TELECOMMUNICATIONS AND RADIO ENGINEERING

UDC 621.391.83(045)
DOI:10.18372/1990-5548.65.14991

[1]A. M. Prodeus,
[2]M. V. Didkovska,
[3]K. A. Kukharicheva,
[4]D. E. Motorniuk

TWO SIMPLIFIED MODELS OF EARLY SOUND REFLECTIONS IN A ROOM

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
E-mails: [1]aprodeus@gmail.com ORCID ID: 0000-0001-7640-0850, [2]maryna.didkovska@gmail.com, [3]katerynakt@gmail.com, [4]daria.motorniuk@gmail.com

Abstract—*Reverberation can be considered as a significant interference if the voice control of the unmanned aerial vehicle is performed by an indoor operator. In this paper, two simplified models of early sound reflections in a room are considered. The first model is a single reflection at a time interval of 0–50 ms. The second model is a set of reflections randomly distributed over the same time interval. For both models, speech intelligibility was calculated according to Speech Transmission Index method and Schroeder's formula. This made it possible to obtain functional dependences of speech intelligibility on the early reflections models parameters. The obtained results for the first model are consistent with the results of earlier studies, which confirm the validity of the method used. An analysis of the statistical properties of the second model showed that an increase in the number of reflections from 2 to 8 results in a decrease, in average, of speech intelligibility from "excellent" to "good" for reverberation times 0.4–1 s. Increasing the reflections number to 18 leads to the reduced speech intelligibility close to the border between "good" and "fair". Further increasing of the reflections number almost does not improve speech intelligibility. It was also found that there is a much larger scatter in speech intelligibility values for a small number of early reflections than for a large number of ones. Analysis of this phenomenon allows one to conclude about the danger of strong early reflections located near the end of the 0–50 ms interval.*

Index terms—Early sound reflections; speech intelligibility; room impulse response model; model parameters; number of reflections.

I. INTRODUCTION

The performance of voice control systems for unmanned aerial vehicles (UAVs), which include automatic speech recognition (ASR) systems, significantly depends on the type and intensity of sound interference generated by the surrounding space. Reverberation can be considered as a significant interference if the voice control of the UAV is carried out by an operator located inside the room. As a result of the multiple sound reflections from the ceiling, floor, walls and other surfaces, the speech signal at the microphone output differs from the pure speech signal which negatively affects speech intelligibility and, as a result, the performance of the ASR system. Thus, the problem of the influence of reverberation on speech intelligibility in UAV voice control systems is relevant.

Research of the influence of early sound reflections on speech intelligibility has been going on for a long time. It is shown in [1] that the human auditory system integrates the action of early sound reflections with the action of direct sound if reflections arrive at the listener with a delay of no more than 30–40 ms relative to direct sound. In paper [2], it was found that early reflections within a delay of 0–95 ms have a positive effect on speech intelligibility while later reflections impair intelligibility, acting as noise interference. This fact allowed the C95 parameter to be used as a measure of speech intelligibility [3]. In paper [4], the C50 clarity index, similar to the C95 parameter, is considered. It is also noted that C50 is much better than the reverberation time T_{60} characterizes the effect of reverberation on the ASR systems performance. It was shown in [3] that Speech Transmission Index (STI) is no less effective measure of speech intelligibility in comparison with C50.

Since different parts of the room impulse response (RIR) affect speech intelligibility in different ways, researchers often use the technique of "cutting out" the investigated part of the RIR and convolving this part with speech signal. This technique was used, for example, in [5] when analyzing the effect of early and late reflections on the ASR systems operation. A similar technique was used in [6] when analyzing the effect of reverberation on the auditory system of people with cochlear implants, and also in [7] when analyzing the perception of reverberated speech by listeners with normal hearing.

However, instead of cutting out the desired part of the RIR, one can synthesize this part using computer simulation [8], [9], [10]. In this approach physical [2] or mathematical [11], [12] models of real rooms are usually used.

At the same time, use of too accurate and complex models of real rooms can be unproductive when researchers are forced to deal with the revealed phenomena without giving convincing explanations. Unfortunately, the brilliant works [11], [12] devoted to the early reflections analysis are not free from this shortcoming.

Therefore, the object of this paper is to analyze two simplified early reflections models that can help to explain the effect of early reflections on indoor speech intelligibility.

II. PROBLEM STATEMENT

The problem to be solved consists in assessing the intelligibility of speech distorted by early reflections. The model of RIR in the form of pulse stream

$$h(t)=\left[\delta(t)+\sum_{n=1}^{N}a_n\delta(t-t_n)\right]e^{-\alpha t}, \qquad (1)$$

where $\delta(t)$ is Dirac delta function, a_n and t_n are random variables, N is a number of pulses in a time interval, T_{60} is a reverberation time, $\alpha=\ln(10^3)/T_{60}$ is a parameter that determines the rate of RIR decay, seems to be adequate. Note that model (1) is part of a more general RIR model called "Hybrid Energy Decay Curve" [13].

In box shaped rooms, the pulse stream (1) density d increases over time according to the quadratic law [14]:

$$d(t)=4\pi c^3t^2/V. \qquad (2)$$

In this paper, two mathematical models of early reflections that are simpler than (1) – (2) model are presented. The first model is a single reflection, and the second model is a constant-density pulses stream.

In this work, the STI is used as a measure of speech intelligibility distorted by early reflections. The STI estimates are in good agreement with the results of subjective speech intelligibility assessment [3].

Since the situation with a single reflector has been studied well enough [1], consideration of the first model will make it possible to judge the correctness of the chosen method for analyzing early reflections. As for the model in the form of a constant-density pulses stream, it can be considered as the first approximation to the complete model (1) – (2).

III. ORGANIZATION OF STUDIES

According to the modulation method [15], [16], the STI is estimated as follows:

$$STI=\sum_{k=1}^{7}\alpha_k\cdot MT_k-\sum_{k=1}^{6}\beta_k\cdot\sqrt{MT_k\cdot MT_{k+1}}, \qquad (3)$$

$$MT_k=\frac{1}{14}\sum_{i=1}^{14}T_{ki},$$

$$T_{ki}=\begin{cases}0, & E_{ki}<-15,\\ (E_{ki}+15)/30, & -15\le E_{ki}\le+15,\\ 1, & E_{ki}<-15,\end{cases}$$

$$E_{ki}=10\lg\frac{m_{ki}}{1-m_{ki}},$$

$$m_{ki}=\left|\int_0^\infty h_k^2(t)\exp(-j2\pi F_it)dt\right|\bigg/\int_0^\infty h_k^2(t)dt, \qquad (4)$$

where $h_k(t)$ is the result of filtering of $h(t)$ by kth bandpass filter (7 octave filters with central frequencies from 125 Hz to 8 kHz are used); F_i is a modulation frequency (in practice, 14 values of F_i are used, in a range from 0.63 Hz to 12.5 Hz); α_k and β_k are weights and redundancy factors respectively [15].

Single reflection RIR model can be described by a simple expression

$$h(t)=\delta(t)+g\delta(t-t_0), \qquad (5)$$

where t_0 is a time delay, g $(0\le g\le 1)$ is a reflection level. Plots of function (5) and the corresponding amplitude frequency response (AFR)

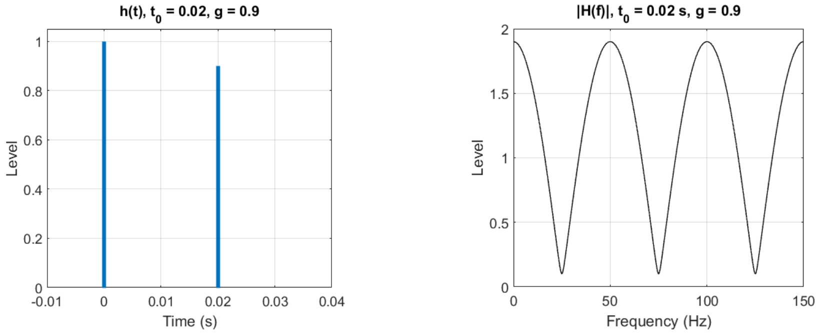
$$|H(f)|=\sqrt{1+g^2+2g\cos 2\pi f t_0},$$

are shown in Fig. 1 for t_0 = 0.02 s, g = 0.9.

Similar plots for the second RIR model, described by (1), for T_{60} = 1 s and a constant density of reflectors d =160 Hz, are shown in Fig. 2.

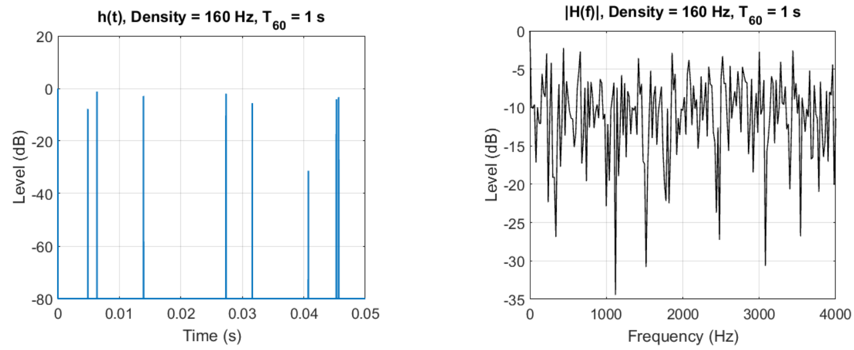
Amplitude frequency response (AFR) plots shown in Figs 1 and 2 clearly demonstrate that the main reason for the speech signal quality degradation is the significant unevenness of the room frequency response. Since the concepts of the speech quality and intelligibility are not unambiguously related, the plots in Figs 1 and 2 are insufficient to explain the reasons for the speech intelligibility degradation. To compensate this, one can use (3) – (4) for speech intelligibility computation.

For the first RIR model, one can write



a) b)

Fig. 1. RIR (a) and AFR (b) for the first model

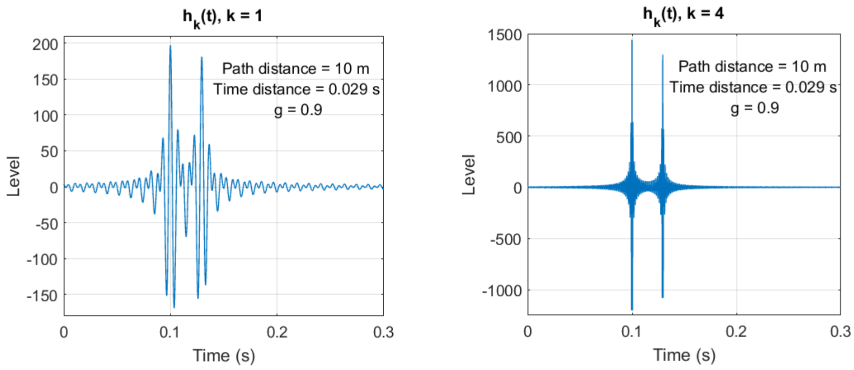


a) b)

Fig. 2. RIR (a) and AFR (b) for the second model

$$h_k(t)=\int_{-f_{k2}}^{-f_{k1}} H(f)e^{j2\pi ft}df+\int_{f_{k1}}^{f_{k2}} H(f)e^{j2\pi ft}df=2\left[f_{k2}\,\mathrm{Sa}(2\pi f_{k2}t)-f_{k1}\,\mathrm{Sa}(2\pi f_{k1}t)\right] \\ +2g\left[f_{k2}\,\mathrm{Sa}\left(2\pi f_{k2}(t-t_0)\right)-f_{k1}\,\mathrm{Sa}\left(2\pi f_{k1}(t-t_0)\right)\right]. \tag{6}$$

where f_{k1} and f_{k2} are lower and upper cutoff frequencies of the kth bandpass filter with the rectangular shape frequency response $H(f)$, $\mathrm{Sa}(x)=\sin x/x$. Function (6) form for $k=1$ and $k=4$ for $t_0=0.029$ c and $g=0.9$ is shown in Fig. 3 with a shift along the time axis by 0.1 s. Such a shift must be done in calculations using relation (4) in order for the left tail of the function $\mathrm{Sa}(x)$ to be taken into account.



a) b)

Fig. 3. Function (6) form for $k=1$ (a) and $k=4$ (b)

Generalization (6) to the second RIR model case can be written as:

$$h_k(t) = 2\left[f_{k2}\,\mathrm{Sa}(2\pi f_{k2}t) - f_{k1}\,\mathrm{Sa}(2\pi f_{k1}t)\right]$$
$$+2\sum_{n=1}^{N} a_n\left[f_{k2}\,\mathrm{Sa}(2\pi f_{k2}(t-t_n)) - f_{k1}\,\mathrm{Sa}(2\pi f_{k1}(t-t_n))\right]. \quad (7)$$

As one can notice, STI calculations are easy to implement using (3), (4), (6), and (7).

IV. Simulation Results

A. Single Reflection Model

The results of simulation the dependence of STI on the direct and reflected sound paths (times) of arrival difference are shown in Fig. 4 for different relative levels of reflected sound. As one can see, in the case of strong reflection ($g = 0.9$), if the difference in arrival times is close to 30 ms, the STI rating decreases by 1 point (Table I) from “excellent” to “good”. A further time difference increase to 40 ms leads to a decrease of speech intelligibility by another 1 point, down to the “fair” rating. At the same time, the STI value is close to 1 if the time difference does not exceed 10–13 ms.

TABLE I. STI Score and Intelligibility Rating [16]

STI	Intelligibility rating
> 0.75	Excellent
0.60–0.75	Good
0.45–0.60	Fair
0.30–0.45	Pour
< 0.30	Bad

In the case of reflections of moderate strength ($g = 0.5$), the intelligibility rating decreases from “excellent” to “good” if the path difference is close to 13.6 m (40 ms time difference).

B. Constant Density Pulses Stream Model

Dependences of the mean STI value on the density of early reflections in the interval 0–50 ms are shown in Fig. 5. Averaging was performed over 100 STI estimates for 0.4 s, 0.7 s and 1 s reverberation times. The ends of the vertical lines correspond to the boundaries of the 95% confidence interval.

Two cases are considered in Fig. 5. In the first case (Fig. 5a) the density was set to 40 Hz, 375 Hz, 1030 Hz and 2060 Hz. One observes that the average intelligibility is minimal at a density close to 375 Hz. Shown in Fig. 5b, a similar graph for a density from 40 Hz to 320 Hz allows one to conclude that really the minimum of the average intelligibility is near $d \approx 375$ Hz.

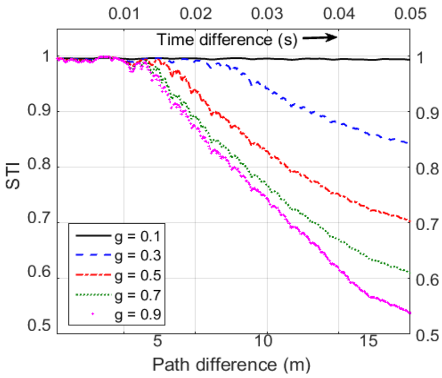
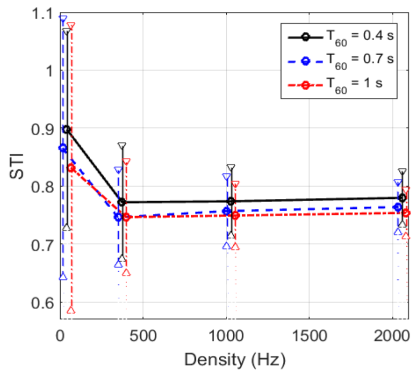
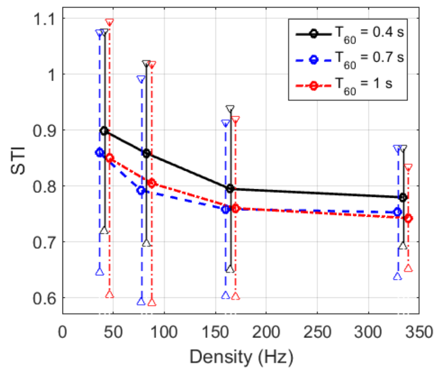


Fig. 4. STI estimates for the first RIR model



a)



b)

Fig. 5. STI estimates for the second RIR model

Shown in Fig. 5 graphs can be interpreted also in terms of the number N of reflections in the time interval 0–50 ms. Indeed, the rapid intelligibility decrease with density increase from 40 Hz to 160 Hz (Fig. 5a) corresponds to an increase in the average number of reflections from 2 to 8. The slow decline in intelligibility continues to 360–375 Hz (18–19 reflections). A further increase of N from 18 to 100 even somewhat increases the average intelligibility. However, this increase is so insignificant that it can be said that speech intelligibility has stabilized at a minimum level.

It follows from the graphs in Fig. 5 that T_{60} increase from 0.4 s to 1 s leads to an insignificant speech intelligibility degrade. This result can also be explained by the properties of the first model. Indeed, an increase of T_{60} leads to an increase of the late reflections influence and, as a consequence, speech intelligibility degrades.

Summarizing the results obtained for the second model, one can come to the following conclusions:

- the average number reflections increase from 2 to 8 leads to the rapid average speech intelligibility degrade from "excellent" to "good" for a reverberation time from 0.4 s to 1 s;
- there is a minimum of average speech intelligibility which is close to the border between "good" and "fair" when the average number of reflections is close to 18;
- the further increasing of average reflections number to 100 results in a barely noticeable speech intelligibility enhancement;
- for a small number of early reflections (from 2 to 8), speech intelligibility can vary over a wide range; this range is noticeably narrowed with increasing of reflections quantity;
- an increase of T_{60} from 0.4 s to 1 s leads to an insignificant speech intelligibility decrease.

V. Discussion

The obtained results for the first model are in good agreement with the well-known conclusion [1] that the first strong reflection insignificantly worsens speech intelligibility provided that the delay of the reflected sound relative to the direct sound does not exceed 30–40 ms. In addition, these results are consistent with the conclusion [17] that the harmful effect of the first reflection increases with an increase in its amplitude and delay.

A good agreement is also for the second model in the case of small (from 2 to 8) number of reflections in the 0–50 ms interval. This number of reflections is close to the real situation when a listener in a box shaped room perceives 6 primary reflections. A further increase of reflections number is possible by adding special reflectors, as well as by adding secondary and tertiary reflections.

Interestingly, there is a much larger scatter in speech intelligibility values for a small number of reflections than for a large number of reflections. This result can be explained using the first model results (Fig. 4). Indeed, with a small number of reflections, extreme situations are possible, when most of the strong reflections are concentrated either at the beginning of the 0-50 ms interval (Fig. 6a) or at its end (Fig. 6b). In the first case, the intelligibility is high, while in the second case, the intelligibility is low. With a large number of reflections, such situations are unlikely, so the scatter of the intelligibility values is relatively small.

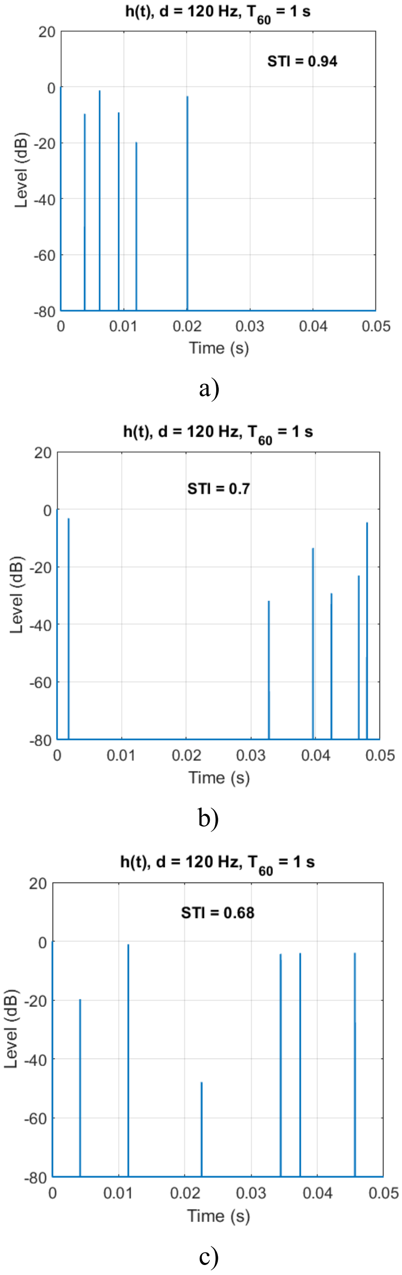


Fig. 6. RIRs for case of small number of reflections: (a) all reflections are concentrated at the beginning of 0–50 ms time interval; (b), (c) strong reflections are concentrated at the end of the time interval

Note that extremely high and low speech intelligibility values can also be obtained with a relatively uniform distribution of reflections in the 0-50 ms interval. In this case, the location of the strong reflections is important. Plot in Fig. 6c shows a case where the reflections are relatively evenly distributed, but the strong reflections are concentrated at the end of the 0–50 ms interval, resulting in reduced speech intelligibility.

VI. CONCLUSION

In this paper, two models of early sound reflections in a room are considered. The analysis of the properties of the first model, that assumes a single reflection in the time interval 0–50 ms, made it possible to obtain STI dependences on the reflected sound delay and strength. The dependences are consistent with the results of earlier studies, which confirm the validity of the method used.

Due to the second model, random reflections are distributed over the time interval 0-50 ms. Analysis of the properties of this model showed that for a small (up to 8) number of reflections the speech intelligibility varies from an excellent ($\mathrm{STI} \approx 1$) to fair ($\mathrm{STI} \approx 0.6$). High STI values can be obtained when the strongest reflections are concentrated at the beginning of the 0–50 ms interval. Conversely, the lowest STI values occur when the strongest reflections are at the end of the interval.

For T_{60} values from 0.4 s to 1 s, an increase of the number of early reflections from 2 to 8 leads to an average intelligibility degradation from "excellent" ($\mathrm{STI} = 0.85 - 0.9$) to "good" ($\mathrm{STI} = 0.76 - 0.8$). For 18 reflections, average intelligibility is reduced ($\mathrm{STI} = 0.75 - 0.78$). Increasing of the reflections number further to 100 results in a slight speech intelligibility increase. Note that this increase is insignificant and does not affect speech intelligibility in practice.

In the future, it is reasonable to consider more realistic models of early reflections that assume the variable density pulses stream and adjustable delay between the direct sound and the first reflection.

REFERENCES

[1] H. Haas, "Uber den Einfluss eines Einfachechos auf die Horsamkeit von Sprache," *Acustica*, vol. 1, pp. 49–58, 1951.

[2] J. Lochner and J. Burger, "The influence of reflections on auditorium acoustics," *J. Sound Vib.*, vol. 1, pp. 426–454, 1964. https://doi.org/10.1016/0022-460X(64)90057-4

[3] K. Jacob, "Correlation of Speech Intelligibility Tests in Reverberant Rooms with Three Predictive Algorithms," *Proc. of 85th Convention of the Audio Engineering Society*, Los Angeles, pp. 1020–1030, Nov. 3-6, 1988.

[4] F. Xiong, S. Goetze, and B. Meyer, "Estimating Room Acoustic Parameters for Speech Recognizer Adaptation and Combination in Reverberant Environments," *Proc.-ICASSP, IEEE International Conference on Acoustics*, Speech and Signal Processing, May 2014. https://doi.org/10.1109/ICASSP.2014.6854659

[5] H. Golzer and M. Kleinschmidt, "Importance of early and late reflections for automatic speech recognition in reverberant environments," *Proc. Elektronische Sprachsignalverarbeitung (ESSV)*, 2003.

[6] Y. Hu and K. Kokkinakis, "Effects of early and late reflections on intelligibility of reverberated speech by cochlear implant listeners," *J. Acoust. Soc. Am.*, vol. 135(1), January 2014. https://doi.org/10.1121/1.4834455

[7] A. Prodeus, M. Didkovska, D. Motorniuk, and O. Dvornyk, "The Effects of Noise, Early and Late Reflections on Speech Intelligibility," *Proc. 2020 IEEE 40th Int. Conf. on Electronics and Nanotechnology (ELNANO)*, pp. 488–492. https://doi.org/10.1109/ELNANO50318.2020.9088854

[8] M. R. Schroeder, B. S. Atal, and C. Bird, "Digital computers in room acoustics," *in Proceedings of the 4th International Congress on Acoustics*, ICA, Copenhagen, p. M21, 1962.

[9] J. Rindel, "The Use of Computer Modeling in Room Acoustics," *Journal of Vibroengineearing*, no. 3(4), pp. 219–224, 2000.

[10] M. Vorlander, "Computer simulations in room acoustics: Concepts and uncertainties," *J. Acoust. Soc. Am.*, vol. 133 (3), pp. 1203–1213, March 2013. https://doi.org/10.1121/1.4788978

[11] J. Bradley, H. Sato, and M. Picard, "On the importance of early reflections for speech in rooms," *J. Acoust. Soc. Am.*, vol. 113, no. 6, pp. 3233–3244, 2003. https://doi.org/10.1121/1.1570439

[12] Arweiler, J. Buchholz, and T. Dau, "Speech intelligibility enhancement by early reflections," *Proc. of ISAAR*, Marienlyst, Denmark, August 2009.

[13] K. D. Jacob, T. K. Birkle, C. B. Ickler, "Accurate Prediction od Speech Intelligibility without the Use of In-Room Measurements," *J. Audio Eng. Soc.*, vol. 39, no. 4, pp. 232–242, April 1991.

[14] H. Kuttruff, Room Acoustics. Fifth Edition. Spon Press an imprint of the Taylor & Francis Group, London and New York, 2009.

[15] H. J. M. Steeneken, and T. Houtgast, "Validation of the revised STIr method," *Elsevier Speech Communication*, vol. 38, pp. 26–37, 2002. https://doi.org/10.1016/S0167-6393(02)00010-9

[16] H. Steeneken, "Forty years of speech intelligibility assessment (and some history)," *Proc. of the Institute of Acoustics*, vol. 36, Pt.3, 2014.

[17] A. Warzybok, J. Rennies, S. Doclo, and B. Kollmeier, "Influence of early reflections on speech intelligibility under different noise conditions," *Proc. of Forum Acusticum 2011*, June 27-July 1, 2011, Aalborg, Denmark.

Received August 21, 2020

Prodeus Arkadiy. orcid.org/0000-0001-7640-0850. Doctor of Engineering Science. Professor.
Department of Acoustic and Multimedia Electronic Systems, 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: 178.
E-mail: aprodeus@gmail.com

Didkovska Maryna. Candidate of Science (Engineering). Associate Professor.
Department of Mathematical Methods of System Analysis, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (2002).
Research interests: system analysis and management.
Publications: 76.
E-mail: maryna.didkovska@gmail.com

Kukharicheva Kateryna. Post-graduate Student.
Department of Acoustic and Multimedia Electronic Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (2017).
Research interests: digital signal processing.
Publications: 4.
E-mail: katerynakt@gmail.com

Motorniuk Daria. Post-graduate Student.
Department of Acoustic and Multimedia Electronic Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine.
Education: Kyiv Polytechnic Institute, Kyiv, Ukraine, (2016).
Research interests: digital signal processing.
Publications: 4.
E-mail: daria.motorniuk@gmail.com

А. М. Продеус, М. В. Дідковська, К. А. Кухарічева, Д. Є. Моторнюк. Дві спрощені моделі ранніх відбиттів звуку в приміщенні

Реверберація може бути значною перешкодою, якщо голосове керування безпілотним літальним апаратом виконує оператор, що знаходиться в приміщенні. В даній роботі розглядаються дві спрощені моделі ранніх відбиттів звуку в кімнаті. Перша модель – це одноразове відбиття в межах інтервалу часу 0–50 мс. Друга модель – це набір випадкових рівномірно розподілених відбиттів на тому самому інтервалі часу. Для обох моделей розбірливість мови розраховано за методом Speech Transmission Index та із використанням формули Шредера. Це дозволило отримати функціональні залежності розбірливості мови від параметрів моделей ранніх відбиттів. Отримані результати для першої моделі узгоджуються з результатами попередніх досліджень, що підтверджує обгрунтованість використовуваного методу. Аналіз статистичних властивостей другої моделі показав, що збільшення кількості відбиттів з двох до восьми призводить до зменшення в середньому розбірливості мови від «відмінно» до «добре» для часу реверберації 0,4–1 s. Збільшення кількості відбиттів до вісімнадцяти призводить до зниження розбірливості мови до межі між градаціями «добре» та «задовільно». Подальше збільшення кількості відбиттів майже не покращує розбірливість мови. Також було виявлено, що при невеликій кількості відбиттів розкид оцінок розбірливості є значно більшим, ніж при великій кількості відбиттів. Аналіз цього явища дозволяє зробити висновок про небезпеку сильних ранніх відбиттів, розташованих поблизу кінця інтервалу 0–50 мс.
Ключові слова: ранні відбиття звуку; розбірливість мови; модель імпульсної характеристики приміщення; параметри моделі; кількість відбиттів.

Продеус Аркадій Миколайович. orcid.org/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: 178.
E-mail: aprodeus@gmail.com

Дідковська Марина Віталіївна. Кандидат технічних наук. Доцент.
Кафедра математичних методів системного аналізу, Національний технічний університет України «Київський політехнічний інститут ім. І. Сікорського», Київ, Україна.
Освіта: Київський політехнічний інститут, Київ, Україна, (2002).
Напрямок наукової діяльності: системний аналіз та менеджмент.
Кількість публікацій: 76.
E-mail: maryna.didkovska@gmail.com

Кухарічева Катерина Андріївна. 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, (2017).
Напрямок наукової діяльності: цифрова обробка сигналів.
Кількість публікацій: 4.
E-mail: katerynakt@gmail.com

Моторнюк Дар'я Євгенівна. 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, (2016).
Напрямок наукової діяльності: цифрова обробка сигналів.
Кількість публікацій: 4.
E-mail: daria.motorniuk@gmail.com

А. Н. Продеус, М. В. Дидковская, К. А. Кухаричева, Д. Е. Моторнюк. Две упрощенные модели ранних отражений звука в помещении

Реверберацию можно рассматривать как существенную помеху, если голосовое управление беспилотными летательными аппаратами осуществляется оператором, находящимся в помещении. В данной статье рассматриваются две упрощенные модели ранних отражений звука в помещении. Первая модель представляет собой однократное отражение на временном интервале 0–50 мс. Вторая модель представляет собой набор случайных равномерно распределенных отражений на том же интервале времени. Для обеих моделей разборчивость речи рассчитывалась методом индекса передачи речи, с использованием формулы Шредера. Это позволило получить функциональные зависимости разборчивости речи от параметров моделей реверберации. Полученные результаты для первой модели согласуются с результатами более ранних исследований, что подтверждает правомерность использованного метода. Анализ статистических свойств второй модели показал, что увеличение количества отражений с двух до восьми приводит к снижению разборчивости речи в среднем с «отлично» до «хорошо» при времени реверберации 0,4–1 с. Увеличение количества отражений до восемнадцати приводит к пониженной разборчивости речи, близкой к границе между «хорошо» и «удовлетворительно». Дальнейшее увеличение количества отражений практически не улучшает разборчивость речи. Также было обнаружено, что разброс значений разборчивости речи при небольшом количестве ранних отражений намного больше, чем при большом количестве отражений. Анализ данного явления позволяет сделать вывод об опасности сильных ранних отражений, расположенных вблизи конца интервала 0–50 мс.
Ключевые слова: ранние отражения звука; разборчивость речи; модель импульсной характеристики помещения; параметры модели; количество отражений.

Продеус Аркадий Николаевич. orcid.org/0000-0001-7640-0850. Доктор технических наук. Профессор.
Кафедра акустических и мультимедийных электронных систем, Национальный технический университет Украины «Киевский политехнический институт им. И. Сикорского», Киев, Украина.
Образование: Киевский политехнический институт, Киев, Украина, (1972).
Направление научной деятельности: цифровая обработка сигналов.
Количество публикаций: 178.
E-mail: aprodeus@gmail.com

Дидковская Марина Виталиевна. Кандидат технических наук. Доцент.
Кафедра математических методов системного анализа, Национальный технический университет Украины «Киевский политехнический институт им. И. Сикорского», Киев, Украина.
Образование: Киевский политехнический институт, Киев, Украина, (2002).
Направление научной деятельности: системный анализ и менеджмент.
Количество публикаций: 76.
E-mail: maryna.didkovska@gmail.com

Кухаричева Екатерина Андреевна. Аспирант.
Кафедра акустических и мультимедийных электронных систем, Национальный технический университет Украины «Киевский политехнический институт им. И.Сикорского», Киев, Украина.
Образование: Киевский политехнический институт, Киев, Украина (2017).
Направление научной деятельности: цифровая обработка сигналов.
Количество публикаций: 4.
E-mail: katerynakt@gmail.com

Моторнюк Дарья Евгеньевна. Аспирант.
Кафедра акустических и мультимедийных электронных систем, Национальный технический университет Украины «Киевский политехнический институт им. И. Сикорского», Киев, Украина.
Образование: Киевский политехнический институт, Киев, Украина (2016).
Направление научной деятельности: цифровая обработка сигналов.
Количество публикаций: 4.
E-mail: daria.motorniuk@gmail.com