

تقييم تأثير الضجيج على الإشارة الكلامية/الصوتية

الدكتور المهندس: حسان محمد أحمد

مدرس في قسم هندسة الاتصالات والشبكات - كلية هندسة الحاسوب والمعلوماتية
والاتصالات - الجامعة السورية الخاصة

ملخص

يهدف هذا البحث إلى دراسة وتطوير طريقة تقييم مستوى تشوه الإشارة الكلامية بالاعتماد على نسبة الإشارة إلى الضجيج (Signal-to-Noise Ratio, SNR) واستناداً على نظرية الإدراك النشط (active perception)، ومن ثم بناء الخوارزمية المطلوبة. تستخدم الخوارزمية المقترحة لتقييم مستوى تشوه إشارة التحويل الواحد U (Unitary transform) للإشارات أحادية البعد (1D signals) وحساب السلاسل المتشابهة لأطياف الإشارة والمؤشرات المميزة لتشوهها.

الكلمات المفتاحية: معالجة الإشارة الرقمية، نظرية الإدراك النشط، نسبة الإشارة إلى الضجيج، التحويل الواحد، مستوى التشوه، طيف الإشارة.

Investigation of Noise Influence on Speech/Voice Signal

Dr. Hassan M. Ahmad,

Syrian Private University

Abstract

The aims of this study are evaluating the quality level of a distortion speech signal based on the Signal-to-Noise Ratio (SNR), and active perception theory, and design an algorithm.

The proposed algorithm is used to evaluate the distortion level of speech signal basing on the Unitary transform (U) of 1D signal, and on the calculation of similar strings of signal spectrum and distinctive indexes of signal distortion.

Keywords: DSP, theory of active perception, Signal-to-Noise Ratio [SNR], U-transform, distortion level, signal spectrum.

1. مقدمة

تعدّ مسألة تقييم مستوى تشوّه الإشارة الكلامية/الصوتية (speech\voice signal) إحدى أهم التحديات في معالجة الإشارات الكلامية. تحدث هذه المسألة عادةً أثناء اختبار خوارزميات ترميز (coding algorithms) الإشارة الكلامية وعند دراسة وبحث نظم الإرسال (transmission systems) المختلفة، وكذلك في مسائل تحديد الشخص المتحدث والتعرف على الكلام (Speaker Identification & recognition).

يعدّ التقييم الأولي للمادة الكلامية إحدى أهم مسائل تحديد هوية الشخص المتحدث باستخدام صوته (Voice Identification\verification) والتحقق منها [1]-[19]. وتعدّ نسبة مستوى الإشارة الكلامية الأولية إلى مستوى الضجيج المرافق للإشارة ذاتها (signal-to-noise ratio, SNR) أحد أهم المؤشرات المؤثرة على نوعية ودقة تحديد هوية الشخص المتحدث. ويمكن أن تتغير هذه النسبة وفقاً لتغيرات مستوى الضجيج وتغيرات برامترات (parameters) الإشارة الكلامية للمتحدث [13-14-19]. أجريت أبحاث كثيرة في هذا المجال [2-3-4-15]، من بينها مؤخراً "خوارزمية تقييم نسبة الإشارة إلى الضجيج للإشارات الكلامية" للباحث ستولبوف [19]، والمبنية على أساس حساب تقييمات الطيف الحالي للضجيج (current noise spectrum). بينت نتائج الباحث المذكور قيم مؤشرات تقييم مستوى نسبة الإشارة إلى الضجيج: نسبة عدد الأطياف البعيدة (Remote spectrums) إلى عدد الأطياف في الإشارة، $err_1 = 0.1059$ ، ونسبة عدد الأطياف البعيدة إلى عدد الأطياف في ذات الإشارة من دون الضجيج $err_2 = 0.1065$ مع $SNR = 15dB$.

2. هدف البحث

الهدف الرئيس من هذا البحث هو دراسة وتطوير طريقة لتقييم جودة الإشارة الكلامية المشوّهة بوجود الضجيج، وذلك على أساس نظرية الإدراك النشط (active perception) [20-21]، وبناء الخوارزمية اللازمة لعمل الطريقة.

3. مواد وطرق البحث

3.1 توصيف المسألة

يعدّ المعيار المعروف باسم نسبة الإشارة إلى الضجيج (SNR) أحد أهم المعايير المستخدمة في تقييم مستوى تشوّه الإشارة [5-18-19]. وبالنسبة للإشارات الكلامية، فتعرف هذه النسبة باسم النسبة العالمية لـ SNR، والتي تحدد كالاتي:

$$GSNR = 10 \log \frac{P_s^2}{P_n^2} \quad (1)$$

حيث:

P_s^2 - استطاعة/قدرة (power) الإشارة الكلامية الأولية،

P_n^2 - استطاعة/قدرة (power) إشارة الضجيج.

تعدّ هذه النسبة ملائمة لخصائص الإشارة الكلامية كاملة. ووفقاً لهذه العلاقة، يمكن اعتبار SNR المعيار الأمثل لتقييم المقاطع النشطة فقط من الإشارة الكلامية التي يتم تحليلها. في هذه الحالة، يمكن إعادة كتابة العلاقة (1) كالاتي:

$$SNR = 10 \log \frac{\sum_{n=0}^{l-1} s^2[n] \cdot inf[n]}{\sum_{n=0}^{l-1} v^2[n] \cdot inf[n]} \quad (2)$$

حيث:

$s[n]$ - عينة (sample) الإشارة الأولية (الكلامية) ذات المؤشر n ،

$v[n]$ - عينة إشارة الضجيج ذات المؤشر n ،

$inf[n]$ - المعلومات حول وجود الكلام لأجل العينة ذات المؤشر n من الإشارة.

$P_{s,i}^2$ - استطاعة/قدرة (power) الإشارة الكلامية عند المقطع (frame) i ،

يعدّ الكلام إشارةً شبه ثابتةً التي تتم معالجتها بشكلٍ رئيسي في مقاطع قصيرة،

عادة تكون ذات طول 30 ميلي ثانية تقريباً. بذلك، تكون طريقة حساب نسبة الإشارة إلى

الضجيج (SNR) إحدى أهم الطرائق المستخدمة في تقييم مستوى تشوّهات الإشارة الكلامية [8-10-11].

بالتالي، تُحسب نسبة الإشارة إلى الضجيج (SNR) لأجل مقطع i من إشارة كلامية بالعلاقة التالية:

$$SNR[dB] = 10 \log_{10} \frac{\sum_{n=0}^{N_s-1} s^2[n]}{\sum_{n=0}^{N_v-1} v^2[n]} = 10 \log_{10} \frac{P_{s,i}^2}{P_{v,i}^2} \quad (3)$$

$P_{v,i}^2$ - استطاعة/قدرة إشارة الضجيج عند المقطع i ،

N_s - عدد العينات في الإشارة الأولية،

N_v - عدد العينات في إشارة الضجيج.

لأجل استخدام العلاقة المعطاة يجب أن تتوفر المعلومات عن الإشارة والضجيج.

يمكننا كتابة عملية تشوّه الإشارة بالشكل الآتي:

$$g = \text{DEG}[f] \quad (4)$$

إذ إنّ:

f - الإشارة الكلامية الأولية،

g - الإشارة المشوّهة،

$[\text{DEG}]$ - الأمر (operator) المرتبط بعملية التشويه.

نحدد متطلبات تقييم الجودة كالتالي:

(1) يجب أن تتغير القيمة المحسوبة لمؤشر الجودة مع تغير مستوى تشوّه الإشارة (distortion level).

(2) يجب أن يتم إجراء تقييم قيمة مؤشر الجودة في زمن مقبول (يحدد وفق المسألة، على سبيل المثال: زمن أعظمي 20 s في مسائل التعرف على المتحدث [1]).

3.2 الخوارزمية المقترحة

تتلخص الخوارزمية المقترحة في الخطوات الآتية:

خطوة 1: بناء التمثيل الطيفي (spectral representation) للإشارة باستخدام التحويل الواحد U (Unitary transform) للإشارات أحادية الأبعاد. تتشكل نتيجة تطبيق الخوارزمية على الإشارة الأولية مجموعة M من الأطياف اللحظية، حيث إن M هو عدد مقاطع الإشارة (segments). كل طيف (spectrum) هو مجموعة من المعاملات $M = \{m_1, \dots, m_N\}$ التي تستخدم في تشكيل الطيف (N - عدد المرشحات).

يحسب الطيف لكل مقطع من الإشارة بطول len من العينات.

يتجلى تحويل U في إنشاء تمثيل متعدد المستويات للإشارة بمساعدة مرشحات وولش (Walsh filters) - نظام هارموت (Harmuth system) [6-22-23]، وفقاً للخطوات الآتية:

(a) استخدام مرشحات ذات طول واحد لأجل بناء كل مستوى من مستويات تفكيك (decomposition) الإشارة، والتي تضاعف لاحقاً حتى تصل إلى قياس مقطع الإشارة التي يجري تحليلها.

(b) تنفذ أولاً عملية الترشيح على كامل الإشارة، ومن ثم على كل مقطع منها. رياضياً، يكتب تحويل U كالآتي:

$$D = U(f) \quad (5)$$

حيث:

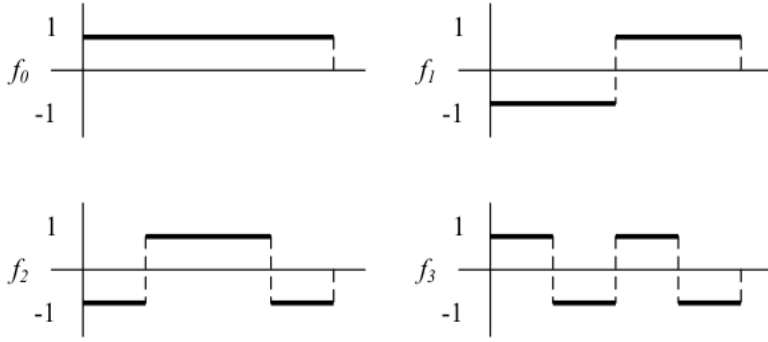
$D = \{D_{ij}\}$ ، و D_{ij} - هو الطيف ذو المؤشر j المتواجد على المستوى ذي المؤشر i

من التفكيك، $i \in [0; J - 1]$ و $j \in [0; M_i - 1]$ ،

J - عدد مستويات التفكيك،

M_i - عدد العناصر على المستوى ذي المؤشر i من التفكيك.

يبين الشكل (1) مرشحات وولش - نظام هارموت المستخدمة في تطبيق الخوارزمية.



الشكل (1): مرشحات وولش - نظام هارموت

خطوة 2: البحث عن السلاسل المتشابهة لأطياف الإشارة:

(1) نختار من المجموعة M سلاسل الأطياف L_i و L_j ($i \neq j$):

$$L_i = \{M_i, \dots, M_{i+c_{len}}\} \quad (6)$$

التي لم يتم استخدامها سابقاً في البحث، حيث:

M_i - الطيف i من المجموعة M ،

c_{len} - طول سلسلة الأطياف.

(2) إذا كانت القيمة المطلقة للفرق بين الأطياف المدرجة في L_1 و L_2 أقل من قيمة

عتبة محددة (thr)، فإن السلسلتين تعدان متشابهتان، وبعد ذلك يشار إلى

السلسلة L_2 على أنها مستخدمة (متكررة) ولا تستخدم في البحث لاحقاً.

(3) تحسب كمية السلاسل المتكررة (d) والفريدة غير المتكررة (nd) لأطياف

الإشارة.

خطوة 3: حساب المؤشر (index) المميز لتشوّه الإشارة.

نستخدم قيمتين لأجل تقييم مستوى تشوّه الإشارة:

(1) نسبة عدد الأطياف البعيدة (remote spectrums) إلى عدد الأطياف في

الإشارة:

$$err_1 = d/(nd + d) \quad (7)$$

حيث: $err_1 \in [0;1]$

(2) نسبة عدد الأطياف البعيدة إلى عدد الأطياف في ذات الإشارة من دون الضجيج:

$$err_2 = d/d_{wn} \quad (8)$$

إذ إنَّ d_{wn} - عدد الأطياف البعيدة بعد معالجة الإشارة من دون الضجيج، و $err_2 \in [0;1]$.

نلاحظ في كلتا الحالتين أن القيم الأقل للمؤشر توافق المستوى الكبير لتشوّه الإشارة.

تعدّ إحدى سلبيات هذه الطريقة أن الإشارة الأولية خالية الضجيج لا تكون دائماً متاحة.

لأجل المقارنة، نعرض النتائج الحسابية للمؤشرين المذكورين والتي حصلنا عليها بنتيجة التجارب العملية.

3.3 التجربة والنتائج ومناقشتها

أجريت التجارب الحسابية وتسجيل الإشارات باستخدام بيئة ماتلاب [9-12].

خلال التجربة، تم استخدام الأنواع الآتية من الضجيج:

- ضجيج جمعي عادي (Additive normal noise).
- ضجيج جمعي منتظم (Additive uniform noise).

وتم أيضاً استخدام الإشارات الاختبارية الآتية:

(1) تسجيلات صوتية لكلام 10 متحدثين (مثال: الشكل (2))، حيث تمت معايرة

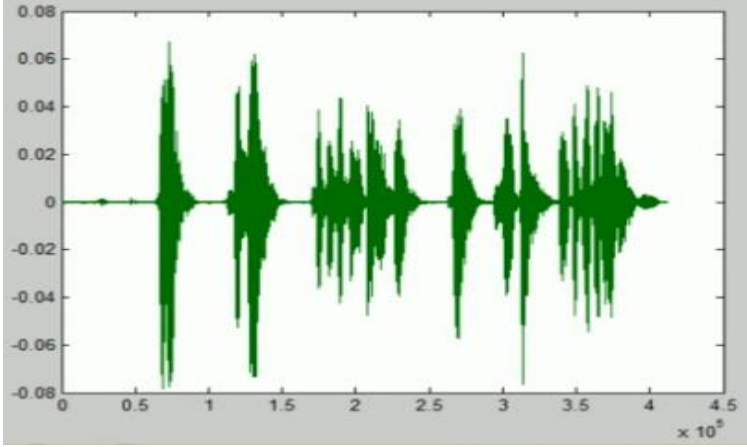
سعة/مطال (amplitude) كل تسجيل على المقطع (segment) [1; 0]، الشكل

(3). وتجدر الإشارة هنا إلى أنه، إذا لم تحصل المعايرة المشار إليها، فإن الضجيج

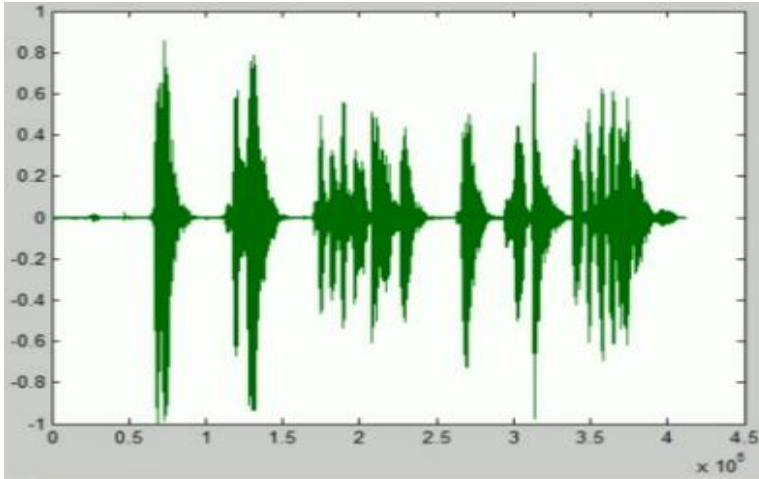
سيؤثر على الإشارة بأشكال مختلفة، وبالتالي فإن مقارنة نتائج عمل الخوارزمية على

إشارات مختلفة ستكون مستحيلة.

(2) إشارة جيبيية (sine) بتردد 50 هرتز.



الشكل (2): تسجيل صوتي لكلام متحدث

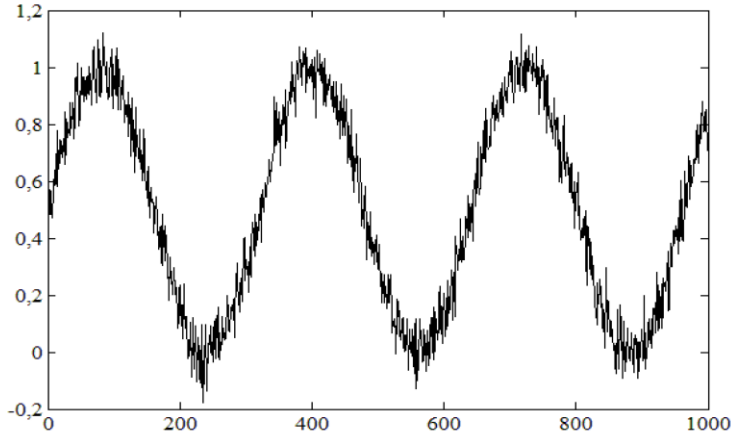


الشكل (3): الإشارة الكلامية بعد المعايرة

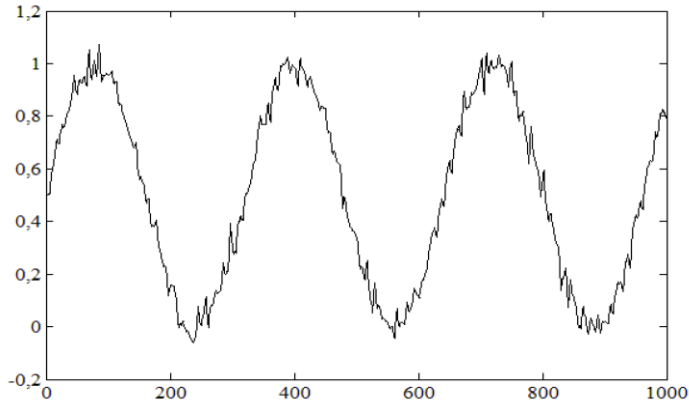
تقدم فيما يلي الطريقة المقترحة النتائج التالية:

1. يوجد طول أصغري (minimum) للإشارة $len_f > len$ ضروري لإجراء تقييم جودة الإشارة. إذا لم تحقق الإشارة التي يتم تحليلها الشرط المشار إليه فإن قيمة المؤشر المحسوبة ستكون دائماً مساوية للصفر، وبغض النظر عن مستوى التشويه. بينت التجارب التي أجريت في سياق البحث أن القيمة الوسطى للطول len_f ، بالنسبة للإشارة الكلامية، قد بلغت 0.5 ثانية.

2. نفترض وجود إشارتين مشوهتين g_1 و g_2 ، واللتين تم الحصول عليهما على أساس الإشارة f بمساعدة عملية الأمر $DEG[0]$. لأجل تشكيل الإشارتين g_1 و g_2 تم استخدام الضجيج ذاته، لكن مع جذر مربع متوسط انحراف مختلف (Root-Mean-Square Deviation (RMSD))، أي $\sigma_1 < \sigma_2$.
 بنتيجة تطبيق الخوارزمية المقترحة على كل من الإشارات، نحصل على تقييم مستوى تشوّه كل إشارة err_{g_1} و err_{g_2} ، حيث $err_{g_1} > err_{g_2}$.
3. تم استخدام مجموعات مختلفة لقيم البارامترات خلال مرحلة اختبار الخوارزمية. لأجل الاستخدام العملي للخوارزمية، وبنتيجة التجارب نقترح البارامترات الآتية:
 $c_{len} = 1$ ، $N = 4$ ، $thr = 0.1$ و $len = 32$.
 تم اختيار الطول len للمقطع الذي تجرى معالجته مساوياً 32 عينة، وذلك لأن نظام السمع يقيّم التغيرات التي تحصل في النبضة الصوتية على مجال زمني أقل من 1 ميلي ثانية (1ms) [2-17]، مع الأخذ بالاعتبار أن 2 ميلي ثانية (2ms) توافق 32 عينة عند تردد أخذ العينات (sampling frequency) من 16 كيلوهرتز (16 kHz).
 يمكن أن يتغير طول الإشارة (c_{len}) التي يجري تحليلها ومعالجتها وفقاً لدقة تقييم مستوى التشوّه المطلوبة.
4. حصلنا من خلال التجربة على أن قيم المؤشر err_1 (err_2) تنخفض مع زيادة تشتت الضجيج والإشارة المشوّهة (noise dispersion) لأجل كل أنواع الضجيج المستخدمة.
5. يرتبط زمن حساب مؤشر التقييم بطول الإشارة وقيم البارامترات المختارة. على سبيل المثال، لأجل البارامترات المقترحة أعلاه وطول إشارة من 50000 عينة، فإن زمن التقييم هو 54 ثانية.
- يبين الشكل (2) إشارة دورية ($SNR = 15\text{ dB}$) والتي تم تشوّيها بضجيج عادي، بينما يبين الشكل (3) الإشارة ذاتها بعد الترشيح والاستعادة.



الشكل (4): إشارة دورية مشوّهة

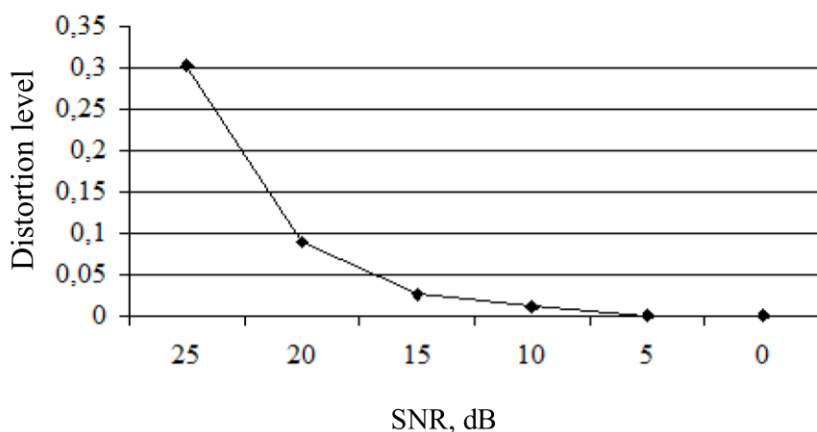


الشكل (5): الإشارة المستعادة

يبين الجدول (1) نتائج حساب المؤشرات err_1 و err_2 لكل من الإشارة الدورية المشوّهة والإشارة الدورية ذاتها بعد الاستعادة، (انظر الشكل(6))، في حال ضجيج عادي. ومن خلال المقارنة، نلاحظ أن قيمة هذه المؤشرات متساوية تقريباً.

الجدول (1): إشارة دورية (sin = 50Hz)، ضجيج عادي

0.609	0.340	0.199	0.109	0.063	0.034	RMSD للضجيج
الإشارة المشوّهة (distortion signal)						
0	5	10	15	20	25	SNR, dB
0	0.001	0.012	0.026	0.090	0.303	err_1
0	0.001	0.013	0.027	0.091	0.305	err_2
الإشارة المستعادة (recovered signal)						
1.947	7.811	12.853	16.301	22.348	26.850	SNR, dB
0.028	0.042	0.054	0.106	0.182	0.419	err_1
0.029	0.043	0.055	0.107	0.183	0.422	err_2



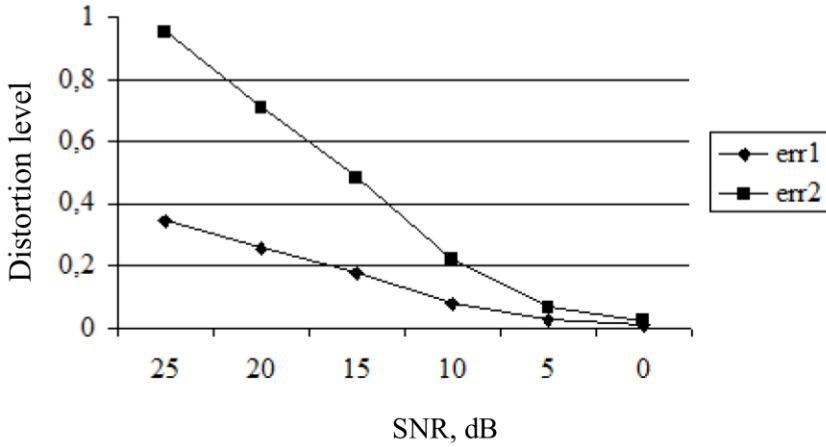
الشكل (6): العلاقة بين قيمة المؤشر $err_1 \approx err_2$ ومستوى تشوّه الإشارة في حال ضجيج عادي

يقدم الجدول (2) نتائج دراسة وبحث الإشارة الكلامية قبل وبعد التشويّه، (انظر

الشكل (7)).

الجدول (2): إشارة كلامية، ضجيج منتظم

0.222	0.121	0.069	0.040	0.024	0.012	RMSD للضجيج
الإشارة المشوّهة (distortion signal)						
0	5	10	15	20	25	SNR, dB
0.007	0.024	0.079	0.175	0.257	0.345	err_1
0.021	0.067	0.219	0.484	0.710	0.952	err_2
الإشارة المستعادة (recovered signal)						
1.618	5.966	14.495	18.278	25.483	30.321	SNR, dB
0.033	0.060	0.119	0.224	0.304	0.361	err_1
0.091	0.166	0.330	0.538	0.837	0.994	err_2



الشكل (7): العلاقة بين قيم المؤشرات err_1 و err_2 ومستوى تشوّه الإشارة

لأجل المقارنة، نورد فيما يلي النتائج التي توصل إليها الباحث ستولبوف في هذا المجال [19]، الجداول (3) و(4).

الجدول (3): إشارة دورية ($\sin = 50\text{Hz}$)، ضجيج عادي

الإشارة المشوّهة (distortion signal)					
0	5	10	15	20	SNR, dB
0	0.004	0.014	0.029	0.093	err_1
0	0.006	0.016	0.030	0.094	err_2
الإشارة المستعادة (recovered signal)					
2	8	12	16	22	SNR, dB
0.030	0.045	0.057	0.108	0.185	err_1
0.033	0.047	0.058	0.110	0.187	err_2

الجدول (4): إشارة كلامية، ضجيج منتظم

الإشارة المشوّهة (distortion signal)					
0	5	10	15	20	SNR, dB
0.010	0.027	0.082	0.180	0.260	err_1
0.031	0.077	0.229	0.494	0.720	err_2
الإشارة المستعادة (recovered signal)					
2	5	14	18	25	SNR, dB
0.043	0.070	0.129	0.234	0.310	err_1
0.098	0.176	0.340	0.548	0.847	err_2

وهكذا، فإن الطريقة المقترحة في هذا البحث قد بينت في الواقع انخفاض مستوى تشوّه الإشارة بعد عملية الترشيح والاستعادة، وهو ما يتوافق مع النتائج المشار إليها في الجداول (1) و(2).

تعرض قيم المؤشرات err_1 و err_2 السلوك ذاته، لذلك من الضروري استخدام المؤشر err_1 في التطبيقات العملية لأن حسابه لا يتطلب وجود معلومات عن الإشارة

الخالية من الضجيج. تكون قيم المؤشرات المحسوبة على أساس الطريقة المقترحة متوافقة ومتناسقة مع قيم علاقة الإشارة إلى الضجيج (SNR) للإشارات ذاتها.

5. الاستنتاجات والتوصيات

تم في هذا العمل توصيف خوارزمية تقييم مستوى تشوّه الإشارة. تؤكد نتائج التجارب إمكانية استخدام الخوارزمية المقترحة لأجل حل أي مسألة مفروضة مرتبطة بالإشارة الكلامية، ومنها:

- مسائل النظم الآلية لتأكيد أو تحديد هوية الشخص باستخدام بصمة صوته.
- المسائل الأمنية (التحكم بالدخول الى اماكن العمل و قواعد البيانات والنظم الحاسوبية والمركبات والأسلحة).
- مسائل الفحوص الجنائية (تحليل المكالمات الهاتفية كأدلة في الاجراءات القضائية للتأكد من المتهمين و التعرف عليهم).
- مسائل إدارة الكومبيوتر والاجهزة المنزلية والمعدات الصناعية.

من أهم الميزات الإيجابية للخوارزمية المقترحة: بساطة وسهولة العملية الحسابية، سهولة التطبيق والتنفيذ، وأيضاً إمكانية تقييم مستوى التشوّهات على أساس الإشارة المتاحة فقط.

يمكن العمل مستقبلاً على تطوير هذه الخوارزمية لاستخدامها في حل المسائل المتعلقة بوضوح الكلام وسهولة فهمه، على سبيل المثال: وضوح نطق الكلمات (words intelligibility).

المراجع

1. AHMAD. H, M, **2016** Method of independent text to identify the speaker using his voice, Syria, Journal of Al-Baath University, Volume 38, No. 13, p. 79.
2. ARRAJ. F, DAYOUB, F **2015** Study of the Effect of Modified Spectral Algorithm Modulation Factors and the Length of Time Window in Improving Sound Signals, Tishreen University Journal for Studies and Scientific Research, Engineering Sciences Series, Volume 37, No. 6, P. 120.
3. DONOHO. D, L, 1995 De-noising by soft-thresholding // IEEE Transactions on Information Theory. Vol. 41, No. 3. P. 613–627
4. DONOHO. D, L, JOHNSTONE I M 1996 Ideal denoising in an orthonormal basis chosen from a library of bases // C.R.A.S. Paris. T. 319. Ser. I. P. 1317–1322.
5. FALEK. L, TEFFAHI. H, DJERADI. A **2012** A study method of noise effects on the acoustic parameters of speech signal. Multimedia Computing and Systems (ICMCS), International Conference. IEEE Xplore: 04 October, P.136 - 141.
6. HENNING. F, H 1969 Applications of Walsh Functions in Communications. University of Maryland, IEEE spectrum, November, V. 6 , Issue: 11, P.82-91.
7. HENNING. F, H 1972 Transmission of Information by Orthogonal Functions. Second Edition, Springer-Verlag, New York, Heidelberg, Berlin, 320 P.
8. HESS. W **2012** – Pitch Determination of Speech Signals: Algorithms and Devices. Springer Science & Business Media, Dec 6, 2012 , 698 p.
9. NAN WU, BOFEI WANG **2014** Process and Analysis of Voice Signal by MATLAB. University of Gavle, June 2014, 48 P.
10. PAPADOPOULOS. P, ANDREAS T **2016** Long-Term SNR Estimation of Speech Signals in Known and Unknown Channel Conditions. IEEE/ACM Transactions on Audio, Speech, and Language Processing. Volume: 24, Issue: 12, Dec.2014.
11. PAPADOPOULOS. P, ANDREAS. T, JAMES. G, and SHRIKANTH. N **2014**. A supervised signal-to-noise ratio

- estimation of speech signals. IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP), P.126 - 141.
12. PEREZ-MEANA. H **2007** Advances in Audio and Speech Signal Processing: Technologies and Applications: Technologies and Applications. Idea Group Inc (IGI), Feb 28,2014.
13. SIMON. D, and MOONEN. M **2005** On the Output SNR of the Speech-Distortion Weighted Multichannel Wiener Filter. IEEE, signal processing letters, Vol. 12, N° 12, PP. 809-811.
14. SNR-of-noisy-signal 21 May **2015** <<http://dsp.stackexchange.com/questions/4889/how-do-i-calculate-snr-of-noisy-signal> >.
15. TIWARI. N, PANDEY. P **2015** Speech Enhancement Using Noise Estimation Based on Dynamic Quantile Tracking for Hearing Impaired Listeners. Proc. 21th National Conference on Communications 2015 (NCC 2015), Mumbai, Feb. 27 - Mar. 1, 2015, PP. 609-611.
16. VONDRASEK. M, POLLAK. P 2005 – Methods for Speech SNR estimation: Evaluation Tool and Analysis of VAD Dependency. Prague, Dejvice, Czech Republic, Radio-Engineering, Vol. 14, N° 1, APRIL 2005.
17. ПОЗИН. Н, В 1978 – Элементы теории биологических анализаторов. М.: Наука. Главная редакция физико-математической литературы, Россия, 360 с.
18. РАБИНЕР. Л, ГОУЛД. Б 1978 – Теория и применение цифровой обработки сигналов. М.: Мир, Россия, 848 с.
19. СТОЛБОВ. М, Б **2012** Алгоритмы оценки отношения сигнал/шум речевых сигналов. Научно-технический вестник информационных технологий, механики и оптики, , № 6 (82), Россия.
20. УТРОБИН. В, А 2001 – Информационные модели системы зрительного восприятия для задач компьютерной обработки изображений. Н. Новгород: НГТУ, Россия, 184 с.
21. УТРОБИН. В, А 2004 – Компьютерная обработка изображений. Принятие решений в пространстве эталонов. Н. Новгород: НГТУ, Россия, 228 с.

22. ХАРМУТ. Х, Ф 1975 – Передача информации ортогональными функциями: Пер. с англ. М.: Связь, Россия, 272 с.
23. ХАРМУТ. Х, Ф 1989 – Применение методов теории информации в физике: Пер. с англ. М.: Мир, Россия, 174с.