COMPUTER SCIENCES AND INFORMATION TECHNOLOGIES

UDC 621.391.83(045)
DOI:10.18372/1990-5548.82.19362

[1]**A. M. Prodeus,**
[2]**A. S. Naida**

MEASUREMENT OF REVERBERATION TIME USING A TWO-STAGE ALGORITHM

National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
E-mails: [1]aprodeus@gmail.com ORCID ID: 0000-0001-7640-0850, [2]naida.a.s.2001@gmail.com

Abstract—*The use of voice control of unmanned aerial vehicles is relevant due to the ease of practical use and new opportunities. This technology allows one to simplify the interface, making it more intuitive and natural. However, the quality and intelligibility of speech signals indoors can be significantly impaired by noise and reverberation. Therefore, before using voice technologies, it is desirable to take into account the effect of interferences by preliminary assessment of their parameters. In this paper, an algorithm for estimating the boundary (truncation time) between the informative and non-informative parts of the room impulse response, which allows obtaining believable estimates of the reverberation time, is proposed. The proposed algorithm is two-stage. At the first stage, "rough" envelope of the room impulse response is calculated using the detector-integrator, which allows one to find an approximate value of truncation time and construct an approximate envelope of room impulse response using backward integration method to obtain an approximate estimate of the reverberation time. In the second stage, output data of the first stage are used to refine the truncation time and reverberation time estimates. Experimental tests using recordings of real room impulse responses testify to the efficiency of the proposed algorithm.*

Index Terms—Reverberation time; truncation time; room impulse response; detector-integrator; backward integration method.

I. INTRODUCTION

Voice technology has become an important area of research in various aviation communication systems. Evidence for this statement is at least the fact that numerous groups of voice technology specialists have collaborated on research projects with aeronautical companies such as Euro Control and AENA [1].

Research into voice-controlled unmanned aerial vehicles (UAVs) has recently attracted much interest due to their ease of use and new capabilities. This technology eliminates the need for physical control and makes the interface more intuitive and natural. In [1], a comprehensive overview of UAV systems with voice control is provided, their achievements, problems and future directions are highlighted.

However, the quality and intelligibility of speech signals indoors can be significantly impaired by noise and reverberation. Therefore, before using voice technologies, it is desirable to take into account the effect of interferences by preliminary assessment of their parameters. When measuring the reverberation time (RT), the noise interruption method or the method of analyzing the room impulse response (RIR) are used [2] – [4]. In the first case, the noise signal from the microphone output is subjected to multi-channel band-pass filtering (this is required to obtain information about the RT dependence on frequency). In the second case, shown in Fig. 1, band-pass filter output is subjected to a RIR estimate.

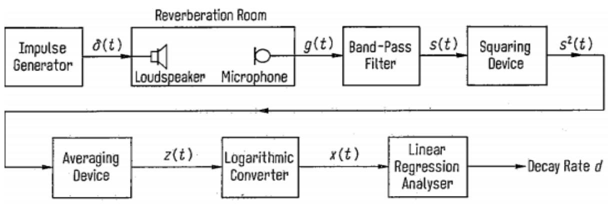


Fig. 1. Block diagram of RT measurement system [2]

Next, the envelope of the signal from the output of each filter is calculated in one of 3 ways [3], [4]. According to the first and second methods, the signal from the output of the band-pass filter is squared and subjected to running average according to an exponential

$$D_1(t) = k\int_{-\infty}^{t} e^{-k(t-\tau)} h^2(\tau)d\tau, \qquad (1)$$

or linear law

$$D_2(t) = \frac{1}{I}\int_{t-I}^{t} h^2(\tau)d\tau, \qquad (2)$$

where $h(t)$ is the RIR, k and I are parameters of running average.

According to [2], the time constant of the exponential averaging should be equal to $T_{60}/30$ or as close as possible to this value, where T_{60} is the reverberation time to be measured. Similarly, averaging time I should be less than $T_{60}/12$. In article [3], eighteen 1/3-octave filters were used during experimental measurements of T_{60} in the frequency band of 100–5000 Hz. Averaging the squared signals from the filter outputs was performed at an interval of 0.1 s, there were 8 such averaging at an interval of 1 s.

It was noted in [4] that during practical measurements, the average interval should be at least 3 times shorter than the reverberation time of the room. It follows from this that before measuring T_{60} one should have a priori information about the possible value of T_{60}, estimating it according to the Sabine and Eyring formulas.

Today, preference is given to the third method of processing the signal from the microphone output, namely, calculation by backward integration method [5]

$$D_s(t) = N\int_t^{\infty} h^2(\tau)d\tau \approx N\int_t^{T_i} h^2(\tau)d\tau, \qquad (3)$$

where N is proportional to the noise power spectral density in the measurement frequency range, T_i is the right border of the informative part of the RIR $h(t)$ (the left border $t=0$ corresponds to the $|h(t)|$ maximum, $|.|$ is module symbol).

Regardless of the calculation method, the resulting envelope is logarithmized and the law of its decay is transformed from exponential to linear. The envelope is then checked for thresholds exceeded at minus 5 dB, minus 15 dB, minus 25 dB, and minus 35 dB, after which the corresponding T_{10}, T_{20} and T_{30} reverberation time estimates are calculated. To calculate the Early Decay Time (EDT), thresholds of 0 dB and minus 10 dB are used.

Expressions (1) – (2) are often associated with the use of analog equipment [2] – [7], although it is clear that nothing prevents their application when using digital equipment. Due to the use of "time-reversed integration", expression (3) is more convenient for the implementation of T_{60} measurements with digital equipment, although, if desired, analog equipment can also be used for calculations.

It should be noted that there is a more significant difference between (1) and (2), on the one hand, and expression (3), on the other hand. The results obtained according to (1) – (2) should be interpreted as average of $h^2(t)$ over the only band-pass filter output, while expression (3) describes the result of averaging the ensemble of outputs. This means that the use of backward integration method allows one to avoid repeating experiments when measuring T_{60}.

However, in reality, the advantage of backward integration method are almost nullified because (3) was obtained in [5] under conditions of absence of background noise. In the presence of background noise, as a rule, the dynamic range of $D_s(t)$ obtained by (3) does not exceed 25 dB [6]. To avoid this shortcoming, it is proposed in [6] to subtract the estimate of the background noise mean level from the $h^2(t)$ before its integration, according to (3).

The theoretical analysis shows a significant dependence of the $D_s(t)$ shape on the choice of the T_i parameter and on the noise level (Fig. 2 [8]).

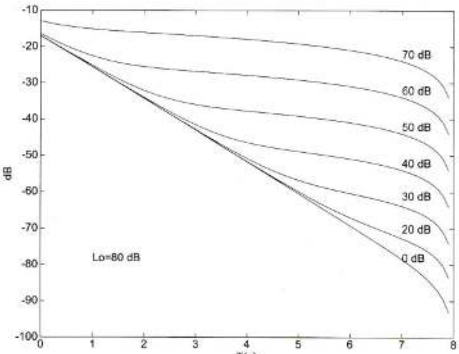


Fig. 2. $D_s(t)$ for signal level L_0 = 80 dB and noise levels from 0 to 70 dB [8]

Therefore, it is proposed in [7] to limit the upper limit of integration in (3) by the value T_i, which is close to half of the expected reverberation time T_{60} (Fig. 3). Other similar examples are given in [9] when overestimated or underestimated values of T_i made it difficult to estimate T_{60}. Thus, the problem of the optimal choice of the T_i parameter is relevant, and a set of studies have been devoted to the search for its solution [10] – [16].

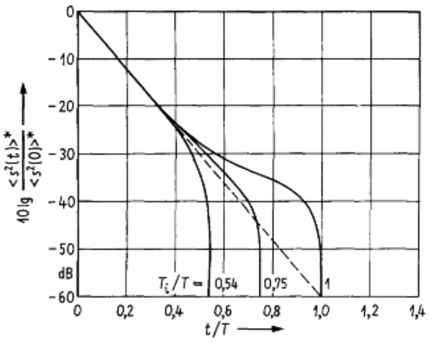


Fig. 3. The influence of the value of T_i on the form of $D_s(t)$ [7]

II. PROBLEM STATEMENT

To date, there are many proposals for choosing the T_i value [6] – [16]. It is appropriate to assume that, in addition to obtaining the "correct" value of T_i, the possibility of simultaneous estimation of EDT, T_{10}, T_{20}, T_{30}, as well as the simplicity of the corresponding calculation algorithm are also important.

For example, in [11], where the calculation of the T_{60} was proposed by the method of nonlinear approximation of the $D_s(t)$, the specified requirements were not met.

In this paper, an attempt is made to eliminate the indicated shortcoming by combining the potential possibilities of (1) – (3).

The proposed algorithm consists of two stages:

- "rough" estimates of EDT*, T^*_{10}, T^*_{20}, T^*_{30}, $T^*_{60} \approx \left(T^*_{10} + T^*_{20}\right)/2$ are obtained using the "rough" curve $D^*_s(t)$, for the construction of which a rough estimate T^*_i obtained using the contour $D_1(t)$ or $D_2(t)$ was used;
- refined estimates of EDT, T_{10}, T_{20}, T_{30} are obtained with the help of the refined curve $D_s(t)$, for the construction of which the refined value of T_i was used; here T_i is the moment of intersection of the line $L = -\left(60/T_{60}\right)t$ with a straight line corresponding to the average level of background noise on the interval $\left(T_i, t_{\text{end}}\right)$ of the function $h^2(t)$, where t_{end} is the end point of the $h^2(t)$.

III. TWO-STAGE ALGORITHM

Let us detail the first stage of the proposed algorithm:

- the contour $D_1(t)$ or $D_2(t)$ is calculated according to (1) or (2) using function $h^2(t)$ previously maximum-normalized;
- contour $D_1(t)$ is decimated with a step Δt forming a discrete sequence $D_1(i\Delta t)$, $i = 1, ..., M$, (further on, to shorten the description, we will not mention similar actions with $D_2(t)$);
- the discrete derivative $d_{1i} = D_1\left(i\Delta t\right) - D_1\left((i+1)\Delta t\right)$ is calculated and normalized by the maximum;
- the "rough" estimate T^*_i of the parameter T_i is calculated as the time value when the maximum-normalized discrete derivative d_{1i} crosses the threshold $tr \approx 0$;
- according to (3), the "rough" curve $D^*_i(t)$ and the "rough" estimates EDT*, T^*_{10}, T^*_{20}, T^*_{30}, $T^*_{60} \approx \left(T^*_{10} + T^*_{20}\right)/2$ are constructed.

Now let us detail the second stage of the algorithm:

- the average background noise level $n_{\text{eff}} = \dfrac{1}{T_{n2} - T_{n1}} \displaystyle\int_{T_{n1}}^{T_{n2}} h^2(\tau)d\tau$ is calculated, where T_{n1} and T_{n2} are chosen within the interval $\left(T^*_i, t_{\text{end}}\right)$;
- the refined value of T_i is calculated as the moment of intersection of the line $L = -\left(60/T_{60}\right)t$ and the level n_{eff} of background noise;
- refined estimates of EDT, T_{10}, T_{20}, T_{30} are obtained using the refined curve $D_s(t)$, constructed using the refined value of T_i.

IV. SET UP OF THE STUDY

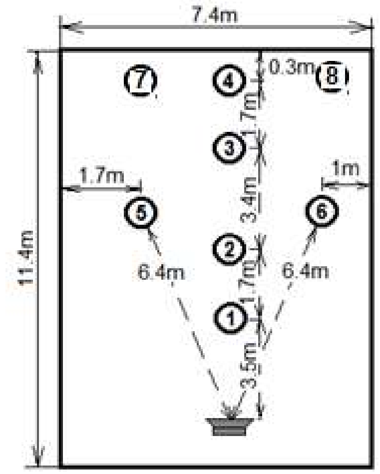
Recordings of a bursting rubber ball sound, made in an auditorium of the National Technical University of Ukraine "Ihor Sikorskyi Kyiv Polytechnic Institute", were used for efficiency testing of the proposed algorithm (Fig. 4a). Recordings have been made for 8 room points (Fig. 4b).

The box-shaped auditorium has a volume of 270 m^3. Artificial Head system was used for measurements (Fig. 4a). It consists of loudspeaker, two microphones near ears of artificial head, external sound card and laptop with the appropriate software.

The duration of the recordings with a sampling frequency of 44100 Hz was 5-6 s. The recordings were made in an empty room at a distances of 3.5-9.5 m from the sound source. Preliminary estimates of the reverberation time were close to 0.8-0.9 s, the noise level in the room was close to 50 dB.

a)

Fig. 4. Auditory and Artificial Head system (a), plan of auditory (b)



b)

Fig. 4. Ending. (See also p. 11)

At the first stage of computer processing of $h(t)$ records, envelope $D_2(t)$ was constructed with the duration $I = 0.1\,\text{s}$ of moving averaging. For further processing of $D_2(t)$, the parameters values $t = 0.1\,\text{s}$, $tr = 10^{-3}$ were adopted.

V. Results of the Study

The shapes of $h^2(t)$ and $D_s(t)$ for the case when the value of T_i is evidently too large and equal to the duration of $h(t)$ record are shown in Fig. 5.

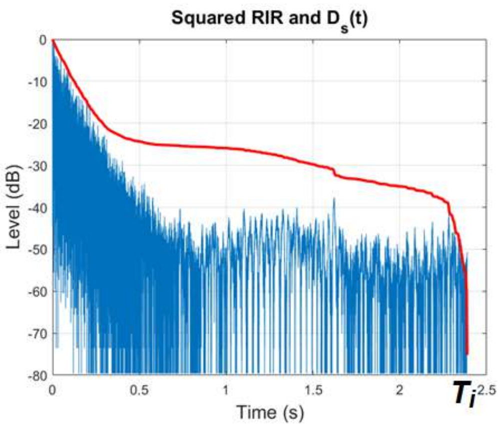
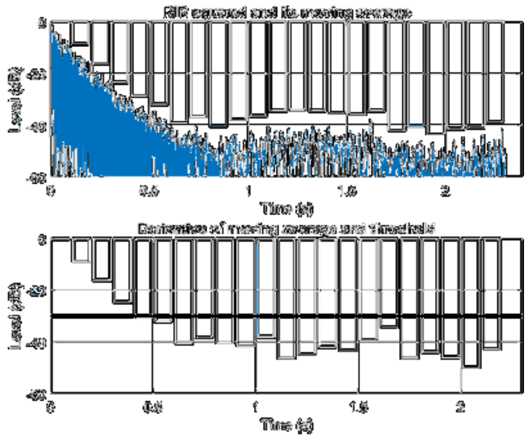


Fig. 5. The shapes of $h^2(t)$ and $D_s(t)$ for too large T_i

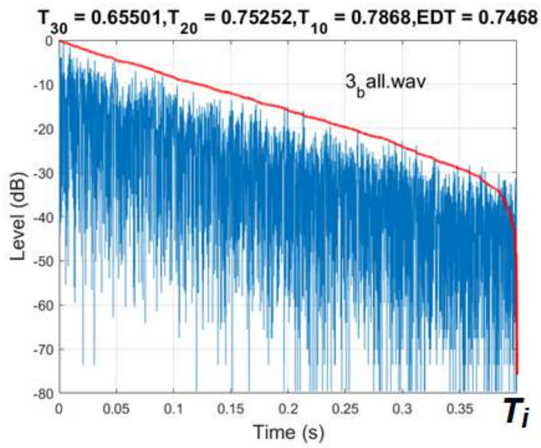
The "rough" estimates of $D_s^*(t)$, EDT*, T_{10}^*, T_{20}^*, T_{30}^*, $T_{60}^* \approx \left(T_{10}^* + T_{20}^*\right)/2$ obtained at the first stage of the algorithm, are shown in Fig. 6. The results of the second stage of the algorithm, where refined estimates of $D_s(t)$ and EDT, T_{10}, T_{20}, T_{30} are obtained, are shown in Fig. 7.

As one can see in Fig. 6a, the shape of the function $D_2(t)$ (bar graph) is similar to the shape of the function $h^2(t)$ contour, and the value of tr (horizontal red line) should be chosen to cross the normalized derivative (relative to its first value) of the $D_2(t)$. The value $tr = 10^{-3}$ was chosen in the case shown in Fig. 6b. Obviously, this choice led to an underestimation of the truncation time T_i (Fig. 6b). Note, however, that the logarithm of the modulus of the derivative is shown in Fig. 6b. This was done only for demonstration purposes. In real calculations, it is advisable to choose the threshold tr much smaller, for example, $tr = 10^{-5}$ or even $tr = 0$.

As can be seen from Fig. 7a, the refined value of T_i is defined as the moment of intersection of the average background noise level (dashed horizontal line) and the line $L = \frac{60}{T_{60}^*}t$ (slanted dashed line), $T_{60}^* \approx \left(T_{10}^* + T_{20}^*\right)/2$. The symbol "V" shows the time borders of the of background noise, and the symbol "X" shows the limits of the background noise area used to estimate the noise level.



a)



b)

Fig. 6. Intermediate (a) and final (b) results of the first stage of an algorithm

VI. Discussion

As can be seen from Fig. 5, the function $D_s(t)$ can hardly be considered a good estimate of the function $h^2(t)$ envelope in the case where the T_i value is equal to the duration of $h^2(t)$. The reason for this is the significant impact of background noise on the results of $D_s(t)$ calculations. As a result, the EDT and T_{10} estimates can still be obtained with sufficient accuracy. At the same time, the unreliability of T_{20} and T_{30} estimates is obvious.

The implementation of the first part of the proposed algorithm (Fig. 6) leads to improved results, but comparing the graphs of Fig. 6b and Fig. 5 shows that the estimate of the value of T_i is significantly underestimated. Nevertheless, the results of the first stage of calculations can be considered acceptable, since further refinement is expected.

The analysis of the second stage results (Fig. 7) shows that the indicated shortcoming of the first stage of calculations was eliminated. Indeed, a comparison of Fig. 7b and Fig. 5 indicates the correct finding of the informative part $h^2(t)$ $(T_i \approx 0.63\text{s})$.

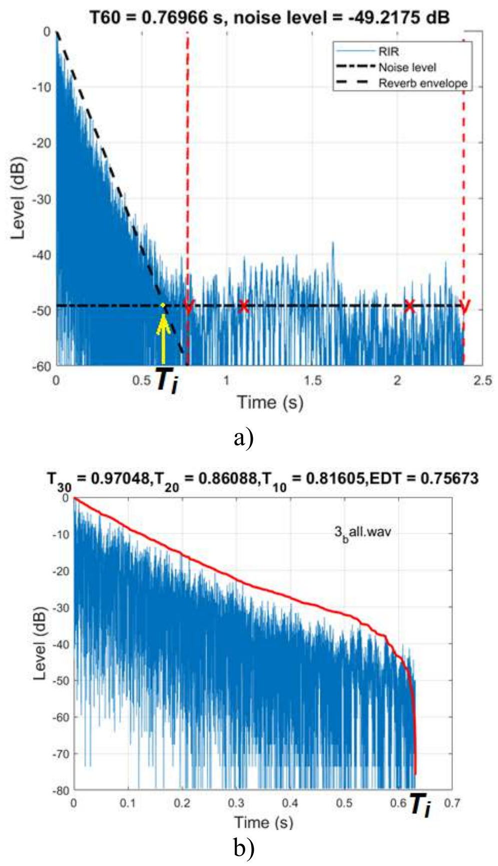


Fig. 7. Intermediate (a) and final (b) results of the second stage of an algorithm

As one can see from Fig. 7a, the "rough" estimate of the reverberation time $T_{60}^* \approx 0.77\text{s}$ is significantly underestimated compared to the final estimate of $T_{60} \approx 0.84$ (Fig. 7b), but this did not prevent one from finding the correct value of $T_i \approx 0.63\text{s}$ and obtaining refined estimates of $D_s(t)$ and EDT, T_{10}, T_{20}, T_{30}.

The analysis of the obtained results shows that the proposed algorithm can be used in practice. However, this algorithm is not free from certain shortcomings, which should be eliminated in the future. In particular, the choice of the value of the threshold *tr* seems insufficiently justified. In the given example, perhaps it would be better to set $tr = 10^{-4}$. Moreover, the assigning the previous value of the value of *tr* causes a desire to somehow avoid this procedure.

Studies show that the results of measurements of the reverberation time remain practically unchanged, if the values of the parameter Δt are in the interval (0.7–1.1) s.

The dependence of T_i and EDT, T_{10}, T_{20}, T_{30} estimates on the position of the microphone in different locations of the room is shown in Table I, where points numbered 1–4 are located along the axis of the room, and points 5–8 are located along the perimeter of the room. The lower two rows of the table shows the estimates of the corresponding mean values and standard deviations.

TABLE I. Estimates of RT in Different Room Locations

No	T_i	EDT	T_{10}	T_{20}	T_{30}
1	0.78	0.64	0.79	0.84	0.95
2	0.81	0.71	0.84	0.91	1.04
3	0.63	0.76	0.82	0.86	0.97
4	0.72	0.77	0.82	0.94	1.06
5	0.76	0.75	0.83	0.95	1.08
6	0.66	0.75	0.82	1.30	1.17
7	0.74	0.78	0.94	1.25	1.21
8	0.78	0.80	0.92	1.09	1.23
m	0.74	0.75	0.85	1.02	1.09
σ	0.06	0.05	0.05	0.18	0.11

At the same time, the computational simplicity of the proposed algorithm, as well as the possibility to obtain sufficiently plausible estimates of parameters T_i, T_{60}, EDT, T_{10}, T_{20}, T_{30} can be an incentive

The results obtained in this article can be used to suppress the effect of reverberation on speech intelligibility, as it was previously proposed to suppress the effect of noise interference [17]. Of course, it is also advisable to use the obtained results in any predictions or measurements that require prior information about the reverberation time.

VII. CONCLUSION

In this paper, an algorithm for estimating the boundary (truncation time) between the informative and non-informative parts of the impulse response of the room, was proposed. The algorithm allows obtaining plausible estimates of reverberation time using the backward integration method. The proposed algorithm is two-stage. At the first stage, the "rough" envelope of the room impulse response is calculated using detector-integrator and approximate estimates of the truncation time and reverberation time are obtained. In the second stage, these results are used to refine the truncation time and reverberation time estimates. Experimental tests using recordings of real room impulse responses testify to the efficiency of the proposed algorithm. The effectiveness of the algorithm should be tested in future studies using models and recordings of RIRs.

REFERENCES

[1] D. Lawrence, A. Pavitra. "Voice-Controlled Drones for Smart City Applications," *Sustainable Innovation for Industry 6.0*, edited by Asha Sharma, et al., IGI Global, 2024, pp. 162–177. https://doi.org/10.4018/979-8-3693-3140-8.ch010

[2] ISO 3382-1975, "Measurement of the Acoustical Characteristics of Rooms for Speech and Music," *International Standards Organization*, Geneva, Switzerland, 1975.

[3] J. Davy, "The variance of decay rates in reverberation rooms," *Acustica*, vol. 43, No.1, 1979, pp. 51–56.

[4] J. Davy, "The variance of impulse decays," *Acustica*, vol.44, no.1, 1980, pp. 51–56.

[5] M. R. Schroeder, "New method of measuring reverberation time," *Journal of the Acoustical Society of America 37*, 1965, pp. 409–412. https://doi.org/10.1121/1.1909343

[6] W. T. Chu, "Comparison of Reverberation Measurements Using Schroeder's Impulse Method and Decay-Curve Averaging Method," *J. Acoust.* Soc. Am., vol. 63, 1978, pp.1444–1450. https://doi.org/10.1121/1.381889

[7] R. Kurer and U. Kurze, "Integrationsverfahren zur Nachhallanswertung," *Acustica* 19, 1967/1968, pp. 313–322.

[8] L. Faiget, C. Legros, R. Ruiz, "Optimization of the lmpulse Response Length: Application to Noisy and Highly Reverberant Rooms," *J. Audio Eng. Soc.*, vol. 46, no. 9, 1998, pp. 741–750.

[9] M. Volander, H. Bietz, "Comparison of Methods for Measuring Reverberation Time," *Acta Acustica United with Acustica*, 80, 1994, pp. 205–215.

[10] A. Lundeby, M. Vorlander, V. Eric, and H. Bietz, "Uncertainties of Measurements in Room Acoustics," *Acustica*, vol. 81, 4, pp. 344–355.

[11] N. Xiang, "Evaluation of reverberation times using a nonlinear regression approach," *J. Acoust. Soc. Am.*, 98 (4), 1995, pp. 2112–2121. https://doi.org/10.1121/1.414460

[12] D. Morgan, "A parametric error analysis of the backward integration method for reverberation time estimation," *J. Acoust. Soc. Am.*, 101 (5), Pt. 1, May 1997, pp. 2686–2693. https://doi.org/10.1121/1.418557

[13] D. Ćirić, and M. Milošević, "Optimal determination of truncation point of room impulse responses," *Build Acoust*, 12(1),15–29, 2005.

https://doi.org/10.1260/1351010053499216

[14] D. D'Orazio, S. De Cesaris, and M. Garai, "Measuring reverberation time using preprocessed energy detection. INTER-NOISE and NOISE-CON Congress and Conference Proceedings," *InterNoise12*, New York City NY, 2012, pp. 2070–2080.

[15] M. Janković, D. Ćirić, and A. Pantić, "Automated estimation of the truncation of room impulse response by applying a nonlinear decay model," *J Acoust Soc Am.*, 139(3):1047–57. Mar. 2016. https://doi.org/10.1121/1.4941657.

[16] M. Chen, and C.-M. Lee, "The Optimal Determination of the Truncation Time of Non-Exponential Sound Decays," *Buildings*, 12, 697, 2022. https://doi.org/10.3390/buildings12050697

[17] A. Prodeus, "Performance measures of noise reduction algorithms in voice control channels of UAVs," *2015 IEEE International Conference Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD)*, Kyiv, Ukraine, 2015, pp. 189–192. https://doi.org/10.1109/APUAVD.2015.7346596

[18] O. Dvornyk, A. Prodeus, M. Didkovska, and D. Motorniuk, "Artificial Software Complex "Artificial Head," Part 1. Adjusting the Frequency Response of the Path," *Microsystems, Electronics and Acoustics*, vol. 22, no. 1, 2020, pp. 56–64. https://doi.org/10.20535/2523-4455.mea.198431

[19] O. Dvornyk, A. Prodeus, D. Motorniuk, and M. Didkovska, "Hardware and Software System "Artificial Head," Part 2. Evaluation of Speech Intelligibility in Classrooms," *Microsystems, Electronics and Acoustics*, vol. 22, no. 3, 2020, pp. 48–55. https://doi.org/10.20535/2523-4455.mea.209928

Received October 19, 2024

Prodeus Arkadiy. ORCID 0000-0001-7640-0850. Doctor of Engineering Science. Professor.
Acoustic and Multimedia Electronic Systems Department, National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute,” Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (1972).
Research interests: digital signal processing.
Publications: more than 180 papers.
E-mail: aprodeus@gmail.com

Naida Anton. Post-graduate Student.
Acoustic and Multimedia Electronic Systems Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (2022).
Research interests: digital signal processing.
Publications: 6.
E-mail: naida.a.s.2001@gmail.com

А. М. Продеус, А. С. Найда. Вимірювання часу реверберації за допомогою двоетапного алгоритму

Застосування голосового керування безпілотними літальними апаратами є актуальним через простоту практичного використання та нові можливості. Ця технологія дозволяє спростити інтерфейс, зробити його більш інтуїтивно зрозумілим і природним. Проте якість і розбірливість мовленнєвих сигналів у приміщенні можуть значно погіршуватися через шум і реверберацію. Тому перед застосуванням голосових технологій бажано врахувати дію завад шляхом попередньої оцінки їх параметрів. У даній роботі запропоновано алгоритм оцінки межі (часу відсікання) між інформативною та неінформативною частинами імпульсної характеристики приміщення, що дозволяє отримати достовірні оцінки часу реверберації. Запропонований алгоритм є двоетапним. На першому етапі за допомогою детектора-інтегратора розраховується «груба» обвідна імпульсної характеристики приміщення, що дозволяє знайти приблизне значення часу відсікання та побудувати наближену обвідну імпульсної характеристики приміщення методом зворотного інтегрування для отримання приблизної оцінки часу реверберації. На другому етапі вихідні дані першого етапу використовуються для уточнення оцінок часу відсікання та часу реверберації. Експериментальні перевірки із записами реальних імпульсних характеристик приміщення свідчать про працездатність запропонованого алгоритму.
Ключові слова: час реверберації; час відсікання; імпульсна характеристика приміщення; детектор-інтегратор; метод зворотного інтегрування.

Продеус Аркадій Миколайович. ORCID 0000-0001-7640-0850. Доктор технічних наук. Професор.
Кафедра акустичних та мультимедійних електронних систем, Національний технічний університет України «Київський політехнічний інститут ім. І. Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (1972).
Напрямок наукової діяльності: цифрова обробка сигналів.
Кількість публікацій: більше 180 наукових робіт.
E-mail: aprodeus@gmail.com

Найда Антон Сергійович. Аспірант.
Кафедра акустичних та мультимедійних електронних систем, Національний технічний університет України «Київський політехнічний інститут ім. І. Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (2022).
Напрямок наукової діяльності: цифрова обробка сигналів.
Кількість публікацій: 5.
E-mail: naida.a.s.2001@gmail.com