

УДК 621.391

DOI: 10.53297/18293336-2021.1-102

## ПРИМЕНЕНИЕ ОКОННОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ В ПРИЛОЖЕНИЯХ ГОЛОСОВОЙ БИОМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ

М.С. Маргарян<sup>1,2</sup>, Б.Ф. Бадалян<sup>2</sup>, О.А. Гомцяян<sup>2</sup>, Д.О. Мосоян<sup>2</sup>

<sup>1</sup> “National Instruments” AM LLC

<sup>2</sup> Национальный политехнический университет Армении

В настоящее время биометрические системы идентификации являются неотъемлемой частью нашей жизни. Сканеры отпечатков пальцев, встроенные в смартфоны, технологии распознавания голоса или речи и прочие инструменты постепенно приходят на замену традиционным методам идентификации и всё чаще проникают в такие крупные сферы, как банковское обслуживание, визовые и миграционные службы, медицина, системы безопасности и т.д. Биометрические системы, по сравнению с традиционными методами, имеют ряд преимуществ: они приспособлены под идентификацию личности без возможности передачи секретного ключа и во многом являются более удобными с точки зрения пользователя. Однако чем активнее ведётся внедрение такого вида систем, тем более остро встаёт вопрос обеспечения информационной безопасности.

По мере интенсивного развития технологий биометрической идентификации и аутентификации в мире проявляется повышенный интерес к разработке инновационных решений по голосовой биометрии. Голос — такая же неотъемлемая черта каждого человека, как и его лицо или отпечатки пальцев. Широкое распространение средств связи открывает большие возможности для применения данного идентификатора. Кроме того, распознавание по голосу весьма удобно для пользователей и требует от них минимум усилий. Системы аутентификации по голосу при записи образца и в процессе последующей идентификации опираются на такие уникальные для каждого человека особенности голоса, как высота, модуляция и частота звука. Эти показатели определяются физическими характеристиками голосового тракта и уникальны для каждого человека.

В данной работе приведена методика анализа голосового сигнала с применением кратковременного Фурье-преобразования с разными оконными функциями. Рассмотрены основные ограничения традиционных методов преобразования Фурье для задач голосовой идентификации. В программном пространстве LabVIEW 2020 смоделирован модуль распознавания голоса.

**Ключевые слова:** биометрические системы идентификации, быстрое преобразование Фурье, оконное преобразование Фурье, парольная фраза, оконная функция, программное обеспечение LabVIEW.

**Введение.** Большинство биометрических голосовых систем безопасности функционируют следующим образом: в базе данных системы хранится цифровой шаблон голоса. Человек, собирающийся получить доступ к компьютерной сети, с помощью микрофона вводит в систему информацию о себе. Поступившие данные сравниваются с образцом, хранимым в базе данных.

Типичная система голосовой идентификации личности состоит из следующих модулей: получение речевого сигнала, выделения образца голоса, выделение отличительных признаков голоса (параметризация), сравнение полученного образца голоса с шаблоном из базы данных и принятие решения "допуск/отказ". Схема верификации голоса представлена на рис. 1.

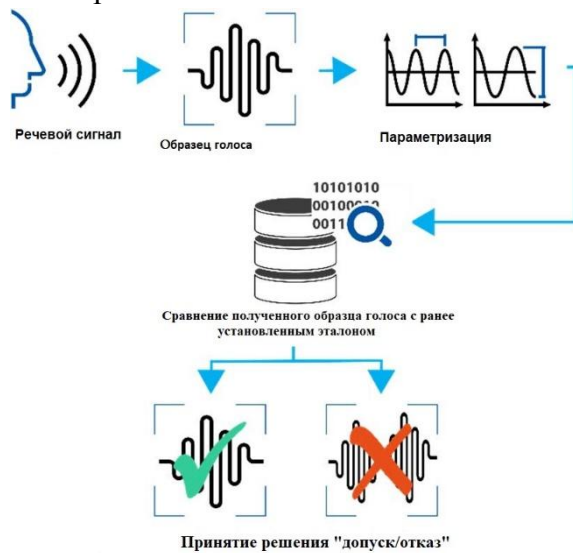


Рис. 1. Схема распознавания голоса

Когда начинается обработка голосового сигнала, он разбивается на дискретные "звуковые кадры", затем происходит преобразование в цифровую модель. После этого выделяется интервал (короткий кадр) композиции, который состоит из дискретных отсчётов (условно будем считать периодическим), к которому будет применяться преобразование Фурье. В результате формируется массив из комплексных чисел, который содержит информацию о фазовом и амплитудном спектрах кадра, которых анализируется [2].

На рис.2 изображена "порезка" звукового сигнала на кадры длины  $N$  с половинным перекрытием. Необходимость в перекрытии вызвана искажением звука в случае, если кадры расположены рядом.

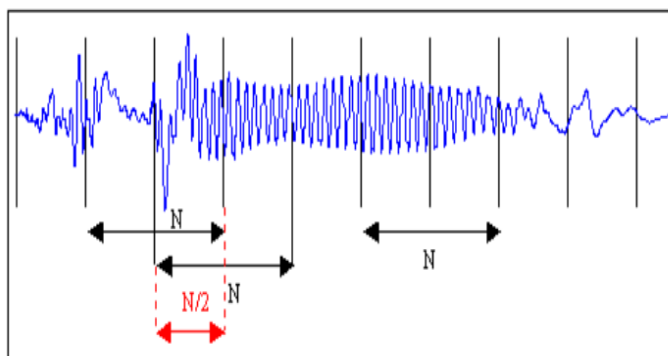


Рис.2. Звуковые кадры

**Целью работы** является применение оконных функций Хемминга, Блэкмана и Кайзера для цифровой обработки голосовых сигналов.

**Методы реализации.** Следует отметить, что некорректное выполнение процедур голосовой идентификации и аутентификации обусловлено такими факторами, как эмоциональное состояние пользователя, сложная акустическая обстановка (шум, помехи, радиоволны), простудные заболевания и естественные изменения голоса. Во всех подобных случаях точный анализ произвольных сигналов (в том числе и нестационарных) преобразованием Фурье практически нереализуем из-за многочисленных недостатков и сложностей [3].

Алгоритмы быстрого преобразования Фурье (БПФ) резко сокращают время спектрального анализа и синтеза, однако не уменьшают погрешности вычислений при заданном числе гармоник.

Для преодоления вышеперечисленных проблем в приложениях голосовой биометрической идентификации наиболее предпочтительно применение оконного или кратковременного преобразования Фурье.

**Оконное преобразование Фурье** - это разновидность преобразования Фурье, определяемая следующим образом:

$$F(t, \omega) = \int_{-\infty}^{\infty} f(\tau)W(t - \tau)e^{-i\omega t} dt, \quad (1)$$

где  $W(t - \tau)$  - некоторая оконная функция. В случае дискретного преобразования используется аналогичная формула:

$$F(m, \omega) = \sum_{n=-\infty}^{\infty} f[n] \omega[n - m] e^{-i\omega t}.$$

Выбор оконной функции - непростая задача. В соответствии с принципом неопределенности Гейзенберга нельзя иметь неограниченно хорошее разрешение и во временной, и в частотной областях [4]. Поэтому если в оконном преобразовании применять короткие окна, то можно достичь хорошего временного разрешения, теряя при этом в частотном разрешении. При длинных окнах ситуация меняется на противоположную. Каждая оконная функция имеет свои особенности для различных приложений.

Окно Хемминга используется для уменьшения уровней боковых лепестков в сигнале, а также обеспечения меньшей ширины главного лепестка и резкой переходной полосы. **Окно Хемминга** имеет лучшую избирательность для большого сигнала и описывается формулой

$$\omega(n) = 0,53836 - 0,46164 \cos\left(\frac{2\pi n}{N-1}\right), \quad (2)$$

где  $N$  - ширина окна;  $\omega(n)$  = оконная функция.

Для описания **окна Блэкмана** используется формула

$$\omega(n) = a_0 - a_1 \cos\left(\frac{2\pi n}{N-1}\right) + a_2 \cos\left(\frac{4\pi n}{N-1}\right), \quad (3)$$

где  $a_0 = \frac{1-\alpha}{2}$ ,  $a_1 = \frac{1}{2}$ ,  $a_2 = \frac{\alpha}{2}$ .

Уровень боковых лепестков при использовании данного окна составляет  $-58 \text{ дБ}$  ( $\alpha=0,16$ ).

**Окно Кайзера** описывается формулой

$$\omega(n) = \frac{\left| I_0\left(\beta \sqrt{1 - \left(\frac{2n-N+1}{N-1}\right)^2}\right) \right|}{|I_0(\beta)|}, \quad (4)$$

где  $I_0$  - модифицированная функция Бесселя первого рода нулевого порядка;  $\beta$  - коэффициент, определяющий долю энергии, сосредоточенной в главном лепестке спектра оконной функции. Чем больше  $\beta$ , тем больше доля энергии, шире главный лепесток и меньше уровень боковых лепестков. На практике используются значения от 4 до 9.

**Результаты исследования.** В разработанной системе используется платформа LabVIEW (Laboratory Virtual Instrument Engineering Workbench) 2020.

Графическая среда программирования LabVIEW и основные функции LabVIEW позволяют создавать приложения для сбора данных и управления приборами [5].

В пространстве LabVIEW 2020 разработана программа, позволяющая записывать звуковой сигнал в реальном времени и производить спектральный анализ этих сигналов.

Изначально в реальном времени записывается голос человека. После записи и сбора голоса используется фильтр нижних частот для удаления тишины и фонового шума, и образец голоса кодируется с использованием техники кадрирования.

На рис. 3 представлена часть блок-схемы системы анализа голоса с использованием оконных функций Хемминга, Блэкмана и Кайзера.

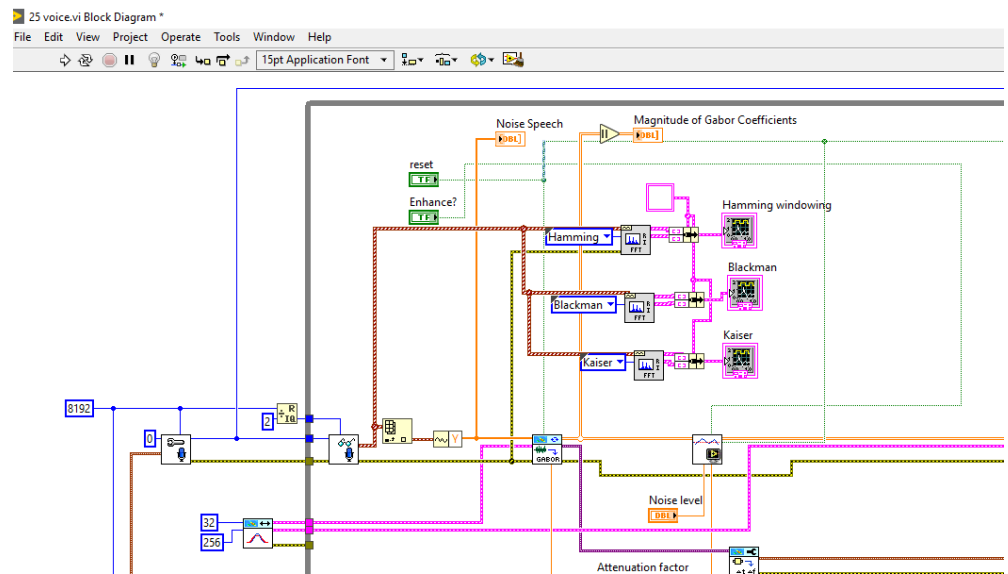


Рис.3. Часть оконных функций блок-диаграммы программы

Каждый сигнал обрабатывается с использованием трех оконных методов.

Ниже - лицевая панель программы, где хорошо виден входной сигнал и анализ этого сигнала с помощью оконных функций (рис.4).

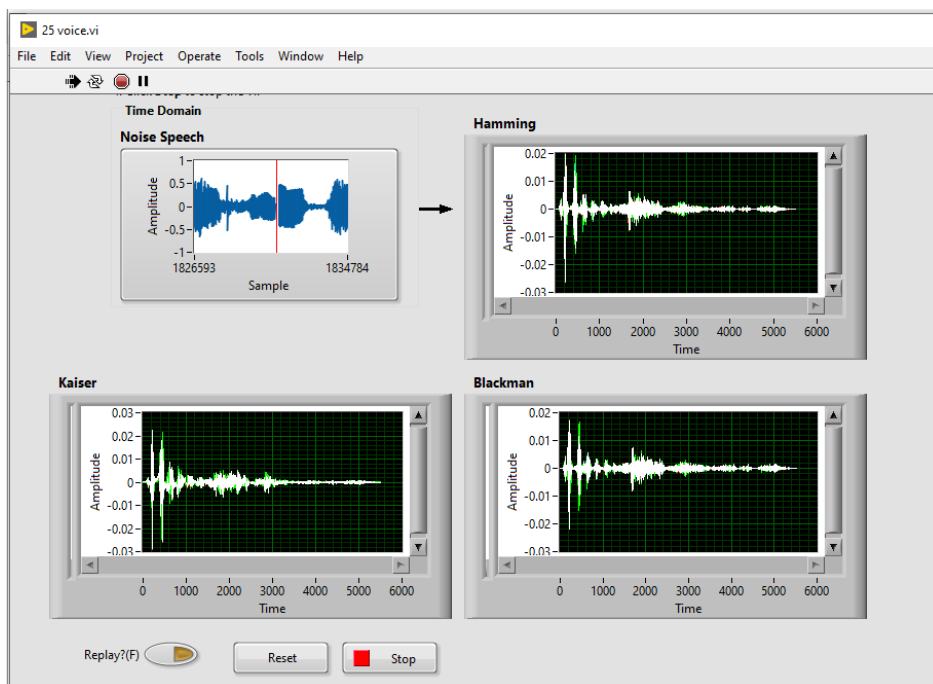


Рис. 4. Результаты применения оконных функций Хемминга, Блэкмана и Кайзера

**Заключение.** На сегодняшний день биометрические технологии успешно конкурируют с традиционными карточными и парольными методами идентификации. Интенсивное развитие и удешевление мультимедийных и цифровых технологий позволяют разрабатывать новые перспективные подходы идентификации личности, среди которых низкой стоимостью реализации и "многообразием" выгодно отличается технология голосовой идентификации.

Для уменьшения вычислительной сложности задачи сравнения уникальных характеристик шаблонов звука предложен метод деления звукового сигнала на кадры определенной длины с последующим умножением на весовые оконные функции.

### Литература

1. **Garcia R.L., Lopez C.A.** Biometric Identification systems // ETSI Telecommunication University of Valladolid.- Valladolid, Spain, 2003.- P. 257-259.
2. **Hartmann, William M.** Signals, Sound, and Sensation.- American Institute of Physics, New York, 1997.

3. **Дьяконов В.П., Абраменкова И.А.** MATLAB. Обработка сигналов и изображений: Специальный справочник.- СПб.: Питер, 2002.- 608 с.
4. **Малла С.** Вэйвлеты в обработке сигналов. - М.: Мир, 2005.- 671 с.
5. **National Instruments.** LabVIEW Fundamentals.- Austin, Texas, 2005. – 165 p.

*Поступила в редакцию 05.03.2021.  
Принята к опубликованию 28.06.2021.*

## **ՖՈՒՐՅԵ ՊԱՏՈՒՀԱՆԱՅԻՆ ՁԵՎԱՓՈԽՈՒԹՅԱՆ ԿԻՐԱՌՈՒՄԸ ՁԱՅՆԱՅԻՆ ԿԵՆՍԱՀԱՓԱԿԱՆ ՆՈՒՅՆԱԿԱՆԱՑՄԱՆ ՀԱՎԵԼՎԱԾՆԵՐՈՒՄ**

**Մ.Ս.Մարգարյան, Բ.Ֆ.Բադալյան, Հ.Ա.Գոմցյան, Դ.Հ. Մոսոյան**

Ներկայումս կենսաչափական նույնականացման համակարգերը մեր կյանքի անբաժանելի մասն են կազմում: Սմարթֆոնների մեջ տեղադրված մատնահետքերի սկաները, ծայնի կամ խոսքի ճանաչման տեխնոլոգիաները և այլ նույնականացման սարքեր աստիճանաբար փոխարինում են նույնականացման ավանդական մեթոդներին և ավելի ու ավելի են թափանցում այնպիսի ոլորտներ, ինչպիսիք են բանկային գործունեությունը, վիզաների և միգրացիոն ծառայությունները, բժշկությունը, անվտանգության համակարգերը և այլն: Կենսաչափական ժամանակակից նույնականացման համակարգերն ունեն մի շարք առավելություններ ավանդական մեթոդների համեմատ, քանի որ դրանք հարմարեցված են անհատի նույնականացման համար: Այնուամենայնիվ, կենսաչափական նույնականացման համակարգերի լայն կիրառությունը պահանջում է նաև տվյալների անվտանգության ապահովման բարձր մակարդակ:

Աշխարհում կենսաչափական նույնականացման տեխնոլոգիաների ինտենսիվ զարգացման հետ մեկտեղ աճում է ծայնային կենսաչափության հիման վրա նույնականացման համակարգերի նորարարական լուծումների նկատմամբ հետաքրքրությունը: Ձայնը յուրաքանչյուր մարդու անբաժանելի մասն է: Հաղորդակցման միջոցների լայն կիրառումը մեծ հնարավորություններ է ստեղծում այս կենսաչափական պարամետրի օգտագործման համար: Բացի այդ, ծայնի ճանաչումը շատ հարմար է օգտվողների համար և նրանցից պահանջում է նվազագույն ջանք: Ձայնի նույնականացման համակարգերը հիմնված են մարդու ծայնի այնպիսի յուրահատուկ հատկությունների վրա, ինչպիսիք են ծայնի բարձրությունը, մոդուլացումը և հաճախականությունը: Այս ցուցանիշները որոշվում են ծայնային տրակտների ֆիզիկական բնութագրերով և բնորոշ են յուրաքանչյուր անձի:

Ներկայացվել է ծայնային ազդանշանի վերլուծության տեխնիկա՝ օգտագործելով Ֆուրյեի կարճաժամկետ ձևափոխության տարբեր պատուհանային ֆունկցիաները: Քննարկվել են Ֆուրյեի ձևափոխության ավանդական մեթոդների հիմնական

սահմանափակումները ձայնի նույնականացման ժամանակ: Ձայնի ճանաչման ծրագիրը մոդելավորվել է LabVIEW 2020 ծրագրային միջավայրում:

**Առանցքային բաներ.** նույնականացման կենսաչափական համակարգեր, Ֆուրյեի արագ ձևափոխություն, Ֆուրյեի պատուհանային ձևափոխություն, գաղտնաբառ, պատուհանային ֆունկցիա, LabVIEW ծրագրային ապահովում:

## USING THE WINDOW FOURIER TRANSFORM IN VOICE BIOMETRIC IDENTIFICATION APPLICATIONS

**M.S. Margaryan, B.F. Badalyan, H.A. Gomtsyan, D.H. Mosoyan**

Currently, biometric identification systems are an integral part of our life. Fingerprint scanners built into smartphones, voice or speech recognition technologies and other tools are gradually replacing the traditional methods of identification and increasingly penetrate into such large areas as banking, visa and migration services, medicine, security systems, etc. Biometric systems have a number of advantages over traditional methods, since they are adapted for personal identification without the possibility of transferring a secret key, and are in many ways more convenient from the user's point of view. However, the more actively the implementation of this type of systems is, the more acute the issue of ensuring information security arises.

With the intensive development of biometric identification and authentication technologies in the world, there is an increased interest in the development of innovative solutions for voice biometrics. The voice is as much an integral part of every person as their face or fingerprints. Widespread communication means open up great opportunities for the use of this identifier. In addition, voice recognition is very convenient for users and requires minimal effort from them. Voice authentication systems rely on voice characteristics unique to each person, such as pitch, modulation and frequency of sound, during sample recording and in the process of subsequent identification. These indices are determined by the physical characteristics of the vocal tract and are unique to each person.

A technique for analyzing a voice signal using a short-term Fourier transform with different window functions is introduced in the present work. The main limitations of the traditional methods of Fourier transform for voice identification problems are considered. A voice recognition module is modeled in the LabVIEW 2020 software space.

**Keywords:** biometric identification systems, fast Fourier transform, windowed Fourier transform, passphrase, window function, LabVIEW.