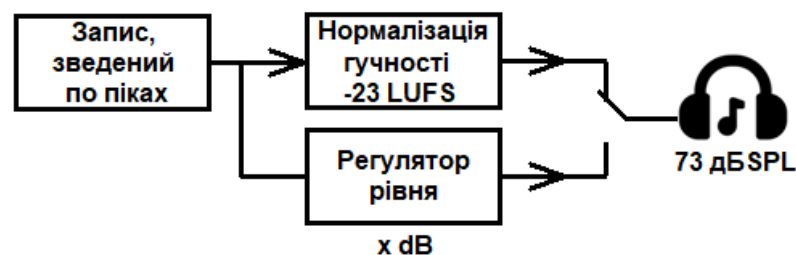


Рисунок 2.9 – Перехід на зведення гучності по LUFS

Такі експерименти мають супроводжуватися додатковою зміною рівня прослуховування, щоб середній акустичний рівень під час зведення залишався однаковим. Рекомендоване значення 73 дBSPL.

На рис.2.10, а показана фонограма «Дольче Габбана» з офіційного альбома В. Сердючки. Фонограма зведена по піковому рівню -5 дБ. Після нормалізації даної фонограми по гучності -23 LUFS, загальний рівень зменшився приблизно на 10 дБ (рис.2.10, б).

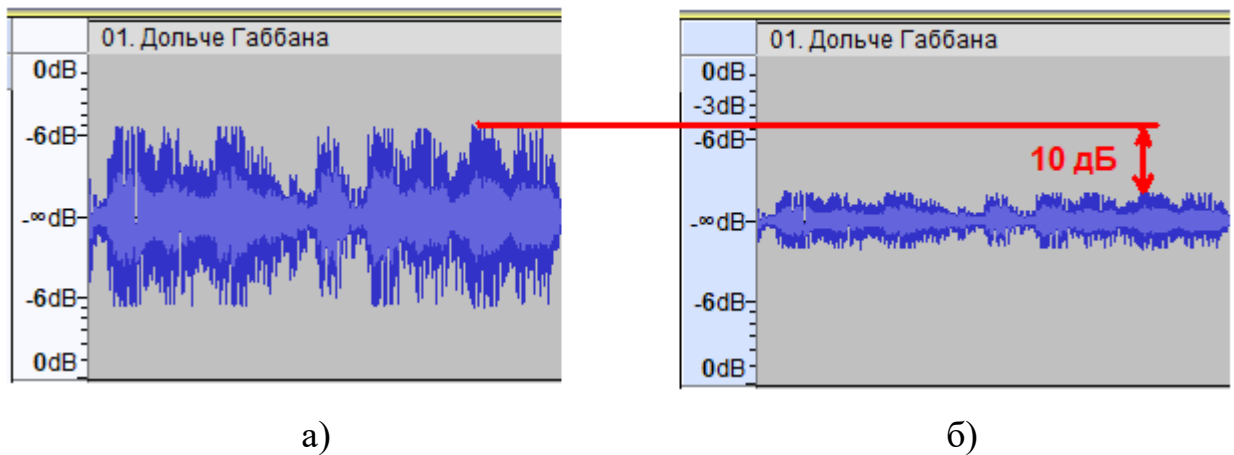


Рисунок 2.10 – Фонограма «Дольче Габбана», нормалізована по пікам (а)
і по -23 LUFS (б)

Для матеріалу іншого жанру, така різниця буде іншою. Так, на рис. 2.11 показано аналогічний експеримент зі звукозаписом чоловічого голосу.

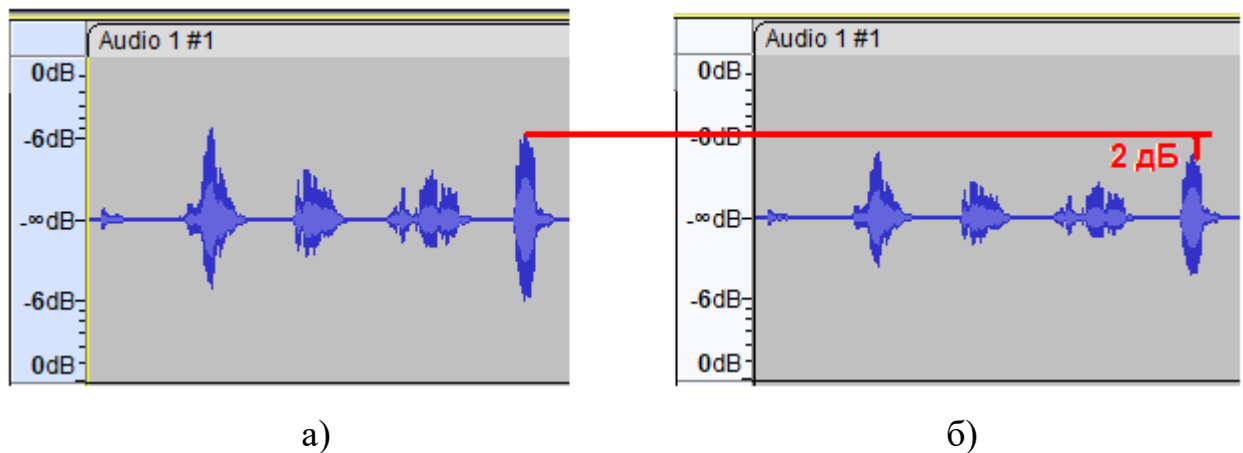


Рисунок 2.11 – Фонограма голосу, нормалізована по пікам (а)
і по -23 LUFS (б)

Нормалізація по -23 LUFS зменшило загальний рівень лише на 2 дБ. Така різниця обумовлена пік-фактором і спектральним складом сигналу. Саме на них реагує наш слух, сприймаючи гучність.

Переваги встановлення рівнів гучності наступні. Більший запас за рівнем можна використати для накладання додаткових звуків і шумів під час трансляції.

Друге рішення установки рівнів доречніше для перших етапів переходу і на початку може підходити для тих, хто працює у прямому ефірі. Зберігаються існуючі вимірювачі, лімітери та практика зведення, а зсув проводиться на виході пульта для отримання цільового рівня гучності -23 LUFS. Вимірювач гучності розташовується після зсуву рівня (рис.2.12), щоб оператори могли зрозуміти точну величину зсуву (яка спочатку будується на припущеннях).

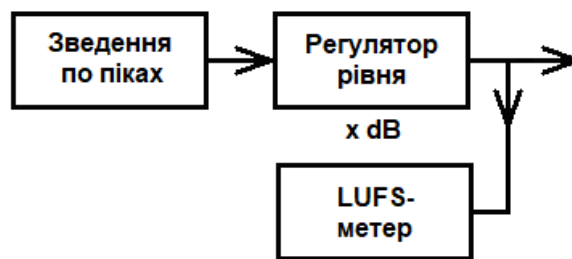


Рисунок 2.12 – Налаштування гучності методом зсуву

Використання вимірювача гучності LUFS паралельно з традиційним RMS вимірювачем – у будь-якому випадку хороша ідея для облегшення переходу. Таким чином можна отримати досвід.

2.2.2 Вимірювання гучності для виробництва та пост виробництва

Вимірювач гучності 'EBU Mode' згідно EBU Tech Doc 3341 має 3 шкали часу [29]:

- миттєва гучність (скорочено "M") – часове вікно: 400 мс;
- короткочасна гучність (скорочено "S") – часове вікно: 3 с;
- інтегрована гучність (скорочено "I") – від 'start' до 'stop'.

'EBU Mode' також визначає дві шкали амплітуди:

- “EBU +9 Scale”, яка має підходити для більшості програм,
- “EBU +18 Scale”, яка може використовуватися для програм з широким діапазоном гучності.

Слід також зазначити наступне:

- обидві шкали можуть показувати або відносний рівень гучності в LU або абсолютний в LUFS;
- '0 LU' в 'EBU mode' дорівнює цільовому рівню -23 LUFS;
- ‘EBU Mode’ не визначає графічний інтерфейс, тому практична реалізація індикатора може бути різною.

На рис.2.13 показана відповідність між шкалами LU і LUFS.

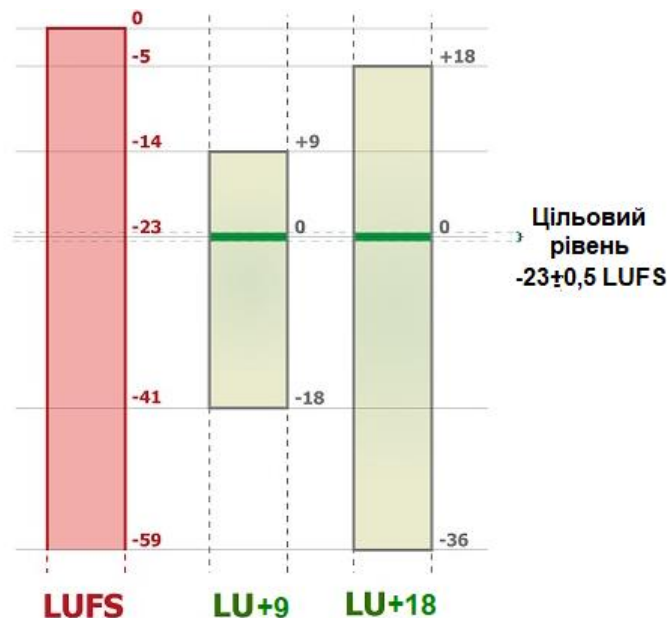


Рисунок 2.13 – відповідність між шкалами LU і LUFS

Якщо фактичний рівень гучності знаходиться в межах допуску ± 0.5 LU (або ± 1.0 LU для прямого ефіру), то подальших дій не потрібно.

У ситуації поствиробництва простий розрахунок посилення поверне програму на цільовий рівень. Для прямих передач не на цільовому рівні коригувальні заходи можуть вживатися далі, у вигляді процесорів гучності, які поступово регулюють інтегрований рівень гучності таких програм у ненав'язливій манері і можуть служити «запобіжною сіткою гучності».

2.2.3 Діапазон гучності (Loudness Range)

Параметр Loudness Range (LRA) визначає зміну гучності у програмі. У минулому це робилося методом «кваліфікованих здогадів» досвідченого персоналу, щоб вирішити, чи вписується чи програма у вікно допуску гучності для цільової аудиторії.

При використанні Loudness Range в кінці періоду вимірювань (зазвичай усієї програми) одна цифра допомагає звукооператору вирішити, чи потрібна подальша динамічна обробка.

Робота з нормалізацією гучності обов'язково означає спостереження та потенційний контроль Loudness Range, так як динамічні можливості розширюються. Це важливо для забезпечення відповідного сигналу для цільової аудиторії і каналу поширення.

У виробництві та поствиробництві може бути створений звуковий контент з високим значенням LRA і максимальним рівнем реальних піків $MPTPL = -1$ dBTP. Але інші платформи можуть вимагати менше значення LRA і менший MPTPL (при збереженні гучності -23 LUFS).

З параметром Loudness Range можна визначати необхідні заходи для потенційної динамічної обробки програми, щоб вмістити її в вікно допуску каналу чи платформи поширення.

Для динамічних програм, що складаються у основному з музики, загальна низькорівнева компресія може дати задовільні результати (рис.2.14):

- низький поріг (< -50 dBFS);
- помірний коефіцієнт стиснення (1:1,2...1:1,5) забезпечують однорідну компресію всього діапазону сигналу.

Залежно від вихідного рівня гучності зсув до цільового рівня -23 LUFS може виконуватися паралельно шляхом регулювання компенсуючого посилення компресора.

Важливо розуміти, що неможливо визначити одне максимальне значення LRA для всіх мовників та програм.

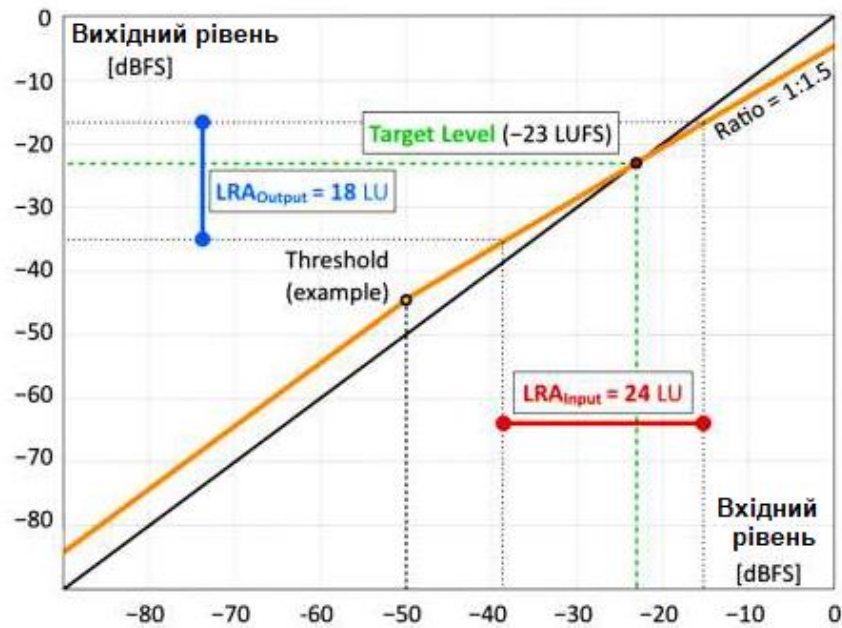


Рисунок 2.14 – Приклад обробки для зменшення Loudness Range компресором із низьким порогом (-50 dBFS) та помірним коефіцієнтом стиснення (1:1.5)

Крім того, LRA дуже корисний як інструмент зведення і не повинен бути перешкодою передачі специфіки програм. Тим не менш, рекомендовані індивідуальні максимальні значення LRA можуть дати хорошу динамічну структуру для різних форматів (наприклад, 5.1 vs. 2.0), жанрів, платформ розповсюдження, а також різного середовища відтворення.

Середні умови прослуховування, вік цільової аудиторії, «зона комфорту прослуховування» для споживача та інші параметри впливають на прийнятність значення LRA для певних програм. Парадигма контролю діапазону гучності починається із загальноприйнятого максимального значення Loudness Range за вищезгаданими принципами і адаптує далі це значення у відповідності з технічними потребами платформ поширення та середовища відтворення.

У жодному разі жоден параметр і відповідне максимально допустиме значення не можуть гарантувати гарний мікс. Це стосується і Loudness Range. Для оцінки якості міксу досвідчені слухачі мають оцінювати програму вухами. LRA дає загальний посібник щодо базових динамічних параметрів

міксу, він може використовуватися для спрямування динамічної обробки в процесорі гучності.

2.3 Висновки по розділу 2

1. Поширена до останнього часу концепція нормалізації піків звукового сигналу до максимально допустимого рівня призвела до однорідних пікових рівнів програм, але дуже різних рівнів гучності. Фактичне коливання залежить від рівня динамічної компресії сигналу. Дана ситуація добре відома по рекламним вставкам, коли звукорежисери перетискають сигнал, сильно збільшуючи його середній рівень.

2. На відміну від цього, нормалізація гучності дозволяє досягти однакової середньої гучності програм. Зі зміною піків залежно від контенту, а також художніх та технічних вимог, слухач може насолоджуватися однорідним середнім рівнем гучності у всіх програмах, не користуючись пультом для частого регулювання гучності.

3. Рекомендацією EBU 128 введені одиниці гучності LUFS, які враховують частотні характеристики слуху. По причині, що звуковий супровід різних жанрів має різний пік-фактор, середня гучність різного контенту буде відмінною. Для вирівнювання суб'єктивної гучності звуку різних жанрів була запропонована цільова величина гучності -23 LUFS, яка враховує пік-фактори різного контенту.

4. По суті є два способи нормалізації гучності: нормалізація аудіосигналу на передавальній стороні, або використання метаданих гучності, а регулювання відбувається на стороні споживача. Нормалізація гучності під час виробництва програм рекомендується через простоту і потенційне підвищення якості.

5. Реально у виробництві мовник стикається з необхідністю нормалізації різного контенту із різних джерел. Йому може потрапляти багато звукових програм без нормалізації гучності. Для таких програм

потрібно мати або автоматичну програмну нормалізацію відразу після завантаження в ефірний сервер, або встановлення процесора гучності на виході центральної апаратної.

Але процесор має обробляти треба тільки ті сигнали, що не нормалізовані на цільовий рівень -23 LUFS. Отже, система мовлення має сигналізувати процесору, коли відтворюється сумісний за гучністю контент, процесор повинен перейти в режим байпас. В іншому випадку процесор може перекомпресувати сигнал.

6. При переході на нові одиниці вимірювання гучності у звукооператора є два варіанти: або переходити на LUFS-індикатор відразу, використовуючи його при зведенні, або зводити по-старому, по піках, а потім нормалізувати до цільового рівня -23 LUFS. Перший спосіб є більш прийнятним, оскільки залишає при зведенні більше простору зверху для збереження динамічного діапазону.

7. Щоби отримати уявлення про те, де знаходяться рівні і яку різниця рівнів LUFS і дБ можна очікувати було зроблено експеримент. Була взята фонограма «Дольче Габбана» з офіційного альбома В. Сердючки, фонограма зведена по піковому рівню -5 дБ. Після нормалізації по гучності -23 LUFS, загальний рівень зменшився приблизно на 10 дБ. Для матеріалу іншого жанру, така різниця буде іншою. Аналогічний експеримент зі звукозаписом чоловічого голосу показав зменшення рівня на 2 дБ. Така різниця обумовлена пік-фактором і спектральним складом сигналу. Саме на них реагує наш слух, сприймаючи гучність.

8. Розглянуті шкали і часові характеристики LUFS-метрів. Для зручності багато програмних продуктів перейшли на спрощену шкалу LU, де -23 LUFS = 0 LU. Існує дві шкали: +9 – для більшості програм, і з розширеним діапазоном гучності +18 – для кінофільмів.

Часові характеристики LUFS-метрів приблизно відповідають аналогічним характеристикам RMS або пікових індикаторів.

9. Крім цільового середнього рівня гучності різні канали або Інтернет-платформи нормують значення діапазону гучності (Loudness Range). Цей параметр допомагає звукооператору вирішити, чи потрібна подальша динамічна обробка. Тобто можна визначати необхідні заходи для того, щоб вмістити програму в вікно допуску каналу чи платформи поширення. Успішно вирішити дану задачу допомагає компресором із низьким порогом (-50 dBFS) та помірним коефіцієнтом стиснення (1:1.5).

3 ПРАКТИЧНІ РЕКОМЕНДАЦІЇ З ВИРІВНЮВАННЯ ГУЧНОСТІ ТЕЛЕВІЗІЙНОГО КОНТЕНТУ

3.1 Мовлення на базі файлів

Оскільки світ мовлення вже перейшов на файлові робочі процеси, тут також необхідно повністю описати концепцію регулювання гучності. Базовий принцип залишається той самий: рекомендується нормалізація гучності та динамічний контроль джерела, особливо для нового контенту. Тим не менш, оскільки метадані є невід'ємною частиною файлових систем, дані рішення в майбутньому все більше будуть спиратися на метадані.

Джерелом мовного файлу, що містить аудіосигнали, може бути процес завантаження, передача із зовнішнього сервера та з файлового архіву.

Для існуючих програм (архівного контенту) є 4 варіанти нормалізації гучності:

- фактична зміна рівня гучності всіх аудіофайлів на цільовий;
- зміна рівня гучності лише «на вимогу»;
- використання результату вимірювання рівня гучності для регулювання рівня відтворення без зміни вихідного рівня гучності;
- транспортування коректних метаданих гучності споживачеві, де проводиться нормалізація.

Вибір рішення залежить від таких факторів як інфраструктура, робочі процеси, управління медіафондами, наявність відповідного обладнання, фінансові ресурси, час тощо.

На початку життя файлу на підприємстві необхідно провести вимірювання, що забезпечують значення

- рівня гучності програми,
- діапазону гучності,
- максимальних рівнів реальних піків (Maximum True Peak Level), тобто трьох характерних параметрів звуку.

Для дуже короткого контенту (<30 с) можна також виміряти та зберегти:

- максимальний рівень миттєвої гучності (Maximum Momentary Loudness Level),
- максимальний рівень короткочасної гучності (Maximum Short-term Loudness Level).

Залежно від результатів цього вимірювання та подальшого методу нормалізації гучності та відповідності допустимому діапазону гучності виробляється схема обробки, що складається з «компонувальних блоків» або «основних завдань».

Тепер докладно розглянемо робочий процес за допомогою загальних блок-схем (рис.3.1 – рис.3.3).



Рисунок 3.1 – Блок обробки гучності програми

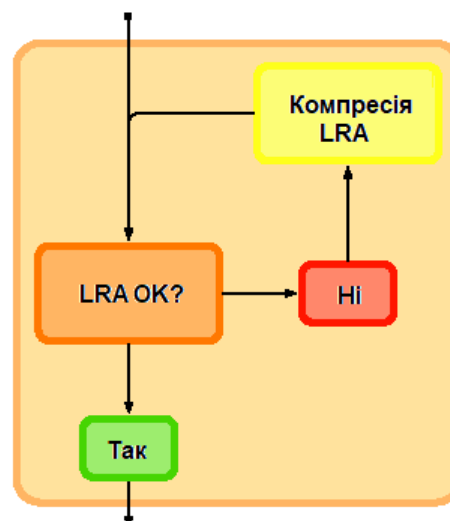


Рисунок 3.2 – Блок обробки діапазону гучності

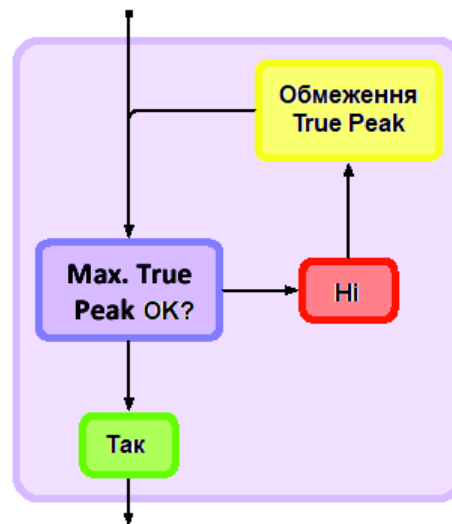


Рисунок 3.3 – Блок обробки Maximum True Peak Level

Три базові блоки, перераховані вище, лежать у центрі будь-якого процесу контролю якості файлів відносно технічних параметрів аудіоконтенту. На початку будь-якої потенційної обробки вимірюються три параметри:

- рівень гучності (Loudness Level – L_K),
- діапазон гучності (Loudness Range – LRA),
- максимальний рівень реальних піків (Maximum True Peak Level – Max TP).

Результат цього первинного вимірювання визначає подальшу обробку.

Можливо кілька різних сценаріїв.

Перший – усі три параметри нормальні (рис.3.4). Це ідеальний результат вимірювання: рівень гучності програми -23.0 LUFS, діапазон гучності LRA в межах, визначених мовником (залежно від жанру та/або платформи розповсюдження), а рівень піків Max TP дорівнює або нижче максимального значення для наміченої системи поширення.



Рисунок 3.4 – Перший сценарій

Другий – рівень гучності програми вище -23.0 LUFS (рис.3.5).



Рисунок 3.5 – Другий сценарій

Цю проблему вирішує проста операція регулювання посилення (зниження рівня):

$$Gain(dB) = -(L_{kв} - L_{кц}), \quad (3.1)$$

де $L_{кв}$ – вимірний рівень гучності;

$L_{кц}$ – цільовий рівень гучності.

Приклад: вимірний $L_{кв} = -19,4$ LUFS; цільовий рівень $L_{кц} = -23,0$ LUFS.

Тоді необхідне посилення буде

$$Gain(dB) = -(-19,4 - (-23,0)) = 3,6 \text{ дБ}. \quad (3.2)$$

Треба мати на увазі, що пікове значення Max TP зменшується на ту саму величину, що і L_k .

Третій – рівень гучності програми нижче -23.0 LUFS (рис.3.6).



Рисунок 3.6 – Третій сценарій

Після підвищення посилення на величину (3.1) слід перерахувати максимальний рівень реальних піків Max TP по формулі

$$MaxTP_{New} = MaxTP_{Old} + Gain(dB), \quad (3.3)$$

оскільки він потенційно знаходиться вище допустимого. Якщо новий Max TP (MaxTP_{New}) перевищує допустиму межу, необхідно виконати обмеження реальних пікових значень, яке має бути виконане згідно з уніфікованим блоком обробки реальних піків (рис.3.3).

Інше рішення, якщо таке обмеження неможливе чи небажано (або потенційно занадто серйозне) – це залишити L_k на початковому низькому рівні і застосувати відповідну установку метаданих гучності (нижче -23, відображаючи вихідний рівень гучності). Це вимагає повнофункціональної системи, що підтримує метаданні (наприклад, Dolby Digital або MPEG 4).

Четвертий – гучність програми нижче -23.0 LUFS, а діапазон гучності ширше внутрішнього допуску для жанру чи каналу поширення (рис.3.7).



Рисунок 3.7 – Четвертий сценарій

Рівень гучності програми L_k можна обробляти відповідно до третього сценарію. Діапазон гучності підлягає обробці компонувальним блоком LRA (рис.3.2) і таким чином потенційно знижується Max TP . Хоча Max TP може перевищувати допустимий ліміт при застосуванні позитивного зсуву посилення до L_k , обробка Max TP може не вимагатися через зниження LRA. Тому необхідно обчислення Max TP під час процесу зниження LRA.

П'ятий – діапазон гучності ширший за допуск для жанру або каналу поширення.



Рисунок 3.7 – П'ятий сценарій

Як показано в розділі 2, компресор з низьким порогом і дуже помірним коефіцієнтом може використовуватися для зменшення LRA (уніфікований

блок діапазону гучності, рис.3.2). Для файлів мають перевагу автоматичні процеси з «цільовим LRA» (Target-LRA). Альтернативно, результат вимірювання LRA може активувати попереднє встановлення динамічного компресора з параметрами, аналогічними переліченим у пп. 2.2.

Max TP може лише зменшитися, тому потенціалу для обробки реальних пікових значень немає.

Шостий – максимальний рівень реальних піків перевищено (рис.3.8).



Рисунок 3.8 – Шостий сценарій

Перевищення рівня -1 dBTP тягне за собою ризик подальших спотворень (наприклад, ЦАП, конвертер частоти дискретизації або кодеку зі зменшенням швидкості потоку). Згідно з уніфікованим блоком Max TP (рис.3.3), обмеження реальних пікових значень застосовується шляхом зниження Max TP. Чи істотно зміниться в результаті цього гучність програми, залежить від числа та розміру піків, яких торкнулася така обробка.

Будь-яка інша комбінація результатів первинного вимірювання L_k, LRA та Max TP входить у процеси, вже описані у вищезгаданих сценаріях.

3.2 Обробка програм різних жанрів

Нормалізації гучності кожної програми до єдиного цільового рівня (-23 LUFS) не може бути ідеальним рішенням, на це є дві причини:

- жодне об'єктивне вимірювання гучності не може дати результат, повністю відповідний слуховим відчуттям;
- завжди будуть індивідуальні вподобання слухачів.

Отже, ідеальне рішення зазвичай неможливе, оскільки сприйняття гучності відрізняється між людьми і залежить від таких факторів як вік,

стать, настрій і т.ін. Треба розуміти, що рекомендації приведення до єдиного рівня гучності забезпечує задовільне прослуховування для різноманітної суміші жанрів більшістю слухачів.

У результаті, наприклад, струнний квартет Шуберта має такий самий інтегрований рівень гучності, як і пісня Леді Гаги, саме -23 LUFS (рис.3.9).

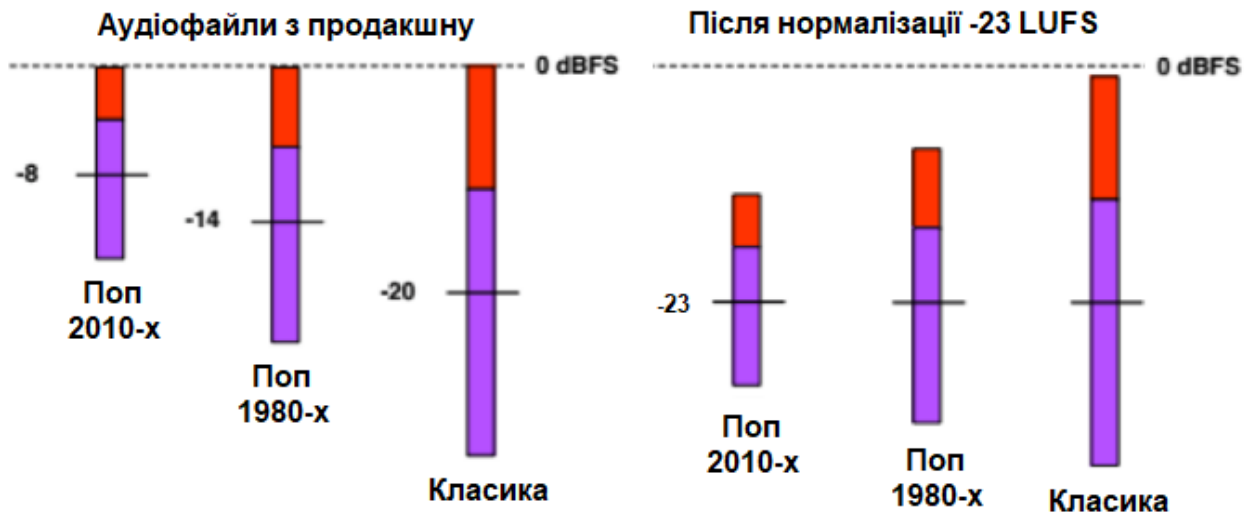


Рисунок 3.9 – Діапазон гучностей музики різних жанрів до (а) і після (б) нормалізації

Хоча це не відображає реальність, але вписує ці елементи у великий масив суміжних програм, це і є мета відстоювання єдиної величини. Слухачі згодні, щоб рівень гучності програм був у так званій «зоні комфорту» близько 8-9 LU, в той час як програми при розповсюдженні завжди асиметричні (наприклад, +3 LU/-5 LU).

У випадках, коли алгоритм об'єктивної гучності не завжди дає ідеальний результат, програма швидше за все залишиться в зоні комфорту. Мовники також повинні враховувати, що аудиторія може регулювати рівень гучності пультом відповідності зі своїми уподобаннями.

Розглянемо особливості нормалізації звука для певних жанрів. Є два жанри, де за певних обставин може бути певна трактовка (також для максимального рівня гучності): реклама та анонси, і музичні програми.

3.2.1 Рекламні вставки (кліпи) та анонси

Цей тип програм найчастіше згадується як подразник слухачів і тому несе головну відповідальність за проблеми гучності. У Великій Британії (правила BCAP – Broadcast Committee of Advertising Practice) та США (акт CALM – Commercial Advertisement Loudness Mitigation), у Німеччині та ще цілому ряді європейських країн для роботи з даним жанром останнім часом запроваджують відповідні закони.

Для контролю динаміки реклами, де існує небезпека раптових різких стрибків гучності, контроль діапазону гучності по LRA не підходить, оскільки обчислення засноване на значеннях короткочасової гучності (інтервал 3 с). Отже, для дуже коротких сплесків дуже мало точок відліку для отримання значної величини LRA.

Альтернативу можна знайти у використанні Maximum Momentary Loudness Level (Max ML – 400 мс) та/або Maximum Short-term Loudness Level (Max SL – 3 с). Ці параметри можуть ефективно використовуватись для обмеження піків гучності, особливо для коротких елементів (<30 с).

Рекомендовані значення близько +8 LU (-15 LUFS) як можливий ліміт для Max ML та +3 LU (-20 LUFS) для Max SL [30]. У будь-якому випадку обидва параметри (Max ML і Max SL) є частиною метаданих Broadcast Wave File Format (BWF) (EBU Tech Doc 3285, version 2.0, 2011 [30]).

Для програм цього жанру, що складаються лише з фонових чи творчо задуманих низькорівневих звуків (наприклад процес приготування кави), можна використовувати рівень гучності нижче за цільовий. Це узгоджується з минулою і теперішню практикою обмеження максимального рівня піків (нині: рівня гучності), але не змушує весь контент залишатися на цьому максимумі. Навмисне низькорівневий звук дає контраст, а це один із найбільш фундаментальних творчих інструментів у будь-якій формі мистецтва.

Короткий хронометраж рекламних кліпів та анонсів, де може ефективно використовуватися цей драматургічний інструмент, навряд чи може вплинути на повсякденний довготривалий середній рівень гучності станції.

Програми, призначені для відтворення нижче за цільовий рівень, вимагають особливої уваги для гарантії проходження автоматичних процесів нормалізації без спотворень. Вони мають бути винятком, а не правилом.

Зрештою, відповідальність за всі ці варіанти та рішення лежить на продюсері, режисері або іншому творчому персоналі відповідно.

3.2.2 Музика

Певні музичні програми, які містять в основному музику, можуть бути:

- з широким діапазоном гучності, наприклад, класична музика,
- або з високим ступенем динамічної компресії як художньої характеристики, наприклад, рок-концерт.

Досвід любителів музики передбачає, що ці програми мають тенденцію до прослуховування з вищим рівнем гучності (у середньому до +2-3 LU), ніж інші жанри. Причинами таких вподобань можуть бути:

- високий потенційний рівень звукового тиску 100...110 дБ (фортисимо симфонічного оркестру, рок-група з потужною звукопідсилювальною апаратурою)

– той факт, що для музики не існує "передніх" та "задніх" звуків – все знаходиться на передньому плані.

Але потенційна диференціація цільового рівня для цих програм може принести більше шкоди, відкривши назад двері до можливості стати голосніше за інших, замість покращення ситуації.

З тих же причин, що і для реклами та анонсів, нормалізація до різного або найвищого цільового рівню не рекомендується. Аудиторія може, як і раніше, користуватися пультом для регулювання (підвищення) рівня гучності

на власний смак. Суміжні програми, наприклад заходів, реклама та анонси, теж зрушаться. Очікується, що це не повинно вивести ці програми з зони комфорту.

3.2.3 Спортивні програми

Спорт – мабуть, один із найбільш проблематичних жанрів щодо встановлення рівнів та нормалізації гучності. Це зумовлено іноді непередбачуваним характером подій. Кілька голів у останні 15 хвилин футбольного матчу, наприклад, можуть значно підвищити інтегрований рівень гучності, даючи значення поза допуском ($-23.0 \text{ LUFS} \pm 1.0 \text{ LU}$).

Загалом, з цим мало що можна зробити, якщо не підготуватися до сильного впливу на динамічні властивості сигналу (за допомогою процесора гучності на виході мікшерного пульта або центральної апаратної).

У будь-якому випадку рекомендується встановити рівень гучності промови коментатора(ів) трохи нижче за цільовий рівень (на -24 LUFS , наприклад), щоб для несподіваного шуму натовпу був запас 1 LU (інтегрована гучність). Якщо подібної реакції публіки не станеться, середній рівень гучності буде нижче цільового рівня, але зазвичай у межах допуску.

Той самий принцип застосовується і в тому випадку, якщо щільність коментаря в різних програмах змінюється. В одній передачі, наприклад, можуть бути два коментатори, які говорять більшу частину часу. А в іншій – лише один, який говорить рідше і тихіше.

Якщо в другій передачі шум натовпу перевищить поріг стробування (але на кілька LU нижче за інтегрований рівень гучності), ці місця «відтягнуть» середню гучність нижче -23 LUFS , коли коментатор не говорить.

Тобто, другу програму потрібно зробити голосніше для виведення на цільовий рівень. В результаті коментарій другої програми сприйматиметься голосніше, ніж два коментатори в першій. Для того, щоб рівень коментаря в

обох випадках був однаковим, оператор/мовник може вважати другу передачу «особливою обставиною» і встановити її рівень гучності нижче -23 LUFS, ефективно здійснивши нормалізацію на базі анкера.

У спортивних передачах з досить тихою атмосферою (наприклад, гольф) відносно стробування всередині інтегрованого вимірювання виключить більшість пауз у коментарі під час рахунку гучності. Така програма повинна легко «приземлитися» у допуск навколо -23 LUFS, якщо рівень коментарів встановлено близько -23 LUFS.

3.2.4 Шоу

У розважальних передачах, наприклад, ігрових чи музичних, передбачуваність подій вище, ніж у спорті, оскільки є концепція розкадрування. Вони схожі на спорт явним анкерним сигналом: ведучий(і). Але й аудиторія завжди грає значну роль, оскільки передає емоції та ажіотаж. Тому аудиторія – не менш важливий сигнал, ніж ведучий. Отже, може бути вигідніше балансувати аудиторію біля цільового рівня – щоб ведучий «літав» вище та нижче.

Точний вибір передніх або анкерних звуків залежить від програми. Для музикальних передач, звичайно, найважливішим сигналом є музика, яка визначатиме рівень гучності програми. Ведучий, можливо, буде нижче -23 LUFS, але це добре, якщо сигнал все одно потрапляє у зону комфорту для слухача (близько +3/-5 LU від цільового рівня).

Якщо очікується, що передача, що мікшується, буде занадто живою і гучною, можна застосувати стратегію тимчасового підвищення рівня моніторингу (1...2 дБ). Це зазвичай допомагає не відхилятися занадто легко вчасно ажіотажу і не регулювати потім рівень гучності, що вийшов за межі допуску.

3.2.5 Кінофільми

Через потенційно дуже динамічний контент кінофільми, представляють найбільшу проблему в інтеграції кіно у програми з балансованою гучністю.

У багатьох випадках звукова доріжка для кінотеатру не підходить для типового домашнього середовища прослуховування. Для прикладу, на рис.3.10 наведена гістограма гучності для фільму «Матриця».

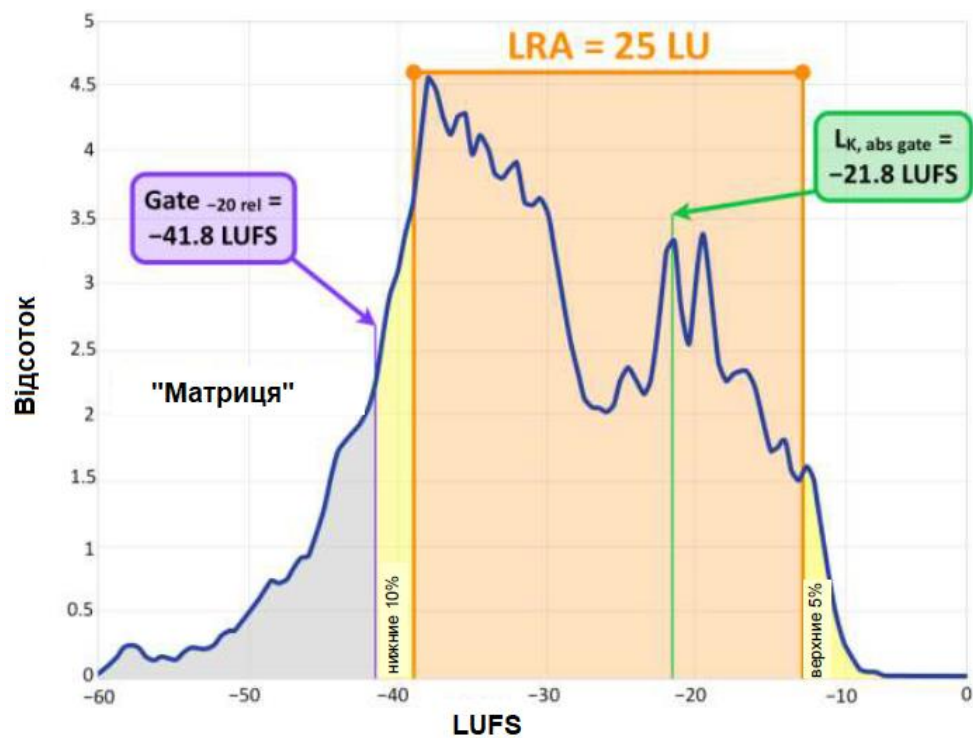


Рисунок 3.10 – Гістограма гучності для фільму «Матриця»

Звичайно, різниця між рівнем гучності програми та максимальним миттєвим або короточасним рівнем гучності є занадто великою для такого варіанта перегляду. Крім того, рівень гучності доріжка часто значно відрізняється від рівня гучності мови.

Така ситуація зустрічається у багатьох кінофільмах, але її можна зустріти і в телесеріалах. Якщо подібний контент проходить звичайний процес нормалізації гучності програм до (-23 LUFS), це може призвести до занадто низького рівня мовлення та занадто гучних звукових ефектів.

Проблема стає більш помітнішою, коли фільм переривається рекламними вставками: різниця рівня мовлення викликає значні перехідні стрибки (рис. 3.11).

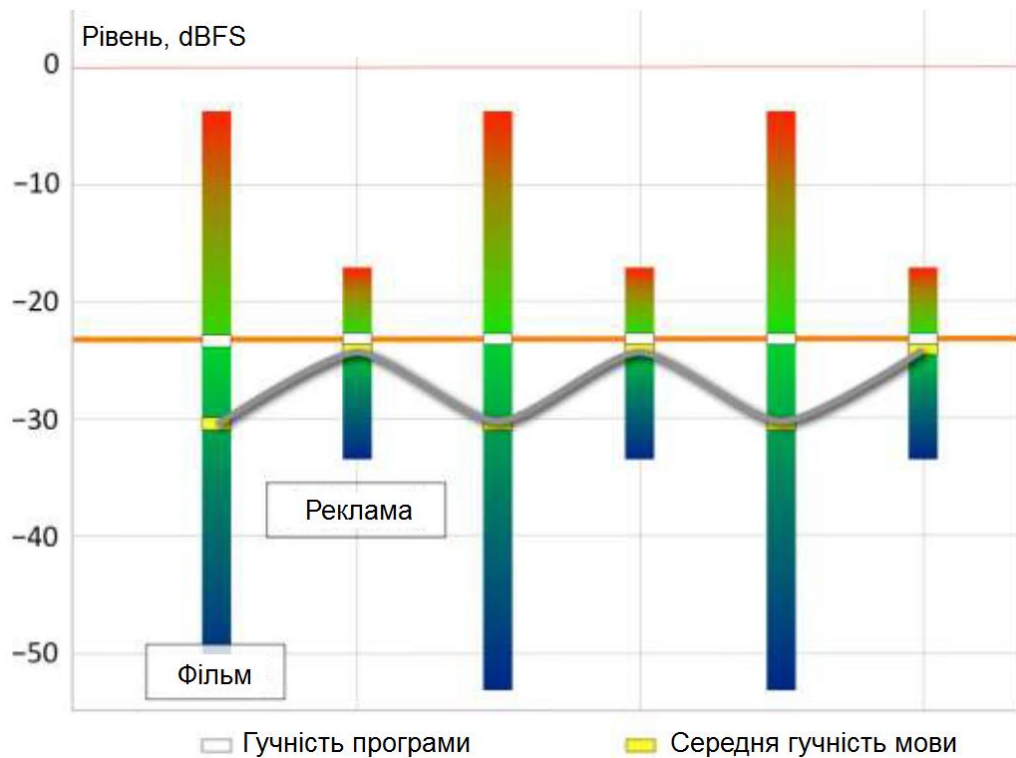


Рисунок 3.11 – Стрибки у гучності мови між фільмом та рекламою

Тому доцільно застосовувати спеціальну динамічну обробку з урахуванням рівня гучності програми, рівня гучності мови, максимального короточасного рівня гучності (та/або максимального миттєвого рівня гучності), а також діапазону гучності.

3.2.6 Використання анкерного сигналу

Для програм з широким діапазоном гучності (>20 LU) або з суттєвою різницею між рівнем гучності програми та рівнем гучності мови (приблизно $>3-5$ LU) можна опціонально використовувати так званий анкерний сигнал для нормалізації гучності, таким чином виконуючи, так би мовити, на кшталт метод стробування сигналу. Цей сигнал може бути мовою чи співом чи, наприклад, певною частиною музичної програми (рис.3.12). Такий анкерний сигнал зазвичай має менший рівень гучності, ніж гучність програми (PL).



Рисунок 3.12 – Вибір анкерного сигналу

Слід зазначити, що вибір анкерного сигналу – активний процес, що вимагає роботи досвідченого оператора. При правильному виконанні вона може допомогти точно відрегулювати гучність програм з широким діапазоном гучності за вибраним анкерним сигналом.

Існує також автоматичне вимірювання одного анкерного сигналу у формі 'Dialogue Intelligence', патентованого алгоритму Dolby Laboratories, який передбачає, що мова – загальний і важливий ний сигнал у мовленні. Алгоритм виявляє, чи є мова в програмі, і після активації вимірює лише гучність протягом мовних інтервалів.

Для програм із вузьким діапазоном гучності різниця між вимірюванням, обмеженим мовою, і вимірюванням усієї програми невелика, зазвичай <1 LU (рис.3.13).

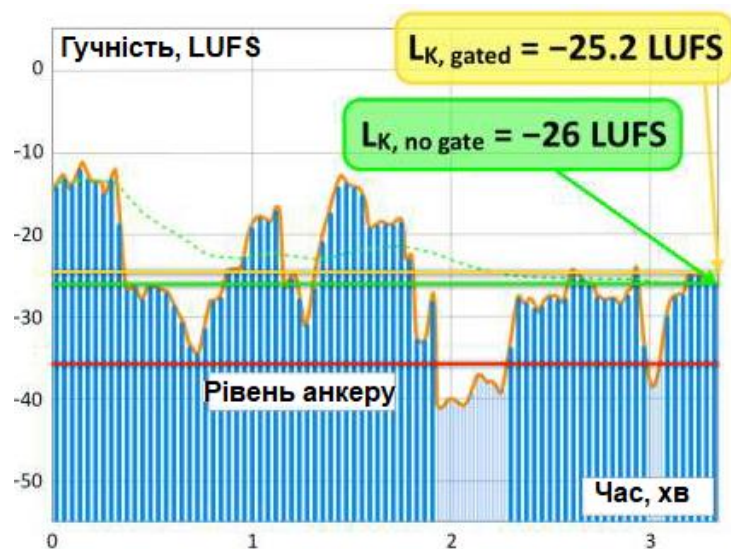


Рисунок 3.13 – Зміна рівня при виборі анкерного сигналу

Для програм з дуже широким діапазоном гучності, наприклад, художніх фільмів ця різниця потенційно збільшується, іноді перевищуючи 8 LU та більше. Автоматичне вимірювання анкерного сигналу покликане допомогти ідентифікувати, яким має бути цільовий рівень.

3.3 Висновок по розділу 3

1. Складено алгоритм процесу нормалізації гучності у медіавиробництві. Він складається з трьох основних блоків: нормалізація середньої гучності, нормалізація діапазону гучності і нормалізація істинних піків сигналу. Корекція середньої гучності здійснюється звичайним регулятором рівня, корекція діапазону гучності – динамічним компресором, корекція піків – ліміттером.

2. Розглянуті найбільш вірогідні сценарії невідповідності нормам трьох ключових параметрів: гучності, діапазону та піків. Складено алгоритм дій звукооператора для кожного із зазначених сценаріїв. Головна ідея полягає в тому, що при корекції середньої гучності вверх оператор має слідкувати, щоби діапазон і піки не вийшли за допустимі межі. В той же час, всі регулювання мають бути помірними, щоби вносити мінімальні корективи в художній замисел звукорежисера.

3. Підсумовані принципи нормалізації гучності для програм різних жанрів. Найменша складність нормалізації звуку музичних програм. Але теж є нюанси. Любителі музики усіх жанрів, від класики до рока, вважають більш комфортними гучніші рівні +2...+3 LU. Але це може бути дискомфортним для інших слухачів. Тому є концепція зведення музичних програм до стандартного рівня -23 LUFS, а у невеликої кількості любителів музики завжди є пульт керування, де вони зможуть збільшити гучність.

4. Складним жанром є реклама. Для контролю динаміки реклами, де існує небезпека раптових різких стрибків гучності, контроль діапазону гучності по LRA не підходить, оскільки обчислення засноване на значеннях

короткочасової гучності (інтервал 3 с). Для дуже коротких сплесків краще використовувати рівні Max ML з вікном – 400 мс та Max SL з вікном 3 с. Ці параметри можуть ефективно використовуватись для обмеження піків гучності, особливо для коротких елементів (<30 с). Рекомендовані значення - 15 LUFS як ліміт для Max ML та -20 LUFS для Max SL.

5. Ще складнішим жанром з точки зору гучності є шоу та спорт. В шоу масовка запрограмована, тому можна підготуватися до збільшення її рівня. В спорті, якщо гол забили за 5 хвилин до кінця матчу, шум трибун може сильно перевантажити тракт. Отже, треба залишати запас зверху на такі події.

Для програм з широким діапазоном гучності (>20 LU) або з суттєвою різницею між рівнем гучності програми та рівнем гучності мови (приблизно >3-5 LU) можна опціонально використовувати так званий анкерний сигнал для нормалізації гучності. Це, по факту, той сигнал, який має бути опорним в трансляції. В шоу – це голос ведучого, в спорті – голос коментатора. Таким чином реалізується, так би мовити, метод стробування сигналу.

6. Найскладнішим жанром є кінофільми. Для прикладу показана гістограма рівнів з фільму «Матриця». Це дуже широкий динамічний діапазон для будь-якої платформи. Особливо, коли фільм перебивається рекламою. В такому випадку найкращим рішенням буде робота з верхньою і нижньою частинами динамічного діапазону. Зверху застосовуємо м'який ліміттер, знизу – компресор ввєрх, середина діапазону гучностей залишається майже без змін.

ВИСНОВКИ

В даний час телевізійне мовлення є одним із найпопулярніших та найефективніших видів ЗМІ. Зображення на екрані передає дуже великий обсяг інформації. Але загальне враження про телевізійну передачу формується не тільки на основі самої "картинки", поряд з цим величезну роль відіграє якість звукового супроводу.

Звукова доріжка є не менш важливою складовою телевізійної та відеотрансляції, ніж відеоряд. Найбільше роздратування у глядачів викликає необхідність регулювати рівень при переключенні каналів на ТВ або в Інтернеті, а також стрибки гучності всередині однієї програми. Зона комфорту середньої гучності складає всього 8 дБ. У той же час необроблений контент має динамічний діапазон більше 20 дБ.

Гучність звуку є результатом суб'єктивного сприйняття. Крім рівня звукового тиску гучність залежить від спектрального складу сигналу та часових характеристик його обвідної. Тому класичне вимірювання рівня часто не відповідає реальній картини зі сприйняттям гучності. Традиційне лімітування або компресування, які часто використовуються як інструменти регулювання гучності, досить примітивні і не враховують психоакустичних моделей сприйняття.

Метою кваліфікаційної роботи є аналіз методів вимірювання і нормалізації гучності телевізійного або звукового мовлення з метою формулювання зрозумілих рекомендацій для робітників медіаіндустрії, студентів та випускників відповідних навчальних закладів.

В першому розділі виконано аналіз задачі вимірювання і нормалізації гучності в телевізійних і радіотрансляціях. Важливими характеристиками класичних індикаторів рівня є час інтегрування і час повернення. За цими показниками існує широкий спектр індикаторів – від істинно-пікових (True Pick) до середньоквадратичних RMS з часом інтегрування одиниці секунд. Найпоширенішими є VU індикатори з постійною часу інтегрування

300...400 мс. Широка номенклатура індикаторів рівня призводить до появи різного роду комбінованих індикаторів.

Пристрої динамічного регулювання рівня, аналогічно звукорежисеру, використовують інформацію про вимірний рівень сигналу для адаптивної зміни коефіцієнту передачі тракту. При цьому результат їхньої роботи, природно, буде залежати від часових параметрів (Attack і Release) детектора рівня.

Динамічні процесори – нормалізатори, компресори, ліміттери – давно використовуються в радіомовленні. Їх задача – уникнути перевантаження і перемодуляції передавача. Однак часто їх дія призводить до суттєвого зменшення динамічного діапазону трансляції. Якщо для радіопередач, де основний контент – це музика – таке звуження підходить, то для ТВ, де контент набагато різноманітніший, дане рішення не задовольнить.

У другому розділі проведено детальний аналіз методів нормалізації гучності звуку в медіавиробництві.

Поширена до останнього часу концепція нормалізації піків звукового сигналу до максимально допустимого рівня призвела до однорідних пікових рівнів програм, але дуже різних рівнів гучності. Фактичне коливання залежить від рівня динамічної компресії сигналу. Дана ситуація добре відома по рекламним вставкам, коли звукорежисери перетискають сигнал, сильно збільшуючи його середній рівень.

На відміну від цього, нормалізація гучності дозволяє досягти однакової середньої гучності програм. Зі зміною піків залежно від контенту, а також художніх та технічних вимог, слухач може насолоджуватися однорідним середнім рівнем гучності у всіх програмах, не користуючись пультом для частого регулювання гучності.

Рекомендацією EBU 128 введені одиниці гучності LUFS, які враховують частотні характеристики слуху. По причині, що звуковий супровід різних жанрів має різний пік-фактор, середня гучність різного контенту буде відмінною. Для вирівнювання суб'єктивної гучності звуку

різних жанрів була запропонована цільова величина гучності -23 LUFS, яка враховує пік-фактори різного контенту.

По суті є два способи нормалізації гучності: нормалізація аудіосигналу на передавальній стороні, або використання метаданих гучності, а регулювання відбувається на стороні споживача. Нормалізація гучності під час виробництва програм рекомендується через простоту і потенційне підвищення якості.

Реально у виробництві мовник стикається з необхідністю нормалізації різного контенту із різних джерел. Йому може потрапляти багато звукових програм без нормалізації гучності. Для таких програм потрібно мати або автоматичну програмну нормалізацію відразу після завантаження в ефірний сервер, або встановлення процесора гучності на виході центральної апаратної.

Але процесор має обробляти тільки ті сигнали, що не нормалізовані на цільовий рівень -23 LUFS. Отже, система мовлення має сигналізувати процесору, коли відтворюється сумісний за гучністю контент, процесор повинен перейти в режим байпас. В іншому випадку процесор може перекомпресувати сигнал.

При переході на нові одиниці вимірювання гучності у звукооператора є два варіанти: або переходити на LUFS-індикатор відразу, використовуючи його при зведенні, або зводити по-старому, по піках, а потім нормалізувати до цільового рівня -23 LUFS. Перший спосіб є більш прийнятним, оскільки залишає при зведенні більше простору зверху для збереження динамічного діапазону.

Щоби отримати уявлення про те, де знаходяться рівні і яку різниця рівнів LUFS і дБ можна очікувати було зроблено експеримент. Була взята фонограма «Дольче Габбана» з офіційного альбома В. Сердючки, фонограма зведена по піковому рівню -5 дБ. Після нормалізації по гучності -23 LUFS, загальний рівень зменшився приблизно на 10 дБ. Для матеріалу іншого жанру, така різниця буде іншою. Аналогічний експеримент зі звукозаписом

чоловічого голосу показав зменшення рівня на 2 дБ. Така різниця обумовлена пік-фактором і спектральним складом сигналу. Саме на них реагує наш слух, сприймаючи гучність.

Розглянуті шкали і часові характеристики LUFS-метрів. Для зручності багато програмних продуктів перейшли на спрощену шкалу LU, де $-23 \text{ LUFS} = 0 \text{ LU}$. Існує дві шкали: $+9$ – для більшості програм, і з розширеним діапазоном гучності $+18$ – для кінофільмів.

Крім цільового середнього рівня гучності різні канали або Інтернет-платформи нормують значення діапазону гучності (Loudness Range). Цей параметр допомагає звукооператору вирішити, чи потрібна подальша динамічна обробка. Тобто можна визначати необхідні заходи для того, щоб вмістити програму в вікно допуску каналу чи платформи поширення. Успішно вирішити дану задачу допомагає компресор із низьким порогом (-50 dBFS) та помірним коефіцієнтом стиснення (1:1.5).

Складено алгоритм процесу нормалізації гучності у медіавиробництві. Він складається з трьох основних блоків: нормалізація середньої гучності, нормалізація діапазону гучності і нормалізація істинних піків сигналу. Корекція середньої гучності здійснюється звичайним регулятором рівня, корекція діапазону гучності – динамічним компресором, корекція піків – ліміттером.

Розглянуті найбільш вірогідні сценарії невідповідності нормам трьох ключових параметрів: гучності, діапазону та піків. Складено алгоритм дій звукооператора для кожного із зазначених сценаріїв. Головна ідея полягає в тому, що при корекції середньої гучності вверх оператор має слідкувати, щоби діапазон і піки не вийшли за допустимі межі. В той же час, всі регулювання мають бути помірними, щоби вносити мінімальні корективи в художній замисел звукорежисера.

Підсумовані принципи нормалізації гучності для програм різних жанрів. Найменша складність нормалізації звуку музичних програм. Але теж є нюанси. Любителі музики усіх жанрів, від класики до рока, вважають більш

комфортними гучніші рівні +2...+3 LU. Але це може бути дискомфортним для інших слухачів. Тому є концепція зведення музичних програм до стандартного рівня -23 LUFS, а у невеликої кількості любителів музики завжди є пульт керування, де вони зможуть збільшити гучність.

Складним жанром є реклама. Для контролю динаміки реклами, де існує небезпека раптових різких стрибків гучності, контроль діапазону гучності по LRA не підходить, оскільки обчислення засноване на значеннях короткочасової гучності (інтервал 3 с). Для дуже коротких сплесків краще використовувати рівні Max ML з вікном – 400 мс та Max SL з вікном 3 с. Ці параметри можуть ефективно використовуватись для обмеження піків гучності, особливо для коротких елементів (<30 с). Рекомендовані значення - 15 LUFS як ліміт для Max ML та -20 LUFS для Max SL.

Ще складнішим жанром з точки зору гучності є шоу та спорт. В шоу масовка запрограмована, тому можна підготуватися до збільшення її рівня. В спорті, якщо гол забили за 5 хвилин до кінця матчу, шум трибун може сильно перевантажити тракт. Отже, треба залишати запас зверху на такі події.

Для програм з широким діапазоном гучності (>20 LU) або з суттєвою різницею між рівнем гучності програми та рівнем гучності мови (приблизно >3-5 LU) можна опціонально використовувати так званий анкерний сигнал для нормалізації гучності. Це, по факту, той сигнал, який має бути опорним в трансляції. В шоу – це голос ведучого, в спорті – голос коментатора. Таким чином реалізується, так би мовити, метод стробування сигналу.

Найскладнішим жанром є кінофільми. Для прикладу показана гістограма рівнів з фільму «Матриця». Це дуже широкий динамічний діапазон для будь-якої платформи. Особливо, коли фільм перебивається рекламою. В такому випадку найкращим рішенням буде робота з верхньою і нижньою частинами динамічного діапазону. Зверху застосовуємо м'який ліміттер, знизу – компресор вверх, середина діапазону гучностей залишається майже без змін.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Bartholomew Elias. Sound Basics: A Primer in Psychoacoustics. – United States Air Force Research Laboratory, 1998. – 198 p.
2. Ballou G. Handbook for Sound Engineers (Fourth ed.). – Burlington: Focal Press, 2008. – 43 p.
3. Christopher J. Plack. The Sense of Hearing. Lawrence Erlbaum Associates, 2005. – 267 p.
4. ISO 226:1987. Acoustics. Normal equal-loudness level contours.
5. ISO 226:2003. Acoustics. Normal equal-loudness level contours.
6. Alec Nisbett. Sound Studio: Audio techniques for Radio, Television, Film and Recording. 7th Edition, 2004. – 400 p.
7. Walter Fischer. Digital video and audio broadcasting technology. A practical engineering guide. Third Edition. Shpringer, 2010. – 828 p.
8. Michael C. Keith. The radio station. 6th ed. Burlington, MA: Elsevier/Focal Press, 2004. – 188 p.
9. Complete broadcasting industry guide: television, radio, cable & streaming. Amenia, NY: Grey House Publishing, 2020. – 732 p.
10. Mohammed S. Nick, Global radio: from shortwave to streaming. Lanham, Maryland : Lexington Books, 2019. – 137 p.
11. Amanda D. Lotz. The television will be revolutionized. Second edition. New York : New York University Press, 2014. – 331 p.
12. Radio's second century : past, present, and future perspectives / edited by John Allen Hendricks. New Brunswick, New Jersey: Rutgers University Press, 2020. – 301 p.
13. The Routledge companion to radio and podcast studies / edited by Mia Lindgren and Jason Loviglio. Abingdon, Oxon ; New York, NY: Routledge, 2022. – 387 c.
14. В.Н. Олейников, О.В. Зубков, В.М. Карташов, И.В. Корытцев, С.И. Бабкин, С.А. Шейко, И.С. Селезнев. Экспериментальная оценка

эффективности алгоритмов пеленгования беспилотных летательных аппаратов по акустическому излучению. Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2019. – Вып. 199. – С. 29 – 37.

15.V. Kartashov, V. Oleynikov, I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin, I. Selieznov. Use of Acoustic Signature for Detection, Recognition and Direction Finding of Small Unmanned Aerial Vehicles. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). 2020. 4 p.

16.Kartashov V.M., Oleynikov V.N, Zubkov O.V., Koryttsev I.V., Babkin S. I., Sheiko S.A., Kolendovskaya M.M. Spatial-temporal Processing of acoustic Signals of Unmanned Aerial Vehicles/ Telecommunications and Radio Engineering. – New York. – 2020. – Vol. 79, №9. – P.769-780.

17.V. Kartashov, V. Oleynikov , I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin. Processing of Wide Band Acoustic Signals During Detection of Unmanned Aerial Vehicles // 2020 IEEE Ukrainian Microwave Week (UkrMW). Kharkiv, Ukraine, September 21 - 25, 2020. Volume 1 on 2020 IEEE 12th International Conference on Antenna Theory and Techniques (ICATT). pp. 35-39.

18.V.M. Kartashov, G.I. Sidorov, S.A. Sheiko, M.M. Kolendovskaya, O.Yu. Sergienko. Principles of construction and assessment of technical characteristics of multi-frequency atmospheric sodar in the humidity measurement mode. Telecommunications and Radio Engineering. Vol. 79. N.4. 2020. – pp. 323-333.

19.S. Sheiko. Study of the method for assessing atmospheric turbulence by the envelope of sodar signals // Eastern-European Journal of Enterprise Technologies. – 2/5 (92). – April, 2018. – p. 33–40.

20.Сідоров Г.І., Шейко С.О., Шаповалов С.В., Полонська А.С., Дмитренко А.І. Акустичний метод вимірювання турбулентного стану атмосферного прикордонного шару // Радиотехніка: Всеукр. міжвід. наук.-техн. зб. 2018. – Вип. 192. – С. 46–50.

21.Valerii V. Semenets, V. M. Kartashov, V. I. Leonidov. Registration of refraction phenomenon in the problem of acoustic sounding of atmosphere in

airports zone. Telecommunications and Radio Engineering. Volume 77, Issue 5, 2018. – P. 461-468.

22. Бабак К.В. Технічні аспекти створення електронної музичної композиції // 27-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті». Зб. матеріалів форуму. Т. 3. – Харків: ХНУРЕ. 2023. – с. 57-58.

23. Свірідок М.С. Технічні аспекти створення музичної композиції // 27-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті». Зб. матеріалів форуму. Т. 3. – Харків: ХНУРЕ. 2023. – с. 104-105.

24. Курдиш В.В. Алгоритм синхронізації звуку і відео в інтерв'ю // 27-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті». Зб. матеріалів форуму. Т. 3. – Харків: ХНУРЕ. 2023. – с. 129-130.

25. Древальський Р.В. Дослідження методу корекції звуку для компенсації впливу приміщення /25-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті». Зб. матеріалів форуму. Т. 3. – Харків: ХНУРЕ. 2021. – с. 119-120.

26. Удовік Д.В. Дослідження методів зменшення еквівалентної реверберації в звукозаписі: кваліфікаційна робота на здобуття освітнього ступеня магістра. – Х.: ХНУРЕ. – 2022 р. – 65 с.

27. Тарусін В.Ю. Дослідження методів компенсації спотворень в звукових трактах: кваліфікаційна робота на здобуття освітнього ступеня магістра. – Х.: ХНУРЕ. – 2022 р. – 78 с.

28. Мезенцев І.О. Дослідження алгоритмів автоматизованої еквалізації звукозапису голосу: кваліфікаційна робота на здобуття освітнього ступеня магістра. – Х.: ХНУРЕ. – 2022 р. – 69 с.

29. EBU Tech Doc 3341. Loudness Metering: 'EBU Mode' Metering to Supplement EBU R 128 Loudness Normalization.

30. EBU Tech Doc 3285. Specification of the Broadcast Wave Format (BWF). A format for audio data files in broadcasting.