



МІКШУВАННЯ МОВНИХ СИГНАЛІВ У МУЛЬТИМЕДІЙНИХ СИСТЕМАХ РЕАЛЬНОГО ЧАСУ

Мельник А.О ¹⁾, Коркішко Т.А ²⁾, Шевчук Р.П ³⁾

¹⁾ Кафедра електронних обчислювальних машин Державного університету “Львівська політехніка”,
aomelnyk@polynet.lviv.ua;

²⁾ Samsung Advanced Institute of Technology, k.tymur@samsung.com;

³⁾ Кафедра комп’ютерних наук Тернопільського державного економічного університету, rsh@tanet.edu.te.ua

Резюме: У роботі розглянуто базові принципи мікшування та наведено структуру одного з найпростіших методів мікшування мовних сигналів. Сформовано вимоги до мовних сигналів, які подаються на вхід мікшера. На основі цих вимог та на основі аналізу недоліків відомих методів мікшування запропоновано метод багатоступінчастого мікшування, який дозволяє міксувати значення відліків в міру надходження блоків даних у блок мікшування, а також міксувати значення відліків з блоків даних, що були одержані шляхом декомпресії стиснених мовних сигналів різних форматів.

Ключові слова: Мікшування, багатоступінчасте мікшування, мікшування мовних сигналів, структура блоку мікшування

1. ВСТУП

Широке впровадження мультимедійних систем реального часу, таких як IP-телефонія та мультимедіа-конференції, вимагає від мережного обладнання забезпечення якісної взаємодії великої кількості учасників мультимедійного сеансу з ефективним використанням пропускних спроможностей цифрових каналів зв’язку. Одним із методів підвищення ефективності використання цифрових каналів зв’язку у мультимедійних системах реального часу є мікшування - процес змішування цифрових потоків, які генеруються активними учасниками мультимедійного сеансу з подальшою нормалізацією сформованого потоку [1,2]. Об’єктами мікшування у мультимедійних системах реального часу є потоки мовних сигналів. Мікшування мовних сигналів виконується блоком мікшування (мікшером) [4] та дозволяє ущільнити вхідні потоки, що поступають від активних учасників мультимедійних систем до одного вихідного потоку.

2. ПОСТАНОВКА ЗАДАЧІ

Сьогодні існує ряд методів мікшування мовних сигналів, основними відмінностями яких є архітектура мікшера та особливості

мікшування вхідних потоків [1,2,5,6]. До особливостей мікшування можна віднести: 1) формат сигналів, що подаються на вхід мікшера; 2) спосіб виконання мікшування. Спільним для всіх методів мікшування мовних сигналів є те, що вони працюють з декомпресованими мовними сигналами, та на виході мікшера генерують один вихідний потік.

Основною вимогою, що ставиться до мовних сигналів, які подаються на вхід мікшера, є приведення їх до одного формату. Однак структурна та семантична відмінність алгоритмів стиснення мовних сигналів, які використовуються у мультимедійних системах стандарту H.323 [7] та SIP [8], не дозволяє одержати однозначного формату мовного сигналу після декомпресії [9-13]. Дана особливість алгоритмів стиснення практично не враховується при побудові методів мікшування.

Тому постає задача побудови нового методу мікшування, який би враховував особливості сигналів, що були одержані шляхом декомпресії стиснених мовних сигналів різних форматів.

3. АНАЛІЗ ЛІТЕРАТУРИ

У роботі [5] пропонується та експериментально досліджується метод частково розподіленого мікшування, який дає змогу оптимізувати потоки даних, що передаються

згідно з протоколом TCP. Результати досліджень, проведених у роботі [5] показали, що даний метод з успіхом можна використовувати на перевантажених мережах, у яких спостерігається високий процент втрати пакетів. Особливостями даного методу є його орієнтація на ієрархічну архітектуру багатоабонентських мультимедіа-конференцій [1,9-11] та використання часових інтервалів при визначенні часу мікшування.

У роботі [1] розглядаються базові принципи мікшування в розрізі там же запропонованих алгоритмів бінарного та множинного мікшування. Як і у роботі [5], дані алгоритми здійснюють мікшування через визначені інтервали часу.

У роботі [2] пропонується метод мікшування, який базується на організації логічних черг в розподіленому буфері мікшування. Алгоритм реалізації цього методу передбачає обробку мовних сигналів, які були дискретизовані з однією частотою.

У роботі [6] пропонується метод і система для мікшування мовних потоків засобами звукової карти комп'ютера. Даний метод має ряд недоліків, описаних в [6], та майже не використовується для вирішення задач мікшування мовних сигналів у мультимедійних системах реального часу.

Проаналізувавши наведені вище роботи, можна зробити висновок, що методи мікшування, які приводяться у цих роботах, орієнтовані на опрацювання мовних сигналів одного формату. Крім того, мікшування мовних сигналів на основі часових інтервалів, що використовується у методах, представлених у роботах [1,5], не завжди дозволяє отримати всі пакети від активних учасників мультимедійного сеансу.

4. МЕТА РОБОТИ

Метою роботи є розробка нового методу мікшування.

5. ОГЛЯД БАЗОВИХ ПРИНЦИПІВ МІКШУВАННЯ

Найпростіший мультимедійний сеанс між трьома активними учасниками, що спілкуються в реальному часі, передбачає одержання кожним із них комплексного мовного потоку, який є сумою лінійних комбінацій впливових звукових хвиль кожного учасника.

Нехай $L_{i,j}(q)$ – блок даних із інкапсульованими значеннями лінійних відліків q , який надходить на вхід блоку мікшування (де i – номер блоку в каналі; j – номер каналу; $i=1, \dots, k$; $j=1, \dots, N$; k –

максимальна кількість блоків даних у j -му каналі; N – кількість каналів).

Якщо припустити, що у кожен канал блоку мікшування надходять блоки даних тільки від одного активного учасника мультимедійного сеансу, тоді для одержання мікшованого потоку з блоків даних, які надійшли на вхід блоку мікшування, нам необхідно використати вираз:

$$M_i(q) = \sum_{i=1}^k \sum_{j=1}^N L_{i,j}(q) \quad (1)$$

де $M_i(q)$ – мікшований потік, що складається з арифметичної суми i -тих значень лінійних відліків q з блоків даних $L_{i,j}(q)$.

Крім того, для кожного активного учасника P_M ($M=1, \dots, S$; S – кількість активних учасників мультимедійного сеансу) необхідно з мікшованого потоку $M_i(q)$ виключити його власний блок даних:

$$P_M(L_{i,j}(q)) = M_i(q) - L_{i,M}(q). \quad (2)$$

Вирази (1) та (2) є базою для побудови будь-якого методу мікшування [2].

Припустимо що всі учасники P_M генерують свої медіа-пакети із сталим періодом per ($per=const$). Тоді найменший інтервал, протягом якого два незалежні учасники гарантовано згенерують свої медіа-пакети, рівний $per/2$. Для P_M активних учасників, таким інтервалом буде $per=per/(per/2^{P_M-1})$. Мікшування лінійних відліків з двох незалежних блоків даних, що поступають у різні канали блоку мікшування, регламентується правилом [1]: правило мікшування: значення лінійних відліків q з блоків даних $L_{1,1}(q)$ та $L_{1,2}(q)$ можуть мішуватись, якщо $|g(L_{1,1}(q)) - g(L_{1,2}(q))| < per$, де $g(L_{1,1}(q))$ і $g(L_{1,2}(q))$ - часи генерації медіа-пакетів учасниками мультимедійного сеансу, у яких містяться блоки даних $L_{1,1}(q)$ та $L_{1,2}(q)$.

Більшість мультимедійних систем передають медіа-пакети згідно з протоколом RTP [3], який дозволяє передавати пакети в реальному часі. У цьому протоколі визначається тип блоку даних, проводиться нумерація посилок, присвоюються часові відмітки часу ? та проводиться моніторинг процесу передачі пакету. При цьому до вихідного блоку даних, який формується блоком мікшування, додається заголовок RTP, у якому модифікуються наступні поля [12]:

- ідентифікатор джерела ідентифікує проміжне джерело генерування RTP-пакетів, та створює список SSRC-ідентифікаторів джерел, блоки даних яких сформували комплексний потік;

- 32-бітова RTP-відмітка часу фіксує час створення пакету джерелом;
- ідентифікатор джерела синхронізації визначається 32-бітовим числовим SSRC-ідентифікатором та визначає джерело генерування RTP-пакетів;
- порядковий номер пакета (15 біт).

Дані поля RTP-заголовку дозволяють правильно розмістити одержані блоки даних у буфері та забезпечують їх синхронізацію.

При роботі найпростішого методу мікшування блоки даних $L_{i,j}(q)$ із інкапсульованими значеннями лінійних відліків q надходять у канали блоку мікшування з блоків декомпресії медіа-пакетів та зберігаються у буферах блоку мікшування, який організований в режимі FIFO. Кожен буфер у блоці мікшувань помічається відповідно до відмітки RTP-часу. Кількість буферів у блоці мікшування встановлюється динамічно та лінійно залежить від кількості учасників мультимедійного сеансу. Буферизація отриманих $L_{i,j}(q)$ блоків даних від P учасників мультимедійного сеансу виконується до того часу, поки, як мінімум, не отримано по i -му блоку даних з кожного каналу (рисунок 1).

Мікшування значень лінійних відліків з i -тих блоків даних відбувається через кожен вихідний інтервал пакетування, який визначається таймером і при виконанні правила мікшування.

Блок мікшування змішує блоки даних $L_{i,j}(q)$ у комплексний потік $M_i(q)$ за допомогою простого додавання значень лінійних відліків q з кожного i -го блоку даних зі всіх j -тих каналів. Інтервали пакетування можуть зменшуватись та збільшуватись залежно від зміни часу спрацювання таймера. Після одержання блоку даних $M_i(q)$ проводиться процедура виключення власного блоку даних для кожного активного учасника з комплексного потоку, згідно з виразом (2), яка проводиться у комірках A_1, \dots, A_S модуля виключення власного потоку учасника. Результатом роботи блоку мікшування є генерування пакету із RTP-заголовком (рисунок 2), який відправляється у блоки компресії.

Розглянутий метод має ряд недоліків, хоч на ньому і базуються більшість методів мікшування [1].

Необхідно відзначити, що наведений вище метод мікшування відносять до методів одноступінчатого мікшування. Ці методи змішують значення лінійних відліків тільки у випадку, коли одержано по i -му блоку даних з кожного j -го каналу [11]. Крім названих, існують методи багатоступінчастого мікшування. У цих методах процес мікшування відбувається, коли хоч би у два канали блоку мікшування поступило по i -му блоку даних.

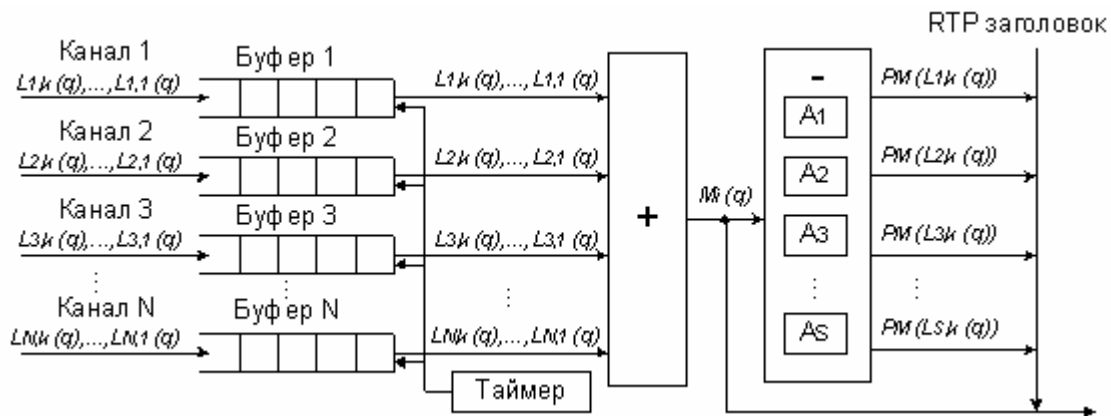


Рис. 1- Структура блоку мікшування, який реалізує один з найпростіших методів мікшування

0	2	3	4	7	8	16	31	Блок даних
V=2	P	X	СС	M	PT	Порядковий номер		
RTP відмітка часу								
Ідентифікатор джерела синхронізації								
Ідентифікатори джерел								

Рис. 2 – Структура медіа-пакету на виході блоку мікшування

Питання вибору оптимального методу мікшування для конкретної мультимедійної системи є достатньо складною задачею, оскільки її вирішення визначається рядом чинників: зокрема, типом фізичного середовища, способом реалізації мультимедіа системи, способом управління мережею, завантаженістю та пропускною спроможністю мережі. Лише детально проаналізувавши кожен із цих чинників, можна давати рекомендації по вибору методу мікшування та засобу його реалізації.

Відомі сьогодні методи мікшування можуть реалізовуватись наступним чином:

- мікшером, який вмонтований у трансодері шлюзу чи пристрої управління багатоабоненськими мультимедіа-конференціями;
- звуковою картою комп'ютера;
- мікшером, вмонтованим у термінал мультимедійної системи;
- операційною системою;
- програмно, на рівні додатків.

Необхідно відзначити, що мікшування медіа-пакетів має ряд недоліків, які прийнято розглядати, зважаючи на якість мовного сигналу, одержаного приймачем. Зважаючи на цей факт, до основних недоліків мікшування відносять [5]:

- мовний сигнал після мікшування є більш зашумленим (або/і має менший динамічний діапазон), оскільки включає шуми всіх мовних сигналів, що змішувались з ним;
- мікшування не працює з сигналами, які відрізняються параметрами;
- у процесі мікшування створюється додаткова затримка на шляху від джерела до приймача;
- вводиться додаткове обладнання в мережу, до якого висувуються вимоги обробки медіа-пакетів в реальному часі.

Однак поряд із цими недоліками виділяють і переваги мікшування, головна з яких – суттєве зменшення кількості потоків даних, що передаються між активними учасниками мультимедійного сеансу в реальному часі, що дозволяє ефективно керувати потоками даних в мережі.

6. МЕТОД БАГАТОСТУПІНЧАСТОГО МІКШУВАННЯ

Задача побудови нових методів мікшування зумовлена, в першу чергу, сьогоднішніми умовами розвитку мультимедійних систем реального часу. У протоколах підтримки мультимедійних сеансів зв'язку спостерігається

тенденція до використання декількох алгоритмів стиснення мовних сигналів. Зокрема, у протоколі H.323 регламентовано використання п'яти алгоритмів стиснення мовних сигналів [7], у протоколі SIP – чотирьох [8]. Однак, відомі нам методи мікшування не пристосовані до змішування значень відліків з блоків даних, що були одержані шляхом декомпресії стиснених мовних сигналів різних форматів. Основними відмінностями між відліками у таких блоках даних є частоти квантування та бітрейд. Крім того, блоки даних $L_{i,j}(q)$, що поступають з різних блоків декомпресії, часто відрізняються розміром.

Говорячи про структуру та принцип роботи методів мікшування, необхідно відзначити, що методи багатоступінчастого мікшування характеризуються кращою швидкістю, яка одержується за рахунок відсутності затримок, зв'язаних з часом очікування i -тих блоків даних [11]. Однак алгоритми виконання цих методів слабо дослідженні та мало описані в літературі.

З врахуванням цих особливостей та дотримуючись представлених вище принципів мікшування ми пропонуємо новий метод багатоступінчастого мікшування на базі пам'яті довільним доступом (адресної пам'яті).

Структура блоку мікшування, який реалізовує даний метод, представлена на рисунку 3.

Як видно з рис. 3, блоки даних $L_{i,j}(q)$ із інкапсульованими значеннями лінійних відліків q через канали передаються у модуль корекції частоти дискретизації, у якому в разі необхідності (якщо частота дискретизації відліків не рівна 8 КГц), застосовується алгоритм передискретизації [13]. Далі блоки даних передаються у модуль корекції бітрейду, у якому значення всіх лінійних відліків q з кожного блоку даних приводяться до 16-бітового формату, та узгоджується розмір блоку даних $Lb(L_{i,j}(q))$. З модуля корекції бітрейду блоки даних відправляються в модуль пам'яті з довільним доступом. У модулі пам'яті для кожного i -го номеру блоку, що надходить з будь-якого каналу, виділяється комірка пам'яті об'ємом $Lb(L_{i,j}(q))$ байт, якій присвоюється відповідна адреса. Процес мікшування блоків даних відбувається за принципом роботи методів багатоступінчастого мікшування та здійснюється у модулі пам'яті. Це означає, що значення відліків q з кожного i -го номеру блоку сумуються у комірках пам'яті по мірі їх надходження у модуль пам'яті. Усі комірки модуля пам'яті помічаються відповідно до відмітки RTP-часу. Пристрій керування проводить розподіл блоків даних у комірки модуля пам'яті. Функцією лічильника та

таймера, що знаходяться у модулі керування, є визначення моменту часу, коли необхідно передати сформований блок даних $M_i(q)$ з комірки модуля пам'яті у модуль виключення власного блоку даних учасника мультимедійного сеансу. Після одержання блоку даних $M_i(q)$ проводиться процедура виключення власного

блоку даних для кожного активного учасника з комплексного потоку, згідно з виразом (2), у комірках A_1, \dots, A_S модуля виключення власного потоку учасника. Результатом роботи блоку мікшування є генерування пакету із RTP-заголовком, який відправляється у блоки компресії.

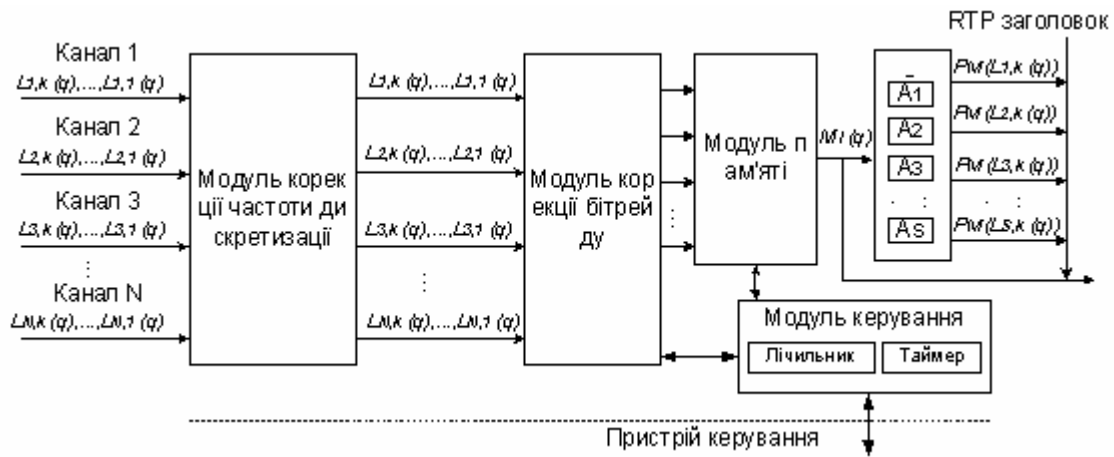


Рис. 3 – Структура блоку мікшування, який реалізує метод багатоступінчастого мікшування

Розглянемо детальніше функції кожного модуля мікшера:

- Модуль корекції частоти призначений для узгодження частоти дискретизації лінійних відліків q , які надходять у блоки даних на модуль корекції частоти дискретизації. Якщо частота дискретизації відліків не рівна 8 КГц, то до всіх лінійних відліків q , що містяться у блоці даних, застосовується алгоритм передискретизації. Частота 8 КГц була вибрана зважаючи на аналіз декодерів стиснення мовних сигналів, більшість з яких забезпечують саме таку частоту дискретизації [14]. У випадку, коли частота дискретизації відліків рівна 8 КГц, блоки даних передаються до наступного модуля;
- Модуль корекції бітрейду. Цей модуль повинен у випадку необхідності привести формат відліків q у блоках даних до одного бітрейду (16 біт) та забезпечити їх передачу у комірки модуля пам'яті. Крім того, у цьому модулі регулюється розмір блоків даних $L_b(L_{i,j}(q))$. Блоки даних $L_{i,j}(q)$ передаються у комірки модуля пам'яті відповідно до адреси комірки, яка одержується з модуля керування.
- Модуль пам'яті складається з $L_b(L_{i,j}(q))$ однобайтових комірок, які помічаються відповідно до відмітки RTP-часу. Кожній комірниці присвоюється адреса, яка

зв'язується з i -тим номером блоку даних. Відповідно до адреси в комірку передаються блоки даних з лінійними відліками q , які мішкуються в міру надходження. Ще однією функцією модуля пам'яті є передача сформованих блоків у модуль виключення власного блоку даних учасника мультимедійного сеансу. Дана функція активізується після відповідного сигналу з модуля керування.

- Модуль виключення власного блоку даних учасника здійснює виключення власного блоку даних для кожного активного учасника мультимедійного сеансу з блоку даних $M_i(q)$ відповідно до виразу (2), у комірках A_1, \dots, A_S . Після цього блок даних подається на вхід блоків компресії.
- Модуль керування взаємодіє із пристроєм керування транскодера, від якого одержує інформацію про пакети, які поступають на вхід транскодера та про поля із заголовку RTP-пакету. На основі аналізу цієї інформації та інформації, одержаної з модуля корекції бітрейду, модуля пам'яті, модуль керування сигналізує про можливість передачі блоків даних у комірки модуля пам'яті. Модуль керування складається з лічильника і таймера. Основною функцією модуля керування є подання

сигналу модулю пам'яті про можливість передачі i -го блоку даних у модуль виключення власного потоку учасника мультимедійного сеансу. Дана функція виконується на основі спостереження за кількістю активних учасників мультимедійного сеансу та кількістю i -тих блоків даних, які надійшли у модуль пам'яті. Якщо кількість i -тих блоків даних рівна кількості активних учасників P , то модуль керування сигналізує про можливість передавачі сформованого $M_i(q)$ блоку даних з комірки модуля пам'яті у наступний модуль. Для того, щоб не допустити значних затримок, які можуть відбутись через втрату деяких пакетів, у модуль керування введено таймер, який встановлює інтервали пакетування за аналогією з [5].

Переваги розробленого методу:

- метод дозволяє міксувати значення відліків з блоків даних, що були одержані шляхом декомпресії мовних сигналів різних форматів;
- значення відліків мішкуються по мірі надходження блоків даних у блок міксування, що дозволяє уникати черг.

7. ВИСНОВКИ

У роботі проаналізовані відомі методи міксування мовних сигналів, що дозволило виділити їх недоліки та розробити вимоги до побудови нових методів міксування. Розглянуто базові принципи міксування, представлено структуру блоку міксування, який виконує один з найпростіших методів міксування мовних сигналів, показано принцип його роботи. Сформовано вимоги до мовних сигналів, які подаються на вхід мікшера. На основі цих вимог та на основі аналізу недоліків відомих методів міксування запропоновано метод багатоступінчастого міксування. У роботі наведена структура мікшера, який реалізує запропонований метод міксування, та визначено функції кожного модуля у цій структурі. Основними перевагами розробленого методу є:

- метод дозволяє міксувати значення відліків з блоків даних, що були одержані шляхом декомпресії мовних сигналів різних форматів;
- значення відліків мішкуються по мірі надходження блоків даних у блок міксування, що дозволяє уникати черг.

Основним недоліком даного методу, на нашу думку, є введення у структуру блока міксування

додаткових модулів, що дещо збільшує затримки при обробці блоків даних мікшером.

Даний метод рекомендується використовувати у мультимедійних сеансах із багатьма активними учасниками, оскільки він, на відміну від методів одноступінчастого міксування, дозволяє в реальному часі міксувати медіа-пакети учасників сеансу, не очікуючи прибуття всіх медіа-пакетів активних учасників.

Метод рекомендується реалізовувати на базі мікшера, який вмонтований у транскодер шлюзу чи пристрою управління багатоабоненськими мультимедіа-конференціями.

8. НАПРЯМКИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Для одержаного методу багатоступінчастого міксування планується розробити алгоритм методу та провести серію обчислювальних експериментів для того щоб доказати його ефективність.

Додатково, актуальною є задача реалізації одержаного методу у структурі багатоканального транскодера, запропонованого у [4].

9. ЛІТЕРАТУРА

- [1] Rangan P. V., Harrick M., Ramanathan V. S. Communication Architectures and Algorithms for Media Mixing in Multimedia Conferences, IEEE/ACM Transactions on Networking, Vol.1, No.1, February 1993.
- [2] Agusthn JG, Hussein AW. Audio mixing for interactive multimedia communications. In: Wang P, ed. Proc of the JCIS'98. NC: Research Triangle, 1998. 217-220.
- [3] Schulzrinne, H., Casner, S., Frederick, R., Jacobson, V., RTP: A Transport Protocol for Real-Time Applications, IETF RFC 1889, January 1996.
- [4] A. Melnyk., R. Shevchuk. Transcoding of Formats of Compressed Speech Signals // Proceedings of the 8-th International Conference CADSM'2005. - Lviv-Polyana, Ukraine, 23 - 26 February 2005, P. 151-153.
- [5] Radenkovic, M., Greenhalgh, C. Multi-party Distributed Audio Service with TCP Fairness // Proceedings of the Sixth conference of the UK_VRSIG, Salford, UK, September 1999.
- [6] R. Heddle., F. Yerrace., G. Dahl. Method and system for mixing audio streams in a computing systems // United States Patent № 5,703,794
- [7] International Telecommunication Union, "Packet based multimedia communication systems", Recommendation H.323 /

Telecommunication Standardization Sector of ITU, Geneva, Switzerland, Feb., 1998.

- [8] Б.С. Гольдштейн, А.В. Пинчук, А.Л. Суховицкий. IP-телефония // М.: Радио и связь.- 2001. – 336 с.
- [9] Т. Коркішко, Р. Шевчук. Аналіз архітектур багатоабонентських мультимедіа-конференцій // Матеріали третьої міжнародної науково-технічної конференції “Проблеми інформатики і моделювання.”. – Х.: НТУ “ХПІ”, 2003. – 44 с.
- [10] Rangan P. V., Harrick M., Ramanathan V. S. Hierarchical Conferencing Architectures for Inter-Group Multimedia Collaboration. In Proceedings of the Conference on Organizational Computing Systems (COCS’91), SIGOIS Bulletin, Vol. 12, No. 2-3, pages 43–55, November 1991.
- [11] Rangan P. V., Harrick M., and etc. Optimal Communication Architectures for Multimedia Conferencing in Distributed Systems. ICDCS 1992: 46-53.
- [12] Семенов Ю. А. “Сети Интернет. Архитектура и протоколы”, СИРИНЪ, 1998. – 424 с.
- [13] А.Б. Сергиенко. Цифровая обработка сигналов. – СПб.: Питер, 2003. – 604 с.
- [14] Мельник А.О., Шевчук Р.П. Порівняльний аналіз алгоритмів стиснення мовних сигналів // Вісник "Комп'ютерні системи та мережі" НУ “Львівська політехніка” . – Львів, 2005.



Тимур Коркішко, закінчив Державний університет “Львівська політехніка” за спеціальністю “Комп'ютерні та інтелектуальні системи і мережі” у 1997 році. У 2003 році отримав науковий ступінь кандидата технічних наук. З 2002 року працює старшим викладачем кафедри комп'ютерних наук Тернопільської академії народного господарства. Викладає курси: мови програмування, основи автоматизованого проектування засобів обчислювальної техніки. Області наукових інтересів: високошвидкісна криптографія, методологія розробки криптографічних процесорів і акселераторів, мови програмування. Автор більш як 25 наукових праць.



Шевчук Руслан Петрович, у 2003 році закінчив магістратуру ІКІТ Тернопільської академії народного господарства за спеціальністю “Комп'ютерні системи і мережі”. З 2003 року викладач кафедри комп'ютерних наук. З 2004 аспірант кафедри комп'ютерних наук. Наукові інтереси: алгоритми стиснення мовних сигналів, інтернет телефонія, мультимедіа технології. Автор 12 наукових праць.
Email: rsh@tanet.edu.te.ua



Мельник Анатолій Олексійович, доктор технічних наук, професор, завідувач кафедри ЕОМ НУ “Львівська Політехніка”. З 2004 року працює на кафедрі “Комп'ютерних наук” ТДЕУ за сумісництвом. Наукові інтереси: архітектура

універсальних та спеціалізованих комп'ютерів і комп'ютерних систем, в тому числі багатопроцесорних комп'ютерних систем на основі багатоканальних конвеєрних процесорів з спільним та з розподіленим масивом багатопортової пам'яті, теоретичні основи їх побудови і методи високорівневого проектування, а також дослідження, синтез і реалізація комп'ютерних засобів для вирішення задач складної обробки потоків даних високої інтенсивності, включаючи задачі моделювання, управління, цифрової обробки сигналів та зображень. Автор понад 200 наукових праць, в тому числі 86 авторських свідоцтв та патентів.

Email: aomelnyk@polynet.lviv.ua

MIXING SPEECH SIGNALS FOR THE REAL-TIME MULTIMEDIA SYSTEMS

A. Melnyk ¹⁾, T. Korkishko ²⁾, R. Shevchuk ³⁾

¹⁾ Computer Engineering Department, Lviv Polytechnic National University, S. Bandery Str., 12, 79046, Lviv, Ukraine, E-mail: aomelnyk@polynet.lviv.ua

²⁾ Samsung Advanced Institute of Technology, E-mail: k.tymur@samsung.com

³⁾ Computer Science Department, Ternopil State Economic University, 3 Peremoga Square, 46004, Ternopil, Ukraine. E-mail: rsh@tanet.edu.te.ua

Abstract: *In this work we considered basic principle of mixing and offered the method of multistage mixing which allows the mixing speech samples on the measure of receipt of data blocks in mixer, and also mixing speech samples, that was got from decompression of compression speech signals of different formats.*

Keywords: *Mixer, multistage mixing, mixing speech signals.*

Wide introduction real-time multimedia systems such as IP-telephony and multimedia conferences requires from the network equipment providing the effective co-operation of a plenty of participants of multimedia session with the optimum use of digital communication channels. One of methods of rise of efficiency of digital communication channels in the real-time multimedia systems is mixing - process of mixing of digital data flows which is generated by the active participants of multimedia session with subsequent normalization of the formed data flow. The data flows of speech signals are the objects of mixing in the real-time multimedia systems. The mixing signals is executed by the block of mixing (mixer) and allowed lower input data flows that act from the active participants of the multimedia systems of to one output data flow.

Today there is the several methods of mixing speech signals, the basic differences of which are architecture of mixer and features of mixing input data flows. In particular, it is possible to take to the features of mixing: 1) format of signals that are given on the mixer input; 2) the device, which carries out mixing; 3) method implementation of mixing; 4) facilities of realization of the mixer. General for all methods of mixing speech signals is the way they work with decompression speech signals and on the output of mixer generate one output data flow. By the basic requirement to the speech signals, which are given on the mixer input, there is adduction of them to the common format. However, structural and semantic difference of algorithms compression of speech signals, which are used in the multimedia systems of standard H.323 and SIP does not allow to get the common format of speech signal after decompression. The given feature of algorithms of compression is not practically taken

into account at the construction of methods of mixing.

Therefore the task of construction of a new method of mixing, which would take into account the resulted features, gets up.

The task of construction of new methods of mixing above all things is conditioned by today's terms of development of the real-time multimedia systems. In protocols of support of multimedia sessions is observed the tendency to use of a few algorithms of compression of speech signals. In particular, in protocol H.323 the use 5 algorithms of compression of speech signals, in protocol SIP – 4. However, the known to us methods of mixing are not adjusted to mixing of samples from the data blocks that were received by decompression of compressed speech signals of different formats. Frequencies of quantum and bit rate are basic differences between counting out in such data blocks. In addition, data blocks that act from different blocks of decompression often differ by their size.

Speaking about a structure and principle of work of methods of mixing, it is necessary to mark that the methods of multistage mixing are characterized by the best fast-acting which turns out due to absence of delays linked in course of time expectation of i data blocks. However, structures and principle of work of these methods are not deeply studied and there is not a sufficient amount of literary sources concerning the topic.

Let us consider $L_{i,j}(q)$ as a data block with the encapsulated samples of q out which acts on the entrance of block of mixing, where i – block number in a channel, j – number of channel, $i=1..k$, $j=1..N$, k – maximal quantity of data blocks in a j channel N – quantity of channels.

If to assume that data blocks only from one active participant of multimedia session acts in every channel of block of mixing, then for the receipt of mixed data flow from the data blocks, which acted on the entrance of block of mixing to us, it is necessary to use expression:

$$Mi(q) = \sum_{i=1}^k \sum_{j=1}^N L_{i,j}(q) \quad (1)$$

where $Mi(q)$ – mixed data flow – linear sum of samples of q out, from the data blocks $L_{i,j}(q)$.

In addition, for every active participant PM ($M=1..S$, S – quantity of active participants of

multimedia session) it is necessary from a mixed data flow $Mi(q)$ to eliminate his own data block:

$$P_M(L_{i,j}(q)) = M_i(q) - L_{i,M}(q) \quad (2)$$

Expressions (1) and (2) is a base for construction of any method of mixing.

Taking into account these features and adhering to the above presented principles of mixing, we offer a new method of multistage mixing on the base of memory with address access.

Structure of block of mixing, which realizes the given method presented on figure 1.

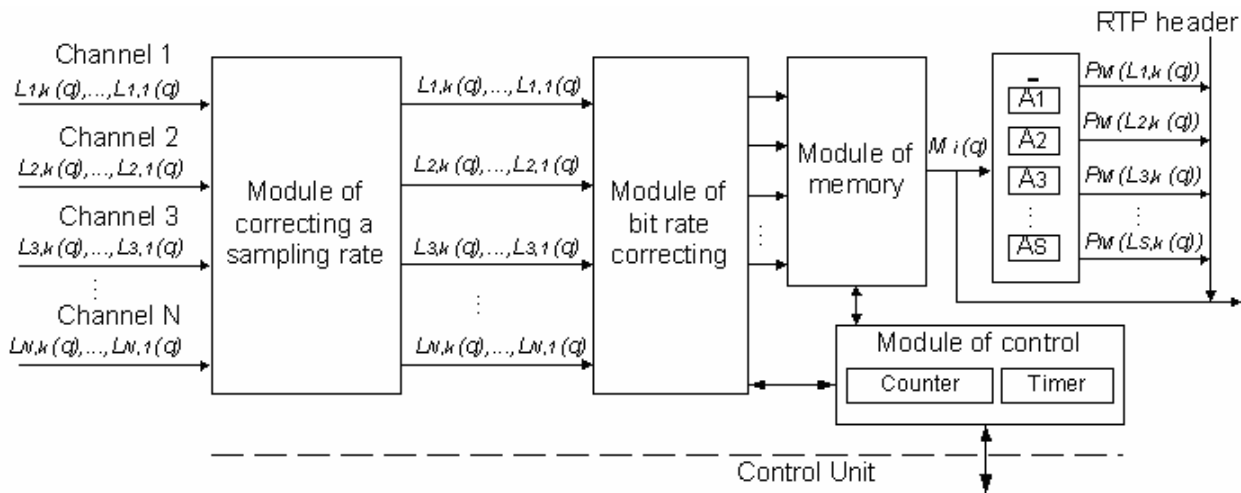


Fig. 1 – Structure of block of mixing, which will realize the method of multistage mixing

CONCLUSIONS

In work we offered method of multistage mixing and presented the structure of mixer, which will realize the given method. Main advantages of the developed method are:

1. The method allows mixing samples out from the data blocks that was got by decompression of compressed speech signals of different formats.

2. Samples mixing out on the measure of receipt of data blocks in the block of mixing, that allows to avoid turns.

We recommended the use of the given method in multimedia sessions with many active participants, because it, unlike the methods of monostage mixing, allows to mix media-packets of participants of session in the real time without expecting the arrival of all media-packets of active participants.