ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 2010. No. 8(104)
ELEKTRONIKA IR ELEKTROTECHNIKA

T 180 TELECOMMUNICATIONS ENGINEERING
TELEKOMUNIKACIJŲ INŽINERIJA

The Tool for Quality Estimation of Short Voice Segments

A. Anskaitis, A. Kajackas
Telecommunications Engineering Department, Vilnius Gediminas Technical University, Naugarduko str. 41, LT-03227 Vilnius, Lithuania, e-mails: aurimas.anskaitis@el.vgtu.lt, algimantas.kajackas@el.vgtu.lt

Introduction

The quality of transmitted speech depends on many technical factors such as channel bandwidth, signal level, echo, delay, signal-to-noise ratio, and codec type. In mobile networks there is one major additional factor affecting speech quality – a packet loss. The same applies to VoIP applications.

It is widely recognized that objective and reliable voice quality measure is PESQ (Perceptual Evaluation of Speech Quality) algorithm [1]. This algorithm is defined in ITU Recommendation P.862 [1] and it is the most popular algorithm for voice quality evaluation. The PESQ algorithm implements the base model of human voice perception. Some benchmark tests of PESQ have yielded an average correlation of 0.935 with the corresponding subjective MOS values [1]. It is also important that packet loss and packet loss concealment in CELP codecs are the factors for which PESQ had demonstrated acceptable accuracy ([1], Table 1/P.862).

There are some papers analyzing the impact of lost frames on voice quality using PESQ [3, 4]. In these works the dependence of PESQ score on Frame Error Rate is investigated.

ITU specification [2] recommends using relatively long (8 – 30) s voice segments with PESQ. Systemic analysis of voice quality using PESQ raises some problems because signal of required duration contains a big number (400 – 1500) of 20 ms voice frames. However, each and every frame has distinct impact on speech quality [5, 6]. When there are a few lost frames in signal segment during single measurement, the measured quality become a random number. This number shows an aggregated impact of many by coincidence lost frames.

Thorough investigation of the impact of series of lost frames requires simulation of frame errors in all possible positions of selected voice segment. The number of experiments increases with the signal length and the number of lost frames in a signal. This problem could be solved using short signals, 1 s long for example.

Analysis of the importance of lost frames in a short signal allows paying more attention to intelligibility problem. Such analysis could help for creation of packet marking [7] and means of selective coding of voice frames [8].

By pursuing experiments with short signals and using conventional PESQ measurer we have perceived that the relative uncertainty of $\Delta MOS_{\mathrm{PESQ}}$ measurements caused by changes of position of measurement window repeats with the period of 16 ms (this equals halve of PESQ phoneme) and evaluated by several experiments was estimated in the rage ±(0.2 … 0.55) [6].

While looking for the methods of reduction of uncertainty of measurements, we proposed to compose special signal for quality measurements. This signal is formed periodically repeating short voice signal (until 1 s) under investigation [5, 6]. This way the signal can be prolonged to period of 8 – 30 s, which is recommended by PESQ standard.

The purpose of the work is creation and verification of the tool for quality estimation of short voice segments. Verification here is performed by comparing results of measurements of proposed tool with standard PESQ. Such tool is required when testing other voice quality evaluation methods, for example, when investigating quality degradation due to lost frames.

PESQ algorithm

Further analysis relies heavily on the general structure of PESQ algorithm. PESQ algorithm carries out the following steps in its operation [1]:

1. Level and time alignment pre-processing.
2. Fitting of perceptual model.
3. Calculation of the disturbance densities for every 32 ms phoneme.
4. Aggregation of the disturbance densities over frequency and emphasis. Symmetric and asymmetric disturbances – D_n^{s} and D_n^{a} are calculated for every phoneme.
5. Nonlinear aggregation of the disturbances D_n^{s} and D_n^{a} within split second intervals. For every 320 ms long syllable are calculated symmetric and asymmetric syllable

disturbances L_i^{s} and L_i^{a}.

6. The last one aggregation over the duration of the speech signal is performed using typical mean square averaging algorithm. Symmetric and asymmetric quality degradations d^{s} and d^{a} are calculated. The number N of aggregated disturbances L_i^{s} and L_i^{a} equals the number syllables in PESQ measurement window T. When only one syllable contains distortions, overall quality degradation is reduced significantly with the increase of signal length

$$d \approx L / \sqrt{N} . \tag{1}$$

7. The final PESQ score is calculated as follows

$$V_{\mathrm{PESQ}} = 4.5 - 0.1 \cdot d^{\mathrm{s}} - 0.0309 \cdot d^{\mathrm{a}}. \tag{2}$$

It can be seen from this short presentation of PESQ that the first three steps are devoted for calculation of the disturbances D_n^s and D_n^a. The final PESQ score is a linear combination of the average disturbance value and the average asymmetrical disturbance value. Aggregation of the disturbances is done on two levels – using phonemes and syllables.

For the purpose of the paper it is important that the final step is aggregation of the disturbances over the duration of the speech signal. Aggregation is always some king of averaging operation.

It is known [3, 5, 6, 9], that each and every voice frame has different impact on overall speech quality. Averaging over the duration of the long (8 – 30 s) speech signal calculates cumulative impact of all lost frames. When signals used for quality measurement are short, averaging does not come into play. Because of this measurement results depend on initial conditions.

Composite measurement signal

While trying to adapt PESQ for short signals, we offered to use special composite signal which is constructed of periodically repeated measurement signal.

This design is based on the provision of ITU-T Recommendation [1]: “Real speech test signal may be constructed by concatenating short fragments of real speech while retaining a representative structure of speech and silence”.

Composite signal for measurements is constructed as follows:

1. A segment of original voice signal $v_{\mathrm{T}}(t)$ of relatively short duration T_{w} shall be prepared

$$v_{\mathrm{T}}(t) = \begin{cases} 0, & t < 0, \\ v(t), & 0 \le t \le T_{\mathrm{w}}, \\ 0, & t > T_{\mathrm{w}}, \end{cases} \tag{3}$$

where T_{w} – initial measurement window, $T_{\mathrm{w}} < 1$s.

In experiments the i-th frame of voice signal $v(t)$ is depicted as code V_i and fit the i-th data packet C_i

$$v(t) \Rightarrow V_i , \tag{4}$$

where $t_0 + (i-1)T_{\mathrm{f}} \le t \le t_0 + iT_{\mathrm{f}}$.

The choice of the initial voice signal and measurement window T_{w} length is arbitrary. The following logic was applied while choosing this parameter. Subjective evaluation of voice distortion is possible when signal contains a word or few words. Thus the shortest duration T_{w} should allow the calculation of at least one symmetric and asymmetric syllable disturbances L_i^{s} and L_i^{a}. Such duration equals 1.5 PESQ syllable, that is $T_{\mathrm{w}} \ge 1.5 \cdot 320$ ms long.

The other criterion for selection of T_{w} is the duration of a word in the investigated language. The articulatory durations for words and syllables are estimated in [10]. The average duration of isolated words is 499 ms (Stdev. is 104 ms) and the average duration for words in memorized sequences is 273 ms. (Stdev. is 69 ms). By composing of prolonged test signal for experiments we used T_{w} equal to 0.5 or 1 s.

2. The composite measurement signal consisting of periodically J times replicated original voice segments $v_{\mathrm{T}}(t)$ shall be constructed as

$$v_{\Sigma}(t) = \begin{cases} v_{\mathrm{T}}(t\text{-}jT_{\mathrm{r}}), \text{if } (j-1) \cdot T_{\mathrm{r}} \le t \le j \cdot T_{\mathrm{r}}, \\ 0, \text{if } t \ge JT_{\mathrm{r}} \text{ or } t < 0, \end{cases} \tag{5}$$

where T_{r} is a replication period, and $T_{\mathrm{r}} \le t \le JT_{\mathrm{r}}$, j = 1,…, J.. The replication period in (5) shall be chosen as $T_{\mathrm{r}} \ge T_{\mathrm{w}}$. This way it is possible to write

$$T_{\mathrm{r}} = T_{\mathrm{w}} + \Delta T , \tag{6}$$

where ΔT is the additional increase of replication period. In this manner we achieve randomization of position of lost frame within extended measurement window. PESQ algorithm has time structure with a period of 16 ms, so the randomization of positions on lost frames in the measurement window will occur when repetition period is not iterative to 16 ms or namely, not iterative to period of variations.

3. Formula (5) describes the composite reference signal $v_{\Sigma}(t)$. To investigate value of i-th frame within the measurement window T_{w}, the degraded composite signal is constructed in two stages.

In the first stage signal of duration T_{w} is created by deletion of i-th frame according to rule described follows

$$v_T^{\mathrm{d}}(t) = \begin{cases} v_{\mathrm{T}}(t), \text{when } t_0 \le t \le t_0 + (i-1) \cdot T_{\mathrm{f}}, \\ 0, \text{when } t_0 + (i-1) \cdot T_{\mathrm{f}} \le t \le t_0 + i \cdot T_{\mathrm{f}}, \\ v_{\mathrm{T}}(t), \text{when } t_0 + i \cdot T_{\mathrm{f}} \le t \le t_0 + I \cdot T. \end{cases} \tag{7}$$

In the second stage composite degraded signal $v_{\Sigma}^{\mathrm{d}}(t)$ is made according to formula (4) by replacing the original signal $v_{\mathrm{T}}(t)$ with degraded one $v_{\mathrm{T}}^{\mathrm{d}}(t)$.

In reality, when forming signals according to (5) and (7) formulas, operations are performed not on signals $v_{\mathrm{T}}(t)$ and $v_{\mathrm{T}}^{\mathrm{d}}(t)$ themselves but on their codes V_i as follows from description (4).

The same procedure of signal construction is applied to both original and degraded signals.

Remarks on composite measurement signal

Signal shaped according to (5) present *J* periods of periodic signal. It is known that spectra of some signal and the same periodic prolonged signal are different. It means that properties of such signals are different, too. So, it is natural to ask whether periodical extension changes measured speech quality or not.

Answer to this question is clear – no. It follows from time division method and operations used by PESQ. The basic unit of PESQ algorithm [1] is a phoneme window with size of 32 ms. Overlap between successive phoneme windows is 50 %. The first algorithm step carries integration over frequency to estimate the measures of the perceived disturbance and of the asymmetric disturbance for every 32 ms long phoneme window. At the next step, phoneme disturbance values are aggregated using the nonlinear algorithm.

Experimental tool for quality estimation

The proposed structure of tool for quality of short voice signals estimation is shown in Fig. 1.

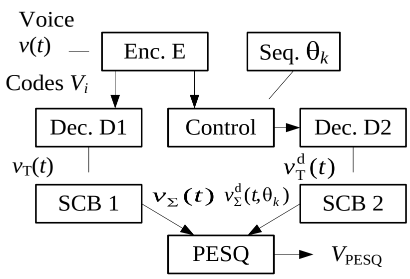

Fig. 1. Tool for quality estimation

Voice signal $v_T(t)$ after encoded in encoder E takes two paths. Signal from the first path is decoded at decoder D1 and extended in signal composition block SCB1. From the signal of the second path modified replica $v_T^d(t)$ is obtained by submitting the encoded signal V_i through the control block, which imitates a loss of random sequence θ_k of frames. Further, down the second path the degraded signal $v_T^d(t)$ is obtained by decoding of bits from control block. Signal composition block SCB2 finally generates signal $v_\Sigma^d(t,\theta_k)$. Both signals constructed according to formulas (3 – 7).

The PESQ measurer then produce voice quality value V_{PESQ} based on comparison of two extended signals: extended reference signal $v_\Sigma(t)$ and extended degraded signal $v_\Sigma^d(t,\theta_k)$.

In this paper, while performing voice quality evaluation experiments, measure of degradation of quality values were used

$$DV_{PESQ} = 4.5 - V_{PESQ} . \qquad (8)$$

Verification and validation

The experiments were performed using AMR-12.2 (3GPP TS 26.090) as encoder E and decoders D1 and D2. The encoding and decoding of the original and degraded signals using codec of the same type eliminates the influence of the codec on voice quality and only the impact of lost frames is evaluated.

The significant peculiarities of AMR coder could be listed as follows:

1. The codec divides input voice signal into 20 ms frames.

2. The error concealment algorithm (3GPP TS 26.091) includes active and passive concealment in the form of parameter repetition/substitution, and attenuation/muting. This approach sometimes can be damaging to speech quality and intelligibility, and can influence variations of the frame value. Therefore the measurement of voice frame value is conditional, dependent on the choice of voice codec.

Uncertainty of measurement results. The first experiment was performed to verify the uncertainty of measurements of short voice signals. In experiments, measurement window was moved along voice signal. The measurements were done using segments of original signals and composite signals. The measurement results:

1. The absolute uncertainty of voice frame quality value V_{PESQ} measurements with measurement window $T_{w.}$= 0.64 s using segments of original signals $v_t(t)$ and $v_t^d(t)$ varied from 0.07 to 1.3.

2. The same experiment performed using composite signals $v_\Sigma(t)$ and $v_\Sigma^d(t)$ gave variation of uncertainty only between 0.001 and 0.09. The uncertainty using composite signals is many times lower.

Impact of measurement window length. The second experiment was performed to verify the impact of measurement window length on voice quality measurement results using composite signals.

This experiment was performed using two measurement windows: $T_{w.}$= 0.64s and $T_{w.}$= 1.28s. Since the second window is longer, we kept identical experiment conditions by choosing the same initial parts of both signals. The frame losses are imitated in the first – identical parts of signals.

The distribution of results of performed measurements $DV_{PESQ}(1.28\,\mathrm{s})$ and $DV_{PESQ}(0.64\,\mathrm{s})/\sqrt{2}$ are shown in Fig. 2. The value of $DV_{PESQ}(0.64\,\mathrm{s})$ in this experiment is multiplied by $1/\sqrt{2}$ because as follows from (1) overall quality degradation is reduced with the increase of signal length. So, if we want to compare measurements of different durations, such scaling of data is mandatory.

The data showed in Fig. 2 certify that the measurement results for two measurement windows are approximate.

The calculated correlation coefficient of between these measurements is 0.92. This means that measurement window $T_{w.}$= 0.64 s is sufficient for voice quality measurement. Such short measurement window is useful for measuring the impact of single lost frame. The small duration of measurement window decreases overall measurement time.

Validation of measurement tool. The purpose of third experiment is comparison of quality measurement results obtained using standard PESQ and modified PESQ method when composite signal is used. This experiment seeks to show that method with composite signal not only improves accuracy of quality measurement but also obtained results are compatible with original PESQ methodology. This experiment is composed of two parts. First part uses standard PESQ and the second uses composite signals.

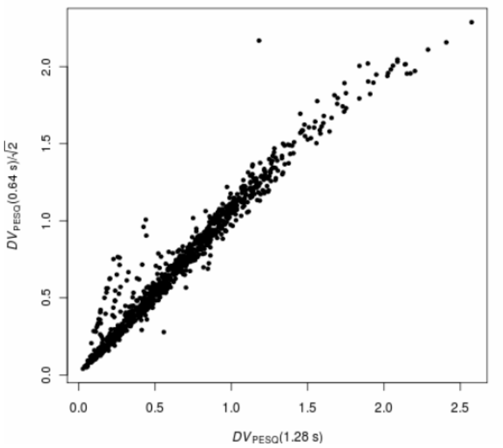

Fig. 2. The impact of measurement window length

Distributions of quality degradation values. In the first part of experiment signal $v_0(t)$ of 8 s duration is sent to encoder E from Fig. 1. This signal is encoded using AMR 12.2 coder and sequence of codes V_i is created.

Generator of lost frames creates random sequence θ_k of these frames with chosen FER. Frame errors indicated by θ_k may be at any place in entire 8 s duration signal. Error insertion block joins together codes V_i and error indication sequence θ_k. The result is sequence of codes V_i^{d} in which some codes are deleted according to θ_k.

Decoders D1 and D2 decodes both sequences V_i and V_i^{d}. The result of decoding are signals $v(t)$ and $v^{\mathrm{d}}(t)$. Both these signals go to PESQ measurer.

Measurements are performed simulating FER of 1, 2 and 4 %. The experiments for each FER were repeated 1000 times. Every time different sequence of frame errors θ_k were generated and used.

Measured degradation of quality value $\mathrm{D}V_{\mathrm{PESQ}}$ together with θ_k are stored in a database.

Distributions of speech quality degradation values were calculated using kernel estimator [11]

$$p(x) = \frac{1}{n \cdot h} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{9}$$

where kernel K is the standard normal distribution; n is the number of points (n = 1000 in our case); x_i – measured quality values we seek to approximate; h is a smoothing parameter calculated as

$$h = 0.9 \cdot \min\left\{\sigma, \frac{R}{1.34}\right\} \cdot n^{-0.2}. \tag{10}$$

There σ is standard deviation and R – interquartile range of data and n – number of data points (1000 in our case).

Speech quality degradation rate density functions of measured $\mathrm{D}V_{\mathrm{PESQ}}(\theta_k)$ values for two voice sentences are presented in Fig. 3: a) – FER=1%, b) – FER=2%, c) – FER=4%. As can be seen the same FER but different pattern of frame losses may cause quite a different degradation of quality values.

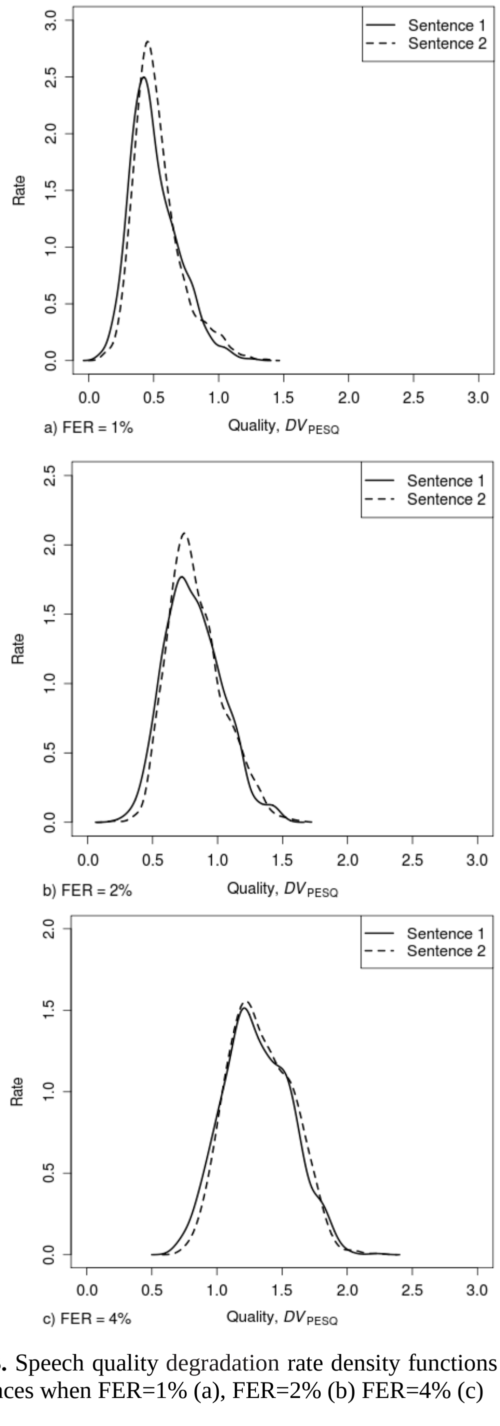

Fig. 3. Speech quality degradation rate density functions for two sentences when FER=1% (a), FER=2% (b) FER=4% (c)

In Fig. 3 presented quality degradation distributions are relatively wide. Calculated speech quality degradation means and standard deviations for two sentences are presented in Table 1.

Table 1. Statistical data of quality measurements

	Sentence 1		Sentence 2	
	Means	Dev.	Means	Dev.
FER=1%	0.513	0.183	0.529	0.178
FER=2%	0.828	0.229	0.844	0.210
FER=4%	1.315	0.238	1.313	0.242

It is obvious from table 1 that statistical data is consistent and sufficient and depend only a bit on the sentence. From presented data it is clear that FER is not sufficient parameter to make conclusions about speech quality.

Validation of modified PESQ. The second part of third experiment is performed using composite signals $v_\Sigma(t)$ and $v_\Sigma^{\mathrm{d}}(t)$. Initial components $v_T(t)$ and $v_{\mathrm{T}}^{\mathrm{d}}(t)$ for formation of composite signals are obtained from signals $v(t)$ and $v^{\mathrm{d}}(t)$ created for the first part of the experiment by dividing the latter signals into shorter (1 s) segments. These signal components are denoted accordingly $v_T(t, j)$ and $v_{\mathrm{T}}^{\mathrm{d}}(t, j)$. Here j =1,…, 8.

In this part of experiment composite signals $v_\Sigma(t)$ and $v_\Sigma^{\mathrm{d}}(t)$ were formed using signal composition blocks SCB1 and SCB2 from Fig. 1.

Quality measurements were performed with PESQ algorithm. Quality values $DV_{\mathrm{PESQ}}(\theta_k, j)$ obtained in this experiment depend not only on positions of lost frames θ_k, but also on the number j of divided signal.

It is important to note, that in this experiment the same copies of degraded signal and the same error vectors θ_k are used as in the first part of this experiment. Such method allows to compare measurement results obtained using original PESQ and results using the proposed method with composite signal construction.

Because overall quality distortion in PESQ is calculated in the last step as an aggregation of disturbances over the duration of the speech signal using typical mean square algorithm, cumulative quality degradation over all 8 s signal is calculated using mean square algorithm

$$DV_{\mathrm{PESQ}}^{\mathrm{composite}}(\theta_k) = \sqrt{\frac{1}{K}\sum_{j=1}^{K} DV^2(\theta_k, j)}. \qquad (11)$$

Scatter plot of $V_{\mathrm{PESQ}}(\theta_k)$ versus $V_{\mathrm{PESQ}}^{\mathrm{composite}}(\theta_k)$ is shown in Fig. 4. It can be seen that the results obtained using composite signal measurement and aggregation of results using mean square algorithm (11) are close to original PESQ results. Root mean square error between these two was 0.089 MOS points. This is quite a small difference and this shows that composite signal method is compatible with original PESQ. Correlation coefficient between data sets is 0.956. Distribution of absolute error between original PESQ quality and cumulative quality from shorter segments is presented in table 2. Presented data are calculated using the first sentence and FER of 2 %. As can be seen from Fig. 4 and Table 2 it is possible to approximate PESQ score in the long signal using PESQ measures of shorter signals which are parts of the long one. Such approximation has acceptable precision in majority of cases. Therefore it is possible to conclude that the method of composite signal expands the applicability of PESQ algorithm for quality measurement of short voice segments.

Table 2. Distribution of absolute difference between quality measurements

Difference of results	Rate of occurrence (%)
Absolute difference < 0.05	34.8
0.05 ≤ Absolute difference < 0.1	30.7
0.1 ≤ Absolute difference < 0.15	20.4
0.15 ≤ Absolute difference < 0.2	8.7
0.2 ≤ Absolute difference	5.4

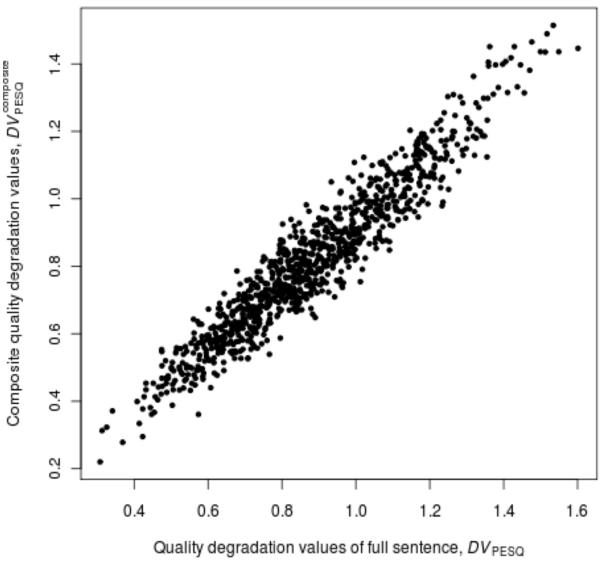

Fig. 4. Scatter plot of quality approximation in a long signal using shorter containing signals

Conclusions

Systemic analysis of impact of packet loss on voice quality using PESQ raises some problems. In the ideal case the impact of every possible packet loss pattern should be estimated. This is computationally infeasible using signal lengths which are recommended for PESQ [2]. The generalized conclusions about the impact of packet loss on voice quality may be deduced after many measurements performed using different combinations of lost packets.

In this paper proposed tool for quality measurement of short segments using composite measurement signal with conventional PESQ algorithm. The composite measurement signal used in the tool is especially valuable when a value of a single voice frame is investigated, because such investigation requires measuring quality of a short signal (0.5 s).

The proposed tool can be used for quality measurement of short (~0.5 s) voice signals. The verification results obtained using composite signal, are compatible with those generated by standard PESQ.

The proposed tool with composite measurement signal reduces delay time required for gathering of measurement signal and reduces the amount of experiments because the number of voice frames is smaller in short sentences. Processing delay may become important in real time calculations of packet quality values.

Acknowledgement

This research was supported by the Research Council of Lithuania.

References

1. ITU–T Rec. P.862, "Perceptual evaluation of speech quality (PESQ): An objective method for end–to–end speech quality assessment of narrow–band telephone networks and speech codecs", 2001.
2. ITU–T Rec. P.862.3 – Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2
3. **Hoene C., Dulamsuren–Lalla E.** Predicting performance of PESQ in case of single frame losses //Measurement of Speech and Audio Quality in Networks Workshop (MESAQIN), Prague, CZ, June 2004.
4. **Sun L., Ifeachor E.** New models for perceived voice quality prediction and their applications *in play– out* buffer optimization for VoIP networks // Proc. IEEE Communication, 2004. – Vol. 3. – P. 1478–1483.
5. **Kajackas A., Anskaitis A., Guršnys D.** Peculiarities of Testing the Impact of Packet Loss on Voice Quality // Electronics and Electrical Engineering. – Kaunas: Technologija, 2008. – No. 2(82). – P. 35–40.
6. **Kajackas A., Anskaitis A.** An Investigation of the Perceptual Value of Voice Frames // Informatica, 2009. – Vol. 20. – No. 4. – P. 1–12.
7. **De Martin J. C.** Source–driven Packet Marking for Speech Transmission over Differentiated–Services Networks // Proceedings of IEEE ICASSP, 2001. – Vol. 2. – P. 753–756.
8. **Sanneck H., Long Le N. T., Haardt M., Mohr W.** Selective Packet Prioritization for Wireless Voice over IP // Proceedings of WPMC, 2001. – Vol. 2. – P. 621–629.
9. **Hoene C., Rathke B., Wolisz A.** On the Importance of a VoIP Packet // Proc. of ISCA Tutorial and Research Workshop on the Auditory Quality of Systems, 2003.
10. **Mueller S. T., Seymour T. L., Kieras D. E., Meyer D. E.** Theoretical Implications of Articulatory Duration, Phonological Similarity, and Phonological Complexity in Verbal Working Memory // Journal of Experimental Psychology: Learning, Memory, and Cognition, 2003. – Vol. 29. – No. 6. – P. 1353–1380.
11. **Silverman B. W.** Density Estimation for Statistics and Data Analysis. – London: Chapman and Hall, 1996.

Received 2010 05 26

A. Anskaitis, A. Kajackas. The Tool for Quality Estimation of Short Voice Segments // Electronics and Electrical Engineering. – Kaunas: Technologija, 2010. – No. 8(104). – P. 97–102.

It is recognized that objective and reliable voice quality measure is PESQ algorithm. Systemic analysis of impact of packet loss on voice quality using PESQ raises some problems because each and every packet has distinct impact on speech quality and long signal as required in ITU specification [2] contains a lot of positions for possible packet losses. When there are a few lost packets in single measurement, the quality measured become a random number. The generalized conclusions about impact of packet loss on voice quality may be educed after many measurements performed using different combinations of lost packets. In this paper proposed tool for quality measurement of short segments using composite measurement signal with conventional PESQ algorithm. It is shown that the proposed tool is compatible with existing PESQ algorithm. The measurement results obtained using composite signals are compatible with those generated by standard PESQ. The proposed tool with composite measurement signal reduces the amount of experiments because the number of voice frames is smaller in short sentences. Ill. 4, bibl. 11, tabl. 2 (in English; abstracts in English, Russian and Lithuanian).

А. Анскайтис, А. Каяцкас. Средство для оценки качества коротких отрезков речевого сигнала // Электроника и электротехника. – Каунас: Технология, 2010. – № 8(104). – С. 97–102.

Общепризнано, что PESQ представляет собой достаточно совершенный алгоритм измерения качества речевого сигнала. Проведенные многочисленные эксперименты показывают, что различные стертые пакеты различно влияют на качество сигнала. При проведении одноразовых измерений поучаемые данные должны рассматриваться как случайные оценки. Для получения представительных выводов необходимо провести множество экспериментов, перебирая притом множество различных стираемых пакетов. В данной работе предложен алгоритм оценки качества коротких отрезков речевого сигнала, для этой цели адаптируя стандартный PESQ алгоритм. Показано, что предложенное средство оценки качества обеспечивает результаты измерения мало отличающиеся от тех, которые получаются при применении стандартного PESQ алгоритма. Предложенное средство создает новые возможности исследовать влияние отдельных пакетов и сокращает объем измерительных экспериментов. Ил. 4, библ. 11, табл. 2 (на английском языке; рефераты на английском, русском и литовском яз.).

A. Anskaitis, A. Kajackas. Trumpų balso atkarpų kokybės vertinimo priemonė // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2010. – Nr. 8(104). – P. 97–102.

PESQ yra plačiausiai taikomas balso kokybės matavimo algoritmas. Atliekant matavimo eksperimentus su šiuo algoritmu, nesunku įsitikinti, kad skirtingi ištrinti balso paketai bei skirtingas ištrinamų paketų išsidėstymas laiko ašyje skirtingai veikia balso kokybės įverčius. Kai atliekami vienkartiniai kokybės matavimai, gaunami duomenys vertintini tik kaip atsitiktiniai dydžiai, nes jų rezultatai atspindi matavimo metu ištrintų konkrečių paketų įtaką. Norint gauti apibendrintas išvadas, būtina atlikti daug eksperimentų ištrinant skirtingus, skirtingai išdėstytus paketus ir sukaupti išsamią statistiką. Šiame darbe pateikiamas naujas trumpų balso atkarpų kokybės vertinimo algoritmas bei sukurta priemonė tam tikslui pritaikant standartinį PESQ matuoklį. Parodoma, kad sukurta priemonė užtikrina matavimų rezultatus, artimus tiems, kurie gaunami atliekant matavimus su standartiniu PESQ. Sukurta matavimo priemonė daug kartų sumažina matavimo eksperimentų skaičių ir taip paspartina tyrimus. Il. 4, bibl. 11, lent. 2 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).