

Список літератури

1. *Contract PCI board breaks bandwidth barriers.* / *Electronic products.* – 1992 – Febr. 1999. – P.86.
2. Ковальчук С.П. Підходи до формування панорамних зображень за допомогою кількох телевізійних камер на основі матриць ПЗЗ: Зб. наук. праць (радіоелектронне обладнання). Вип. 2. –К: КІ ВПС, – 1998. – С.37–42.
3. Панорамна телевізійна система високого розрізнення / С.І. Мірошниченко, Е.О. Жилко, В.В. Кулаков, С.П. Ковальчук, А.О. Невгасимий. – № 98031578; Заяв. 16.04.98.
4. *Цифровая обработка сигналов* : Справ. / Под ред. Л.М. Гольденберг, Б.Д. Матюшкин, М.Н. Поляк. – М.: Радио и связь, 1985. – С.312.
5. *Гудмен Дж.* Статистическая оптика. – М.: Мир, 1988. – С. 343-345.

Стаття надійшла до редакції 06.07.01.

УДК 534.782.001:621.39

ББК 2 8 74-013.2

О.І. Давлет'янц, О.В. Коломієць

СТИСНЕННЯ МОВНИХ СИГНАЛІВ ПІСЛЯ ДЕЛЬТА- ТА ІМПУЛЬСНО-КОВОДОЇ МОДУЛЯЦІЙ

Проведено класифікацію методів цифрового представлення мовних сигналів і дано їх стислі характеристики. Запропоновано алгоритми стиснення об'єму цифрових даних після дельта- та імпульсно-кодової модуляції.

По багатьом каналах зв'язку інформація передається в мовній формі. Широке впровадження цифрових каналів зв'язку і систем реєстрації мовних сигналів привело до необхідності стиснення цифрового представлення цих сигналів. Зменшення об'єму цифрових даних, що описують мовний сигнал, еквівалентно збільшенню пропускної здатності цифрового каналу або зниженню необхідної ємності запам'ятовуючого пристрою для реєстрації мовного сигналу заданої тривалості.

Сучасний рівень розвитку техніки дозволяє формувати (синтезувати) мовні повідомлення технічними засобами. Оператору при цьому приділяється лише функція контролю.

Розглянемо основні методи представлення мовних повідомлень у цифровій формі [1;2]. Умовно їх можна поділити на три класи.

До *першого класу* відносять методи, що використовують принцип безпосереднього кодування аналогових сигналів. Ці методи застосовують природні мовні сигнали з виходу мікрофона.

Другий клас охоплює методи фонетичного синтезу мови за допомогою формантних синтезаторів.

До *третього класу* належать методи, засновані на використанні параметрів сигналу в частотній області.

Алгоритми, що реалізують різні методи представлення мовних сигналів, зводять до різних питомих (за одиницю часу) об'ємів цифрових даних.

Проведемо короткий аналіз методів і алгоритмів представлення мовних сигналів.

Методи безпосереднього кодування забезпечують збереження форми вихідного сигналу при його відновленні. По суті ці методи використовують принцип цифрового запису аналогового сигналу. При заданій верхній частоті F_n відновлюваного сигналу однозначно визначається і тактова частота дискретизації. Так, при імпульсно-кодовій модуляції частота дискретизації $F_d = (2F_n)$. Стандартна швидкість передачі в типових каналах зв'язку при імпульсно-кодовій модуляції дорівнює 64 Кбіт/с. При реєстрації мовних сигналів потрібно 64 Кбіт цифрової пам'яті на кожен секунду запису сигналу.

Більш заощадливим є перетворення мовних сигналів за допомогою дельта-модуляції (ДМ). Аналоговий сигнал при ДМ перетворюється в однорозрядний цифровий код. Необхідна швидкість передачі або питомі витрати запам'ятовуючого пристрою для ДМ у порівнянні з імпульсно-кодовою модуляцією при однаковій розбірливості можуть бути знижені в два-три рази.

Подальше стиснення ДМ сигналу можна здійснювати введенням елементів адаптації в модулятор. Керованими в адаптивному модуляторі можуть бути як крок квантування, так і інтервал дискретизації.

При запису ДМ сигналу в цифрову пам'ять мінімального об'єму, що забезпечує його максимальне стиснення, потрібне його однозначне зворотне перетворення. Стиснення сигналу корисне і при його трансляції по каналу зв'язку, оскільки при цьому можна використовувати добре розроблені методи перешкодостійкого кодування, що забезпечує не тільки виявлення, але і виправлення помилок. Введена надмірність перешкодостійкого кодування може бути навіть меншою ніж виключеної природної надмірності ДМ сигналу.

Методи фонетичного синтезу за допомогою аналогових формантних синтезаторів - найекономніші за витратами пам'яті на одиницю часу відтвореного сигналу (100-600 біт/с) [2]. Мовний сигнал формується з окремо генерованих фонем. У пам'яті ЕОМ повинні зберігатися не тільки основні фонемі, але і численні їх варіації – алофони.

Основним недоліком методу є неприродність звучання синтезованої мови. Цей недолік не є визначальним при діалозі з ЕОМ або роботом. Однак, наприклад, у системі керування повітряним рухом, особливо при використанні каналів зв'язку, схильних до впливу перешкод, використання розглядуваного методу синтезу мовних сигналів небажано. Сприйняття «роботизованої» мови створює додаткову напруженість оператора, що не сприяє підвищенню безпеки польотів.

Методи відновлення звуків мови в частотній області займають за об'ємом даних (1200- 5000 біт/с) середнє значення між першими двома методами.

Методом цього класу, що найбільш активно розвивається, є так званий метод лінійного предиктивного кодування (ЛПК), за яким вибірки мовного сигналу розглядаються як лінійні комбінації минулих вибірок. Збережені в пам'яті параметри сигналу є результатом аналізу спектрального складу мови. Предиктивні параметри (вихідні дані ЛПК) використовуються в цифрових фільтрах при відновленні мови. Звичайна кількість предиктивних коефіцієнтів - понад 10. Якість відновленого сигналу при ЛПК наближається до природного.

Алгоритми ЛПК вимагають величезного об'єму обчислень. У реальному часі реалізація алгоритму ЛПК можлива при використанні обчислювальних засобів великої продуктивності. Дана обставина обмежує область застосування алгоритмів ЛПК.

Порівняно просто реалізований алгоритм стиснення цифрових даних може бути отриманий послідовним виконанням процедур ДМ і кодування серій елементів ДМ сигналу.

Сигнал після ДМ являє собою випадкову послідовність бінарно-квантованих відліків. При належно обраних параметрах модулятора максимальна довжина серії відліків одного рівня обмежена [3]. При цьому істотно розрізняються також частоти появи серій з різним розташуванням елементів. У табл.1 наведені найчастіші комбінації восьмибітних відрізків мовного сигналу після дельта-модуляції. Коди комбінацій подані в послідовності, обумовленій їхніми рангами.

Визначивши частішим послідовностям найбільш короткі коди, можна одержати зниження об'єму цифрових даних.

Реалізація ДМ сигналу при експериментальному дослідженні запропонованого варіанта стиснення цифрових даних розбита на послідовність восьмибітових блоків. Аналіз показав, що більше половини блоків відносяться до десяти типів. Для скорочення надмірності кожний восьмибітовий блок кодувався словом, що складається з двох частин – префікса і суфікса. Префікс визначав номер ряду, до якого відноситься розглянутий блок, а суфікс номер комбінації в даному ряду.

Таблиця 1

Ранг коду	Код восьмибітового слова	Оцінка ймовірності коду	Оцінка функції розподілу
1	10101011	0.1929	0.1929
2	11001101	0.0775	0.2075
3	00110100	0.0387	0.3092
4	11001011	0.0361	0.3453
5	01010110	0.0348	0.3802
6	10101101	0.0331	0.4133
7	00101011	0.0298	0.4432
8	10101100	0.0271	0.4703
9	10110011	0.0231	0.4935
10	10110100	0.0197	0.5132
11	00110011	0.0181	0.5314
12	00101100	0.0179	0.5493

У результаті кодування ДМ сигналу об'єм цифрових даних зменшився на 25 %. Такий вигравш можна вважати істотним, якщо врахувати ту обставину, що сама процедура дельта-модуляції значною мірою знижує об'єм даних (порівняно з імпульсно-ковою модуляцією). Подальшого зниження об'єму даних можна досягти використанням добре відомих алгоритмів стиснення інформації. Зокрема, для зазначених цілей може бути використаний, наприклад, алгоритм кодування Хаффмена [4].

Розглянемо алгоритм стиснення цифрових даних після імпульсно-кової модуляції. Стисненню піддаються блоки даних, що представляють послідовність N K -розрядних відліків (табл. 2).

Для зручності запису аналітичних співвідношень приймемо $N = 2^l$, де l - ціле число.

Скорочення надмірності в розглянутому випадку здійснюється поблочним визначенням мінімально необхідної розрядності відліків. Необхідна розрядність при цьому визначається значенням максимального відліку вибірки.

Таблиця 2

Розряди	8	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	0	0	0	0	0	0	0
	6	1	0	0	0	1	1	0	0	0	0	1	0	0
	5	0	1	0	1	0	1	1	0	1	0	1	0	1
	4	0	1	1	0	0	0	1	1	0	1	0	0	1
	3	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	1	0	1	1	1	1	0	1	0	1	1	0	1
	1	0	0	1	1	0	1	0	1	1	1	0	1	0
Номер відліку	1	2	3	4	5	6	7	8	9	10	11	12	.	N

До проблеми стиснення можна підійти, розглядаючи не стовпці (відліки в табл.2), а розрядні рядки. При цьому замість N K -розрядних будемо мати K N -розрядних слів. Ідея стиснення полягає в тому, що рядки, які містять лише нулі або лише одиниці, можуть бути описані в префіксі і надалі не фіксуватися.

Для опису даних (табл.2) досить використовувати розрядність $k_1=5$ замість попереднього $K=8$. Цей факт може бути відображений у службовому слові, що передує запису (передачі по каналу зв'язку) даного блока. Так при розрядності службового слова $l=3$, а $N=1024$ і $K=8$ економія пам'яті при запису блока складе $C=(K-k_1)N - l = 3069$ біт.

За оцінку ефективності алгоритму використаємо коефіцієнт стиснення цифрових даних

$$\gamma = K 2^l \left(1 + \sum_{k=1}^K m_1 \{S_k\} \right)^{-1}, \quad (1)$$

де $m_1 \{S_k\}$ - математичне сподівання числа біт при записі k -го розрядного рядка; l у знаменнику враховує витрати на службове слово, що відноситься до всього блока даних.

Значення $m_1 \{S_k\}$ при запису k -го розрядного рядка визначиться співвідношенням

$$m_1 \{S_k\} = \sum_{i=0}^{n_1} P_{ik} (l + il) + \left(1 - \sum_{i=0}^{n_1} P_{ik} \right) 2^l. \quad (2)$$

Перший член у виразі (2) характеризує розрядні рядки, піддані перетворенню, другий член відноситься до частини цифрових даних, збережених у незмінному вигляді; через P_{ik} позначена ймовірність появи одиниць у k -му розрядному рядку рівно i раз; l - розрядність службового слова в оброблюваному рядку; n_1 - максимально припустиме значення числа одиниць (n), при якому доцільна обробка (при $n > n_1$ запис розрядного рядка здійснюється у незмінній формі).

Використовуючи біноміальний закон, співвідношення (1) і (2) і з огляду на витрати в l біт на додаткове службове слово, що відноситься до всього блока даних, можна одержати вираз для оцінки коефіцієнта стиснення:

$$\gamma = K 2^l \left\{ 1 + \sum_{k=1}^K \left[\sum_{i=0}^{n_1} C_N^i P^i(k) (1 - P(k))^{N-i} (l + il) + \left(1 - \sum_{i=0}^{n_1} C_N^i P^i(k) (1 - P(k))^{N-i} \right) 2^l \right] \right\}^{-1}.$$

Як показали експериментальні дослідження, ефективність стиснення залежить від об'єму вибірки N і виду щільності розподілу вихідного сигналу. Виграш тим більший, чим більший коефіцієнт ексцесу, оскільки при більш гострій вершині кривої розподілу старші розрядні рядки в блоці кінцевого розміру з великою ймовірністю містять число одиниць $n < n_1$, що дозволяє згортати ці розрядні рядки в слова меншої розмірності.

Розглянуті алгоритми стиснення не вносять змін у вихідний цифровий сигнал, що дозволяє використовувати їх у системах керування повітряним рухом.

Список літератури

1. *Потапова Р.К.* Речевое управление роботом. – М.: Радио и связь, 1989. – 248 с.
2. *Кейтер Дж.* Компьютеры – синтезаторы речи /Пер. с англ.; Под ред. В.А. Усика. – М.: Мир, 1985. – 237 с.
3. *Давлетьянц А.И., Коломиец А.В.* Статистические характеристики дельта-модулированного сигнала // Статистические методы обработки информации в авиационных радиоэлектронных системах. – К.: КИИГА, 1995. – С. 23, 24.
4. *Мастрюков Д.* Алгоритмы сжатия информации. Сжатие по Хаффмену // Монитор. – 1993. – №7,8. – С. 14–20.

Стаття надійшла до редакції 06.07.01.