

Об'єктивне та суб'єктивне оцінювання якості музичних сигналів обмежених смугою частот

Котвицький І. В., ORCID [0000-0003-2800-0080](https://orcid.org/0000-0003-2800-0080)

e-mail igorktvzk@gmail.com

Оса А. В., ORCID [0000-0002-5973-6288](https://orcid.org/0000-0002-5973-6288)

e-mail andriibee796@gmail.com

Національний технічний університет України

"Київський політехнічний інститут імені Ігоря Сікорського" kpi.ua

Київ, Україна

Реферат—В роботі виконано аналіз особливостей застосування деяких об'єктивних показників якості звукових сигналів, таких як сегментне співвідношення сигнал-шум, логарифмічно-спектральні спотворення і барк-спектральні спотворення для оцінки якості музичних сигналів з обмеженою смугою частот. Отримані результати об'єктивного оцінювання зіставлені з оцінками якості, отриманими шляхом суб'єктивного тестування.

Бібл. 9, рис. 3.

Ключові слова — якість музичного сигналу; об'єктивна оцінка; суб'єктивна оцінка

І. ВСТУП

В роботі [1] шляхом обчислення коефіцієнтів кореляції та побудови карт відповідності, зіставлені результати об'єктивного і суб'єктивного оцінювання якості мовних сигналів, спотворених обмеженням смуги частот. При цьому передбачалося, що мовні сигнали займають смугу частот, що не перевищує 11 кГц. Але, для якісного відтворення музичних сигналів необхідний значно ширший частотний діапазон, що досягає 20 кГц. У зв'язку з цим становлять інтерес дослідження відповідності суб'єктивних і об'єктивних показників якості для музичних сигналів, що є основною метою даної роботи.

Деякі особливості двох простих в обчислювальному плані мір якості, таких як сегментне відношення сигнал-шум SSNR (Segmental Signal to Noise Ratio) і логарифмічно-спектральні спотворення LSD (Logarithmic Spectral Distortion), розглянуті в роботах [2]–[7]. При цьому показано, що особливості не монотонності розподілу показників LSD та SSNR можуть призводити до отримання результатів, які погано узгоджуються з результатами суб'єктивних випробувань. Оскільки в згаданих роботах аналізу піддавалися лише мовні сигнали, іншою метою даної роботи є дослідження особливостей показників SSNR і LSD стосовно музичних сигналів.

ІІ. СУБ'ЄКТИВНЕ І ОБ'ЄКТИВНЕ ОЦІНЮВАННЯ ЯКОСТІ СИГНАЛІВ

У даній роботі використаний метод суб'єктивного тестування, іменованій «оцінкою ступеня деградації» (Degradation Category Rating – DCR). Відповідно до цього методу, слухачі на слух порівнюють якість спотвореного і еталонного мовних сигналів, виставляючи оцінки за п'ятибальною шкалою, іменованою

«шкалою середньої деградації» (Degradation Mean Opinion Score – DMOS) [8].

В якості об'єктивних заходів якості мовного сигналу використані сегментне відношення сигнал-шум SSNR, логарифмічні спектральні спотворення LSD і барк-спектральні спотворення BSD (Bark Spectral Distortion) [3].

$$SSNR = \frac{1}{M} \sum_{m=1}^M 10 \lg \left[\frac{\sum_{n=R(m-1)+1}^{Rm} x^2(n, m)}{\sum_{n=R(m-1)+1}^{Rm} [x(n, m) - y(n, m)]^2} \right],$$

де $x(n, m)$ та $y(n, m)$ – n -ті вибірки m -го фрейму чистого сигналу $x(n)$ і спотвореного сигналу $y(n)$, відповідно; M – кількість фреймів; R – кількість вибірок в фреймі.

Міра LSD належить до групи показників частотної області:

$$LSD = \frac{2}{RM} \sum_{m=1}^M \sum_{j=1}^J |G\{X(j, m)\} - G\{Y(j, m)\}|,$$

$$G\{X(j, m)\} = \max\{20 \lg(|X(j, m)|), \delta\},$$

$$\delta = \max_{l, k} \{20 \lg(|X(l, k)|)\} - 50,$$

де $X(j, m)$ та $Y(j, m)$ – дискретні перетворення Фур'є m -го фрейму сигналів $x(n)$ і $y(n)$, відповідно, j – номер частотної вибірки, J – кількість частотних вибірок.



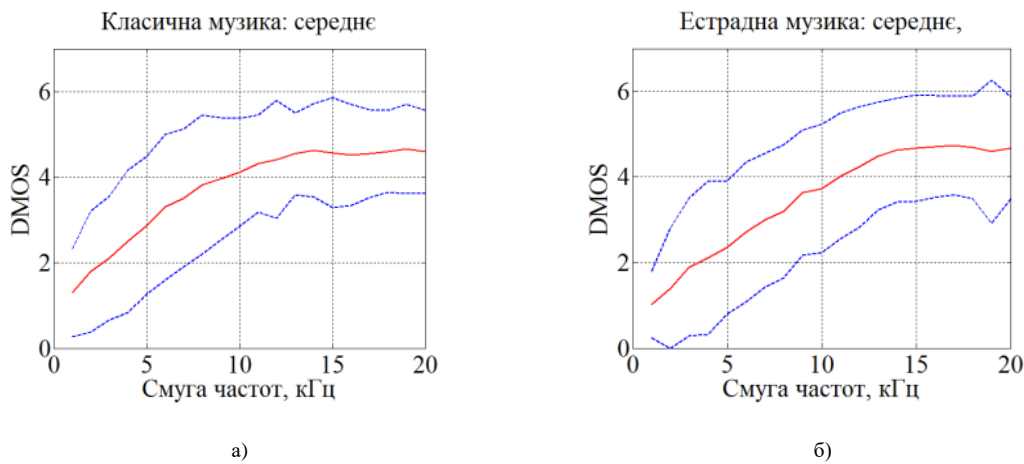


Рис. 1 Результати суб'єктивного оцінювання за шкалою DMOS

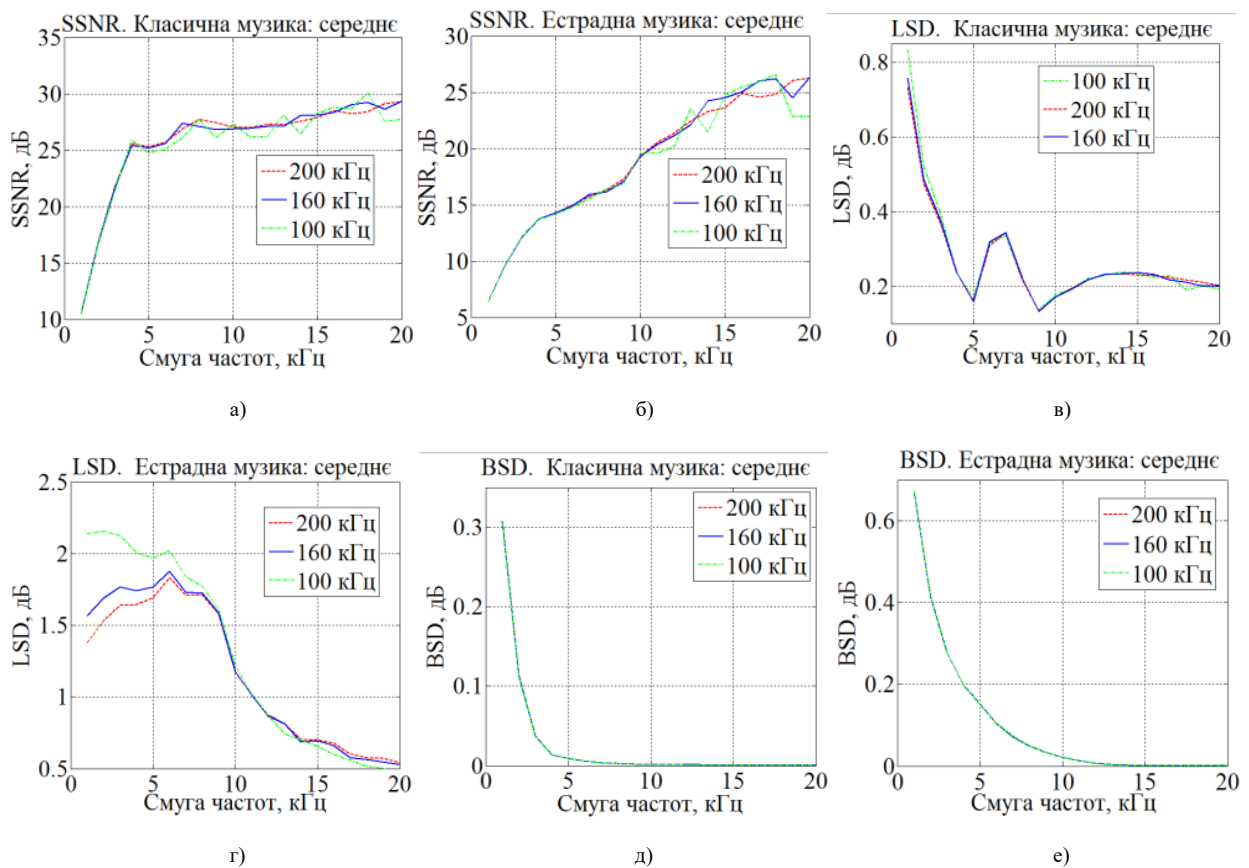


Рис. 2 Залежності SSNR (Δf) (а) та (б), LSD (Δf) (в) та (г), BSD (Δf) (д) та (е)

Міра BSD відноситься до групи перцептуальних показників, які враховують особливості слухової системи людини:

$$BSD = \frac{\sum_{m=1}^M \sum_{k=1}^K [B_x(k, m) - B_y(k, m)]^2}{\sum_{m=1}^M \sum_{k=1}^K [B_x(k, m)]^2},$$

де $B_x(k, m)$ і $B_y(k, m)$ – барк-спектри m -го фрейму сигналів $x(n)$ і $y(n)$, відповідно, k – номер критичної смуги частот.

III. ОРГАНІЗАЦІЯ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

При експериментальних дослідженнях використані 16-секундні фрагменти музичних творів категорій «естрадна» і «класична» музика. Вимоги до



формату файлів – WAVE з частотою дискретизації 44,1 кГц. Вимоги до обладнання – ноутбук із зовнішнім аудіо інтерфейсом, що підтримує програвання сигналу з частотою дискретизації 44,1 кГц та навушники.

Нижче наведені назви файлів і відповідних композицій.

A. Зразки класичної музики:

- Бах - Концерт Мі мажор (Алегро) # 1;
- Бах - Концерт Ре мажор (Алегро) # 1;
- Бах - Концерт Ре мажор (Алегро) # 3;
- Бах - Концерт Ре мінор (Алегро) # 3.

B. Зразки естрадної музики:

- 3 Doors Down - I'm Here without you baby;
- Hoobastank - Reason;
- Stop and stay - One republic;
- Lived a lie - You me at six.

Сптворені фрагменти музичних сигналів формувалися шляхом почергового пропущення їх через нерекурсивні фільтри низької частоти, синтезовані методом Ремеза [9]. Характеристики фільтрів:

- частота зрізу змінювалася від 1 до 20 кГц з кроком 1 кГц;
- розмір перехідної зони становив 5% від смуги пропускання;
- нерівномірність АЧХ в смузі пропускання становила 1 дБ;
- загасання в смузі придушення становило мінус 70 дБ.

Суб'єктивне оцінювання якості музичного сигналу виконувалося 30-ю юнаками і дівчатами у віці 20-25 років.

IV. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

На рис. 1 наведені усереднені результати (суцільна червона лінія) суб'єктивного оцінювання, які свідчать про практично монотонне зростання якості музичного сигналу зі збільшенням смуги частот. Штриховими синіми лініями показані межі 95%-го довірчого інтервалу.

З наведених результатів випливає, що для слуху людини досить смуги частот 14-15 кГц, в області якої якість музичних композицій має найвищу оцінку за шкалою DMOS. При цьому слід зазначити значний розкид думок слухачів щодо якості матеріалу, який прослуховується.

Результати об'єктивного оцінювання якості цих же музичних сигналів наведені на рис. 2. З огляду на зазначену в [5] чутливість міри SSNR до значення частоти дискретизації, в даній роботі дослідженню

піддавалися як сигнали з вихідною частотою дискретизації 44,1 кГц, так і ці ж сигнали, піддані переддискретизації до частот 100 кГц, 160 кГц і 200 кГц. Як випливає з наведених результатів, підвищення частоти дискретизації благотворно впливає на оцінки показника SSNR, дозволяючи усунути алогічні порушення монотонності залежності SSNR від смуги частот Δf . Дещо несподіваним виявився факт позитивного впливу процедури інтерполяції на точність вимірювань показника LSD: завдяки підвищенню частоти дискретизації вдалося виявити значне порушення (для значень смуги частот 1-5 кГц), монотонної залежності кривої LSD від смуги частот Δf . Що стосується оцінок показника BSD, чутливість до частоти дискретизації не виявлена.

Як випливає з графіків рис. 2, всі розглянуті об'єктивні міри виявилися чутливими до спектральних особливостей музичних сигналів. Проте поведінка графіків $LSD(\Delta f)$ відрізняється вираженими порушеннями монотонності. Так, для класичної музики характерно порушення монотонності залежності $LSD(\Delta f)$ в інтервалі значень $\Delta f = 0,5-1,4$ кГц, а для естрадної музики — в інтервалі $\Delta f = 1-5$ кГц. При цьому графік $LSD(\Delta f)$ для класичної музики практично не залежить від частоти дискретизації, тоді як у випадку естрадної музики форма $LSD(\Delta f)$ істотно залежить від частоти дискретизації. Дослідження показали, що однією з причин порушення монотонності залежності $LSD(\Delta f)$ є не ідеально прямокутна форма АЧХ використовуваних НЧ фільтрів. Дійсно, наведені на рис. 3 графіки, отримані при виконанні фільтрації в частотній області з АЧХ ідеально прямокутної форми, є монотонними.

Відзначимо, що при використанні НЧ фільтрів з прямокутними АЧХ значення $LSD(\Delta f)$ також виявилися залежними від вибору частоти дискретизації, як і при використанні «неідеальних» НЧ фільтрів. Причини цього феномену доцільно дослідити у майбутньому.

Зауважимо, що хоча в суб'єктивних випробуваннях брали участь 30 осіб, дану вибірку навряд чи можна вважати представницькою, тому у подальшому її обсяг слід збільшити. Крім того, досить обмеженою є вибірка музичних творів, що тестувалися. З огляду на перелічені особливості, отримані в даній роботі, результати слід розглядати як попередні.

ВИСНОВКИ

Отримано залежності об'єктивних і суб'єктивних оцінок якості музичних сигналів, обмежених за смугою частот, від ширини цієї смуги частот. Показано, що інтерполяція сигналів, що тестуються, дозволяє підвищити достовірність оцінювання з використанням показників SSNR і LSD. Крім того, показано, що однією з можливих причин порушення монотонності залежності LSD від смуги частот є не ідеальна прямокутна форма АЧХ використаних НЧ фільтрів.

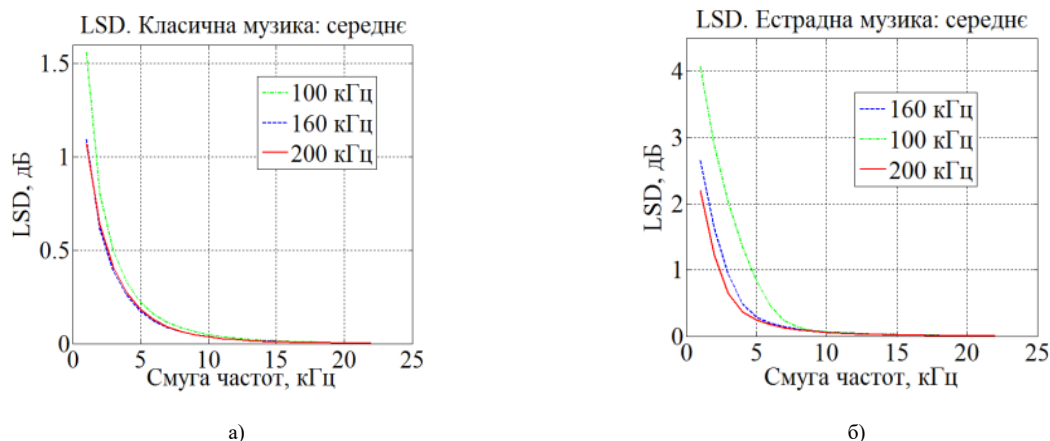


Рис. 3 Залежність LSD (Δf) для фільтрів з прямокутною АЧХ

Надалі доцільно збільшити обсяги вибірок музичних творів, збільшити кількість слухачів та провести апроксимацію отриманих карт відповідності суб'єктивних показників за шкалою DMOS, що отримані експериментально та об'єктивних показників, що досліджувались в цій роботі

ПОДЯКА

Особлива подяка моєму науковому керівнику професору [Аркадію Миколайовичу Пролеусу](#). Також хочу подякувати студентам за участь у суб'єктивних дослідженнях.

ПЕРЕЛІК ПОСИЛАНЬ

- [1] Y. O. Mytai, K. S. Zamsha, B. V. Lozinskiy, O. S. Stepanovska, and A. M. Prodeus, "Objective and subjective assessment of bandlimited signaling speech quality," *Electron. Commun.*, vol. 21, no. 1, pp. 18–26, Nov. 2016. DOI: [10.20535/2312-1807.2016.21.1.82250](https://doi.org/10.20535/2312-1807.2016.21.1.82250)
- [2] A. N. Prodeus, V. S. Didkovskiy, and I. V. Kotvytskiy, "Pro otsinku yakosi movlennia z vykorystanniam indikatoriv loharifmichnykh spektralnykh spotvoren [On the estimation of speech quality using the indicator logarithmic spectral distortions]," in *5th International Scientific and Technical Conference Information Problems in the Theory of Acoustic, Radioelectronic and Telecommunication Systems, IPST-2016*, 2016, p. 14.
- [3] A. Prodeus, "On Some Features of Log-Spectral Distortion as Speech Quality Measure," *Autom. Softw. Dev. Eng. J.*, vol. 1, 2016. URL: <https://asdej.xyz/asde-journal-vol-1/>
- [4] N. V. Bogdanova and A. M. Prodeus, "Objective quality evaluation of speech band-limited signals," *Electron. Commun.*, vol. 19, no. 6, pp. 58–65, Dec. 2014. DOI: [10.20535/2312-1807.2014.19.6.113479](https://doi.org/10.20535/2312-1807.2014.19.6.113479)
- [5] A. Prodeus, "Reducing Sensitivity of Segmental Signal-to-Noise Ratio Estimator to Time-Alignment Error," *Int. J. Electr. Electron. Sci.*, vol. 2, no. 2, pp. 31–36, 2015. URL: <http://www.aascit.org/journal/archive2?journalId=915&paperId=2822>
- [6] S. R. Quackenbush, T. P. Barnwell, and M. A. Clements, *Objective measures of speech quality*. Englewood Cliffs, N.J.: Prentice Hall, 1988. ISBN: [978-0136290568](https://www.isbn-international.org/product/978-0136290568)
- [7] N. V. Bogdanova and A. M. Prodeus, "Otsinka vplivu nelineinosti fazovoi chastotnoi kharakterystyky traktu na yakist muzychnykh signaliv [Assessment of the impact of system phase response non-linearity on the music signals quality]," *Electron. Commun.*, vol. 20, no. 4, pp. 29–35, 2015. DOI: [10.20535/2312-1807.2015.20.4.69901](https://doi.org/10.20535/2312-1807.2015.20.4.69901)
- [8] N. Côté, *Integral and Diagnostic Intrusive Prediction of Speech Quality*. Springer-Verlag Berlin Heidelberg, 2011. ISBN: [978-3-642-18462-8](https://www.isbn-international.org/product/978-3-642-18462-8)
- [9] "Matlab Auditory Toolbox (version 2.0)." [Online]. Available: https://www.mathworks.com/matlabcentral/newsreader/view_thread/8417.

Надійшла до редакції 02 вересня 2017 року

УДК 004.934

Объективное и субъективное оценивание качества музыкальных сигналов ограниченных полосой частот

Котвицкий И. В., ORCID [0000-0003-2800-0080](https://orcid.org/0000-0003-2800-0080)
e-mail igorktvzk@gmail.com



Оса А. В., ORCID [0000-0002-5973-6288](https://orcid.org/0000-0002-5973-6288)

e-mail andriibee796@gmail.com

Национальный технический университет Украины

"Киевский политехнический институт имени Игоря Сикорского" kpi.ua

Киев, Украина

Реферат—В работе выполнен анализ особенностей применения некоторых объективных показателей качества звуковых сигналов, таких как сегментное отношение сигнал-шум, логарифмически-спектральные искажения и барк-спектральные искажения для оценки качества музыкальных сигналов с ограниченной полосой частот. Полученные результаты объективного оценивания сопоставлены с оценками качества, полученными путем субъективного тестирования.

Библ. 9, рис. 3.

Ключевые слова — качество музыкального сигнала; объективная оценка; субъективная оценка.

UDC 004.934

Frequency band limited musical signals objective and subjective quality assessment

I. V. Kotvytskyi, ORCID [0000-0003-2800-0080](https://orcid.org/0000-0003-2800-0080)

e-mail igorktvzk@gmail.com

A. V. Osa, ORCID [0000-0002-5973-6288](https://orcid.org/0000-0002-5973-6288)

e-mail andriibee796@gmail.com

National technical university of Ukraine "Igor Sikorsky Kyiv polytechnic institute" kpi.ua

Kyiv, Ukraine

Abstract—In this paper, analysis of features of some objective indicators of the audio signals quality, such as the segmental signal-to-noise ratio, log-spectral distortions and bark-spectral distortions, for estimating the quality of musical signals with a limited frequency band has been presented. It has been shown that the interpolation of the test signals allows increasing the reliability of the evaluation using segmental signal-to-noise ratio and log-spectral distortions indices. Results of objective and subjective quality estimation of speech signals distorted due to bandwidth limitation are compared in the earlier work by means of correlation coefficients calculation and construction of correspondence maps. It was assumed that speech signals occupy a frequency band not exceeding 11 kHz. Meanwhile for high-quality musical signal reproduction, a much wider range of frequencies up to 20 kHz is required. In this regard, it is interesting to study the relevance of subjective and objective indicators of quality for musical signals, which is the main purpose of this work. Experimental studies used 16 second fragments of musical compositions, categories "pop" and "classical" music. The file format is WAVE with a sampling rate of 44.1 kHz. Increasing of the sampling frequency by means of their interpolation has a beneficial effect on segmental signal-to-noise ratio estimates allowing to eliminate illogical violations of the monotonicity of the segmental signal-to-noise ratio dependence on the frequency bandwidth. Studies have shown that one of the reasons for the monotonicity violation in the log-spectral distortions dependence is the imperfect frequency response rectangularity of used low-pass filters. Indeed, the log-spectral distortions dependences are monotonic when filtering is realized in the frequency domain by means of ideal rectangular frequency response. It should be noted that when using low-pass filters with rectangular frequency response, the log-spectral distortions value also turned out to be dependent on sampling frequency selection as in the case of non-ideal low-pass filters. The reasons for this phenomenon are worth exploring in the future. Although 30 listeners took part in subjective trials, this quantity can hardly be considered as representative, and further its volume should be increased. Note in addition that quantity of the tested music fragments was very limited. Taking into account the pointed features, the results obtained in this paper should be regarded as preliminary. Obtained results of objective evaluation are compared with the quality assessments obtained by subjective testing.

Ref. 9, fig. 3.

Keywords — music signal quality; objective assessment; subjective evaluation.

