

TEXNİKİ VASİTƏLƏRLƏ GENERASIYA OLUNMUŞ FONOQRAMLARDA DANIŞANIN EYNİLƏŞDİRİLMƏSİ PROBLEMLƏRİ

Açar sözlər: süni intellekt, fonogram, neyroşəbəkə, nitqin sintezi, formant, əsas ton, tembr, paradiya.

Keywords: artificial intelligence, phonogram, neural network, speech synthesis, formant, fundamental tone, timbre, paradia.

Anahtar kelimələr: yapay zeka, fonogram, sinir ağı, konuşma sentezi, formant, tuş tonu, tını, paradia.

Hazırda süni intellekt vasitəsi ilə yaradılan fonogramlar mürəkkəb alqoritmlər vasitəsi ilə səsin əsas ton (eləcə də əsas tonun törəmə parametrlərini) və temporal parametrlərini dəyişməklə istənilən səsin imitasiyasının aparılması mümkündür.

Bəzi süni intellekt proqramları vasitəsi ilə (məsələn, Stable Diffusion) şəkildə sintez olunmuş səsə uyğun olaraq şəkil canlanaraq (gözünü qırpır, başını tərpədir) görüntüdəki danışığa uyğun emosiya yaranır.

Neyroşəbəkə vasitəsi ilə səsin və sifətin klonlaşması, sintezi və mətnin nitqə çevrilməsi kimi proseslər real vaxt ərzində həyata keçirilir və bu innovativ nailiyyətlərdən müxtəlif məqsədlər, o cümlədən xoş olmayan niyyətlər üçün istifadə olunur.

Ümumilikdə səsin və nitqin aşağıdakı dəyişdirilmə növləri mümkündür:

- səsin yüksəkliyinin, tembrinin, melodiya və intonasiya kimi prosodik xarakteristikalarının dəyişdirilməsi yolu ilə başverən dəyişikliklər;
- səsin yüksəkliyinin, tembrinin, melodiya və intonasiya kimi prosodik xarakteristikalarının

dəyişdirilməsinə leksik-qrammatik fəndlər və tələffüzdə fonetik-akustik dəyişdirmələr əlavə etməklə baş verən dəyişikliklər (səsin və nitqin “maskalanması”);

- səsin və şifahi nitqin hər hansı məlum “etalon” səsə uyğunlaşdırılması yolu ilə yuxarıda qeyd olunan xarakteristikaların dəyişdirilməsi ilə baş verən dəyişikliklər (səsin və nitqin “parodiyası”, “sintezi”).

Danışanın səsinin və şifahi nitqinin xarakteristikalarının yuxarıda qeyd olunan növ dəyişdirilməsinə informasiya texnologiyasının imkanlarından və danışanın nitq aparatının təbii funksional rejiminin “fizioloji” dəyişdirilməsindən istifadə etməklə nail olmaq mümkündür.

İnformasiya texnologiyasının imkanlarından istifadə etməklə səsin və nitqin dəyişdirilməsinin aşağıdakı üsulları vardır:

- səsin xarakteristikalarını “aparat kompüter texnologiyası” adlanan səsin blyutus-dəyişdirici, skrembler vasitələri ilə real vaxt ərzində dəyişdirilmələri;
- səsin xarakteristikalarını “İP-telefonıya”

kompiuter texnologiyasından istifadə etməklə (giriş signalında səsə tonunu və tembrini dəyişməklə yanaşı “hazır səslər” kolleksiyasından istifadə etməklə həqiqi akustik şəraiti maskalamaqla qatar, bazar, yağış və s. üçün xarakterik səslər əlavə edir) emal edilərək real vaxt ərzində dəyişdirilməsi;

- səsə xarakteristikalarını “səsə stasionar emalı” kompiuter texnologiyasının köməyi ilə (“plaqion” qismində proqram təminatına qoşulmuş xüsusi modul insanın səs traktını modullaşdıraraq səsə xarakteristikalarını dəyişir) emal edilərək dəyişdirilməsi;

- “sintez kompiuter texnologiyası”ndan (“text-to-speech”) istifadə etməklə yığılmış mətnin müxtəlif səslərlə səsləndirilməsi;

- müxtəlif kompiuter texnologiyalarının köməyi ilə “etalon” şifahi nitqin məlum xarakteristikalarına uyğun danışanın şifahi nitqinin dəyişdirilməsi (səsə klonlaşması, səsə parodiyası kimi xidmətlər göstərən internet-sayt xidmətləri).

Danışanın nitq aparatının təbii funksional rejiminin “fizioloji” dəyişdirilməsindən istifadə etməklə səsə və nitqin dəyişdirilməsinin aşağıdakı üsulları mövcuddur:

- danışanın nitq aparatının təbii funksional rejiminin “fizioloji” dəyişdirilməsindən istifadə etməklə səsə yüksəkliyinin, tembrinin, melodiya və intonasiya kimi prosodik xarakteristikalarının dəyişdirilməsi ilə olan dəyişikliklər;

- danışanın nitq aparatının təbii funksional rejiminin “fizioloji” dəyişdirilməsindən istifadə etməklə səsə yüksəkliyinin, tembrinin, melodiya və intonasiya kimi prosodik xarakteristikalarının dəyişdirilməsinə leksik-qrammatik fəndlər və tələffüzdə fonetik-akustik dəyişdirmələr əlavə etməklə olan dəyişikliklər;

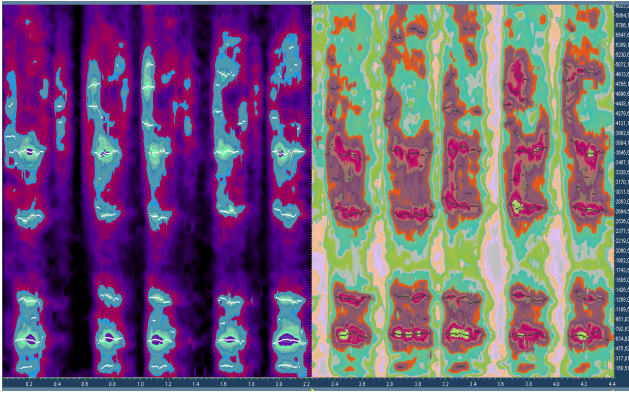
- danışanın nitq aparatının təbii funksional rejiminin “fizioloji” dəyişdirilməsindən istifadə etməklə səsə və şifahi nitqin hər hansı məlum “etalon” səsə uyğunlaşdırılması yolu ilə yuxarıda qeyd olunan xarakteristikaların dəyişdirilməsi ilə olan dəyişikliklər.

Məsələn, danışanın nitq aparatının təbii

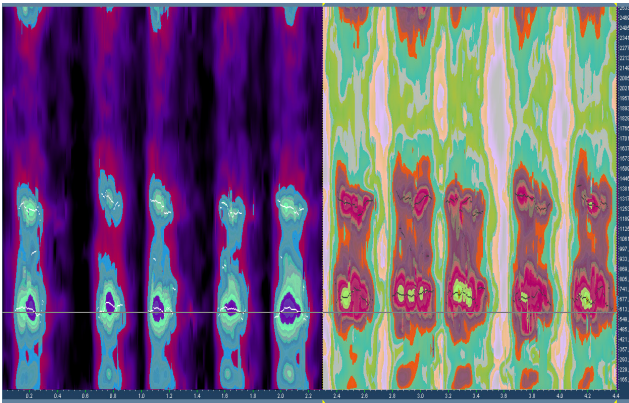
funksional rejiminin “fizioloji” dəyişdirilməsindən istifadə etməklə səsə dəyişdirilməsinin müəyyən olunması zamanı auditiv-linqvistik və instrumental (akustik) analiz zamanı daha çox diqqəti fonetik-linqvistik və formant analizə yönəltmək lazımdır. Qeyd olunan üsulla dəyişdirilmə zamanı səsə prosodik xarakteristikaları ilə yanaşı formantların düzlüyü də dəyişir. Dəyişdirilmənin xüsusi halı olan təqlid zamanı da formantların düzlüyü dəyişir.

Təqlidçinin (ifadəçinin) və təqlid etmək istədiyi diktorun (etalon diktorun) nitqindən götürülmüş eyni sitlər üçün “SİS” və “Justiphone” proqramlar komplekslərinin köməyi ilə formantların hesablanması göstərir ki, təqlidçi həmin səslər üçün yalnız ilk 2-3 formantın yerini dəyişdirə bilər. Çünki hər bir danışan şifahi nitqin səsləndirilməsi zamanı özünün vokal traktının quruluşunu yalnız müəyyən anatomik məhdudiyyətlər daxilində dəyişə bilər və 4-cü və ondan yuxarı formantlara diktor eşitmə yolu ilə “nəzarət” edə bilmir. Ona görə də tədqiqata təqdim olunmuş mübahisəli və müqayisəedici fonogramlardakı danışiq hissəsində uyğun vəziyyət və uyğun emosional halda deyilmiş eyni sözlərdəki eyni vəziyyətdə olan saitlərin formantlarının müqayisəsində ilk 2-3 formant tezliklərinin bərabərliyi, yuxarı formant tezliklərinin isə müxtəlifliyi qeydə alınarsa bu müqayisə olunan diktorlardan birinin başqa şəxsi təqlid etdiyini deməyə əsas verir. Yuxarı tezlik formantlarının nə qədər fərqlənməsi təqlidçinin və təqlid etmək istədiyi diktorun vokal traktının anatomik quruluşlarının nə qədər fərqlənməsindən asılı olacaqdır.

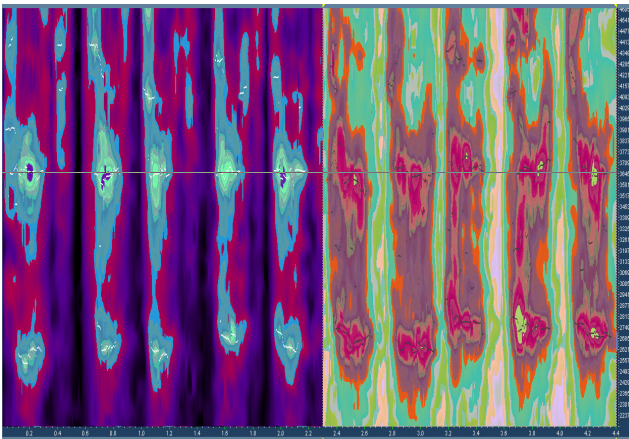
Aşağıdakı şəkildə nitq aparatının təbii funksional rejiminin “fizioloji” dəyişdirilməsindən istifadə etməklə və diktorun təbii tələffüzü zamanı ifadə olunmuş “tad” və “dad” sözlərində qeydə alınan [a] səsənin formant trayektoriyaları göstərilmişdir. Şəkildə birinci hissədə diktorun təbii tələffüzü zamanı səsləndirilən, ikinci hissədə isə digər diktoru təqlidi zamanı səsləndirilən [a] səsənin “Justiphone proqram kompleksi” vasitəsi ilə əldə olunmuş formant trayektoriyaları verilmişdir.



Şəkil 1. [a] səsinin ilk 4 formant (F1, F2, F3 və F4) trayektoriyalarının təsviri.



Şəkil 2. [a] səsinin 1-ci (F1) və 2-ci (F2) formant trayektoriyalarının təsviri.



Şəkil 3. [a] səsinin 4-cü formant (F4) trayektoriyasının təsviri.

Təqdim olunan misaldan göründüyü kimi professional təqlidçi tərəfindən dəyişdirilmiş və təbii [a] səsin 1-ci, 2-ci və 3-cü (F1, F2, F3) formant tezliklərinin fərqli olmasına baxmayaraq 4-cü formant tezlikləri eynilik təşkil edir.

Çoxsaylı nümunələr üzrə aparılmış tədqiqatlar təqlid zamanı müqayisə olunan olunan sait səslərdə ən yaxşı halda ilk üç formant

tezliyinin (F1, F2, F3) bərabərləşdirilməsinin, qəsdən dəyişdirilmə halında isə aşağı formant tezliklərinin fərqləndirilməsinin mümkün olduğunu göstərmişdir.

Göründüyü kimi, formant analizi aparmaqla tədqiq olunan fonoqramda səsin “fizioloji” dəyişdirilməsindən istifadə etməklə qəsdən dəyişdirmə və yamsılama əlamətlərini müəyyən etmək mümkündür. Süni intellekt vasitəsi ilə sintez olunmuş fonoqramlarda səsin akustik parametrlərinin necə və hansı alqoritmlərlə dəyişdirilməsinin müəyyən olunması isə dərin analiz tələb edir və həllini gözləyən məsələlərdəndir.

Səsə və nitqə görə şəxsin eyniləşdirilməsi üzrə tədqiqatda ekspert ilk əvvəl diktorun nitq materialında yamsılama və qəsdən dəyişdirilmə əlamətlərinin olub-olmamasını yoxlamalıdır. Fonoqramın texniki üsulla generasiya olunub-olunmamasının müəyyən olunması üzrə ümumi ekspert metodikasının və bu sahədə kvalifikasiyalı ekspertlərin hazırlanması isə günün əsas məsələlərdəndir.

Texniki üsullarla edilmiş digər dəyişikliklər kimi texniki üsulla səs generasiyanın da fonoqramda öz izini buraxdığına inanaraq dəyişik əlamətlərinin müəyyən olunması üzrə ümumi metodikanın hazırlanmasında səsə və nitqə görə danışanın eyniləşdirilməsi üzrə, montaj üzrə və kompüter texniki ekspertizası üzrə mütəxəssislərdən ibarət tədqiqat qrupunun yaradılmasını günün təxirəsalınmaz məsələsi hesab edirik.

Ədəbiyyat siyahısı:

1. Г.А. Мусаев, Л.П. Алиев. Проблемы исследования цифровых записей телефонных (GSM) разговоров, измененных с помощью разных устройств. «Развитие новых видов и направлений судебной экспертизы»: материалы Всероссийского семинара. ФБУ Южный РЦСЭ Минюста России. – Ростов-на-Дону, 2011. – 130-131 с.

2. Hamlet Musayev, Lətif Əliyev. Təqlid və onun müəyyən edilməsində formantların rolu. Azərbaycan Respublikası, Ədliyyə Nazirliyi, Məhkəmə Ekspertizası, Kriminalistika və kriminologiyanın aktual məsələləri № 39, 2003-cü il, s.101-103.

SPEECH EQUALIZATION PROBLEMS IN TECHNICALLY GENERATED PHONOGRAMS

LATİF ALİYEV

*Forensic Science Center of the Ministry of
Justice Deputy head of the Department of
Manuscripts and Sound Examinations PhD in Physics*

PARVİN NAZARLI

309 to the full secondary school teacher

SUMMARY

Currently, the phonograms created by artificial intelligence are simulated by changing the main tone (as well as the derivative parameters of the main tone) and temporal parameters of the sound by means of complex algorithms. While there is a methodology for determining the signs of intentional change and distortion in the studied phonogram by using “physiological” change of sound by conducting formant analysis, determining how and with what algorithms to change the acoustic parameters of the sound in phonograms synthesized by artificial intelligence is one of the issues that are waiting to be solved.

TEKNİK OLARAK ÜRETİLMİŞ FONOGRAMLARDA KONUŞMA EŞİTLEME SORUNLARI

LATİF ALİYEV

*Adalet Bakanlığı Adli Uzmanlık Merkezi El Yazmaları
ve Ses Sınavı Daire Başkan Yardımcısı*

PARVİN NAZARLI

Tam ortaokul öğretmene 309

ÖZET

Günümüzde yapay zeka tarafından oluşturulan fonogramlar, karmaşık algoritmalar aracılığıyla sesin ana tonu (aynı zamanda ana tonun türev parametreleri) ve zamansal parametreleri değiştirilerek simüle edilmektedir. Formant analizi yapılarak sesin “fizyolojik” değişimini kullanarak incelenen fonogramdaki kasıtlı değişiklik ve distorsiyon işaretlerini tespit etmeye yönelik bir metodoloji mevcutken, yapay zeka tarafından sentezlenen fonogramlarda sesin akustik parametrelerinin nasıl ve hangi algoritmalarla değiştirileceğinin belirlenmesi çözülmeyi bekleyen konulardan biridir.