

## ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ СТИСНЕННЯ МОВНИХ СИГНАЛІВ

© Мельник А.О., Шевчук Р.П., 2004

**Проведено класифікацію методів стиснення мовних сигналів. Виділено переваги та недоліки кожного методу. Визначено основні характеристики алгоритмів стиснення мовних сигналів, за якими їх проаналізовано. Розглянуто принципи побудови алгоритмів стиснення мови. Окреслено сучасний стан галузі стиснення мовних сигналів та виділено основні напрямки її розвитку.**

**In this paper the classification of speech compression algorithms is performed. The advantages and defects of every compression method are selected. Determined basic characteristics of speech compression algorithms, and conducted their analysis. The principles of speech compression algorithms are considered. Shown modern condition industry of speech compression and chosen guidelines of its development.**

**Вступ.** Стрімкий розвиток систем зв'язку міжнародних та національних телекомунікацій характеризується зростанням кількості високошвидкісних цифрових каналів передачі інформації та широкомасштабним підвищенням ефективності роботи пристроїв цифрового стиснення каналів, що зумовлене стрімким зростанням об'єму передач інформації в мережах телекомунікацій. Сьогодні все частіше для підвищення ефективності використання цифрових каналів застосовуються методи стиснення інформації. Необхідно відзначити, що найбільший обсяг інформації, що передається цифровими каналами зв'язку, припадає на мультимедійний трафік [1]. Тому актуальним є стиснення саме мультимедійних потоків даних, зокрема мовних сигналів.

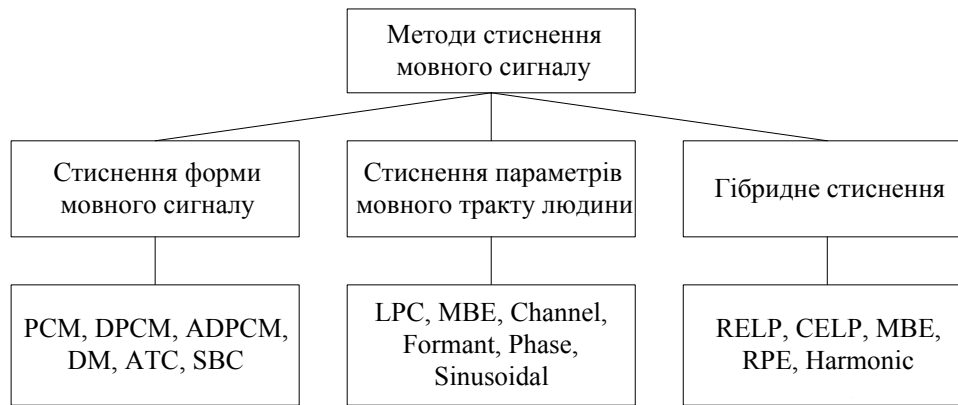
Сьогодні розроблено декілька десятків алгоритмів стиснення мовних сигналів, які структурно та семантично відрізняється один від одного [2]. Сучасні практичні потреби висувають ряд вимог до алгоритмів стиснення мовних сигналів, зокрема, якість мовлення, швидкість передавання сигналів, складність виконання, алгоритмічні затримки, завадостійкість. У зв'язку з цим, останнім часом з'явилося ряд нових алгоритмів стиснення таких мовних сигналів, як CS-ACELP [3], MELP [4], NetCoder [5], що затрудняє остаточний вибір алгоритму для стиснення мовних сигналів.

Тому метою цієї роботи є класифікація та порівняльний аналіз алгоритмів стиснення мовних сигналів за їх основними характеристиками.

**Методи стиснення мовних сигналів.** Задачею будь-якого алгоритму стиснення мовних сигналів є одержання цифрової послідовності, яка потребує мінімальної швидкості передачі і з якої декомпресор зможе відновити вхідний мовний сигнал з мінімальними втратами. Насамперед, слід зазначити, що представлення мови у цифровій формі, яке містить дискретизацію та квантування [3], передбачає втрату деякої частини інформації вхідного мовного сигналу. Відлік сигналу при рівномірному квантуванні представляється 12 бітами, при цьому результуюча швидкість коду буде дорівнювати 96 Кб/с. При нерівномірному квантуванні відлік мовного сигналу представляється 8 бітами, що дає змогу знизити швидкість коду до 64 Кб/с. Частота квантування для більшості алгоритмів стиснення мовних сигналів приймається 8 КГц.

Специфіка методів стиснення мовних сигналів дозволяє поділити їх на три групи (рисунок):

- стиснення форми мовного сигналу;
- стиснення параметрів мовного тракту людини;
- гібридне стиснення.



*Класифікація методів стиснення мовного сигналу та приклади алгоритмів, які їх реалізують*

Перераховані вище методи використовуються залежно від особливостей створення та відтворення мовного сигналу. При створенні мовного сигналу до уваги беруться такі властивості сигналу:

- 1) під час тривалих змін:
  - зміна амплітуди;
  - поділ мови на дзвінки та глухі звуки, а також паузи мовчання;
  - особливості мовної та фонетичної структури.
- 2) під час короточасних змін:
  - короточасна кореляція;
  - особливості структури тону звуку (для дзвінких звуків);
  - особливості структури шуму (для глухих звуків).

При відтворенні мовного сигналу до уваги беруть локальний спектральний динамічний діапазон та слухове маскування.

Якщо не використовується жодна з перелічених властивостей, то до сигналу можна застосувати прості методи стиснення форми сигналу для одержання високої якості мовлення на швидкостях вище 64 Кб/с. Беручи до уваги статичні та динамічні спектральні характеристики мови, можна одержати високу якість мовлення на швидкостях 24–32 Кб/с. Для забезпечення хорошої якості мови на швидкостях нижче 10 Кб/с потрібно враховувати властивості створення мови та шуму. При моделюванні джерела мови, швидкість передачі може бути знижена до 2 Кб/с, однак якість мовлення буде дуже низькою.

**Алгоритми стиснення форми мовного сигналу.** Специфіка функціонування алгоритмів, побудованих згідно з методом стиснення форми мовного сигналу, полягає у відновленні форми вхідного сигналу. В цих алгоритмах використана та обставина, що сусідні значення відліків слабо відрізняються один від одного. Це дає змогу з достатньо високою точністю передбачити значення будь-якого відліку на основі значень декількох відліків, що передували йому.

При побудові алгоритмів згідно з методом кодування форми мовного сигналу названа закономірність використовується двома способами. По-перше, є можливість змінювати параметри квантування залежно від характеристики сигналу. У цьому випадку крок квантування може вимірюватись, що дає змогу звузити динамічний діапазон компресора. Деякі алгоритми передбачають зміну параметрів квантування приблизно у рамках мовних складів, а інші змінюють ці параметри на основі аналізу статичних даних про амплітуду сигналу, одержану за порівняно короткий інтервал часу.

Найпростішим алгоритмом, який належить до цього методу, є імпульсно-кодова модуляція (ІКМ) [8] до переваг якої можна віднести [4]:

- ІКМ – найзагальніший алгоритм стиснення мовних сигналів;
- при використанні ІКМ не важливий степінь кореляції між сусідніми відліками;

- за допомогою ІКМ можна стискати будь-який аналоговий сигнал з обмеженою шириною спектра;
- ІКМ характеризується малою затримкою, яка не перевищує один період вибірки;
- ІКМ компресори та декомпресори найпростіші в реалізації.

Основним недоліком ІКМ є низька ефективність стиснення мовного сигналу.

Сьогодні використовують два різновиди ІКМ: стиснення за  $m$ -законом та стиснення за  $A$ -законом, які відрізняються між собою деталями стиснення.

Складнішим алгоритмом є диференціальна ІКМ (ДІКМ) [7]. До її класифікаційних ознак можна віднести наявність блока лінійного передбачення авторегресійних послідовностей (передбачувача) та використання багаторівневого (більше двох рівнів) квантувальника. ДІКМ при стисненні мовного сигналу враховує кореляцію між відліками. При ДІКМ стисненні по каналу зв'язку передається не кодований відлік, а різниця між поточним відліком і оцінкою даного відліку, одержаною з аналізу попередніх відліків. Системи з ДІКМ забезпечують таку саму якість відновленого сигналу, як і ІКМ та на порядок більшу завадостійкість.

Дельта-модуляція (ДМ) являє собою частковий випадок ДІКМ з використанням однорозрядного квантувальника [9]. При ДМ по каналу зв'язку передають квантовані значення різниці між поточним відліком та його передбаченням, причому квантування проводять лише по двох рівнях. Тому по каналу зв'язку передається тільки один знак рівності. Необхідно відмітити, що ДМ характеризується найкращими параметрами серед всіх методів кодування. Системи, що стискають мовні сигнали за допомогою ДМ, не втрачають працездатності при виникненні поодиноких помилок та характеризуються простотою побудови компресора та декомпресора.

Одним з найпоширеніших алгоритмів, побудованих на принципах стиснення форми мовного сигналу, є адаптивно-диференціальна імпульсно-кодова модуляція (АДІКМ) та її різновиди [10–12]. Метод базується на тому, що в аналоговому мовному сигналі неможливі різкі стрибки інтенсивності. Тому, якщо стискати не саму амплітуду мовного сигналу, а її зміну порівняно з попереднім значенням, то можна використати меншу кількість розрядів. В АДІКМ зміна рівня сигналу стискається до чотиризначного числа, при цьому частота вимірювання амплітуди сигналу зберігається незмінною. Таким шляхом можна одержати прийнятну якість мовлення на швидкостях 16–32 Кб/с.

При використанні алгоритму підсмугового стиснення (SBC) мовний сигнал фільтрують на декілька підсмуг, і кожен підсмуговий сигнал адаптивно стискається [13]. Кількість біт, необхідних для стиснення, визначається відповідно до кількості біт квантування призначених критерієм сприйняття. При стисненні кожного підсмугового сигналу шуми квантування обмежуються своєю підсмугою. Найкращі характеристики спостерігаються в компресорі SBC у разі збільшення кількості частотних піддіапазонів, а також при динамічній зміні кількості біт на вибірку від одного піддіапазону до іншого.

Аналіз принципів побудови та функціонування алгоритмів згідно з методом стиснення форми мовного сигналу дозволяє зробити висновок, що ці алгоритми не залежать від виду сигналів, та можуть ефективно стискати будь-які сигнали. Вони характеризуються стійкістю до широкого діапазону характеристик джерела сигналу, простотою реалізації та порівняно невисокими швидкостями. Недоліком алгоритмів побудованих згідно з цим методом є низька ефективність при роботі з сигналами, в яких спостерігаються різкі стрибки амплітуди.

Для покращання ефективності роботи компресорів форми мовних сигналів пропонується оптимізувати їх для визначених видів сигналів.

**Алгоритм стиснення параметрів мовного тракту людини.** В основі принципів побудови алгоритмів згідно з методом стиснення параметрів мовного тракту людини лежить класична модель створення мови, яка передбачає, що всі голосові зв'язки є джерелом збудження, а голосовий тракт – нелінійним фільтром, який формує спектр мовного сигналу. При цьому параметри джерела та фільтра нестационарні. Компресори, побудовані згідно з методом стиснення параметрів мовного тракту людини, називають вокодерами [14].

На відміну від алгоритмів стиснення, побудованих згідно з методом стиснення форми мовного сигналу, вокодери не використовують інформацію про формування мовного сигналу. Робота вокодерів базується на моделі джерела, з якого одержується інформація про параметри мовного сигналу. Відповідно, результатом стиснення будуть не коди сигналів, а коди параметрів джерела цих сигналів. Тип вокодера залежить від способів знаходження моделі системи та її параметрів.

До інформації, яку одержує компресор в результаті аналізу мовного сигналу та передає декомпресору, відносять: параметри фільтра створення мови, вказівник на голосний/приголосний звук, потужність сигналу збудження та період основного тону для голосних звуків.

За принципами аналізу та синтезу мовних сигналів вокодери поділяють на смугові (канальні), формантні, фонемні, ортогональні, гомоморфні.

Одними з перших алгоритмів стиснення параметрів мовного тракту людини були канальні вокодери [15]. Ці вокодери використовують слабку чутливість слуху людини до незначних фазових зсувів сигналу. Для сегментів мови за допомогою вузькосмугових фільтрів визначається амплітудний спектр. Сигнали з виходів фільтрів детектуються, пропускаються через фільтри низької частоти, дискретизуються та піддаються двійковому стисненню. Таким чином визначаються параметри мовного тракту людини, а також період основного тону збудження та вказівник на голосний/приголосний звук.

Канальні вокодери можуть бути реалізовані в цифровій та аналоговій формі та забезпечують задовільну якість мовлення на швидкостях близько 2,4 Кб/с.

У формантних вокодерах огинаюча спектра мови описується комбінацією формант (резонансних частот голосового тракту) [15]. Основні параметри формант – центральна частота, амплітуда та ширина.

В ортогональних вокодерах огинаюча спектра розкладається в ряд по вибраній системі ортогональних базисних функцій [14]. Обчислені коефіцієнти цього розкладання передаються декомпресорові.

Гомоморфна обробка сигналів є одним з нелінійних методів обробки, який ефективно застосовується до мовних сигналів. З врахуванням моделі системи, що використовується у вокодерах, мовний сигнал можна представити як тимчасову згортку імпульсної перехідної характеристики мовного тракту з сигналом збудження [15]. У частотній області це відповідає добутку частотної характеристики голосового тракту та спектра сигналу збудження.

Вокодери, побудовані згідно з алгоритмом лінійного передбачення, представляють мовний тракт людини за допомогою лінійного фільтра з безперервною імпульсною перехідною характеристикою, в якому кожне чергове значення сигналу може бути отримано як лінійна комбінація деякого числа його попередніх значень [15]. В такому вокодері мовний сигнал ділиться на блоки, для кожного з яких визначаються коефіцієнти фільтра передбачення. Ці коефіцієнти квантуються та передаються декомпресору. Потім мовний сигнал пропускається через фільтр, частотна характеристика якого зворотна частотній характеристиці мовного тракту. На виході фільтра одержуємо помилку передбачення. У результаті таких перетворень набагато виразніше виявляється тривала кореляція в сигналі, що дає змогу точніше визначити частоту основного тону та виділити вказівник на голосний/приголосний звук. Швидкості кодів для вокодерів, побудованих згідно з алгоритмом лінійного передбачення при задовільній якості мовлення, лежать в межах від 1 до 2,4 Кб/с.

Недоліком всіх алгоритмів стиснення мови, побудованих згідно з методом стиснення параметрів мовного тракту людини, є їх синтетична якість мовлення. При цьому навіть суттєве збільшення швидкості коду практично не покращує якості мови. В основному вокодери використовують у спеціалізованих системах зв'язку, де головне не натуральність мовлення, а великий коефіцієнт її стиснення.

**Алгоритми гібридного стиснення.** Алгоритми гібридного стиснення мовних сигналів будуються на основі поєднання властивостей методів стиснення форми мовного сигналу та стиснення параметрів мовного тракту людини [14].

Алгоритми стиснення, побудовані згідно з методом гібридного стиснення, є найновішими та найефективнішими серед сучасних алгоритмів стиснення. Вони використовують найдоступнішу інформацію про мовний сигнал для покращання якості мовлення та зниження швидкості результуючого коду. Зокрема вони використовують замаскований слуховий шум, слухову частоту розширення, слухову фазу нечутливості, складові зміни енергії, довгострокові звукові властивості тракту та інформацію про крок під час квантування [16].

При гібридному методі стиснення, мовний сигнал дискретизується, з метою одержання необхідних параметрів мови. Однак замість безпосереднього стиснення висоти, модуляції та інших характеристик мовного сигналу, ці параметри використовуються для синтезу фрагменту мови, з якого вони були одержані. Синтезований фрагмент мови, звичайно тривалістю 20 мс, порівнюється з початковим зразком. Якщо вони збігаються з прийнятним допуском, то параметри мови залишаються без змін. Якщо ж реальний та синтезований фрагменти відрізняються більш ніж на задану величину, то параметри мови коректуються для досягнення необхідного збігу. Кінцевим етапом цього механізму зворотного зв'язку є виконання процедури аналізу шляхом синтезу [17]: спроба відрегулювати параметри мови для забезпечення синтезу форми сигналу, якомога більш наближеної до початкової форми. Після остаточного визначення значень параметрів мови, компресор зіставляє одержані характеристики з характеристиками кодової книги. При виявленні збігу замість значення параметра використовується його положення в кодовій книзі, що суттєво зменшує обсяг інформації, яку потрібно передати [18]. Описаний вище алгоритм лежить в основі побудови та функціонування сімейства алгоритмів CELP (Code Excited Linear Predictive Coder) [19, 20]. Компресори CELP працюють на швидкостях нижче 4800 Кб/с та забезпечують задовільну якість мовлення. Головним недоліком компресорів типу CELP є їх висока складність та значні обчислювальні потреби.

Алгоритм стиснення мови з регулярним імпульсним збудженням (RELPE) [21] пропускає мовний сигнал через лінійний передбачувач, що дає змогу зменшити кореляцію між відліками вхідного сигналу. Якщо передбачення виконується достатньо добре, то на виході передбачувача буде практично білий шум з рівномірним спектром. Цей білий шум (помилка передбачення) несе всю інформацію про стиснений мовний сигнал, і якщо його пропустити знову через LPC-фільтр [22], то абсолютно точно відновиться початковий мовний сигнал. Проте через наявність у мовному сигналі квазіперіодичних формантних складових, лінійний передбачувач не може усунути довготривалої кореляції з періодом основного тону формант, і вони будуть явно відображатись в спектрі помилки передбачення.

В багатоімпульсних компресорах (MPE) [23] як сигнал збудження беруть послідовність з чотирьох – шести коротких імпульсів. Тимчасове положення кожного з цих імпульсів та їх амплітуди визначаються в процесі процедури аналізу шляхом синтезу до досягнення мінімальних відмінностей між початковим та синтезованим мовними сигналами. Параметри імпульсів збудження, що мінімізують помилку, підбирають послідовно. На практиці достатньо задавати положення імпульсів з кроком близько 1 мс, і точністю амплітуд до 5 %, що забезпечує прийнятну якість мовлення, при швидкості коду близько 10 Кб/с [23].

Аналогічно MPE-компресору, RPE-компресор, використовує як сигнал збудження фіксований набір коротких імпульсів. Однак в цьому компресорі імпульси розташовані регулярно, на однакових відстанях один від одного, тому компресору необхідно визначити лише положення першого імпульсу та амплітуди всіх імпульсів. Отже, декомпресору потрібно передавати менше інформації про положення імпульсів, що дозволяє покращити якість мови синтезованого сигналу [21].

Алгоритми, побудовані згідно з методом гібридного стиснення, характеризуються найбільшою обчислювальною стійкістю серед всіх розглянутих алгоритмів. Основним недоліком цих алгоритмів є їх висока складність виконання.

**Порівняння алгоритмів стиснення мови.** Проведений аналіз основних алгоритмів стиснення мови дозволяє їх порівняти за основними характеристиками.

До переліку основних характеристик алгоритмів стиснення мови включено:

- швидкість передачі мовних сигналів;

- довжина кадру – міра кількості часу, що визначає елементарний відрізок мовного сигналу, який обробляється алгоритмом стиснення;
- затримка – час, необхідний для стиснення мовного сигналу;
- MOS (mean opinion score) – суб’єктивна оцінка якості мовлення, яка одержується опрацюванням оцінок, що даються групами слухачів.

Оцінки інтерпретуються так:

- 4–5 – висока якість мовлення,
- 3.5–4 – прийнятна якість мовлення,
- 3–3.5 – задовільна якість мовлення,
- 2.5–3 – незадовільна якість мовлення, потребує концентрації уваги для розуміння.
- завадостійкість – здатність алгоритму правильно функціонувати при наявності завад;
- можливість передачі одночастотного сигналу;
- метод стиснення мовного сигналу: 1 – стиснення форми мовного сигналу, 2 – стиснення параметрів мовного тракту людини, 3 – гібридне стиснення.

Порівняння алгоритмів стиснення мовного сигналу за їх основними характеристиками наведено в таблиці

### Порівняння характеристик алгоритмів стиснення мовних сигналів

Стандарт	Алгоритм стиснення	Швидкість, Кб/с	Кадр, мс	Затримка, мс	MOS	Завадостійкість	Передача одночастотного сигналу	Метод
1	2	3	4	5	6	7	8	9
ITU G.711	PCM A-law / PCM u-law	64	0,125	0,125/0,75/5	4,15	10	так	1
ITU G.722	SB-ADPCM	64	40	5	4,1	9	так	1
		56	35			9		
		48	30			8		
ITU G.721	ADPCM	32		5	4,1	9	так	1
ITU G.726	ADPCM	40	25	5	3,91	9	так	1
		32	0,125/20	1/5		9		
		24	15	5		8		
		16	10	5		8		
ITU G.728	LD-CELP	16	0,625/10	2,5/3 ... 5	3,69	4	так	3
ITU G.729	CS-ACELP	8	10	10	3,96		так	3
ITU G.729a	CS-ACELP	8	10	10	3,71		так	3
ITU G.723.1	MP-MLQ	6,3	30/24	30/37,5	3,93		так	3
ITU G.723	ACELP	5,3	30/20	30/37,5	3,66			3
INMAR-SAT-M	IMBE	6,4		80	3,1			
	IMBE	3,6						
ETSI GSM	RPE-LTP	13	20		3,3			2
ETSI TETRA	ACELP	4,8			3,4			3
США	MELP	2,4		45	3,5			3
USFS 1015	LPC10e	2,4			2			2
TIA IS-54	VSELP	5,6	20				ні	3
D-AMPS	VSELP	7,95	20		3,3			3
TETRA	ACELP	4,57	30					3
eXpressDSP	RCELP	3,6		30	3,5			3

1	2	3	4	5	6	7	8	9
eXpressDSP	MMBE	2,4	30	45	3,5			3
eXpressDSP	ICELP	4,8	30	60	3,7			3
AudioCodes	NetCoder	6,4	20		3,85			
AudioCodes	NetCoder	7,2	20		3,91			
AudioCodes	NetCoder	8	20		4,1			
DSPSE	Voice-Wave	4,8-12,8					так-ні	
USFS 1016	CELP	4,8					ні	3
USFS 1015	LPC10e	2,4					ні	2

*Примітка.* Пусті клітинки – відсутня у авторів інформація.

До характеристик алгоритмів стиснення мови, які не включені в таблицю, відносять також: складність виконання алгоритму, завантаження ROM та RAM, коефіцієнт стиснення, сумарну затримку, частоту квантування, мовний спектр, можливість та швидкість передачі сигналу модему. Однак в більшості випадків перелічені характеристики залежать від способу виконання алгоритму стиснення мови та суттєво не впливають на вибір алгоритму стиснення.

Проведений аналіз алгоритмів стиснення мовних сигналів показує, що вибір алгоритму для стиснення мови є достатньо складним завданням, вирішення якого залежить від ряду характеристик алгоритму стиснення та вимог системи, в якій алгоритм буде функціонувати. Основною та найбільш істотною характеристикою є якість мовлення. Крім того, до важливих характеристик, на які потрібно зважати при виборі алгоритму, належать швидкість передачі, алгоритмічні затримки, що вносяться алгоритмом при стисненні сигналу, та складність виконання алгоритму.

Слід відзначити, що сучасні засоби стиснення мовних сигналів реалізують в основному два алгоритми, побудовані згідно із стандартами G.723.1 [24] і G.729A [25]. В той самий час об'єктивно кращим в останні декілька років є алгоритм Net-Coder фірми AudioCodes. Його широке розповсюдження стримується тим, що виробники обладнання не хочуть змінювати широковпроваджений стандарт G.729A.

Сьогодні найперспективнішим методом є гібридне стиснення мовних сигналів. Тому актуальним є створення нових алгоритмів низькошвидкісного стиснення мовних сигналів, які дозволять забезпечити прийнятну та високу якість мовлення при низьких швидкостях. Перспективним також є реалізація алгоритмів стиснення мовних сигналів на спеціалізованих процесорах. Особливу увагу необхідно приділити дослідженню та розробці універсальних алгоритмів стиснення мовних сигналів для кожного методу стиснення.

Одержані результати також показують, що для побудови спеціалізованих систем (військові, космічні системи) найкраще підходять алгоритми стиснення, побудовані згідно з методом гібридного стиснення. Для систем загального користування (Інтернет-телефонія, телефонні системи загального призначення, стільниковий зв'язок) використовуються алгоритми, побудовані згідно з методом стиснення форми мовного сигналу та стиснення параметрів мовного тракту людини.

**Висновок.** У роботі проведено класифікацію та аналіз методів стиснення мовних сигналів. Виділено переваги та недоліки кожного методу. Розглянуто принципи побудови основних алгоритмів стиснення мовних сигналів.

Результати аналізу показали, що вибір алгоритму для стиснення мови є достатньо складним завданням, вирішення якого залежить від ряду характеристик алгоритму стиснення та вимог системи в якій алгоритм буде функціонувати. До основних характеристик алгоритмів стиснення, на які потрібно зважати при виборі алгоритму, віднесено: якість мови, швидкість передачі, алгоритмічні затримки, що вносяться алгоритмом при стисненні сигналу, та складність виконання алгоритму.

У роботі також окреслено сучасний стан галузі стиснення мовних сигналів та основні напрямки її розвитку.

1. Гольдштейн В.С., Пинчук А.В., Суховицкий А.Л. IP-Телефония. – М.: Радио и Связь, 2001. – 336 с. 2. Marwan Jabri. Design hits media standards wall // *Electronic Engineering Times*. – 2002. 3. ITU-T Recommendation G.729, Coding of speech at 8 kbit/s using conjugate-structure algebraic code-excited linear prediction (CS-ACELP), March 1996. 4. Lynn M. Supplee Alan. V. McCree. MELP: the new federal standard at 2400 bps. ICASSP 97. 5. [www.audiocodes.com](http://www.audiocodes.com). 6. <http://www.faqs.org/docs/sp>. 7. [http://www.telcenter.ru/russian\\_win/signalling/course\\_2/intro.htm](http://www.telcenter.ru/russian_win/signalling/course_2/intro.htm). 8. Recommendation G.711, “Pulse code modulation (PCM) of voice frequencies”, ITU-T G Series Recommendations, 1988. 9. Jayant N.S. Digital Coding of Speech Waveforms: PCM, DPCM, and DM Quantizers // *Proc. IEEE*. – May 1974. – Vol. 62. – P. 611–632, 10. Recommendation G.726, “40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)”, ITU-T G Series Recommendations, 1990. 11. CCITT Recommendation G.723. Extensions of Recommendation G.721 ADPCM to 24 and 40 kbits/s for DCME Application // *Blue Book*. – Oct. 1988. – Vol. III, Fascicle III. 12. CCITT Recommendation G.721. 32kb/s Adaptive Differential Pulse Code Modulation (ADPCM) // *Blue Book*. – Oct. 1988. – Vol. III, Fascicle III.3. 13. Cox R.V. et al. New Directions in Sub-band Coding // *IEEE Trans. On Selected Areas in Communications, Special Issue on Voice Coding for Communications*. – Feb. 1988. – Vol. 6, No. 2. – P. 391–409. 14. Сапожков М.А., Михайлов В.Г. Вокодерная Связь. – М.: Радио и связь, 1983. 15. Spanias A.S. Speech Coding: a tutorial review // *Proceeding of the IEEE*. – October, 1994. – Vol. 82, No 10. – P. 3071–3075. 16. Li W., Sridhar A., Teng T. Comparison of Speech Coding Algorithms: ADPCM, CELP and VSELP. – Fall, 1999. 17. Gersho A. and Wang S. Recent Trends and Techniques in Speech Coding // *Proc. 24th. Asilomar Conf., Pacific Grove, Nov. 1990*. 18. Хелд Г., Сокращение задержки голоса по IP. – LAN, 2000. – № 07. 19. Schroeder M.R. and Atal B. Code-Excited Linear Prediction (CELP): High Quality Speech at Very Low Bit Rates // *Proc. ICASSP-85. – Tampa, Apr. 1985*. – P. 937. 20. Un C.K. and Magill D.T. The Residual-Excited Linear Prediction Vocoder with Transmission Rate below 9.6 kbits/s // *IEEE Trans. COM-23(12)*. – Dec. 1975. – P. 1466. 21. Kroon P., Deprettere E. and Shuyeter R.J. Regular-Pulse Excitation-A Novel Approach to Effective and Efficient Multi-pulse Coding of Speech // *IEEE Trans. ASSP-34(5)*. – Oct. 1986. 22. <http://www.data-compression.com/speech.html>. 23. Griffin D. and Lim J. Multiband Excitation Vocoder // *IEEE Trans. ASSP-36*. – Aug. 1988. – No. 8. – P. 1223. 24. ITU-T Recommendation G.723.1 Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s. 25. Recommendation G.729 Annex A, “C source code and test vectors for implementation verification of the G.729 reduced complexity 8kbit/s CS-ACELP speech coder”, ITU-T G Series Recommendations, 1996.