

Міністерство освіти і науки України
Національний університет “Львівська політехніка”

Демчук Андрій Богданович

УДК 004.032.6+004.031.43+004.415

**МАТЕМАТИЧНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ПОДАННЯ
ВІДЕОКОНТЕНТУ ДЛЯ ОСІБ З ВАДАМИ ЗОРУ**

01.05.03 – математичне та програмне забезпечення обчислювальних машин
і систем

Автореферат
дисертації на здобуття наукового ступеня
кандидата технічних наук

Львів–2015

Дисертацією є рукопис.

Робота виконана у Національному університеті “Львівська політехніка” Міністерства освіти і науки України.

Науковий керівник: доктор технічних наук, професор
Литвин Василь Володимирович,
Національний університет “Львівська політехніка”,
завідувач кафедри інформаційних систем та мереж

Офіційні опоненти: доктор технічних наук, професор
Сидоров Микола Олександрович,
Національний авіаційний університет,
завідувач кафедри інженерії програмного забезпечення

кандидат технічних наук, доцент
Угрин Дмитро Ілліч,
Чернівецький факультет Національного технічного
університету “Харківський політехнічний інститут”,
завідувач кафедри інформаційних систем

Захист відбудеться “2” липня 2015 р. о 11:00 годині на засіданні спеціалізованої вченої ради Д35.052.05 у Національному університеті “Львівська політехніка” (79013, м. Львів, вул. С. Бандери, 12).

З дисертацією можна ознайомитись у науково-технічній бібліотеці Національного університету “Львівська політехніка” (79013, м. Львів, вул. Професорська, 1).

Автореферат розіслано “28” травня 2015 р.

Учений секретар
спеціалізованої вченої ради,
доктор технічних наук, професор

Р. А. Бунь

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. В Україні проживає близько 600 тисяч людей з вадами зору, з них приблизно шоста частина – тотально незрячі. Відомо, що близько 80% інформації надається людині за допомогою зображення. Тотально незряча особа може отримати лише близько 16% інформації, яка надходить через слуховий орган.

Одним із здобутків людства, який малозрозумілий для осіб з вадами зору – це відеоконтент (фільми, мультиплікація, науково-популярні, навчальні та пізнавальні програми). У час, коли кіноіндустрія має технічні можливості, яких достатньо для втілення будь-якого художнього задуму – перед людьми відкривається величезний світ можливостей сучасного кіномистецтва. Однак всі ці можливості недоступні для осіб з вадами зору. Так, незрячі чують усі слова акторів, звуки навколишнього середовища, процесів на екрані, але їм важко ідентифікувати, кому які слова належать, що відбувається на екрані, яка реакція акторів, котру вони, часто виражають за допомогою рухів чи міміки.

Наукові дослідження в напрямку доповнення відеоконтентів тифлокоментарями, для забезпечення доступу до відеоконтенту осіб з вадами зору, почалися у 80-х роках минулого століття та інтенсивно розвиваються. Основні теоретичні засади аудіоопису наведено у роботах Д. Цінтаса, Д. Аерса, Б. Сігела та С. Хаджеса; С. Ваньшин ввів поняття тифлокоментування на пострадянському просторі, яке близьке поняттю аудіо опису у решті світу; у працях А. Шарковської, Д. Люїса, Р. Богса, Ф. Піко наведено методики побудови адаптованого відеоконтенту та його можливі шляхи розвитку; над питаннями доступності інформації для осіб з вадами зору у відео форматі працювали відомі вчені: Д. Рубенштейн, К. Маркантоніо, М. Кіндер, Д. Кругер, Д. Фелс, Д. Удо, П. Тінг, С. Раї, але адаптація відеоконтенту (пошуку місць для вставлення тифлокоментарів) завжди виконувалася за допомогою людини.

Розвиток науки, комп'ютеризація суспільства та використання мультимедійних технологій зумовили розроблення комп'ютерних систем, які надають можливість доступу особам з вадами зору до різносторонньої інформації. Тому це становище можна подолати шляхом розробки методів та програмних засобів адаптації відеоконтенту для осіб з вадами зору.

Отже, дисертаційне дослідження присвячено розв'язанню актуального науково-практичного завдання розроблення програмно-алгоритмічного комплексу (ПАК) адаптації відеоконтенту для потреб осіб з вадами зору.

Зв'язок роботи з науковими програмами, планами, темами

Дисертаційна робота виконана в межах наукового напряму “Нові комп'ютерні засоби та технології інформатизації суспільства” визначеного пріоритетним у переліку актуальних проблем Міністерством освіти і науки, молоді та спорту України, концепції програми інформатизації НАН України, визначеної пріоритетним напрямом, згідно розпорядження № 146 від 27.02.2004 р. та за тематикою наукових досліджень кафедри інформаційних систем і мереж Національного університету “Львівська політехніка”, зокрема в рамках держбюджетної науково-дослідної роботи “Математичне та програмно-

алгоритмічне моделювання процесів трансляції жестів у текст для спеціалізованих комп'ютерних систем” (номер державної реєстрації 0111U001222). Автор розробив програмно-алгоритмічний комплекс побудови відеоконтенту для осіб з вадами зору, методи функціонування окремих компонент ПАК, які ґрунтуються на теорії координації та асоціативних правилах, що дало змогу формалізувати процес побудови відеоконтенту.

Мета і задачі дослідження. Метою дисертаційної роботи є розроблення математичного та програмного забезпечення адаптації відеоконтенту для осіб з вадами зору.

Для досягнення поставленої мети в дисертації сформульовано і вирішено такі *задачі* дослідження:

- 1) проаналізувати предметну область, визначити нерозв'язані задачі щодо мети дисертаційних досліджень;
- 2) розробити математичне забезпечення адаптації відеоконтенту для осіб з вадами зору на основі процесу тифлокоментування;
- 3) розробити архітектуру програмно-алгоритмічного комплексу адаптації відеоконтенту для осіб з вадами зору;
- 4) дослідити функціонування розробленого програмно-алгоритмічного комплексу адаптації відеоконтенту для осіб з вадами зору на практиці.

Об'єктом дослідження є процес подання відеоконтенту для осіб з вадами зору.

Предметом дослідження є методи та засоби побудови, редагування та адаптації відеоконтенту для осіб з вадами зору.

Методи дослідження. Під час розв'язання сформульованих задач використано теорію спектрального аналізу для пошуку вільних місць у відеоконтенті для вставлення тифлокоментарів; теорію координації – для побудови додаткової звукової доріжки із тифлокоментарями; теорію асоціативних правил – для побудови правил тифлокоментування; уніфіковану мову моделювання UML – для представлення архітектури програмно-алгоритмічного комплексу; об'єктно-орієнтоване та шаблонне програмування – для реалізації програмного забезпечення ПАК адаптації відеоконтенту для осіб з вадами зору.

Наукова новизна одержаних результатів. У дисертаційній роботі розв'язано актуальне наукове завдання подання відеоконтенту у форматі, доступному для сприйняття особам з вадами зору. Наукова новизна одержаних результатів полягає у розробці математичного забезпечення щодо реалізації програмно-алгоритмічного комплексу, призначеного для побудови, редагування та адаптування відеоконтенту для осіб з вадами зору.

Отримано такі нові наукові результати:

вперше:

- розроблено математичне забезпечення подання відеоконтенту для осіб з вадами зору на основі використання спектрального аналізу, теорії координації та асоціативних правил, що дало змогу формалізувати процес тифлокоментування відеоконтенту;

- розроблено подання тифлокоментарів із застосуванням логіки 1-го порядку, котрі слугують основою для тифлокоментаторів, які працюють над побудовою відеоконтенту для осіб з вадами зору, що дало можливість адаптувати відеоконтент до потреб таких користувачів;
- розроблено архітектуру програмного комплексу побудови відеоконтенту для осіб з вадами зору, який містить компоненти тифлокоментування та редагування відеоконтенту, що дало змогу автоматизувати процес побудови відеоконтенту для осіб з вадами зору;

отримав подальший розвиток:

- метод тифлокоментування відеоконтенту шляхом автоматизації процесів (аналіз відеоконтенту, визначення місць для вставки тифлокоментарів, начитування тифлокоментарів, синхронізація створеного звукоряду з відеорядом) адаптації відеоконтенту до потреб осіб з вадами зору;
- метод пошуку ділянок тиші в звукових даних на основі використання перетворювання WaveLet (який ґрунтується на: фільтрації ділянки звуку, перевірці на щільність спектра, перевірці кореляції спектра з еталонним спектром, згладжуванні ділянок масиву), що дало змогу визначати ділянки тиші для вставлення тифлокоментарів у відеоконтент.

Практичне значення одержаних результатів дисертаційної роботи полягає у побудові програмно-алгоритмічного комплексу адаптації відеоконтенту для людей з вадами зору. Зокрема, практично цінними результатами є:

- розроблені методи визначення тиші у звуковій доріжці та координації тифлокоментування можна використати для побудови паралельної звукової доріжки опису відеосюжету;
- узагальнені правила тифлокоментування є придатними для побудови тифлокоментарів певного часового проміжку;
- розроблений програмно-алгоритмічний комплекс “Audio Editor” тифлокоментування можна використати для побудови відеоконтенту, адаптованого для людей з вадами зору.

Основні результати дисертаційних досліджень (а саме побудовані відеоконтенти з тифлокоментарями на основі розробленого програмно-алгоритмічного комплексу, правила побудови тифлокоментарів) використовуються у Комунальному закладі Львівської обласної ради – Львівській спеціальній загальноосвітній школі-інтернаті I–III ступенів № 100 для незрячих; Львівському дитячому дистанційному навчально-консультаційному центрі дітей з особливими потребами на базі НВК “Школа-гімназія «Сихівська»”; Львівському міському та обласному товаристві сліпих (УТОС). Також результати дисертаційних досліджень використано в навчальному процесі Національного університету “Львівська політехніка” при викладанні дисциплін “Об’єктно-орієнтоване програмування”, “Системне програмування” та “Організація баз даних і знань” що підтверджено відповідними актами.

Особистий внесок здобувача. Усі наукові результати, подані у дисертації, здобувач одержав особисто. У друкованих працях, опублікованих у співавторстві, здобувачеві належить: [2, 10] – розроблено підхід до вирішення задачі доступу осіб з вадами зору до відеоконтенту; [5] – аналіз інформаційних технологій адаптації відеоконтенту для осіб з вадами зору; [6, 11] – розроблено методи задання ваг елементів; [7, 17–19] – побудовано модель задачі побудови інтелектуального агента; [8] – побудовано онтологію нозологій для уможливлення доступу до інформаційних ресурсів особам з вадами зору; [9] – проаналізовано методи та засоби адаптації відеоконтенту для осіб з вадами зору; [12] – досліджено проблему доступу осіб з вадами зору до відеоконтенту, визначено метод тифлокоментування як оптимальний для вирішення цієї задачі; [13] – розроблено модель задачі виготовлення відеоконтенту для осіб з вадами зору, систематизовано відповідні правила і способи створення тифлокоментарів; [19] – проаналізовано основні правила обходу дерева синтаксичного розбору речень; [20] – досліджено проблему координації рішень у багаторівневих ієрархічних системах.

Апробація результатів дисертації. Основні наукові та практичні результати роботи оприлюднено та обговорено на міжнародних наукових конференціях, зокрема Міжнародній науковій конференції “Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту” (м. Євпаторія, 2011, 2013 роки); I Міжнародній науково-технічній конференції “Обчислювальний інтелект” (м. Черкаси, 2011 рік); XVIII Всеукраїнській науковій конференції “Сучасні проблеми прикладної математики та інформатики” (м. Львів, 2012 рік); Всеукраїнській науковій конференції “Сучасні проблеми математичного моделювання та обчислювальних методів” (м. Рівне, 2013 рік); Міжнародній науково-практичній конференції “Інформаційні управляючі системи та технології” (м. Одеса, 2013 рік); XIX Міжнародній науково-практичній інтернет-конференції “Проблеми та перспективи розвитку науки на початку третього тисячоліття у країнах СНД” (м. Переяслав-Хмельницький, 2014 рік); Міжнародній конференції: “Наука та сучасність: виклики XXI ст.” (м. Київ, 2014 рік); Міжнародній науково-практичній конференції молодих вчених та студентів “Інформаційні технології, економіка та право: стан та перспективи розвитку” (м. Чернівці, 2014 рік); Міжнародній конференції: “Наука як основа мирного діалогу” (м. Донецьк, 2014 рік); на III Міжнародній науковій конференції “Інформація, комунікація, суспільство 2014” (сmt Славське, Львівська область, 2014 рік), а також на наукових семінарах кафедри інформаційних систем та мереж Національного університету “Львівська політехніка” (2011–2014).

У 2013 р. робота під назвою “Освітній відеоконтент для людей з вадами зору” (автор А. Б. Демчук) експонувалась на стенді Національного університету “Львівська політехніка” під час щорічної загальнодержавної виставки “Інноватика у сучасній освіті”, м. Київ.

Публікації. За результатами дисертаційних досліджень опубліковано 20 наукових праць, зокрема 2 статті у міжнародних періодичних виданнях

[4, 5], 6 статей у фахових наукових виданнях України з технічних наук [1–3, 6–8], з них одне наукове видання Національного університету “Львівська політехніка” [9], та 11 публікацій у матеріалах і тезах міжнародних і всеукраїнських конференцій [10–20].

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** до дисертації обґрунтовано актуальність теми, сформульовано мету та основні задачі досліджень, показано зв'язок із науковими програмами, планами, темами, розкрито наукову новизну. Розглянуто практичну цінність, реалізацію та впровадження результатів роботи. Наведено дані про особистий внесок здобувача, апробацію роботи та публікації.

У **першому розділі** розглянуто сучасний стан наукових досліджень та інформаційних технологій, які зорієнтовані на людей з особливими потребами. Зокрема, проаналізовано можливість адаптації відеоконтенту до потреб осіб з вадами зору. Проведено аналіз сучасного стану соціальної адаптації осіб з вадами зору. Проаналізовано особливості візуального сприйняття та роль слуху у їхньому житті. Встановлено, неможливість повноцінної соціальної адаптації осіб з вадами зору до сучасного інформаційного середовища, без забезпечення доступу до відеоконтенту.

У результаті дослідження встановлено, що однією з найактуальніших проблем для осіб з вадами зору є майже повна недоступність відеоконтенту. Для адаптування відеоконтенту до потреб осіб з вадами зору обрано метод тифлокоментування. Розглянуто історію зародження та розвитку тифлокоментування (аудіодескрипції) за кордоном та в Україні. Розкрито поняття тифлокоментування, його особливостей та методів побудови.

Тифлокоментар – це закадровий опис відеоряду, складений сценаристом і начитаний тифлокоментатором. *Тифлокоментування* – це коментар візуальних елементів (жестів, зовнішнього вигляду акторів, предметів, костюмів, декорацій). Тож за допомогою тифлокоментування, особи з вадами зору можуть уявити весь спектр візуальних прийомів, які використовував режисер та сценарист відеоконтенту. Отже, виникає завдання дослідження, метою якого є розроблення математичного та програмного забезпечення адаптації відеоконтенту до потреб осіб з вадами зору на основі тифлокоментування.

На основі проведеного аналізу встановлено, що для створення відеоконтенту з тифлокоментарем не існує програмних засобів, які б поєднували всі етапи побудови такого виду інформаційного контенту.

Для досягнення поставленої мети, сформульовано задачі дослідження під час розроблення програмно-алгоритмічного комплексу адаптації відеоконтенту для осіб з вадами зору.

У **другому розділі** представлено розроблення математичного забезпечення подання відеоконтенту для осіб з вадами зору на основі використання теорії координації систем (М.Месарович), спектрального аналізу. Побудовано архітектуру програмно-алгоритмічного комплексу адаптації відеоконтенту для осіб з вадами зору. Реалізовано метод пошуку вільних від

діалогів та інших важливих звуків місць у відеоконтенті. Ці знайдені ділянки відеоконтенту слугують для вставки тифлокоментарів.

Розробимо діаграму діяльності вибору тифлокоментаря для опису сюжету. Введемо такі позначення:

S_i – i -й сюжет;

$t(S_i)$ – тривалість i -го сюжету;

$Sem(S_i)$ – семантика i -го сюжету.

Будь-який сюжет S_i описано за допомогою множини тифлокоментарів $F_i = \{f_i^1, f_i^2, \dots, f_i^{n_i}\}$. Тоді, $t(f_i^j)$ – тривалість j -го тифлокоментаря, який описує i -й сюжет, а $Sem(f_i^j)$ – семантика j -го тифлокоментаря, який описує i -й сюжет.

Необхідно вибрати із множини F_i такий j -й тифлокоментар, для якого:

$$t(f_i^j) \leq t(S_i), \quad (1)$$

тоді відповідні семантики майже еквівалентні:

$$Sem(S_i) \cong Sem(f_i^j). \quad (2)$$

Якщо реалізація умови (1) очевидна (увагу лише треба звернути, щоб тифлокоментар не накладався на діалоги сюжету), то для реалізації умови (2) необхідні оцінки експертів.

Для того, щоб тифлокоментар не накладався на діалоги або інші важливі для розуміння сюжету звуки, розділено i -й сюжет на підсюжети, в межах яких немає діалогів: $S_i \supseteq S_1 \cup S_2 \cup \dots \cup S_{m_i}$. Тоді:

$$t(S_i) = t(S_1) + t(S_2) + \dots + t(S_{m_i}). \quad (3)$$

Повинна виконуватись множина умов:

$$t(f_k^j) \leq t(S_k), k = 1, 2, \dots, m_i. \quad (4)$$

З множини F_i вибирають ті елементи, для яких виконується умова (4). Отже, побудовано деяку підмножину тифлокоментарів $F_i \supseteq \tilde{F}_i = \{\tilde{f}_i^1, \tilde{f}_i^2, \dots, \tilde{f}_i^{n_i}\}$. З цієї підмножини вибираємо елемент, для якого виконується умова (2).

Вважаємо, що N експертів (E_1, E_2, \dots, E_N) оцінювали відповідність тифлокоментарів до семантики сюжету. Тоді, $O(E_i, \tilde{f}_j)$ – оцінка i -го експерта j -го тифлокоментаря. Що оцінка вища, то тифлокоментар адекватніше описує відповідний сюжет. Шкала оцінок скінченна. Позначення O_j – сумарна оцінка j -го тифлокоментаря:

$$O_j = \sum_{i=1}^N O(E_i, \tilde{f}_j).$$

Тоді вибираємо той тифлокоментар \tilde{f}_k , для якого:

$$k = \arg \max_j O_j.$$

Отже, отримуємо діаграму діяльності вибору тифлокоментаря для опису i -го сюжету, який наведено на рис. 1.

Процес вставки тифлокоментаря має ієрархічний характер, а саме ієрархію шарів (див. рис. 2). Для математичного забезпечення процесу тифлокоментування використано теорію координації. Розглядаємо дві підсистеми: підсистема верхнього рівня S^0 (тифлокоментування) та нижнього рівня S^1 (вибір місця для вставлення тифлокоментаря). Тобто, підсистема верхнього рівня визначає місце для вставки тифлокоментаря, а координатор запускає модуль начитування тифлокоментаря.



Рис. 1. Діаграма діяльності вибору тифлокоментаря для опису сюжету

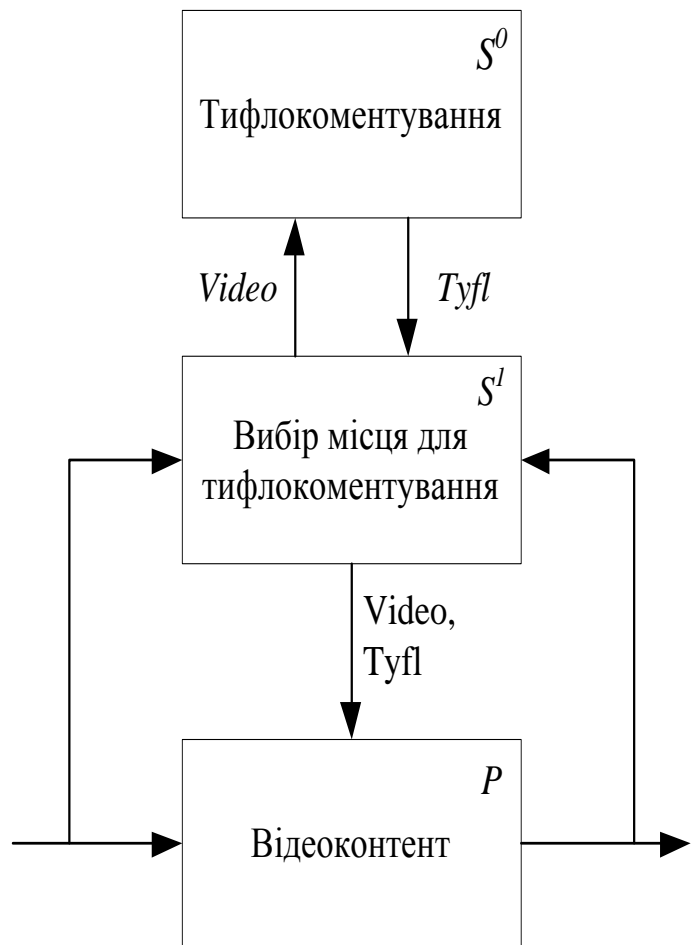


Рис. 2. Ієрархія шарів процесу тифлокоментування

Для пошуку у відеоконтенті ділянок, вільних від діалогів, використано спектральний аналіз. Процес пошуку таких ділянок використовує шість параметрів:

- 1) *порог енергії сигналу, $K, \%$* – це поріг енергії сигналу, в разі перевищення якого розпочинає роботу алгоритм пошуку. Цей параметр необхідний для того, щоб не пропускати шуми та зайві звуки, які тихіші за поріг, а також для

прискорення роботи алгоритму, оскільки для ділянок, які є нижчі від порогу, не виконується спектральний аналіз. Цей параметр бажано виставляти на рівні 15–40 %;

- 2) *рівень щільності спектра, %* – це поріг щільності спектра, перевищуючи який програмно-алгоритмічний комплекс (ПАК) переходить до наступного аналізу. Як виявив практичний аналіз, цей показник залежить від типу звукового сигналу. Для широкосмугових сигналів з рівномірним спектром (як, наприклад, для шуму чи насиченого музичного фрагмента) цей показник низький, для голосового сигналу, де розподілення частотних смуг нерівномірне, – цей параметр вищий. Тобто, цей параметр показує наявність явно виражених певних частотних смуг у спектрі відносно інших. На практиці цей параметр бажано виставляти на рівні 40–60 %. У разі встановлення на 0 % ПАК пропускає цей аналіз, тобто вважає, що весь звуковий сигнал задовольняє вимоги;
- 3) *коефіцієнт кореляції спектра, %* – це поріг кореляції спектра у кожній ділянці звукового сигналу до заданого еталонного спектра. У разі перевищення кореляції порогу ПАК вважає сигнал таким, що відповідає голосовому. На практиці бажано виставляти на рівні 40–70 %;
- 4) *фінальний поріг пошуку, %* – це поріг на результуючому графіку роботи ПАК. Коли сигнал нижчий від цього порогу, ПАК вважає відповідну ділянку тишею та заносить її до списку. Значення за замовчуванням – 50 %;
- 5) *мінімальна довжина тиші, мс* – це такий час, з яким порівнюється довжина знайденої ділянки тиші. Якщо тривалість ділянки тиші менша за цей параметр, то вона видаляється зі списку ділянок тиші;
- 6) *мінімальна довжина корисного сигналу, мс* – це такий час, з яким порівнюється інтервал між сусідніми ділянками тиші. Якщо цей інтервал менший за вказане значення, то ці дві ділянки тиші об'єднуються в одну, так інтервал між ними також стає у складі ділянки тиші. Це необхідно для того, щоб унеможливити раптові короткі перепади сигналу. Цей параметр бажано встановлювати на рівні 50–200 мс. Якщо ПАК вважає тишею дуже короткі слова, то потрібно цей параметр зменшити. Якщо є багато дуже коротких ділянок “корисного” сигналу між ділянками тиші, які насправді не мають корисної інформації, а лише якісь клацання чи шуми, то потрібно цей параметр збільшити.

Архітектура програмно-алгоритмічного комплексу (ПАК) адаптації відеоконтенту до осіб з вадами зору побудована у архітектурному шаблоні MVC, який використовується під час проектування та розробки програмного забезпечення. Архітектура ПАК складається з трьох рівнів – відображення, компонентів та керування даними (рис. 3). Рівень відображення забезпечує за взаємодію тифлокоментатора з ПАК. Рівень компонентів складається із модулів, які призначені для розв'язання задач введення, виведення та опрацювання даних. Компоненти керування даними разом з базами даних відеоконтентів належать до рівня керування даними.

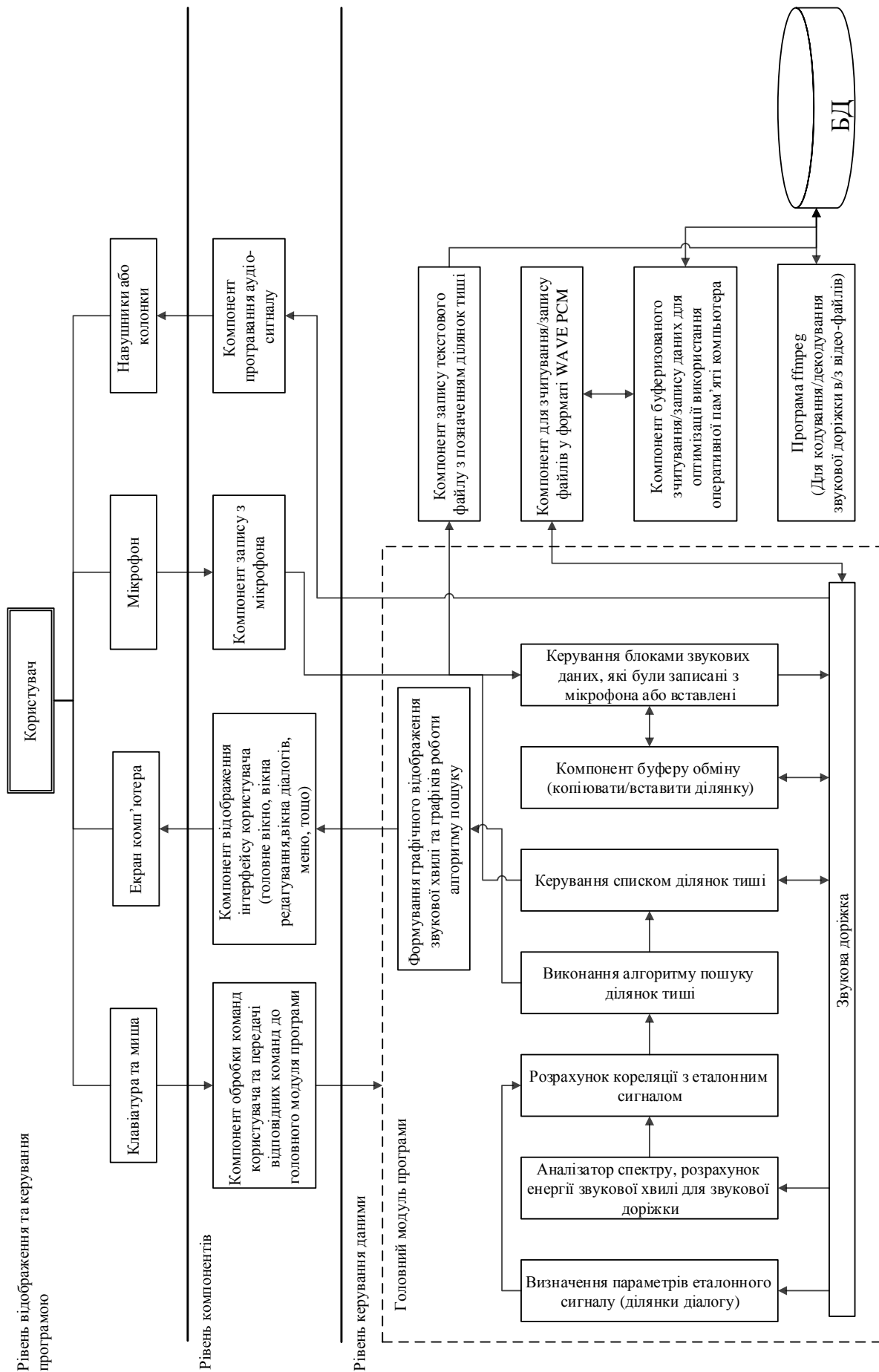


Рис. 3. Архітектура програмного комплексу адаптації відеоконтенту

У третьому розділі описано методи та алгоритми адаптування відеоконтенту шляхом застосування методу тифлокоментування. Наведено алгоритми пошуку місць у відеоконтенті для заповнення їх тифлокоментарями. На основі логіки 1-го порядку та асоціативних правил розроблено правила тифлокоментування для тифлокоментаторів, сценаристів та звукорежисерів адаптації відеоконтенту для осіб з вадами зору.

Розроблено алгоритм адаптації відеоконтенту до потреб осіб з вадами зору, який наведено на рис. 4.

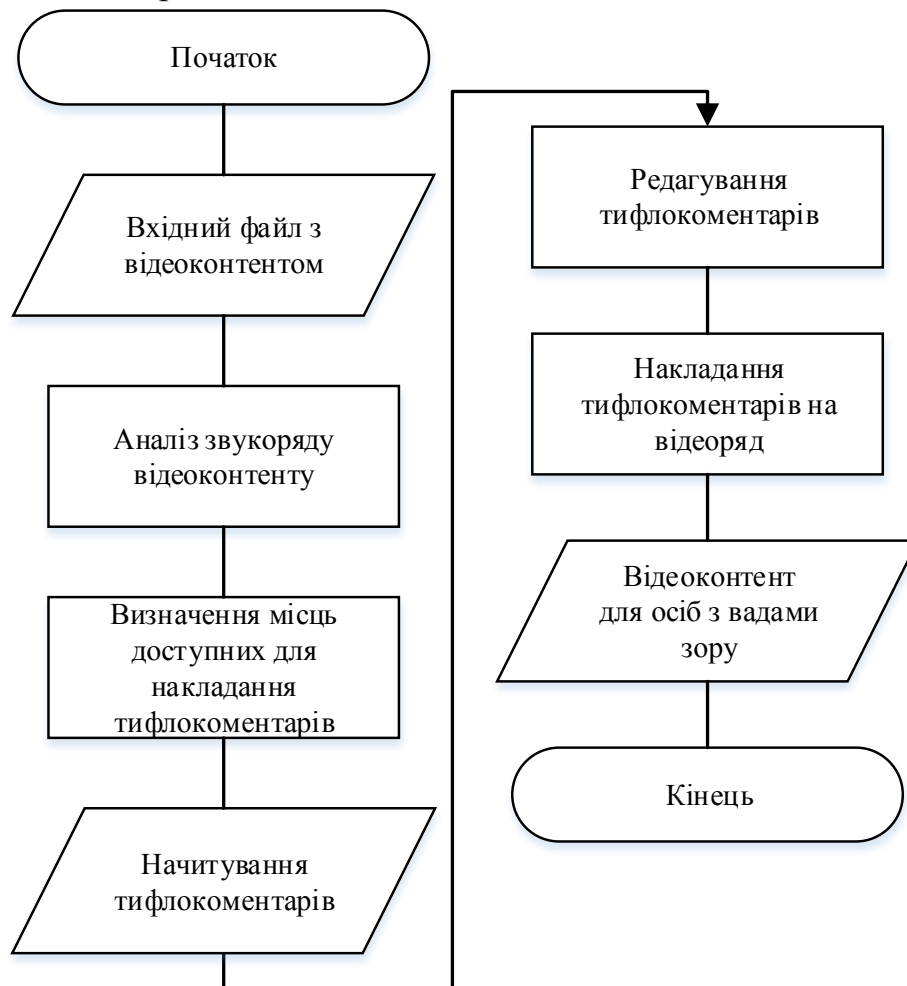


Рис. 4. Блок-схема алгоритму побудови відеоконтенту для осіб з вадами зору

Розроблено алгоритм пошуку у аудіоряді ділянок, які можна використати для накладання тифлокоментарів. Блок-схему алгоритму наведено на рис. 5. Звукова доріжка попередньо розбивається на відрізки завдовжки 10 мс. В оперативній пам'яті комп'ютера створюється масив відповідного розміру до кількості відрізків по 10 мс. Для кожної ділянки завдовжки 10 мс розраховується спектр завдяки перетворенню WaveLet. Для аналізу ПАК бере 1024 семпли звукових даних та рекурсивним методом WaveLet розподіляє на 10 частин довжинами 1, 2, 4, 8, 16, 32, 64, 128, 256 та 512 семплів. Кожна частина відповідає своїй смузі частот. Після цього для кожної частини розраховується енергія. Так отримуємо 10 значень рівня відповідно до частотних смуг для кожної ділянки звуку.

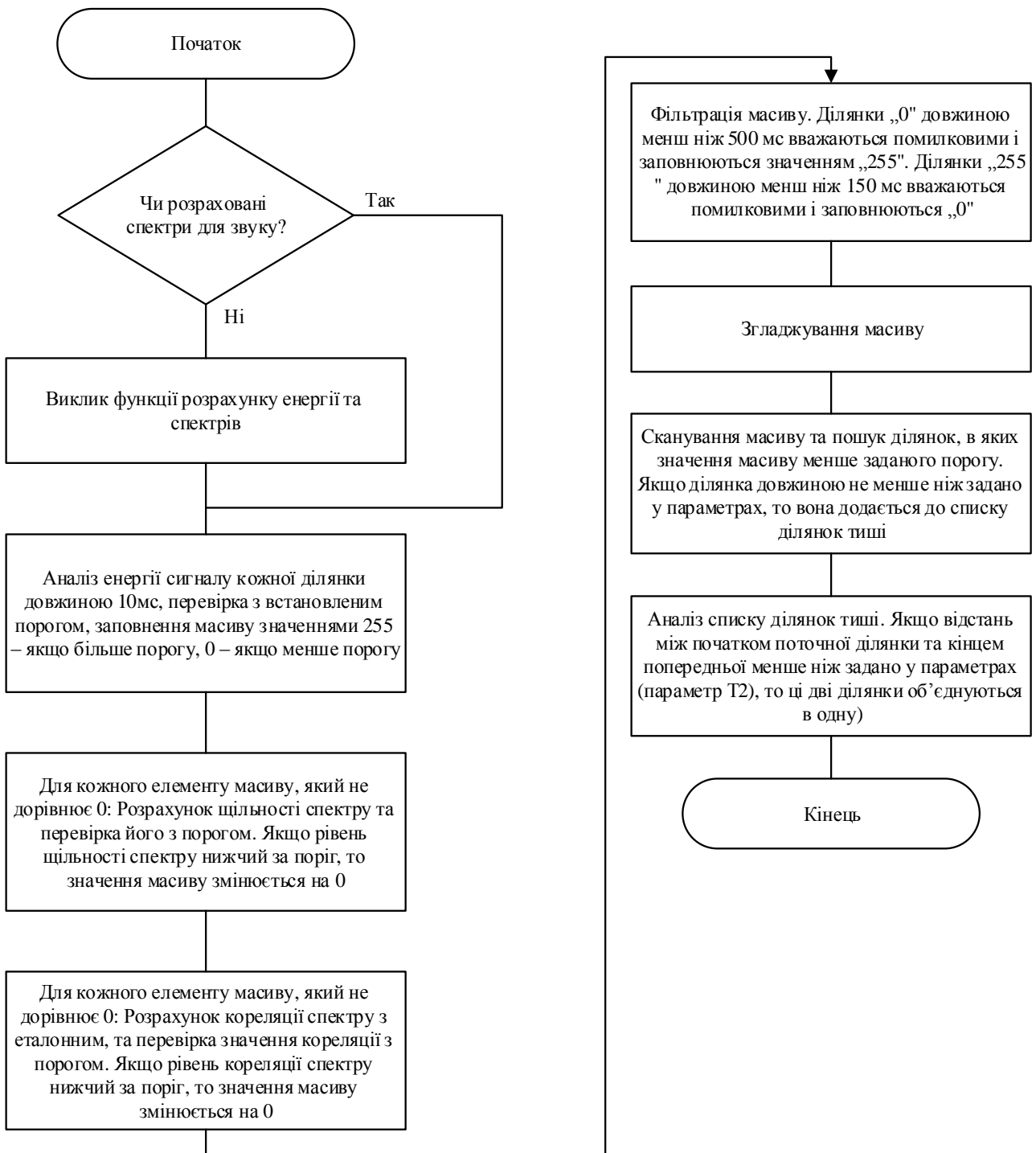


Рис. 5. Блок-схема алгоритму пошуку ділянок доступних для накладання тифлокоментарів

Далі сканується кожна ділянка звукової доріжки відповідно до таких чотирьох етапів:

Етап 1. Фільтрація ділянки звуку за допомогою низькочастотних (НЧ) та високочастотних (ВЧ) фільтрів, пропускаючи діапазон частот від 50 до 300 Гц. Розрахунок енергії отриманого сигналу та порівняння її рівня з прохідним порогом. Так виконується відсіювання явно зайвої інформації, шумів тощо. Якщо ділянка перевірена і значення енергії вище за поріг, то відповідному

елементу масиву присвоюється значення 255, якщо ні – то присвоюється значення 0.

Етап 2. Для кожної ділянки звукової доріжки, якщо відповідний їй елемент масиву дорівнює 255, виконується перевірка на щільність спектра. Щільність спектра розраховується за формулою $SMFdb = -\log_{10}(Gm/Am)$, де Gm та Am – це відповідно середнє геометричне та середнє арифметичне спектра відповідної ділянки.

Середнє арифметичне обчислюється за формулою $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, де n – кількість частотних смуг (у нашому випадку 10 смуг), x_i – значення енергії відповідної частотної смуги. Середнє геометричне відповідно за формулою:

$$G(x_1, x_2, \dots, x_n) = \sqrt[n]{x_1 x_2 \dots x_n} = \left(\prod_{i=1}^n x_i \right)^{1/n}, \text{ де } n \text{ – кількість частотних смуг.}$$

Якщо для конкретної ділянки значення щільності спектра вище за встановлений поріг, то в масиві відповідний елемент залишається таким, що дорівнює 255, а якщо щільність спектра не задовольняє вимогу, то значення масиву для цієї ділянки прирівнюється до 0.

Етап 3. Для кожної ділянки звукової доріжки, якщо відповідний їй елемент масиву дорівнює 255, виконується перевірка на кореляцію її спектра з еталонним спектром (заданим за замовчуванням у ПАК, або знятим користувачем з вибраної ділянки звукової доріжки). Розрахунок кореляції

виконується за формулою: $r_{xy} = \frac{\text{cov}_{xy}}{\sigma_x \sigma_y} = \frac{\sum (X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \sum (Y - \bar{Y})^2}}$, де x та y –

значення кожної частотної смуги для спектрів ділянки та еталону, відповідно.

Порівнюється розрахований коефіцієнт кореляції з встановленим порогом. Якщо перевірка відповідає встановленому значенню, то в масиві залишається значення 255, якщо ні – то відповідне значення масиву дорівнює 0.

Отже, у результаті проходження трьох етапів маємо масив, у якому кожний елемент дорівнює: 0 – якщо не відбулась хоча б одна перевірка; 255 – якщо ця ділянка відповідає всім трьом вимогам). На цьому можна було б завершити перевірку та розпочати пошук ділянок у масиві (елементи масиву, які дорівнюють 0). Але в результаті проходження трьох етапів можуть пройти перевірку короткі сигнали, які мали спектр, подібний до заданого, або навпаки, коротка пауза між словами актора може не бути помічена внаслідок перевірки. Наприклад, за енергією сигналу і в масиві для цієї ділянки буде значення 0. Тому, щоб вирішити цю проблему додано четвертий етап перевірки.

Етап 4. Для видалення раптових значень 0 або 255 виконується згладжування масиву. Спочатку виконується пошук ділянок масиву з значеннями "0", довжина яких не перевищує 500 мс. Такі ділянки вважаються помилковими, значення масиву в них змінюється на 255. Потім виконується

пошук ділянок масиву з значеннями “255”, довжина яких не перевищує 150 мс. Такі ділянки вважаються помилковими, та значення масиву в них змінюється на 0. Виконується згладжування у прямому напрямку, потім виконується згладжування у зворотному напрямку. Тож під час згладжування короткий сигнал (значення 255) дає значно менше значення для цього елемента масиву, але створює залишковий шум завдяки фільтру низької частоти. За зворотного напрямку значення елемента масиву для цієї ділянки є також значно нижчим за 255, і також наявний залишковий шум, але він має протилежний напрямок. Отже, під час розрахунку кореляції між цими двома згладженими масивами отримуємо такий ефект: короткі сигнали практично видаляються завдяки тому, що їхні залишкові шуми після НЧ-фільтрів не перекривались, а для корисного сигналу навпаки – згладжені в прямому і протилежному напрямку значення перекриваються і тому залишаються в масиві (навіть, якщо за рівнем енергії корисний сигнал був менший за короткий паразитний сигнал). Укінці етапу, кореляція між згладженими масивами порівнюється з порогом прийняття рішення, і якщо він не відповідає встановленому параметру, то цей елемент масиву дорівнює 0.

Також у цьому розділі розроблено правила тифлокоментування. З цією метою проведено опитування 100 осіб усіх вікових груп з вадами зору. Після перегляду адаптованого фільму з тифлокоментарями, їм задано десять запитань:

1. Чи запам'ятали Ви назву відеоконтенту?
2. Чи пригадаєте назву студії, яка зняла цей відеоконтент?
3. Чи запам'ятали Ви головних героїв відеоконтенту?
4. Чи зрозуміли Ви сам сюжет, його розвиток протягом відеоконтенту, розв'язку?
5. Чи не було моментів у відеоконтенті, коли Ви втратили зрозумілість того, що відбувається на екрані (незрозумілі звуки, які не були описані тифлокоментарем)?
6. Чи не було моментів, коли слова тифлокоментатора накладались на слова акторів, важливі звуки, і цим заважали сприйняттю відеоконтенту?
7. Чи зрозуміли Ви побут (місцевість, предмети) героїв, які були важливі для сприйняття відеоконтенту?
8. Чи не був тифлокоментар заплутаний, незрозумілий? Чи були терміни, які Ви не зрозуміли?
9. Чи не був тифлокоментар надмірним/недостатнім?
10. Чи задоволені Ви переглядом цього відеоконтенту з тифлокоментарями?

Для опрацювання результатів проведеного опитування використано теорію асоціативних правил. Наприклад, обчислено достовірність правила {ЯКЩО Ви зрозуміли сюжет, ТО чи Ви задоволені переглядом відеоконтенту з тифлокоментарем?}. Достовірність такого правила становить 95 %:

$$\text{conf} = \frac{95\%}{100\%} \cdot 100\% = 95\% .$$

Метою пошуку асоціативних правил є встановлення таких залежностей: якщо трапився деякий набір елементів P , то на підставі цього можна зробити висновок про те, що інший набір елементів Q також повинен трапитися з певною достовірністю $\forall x P(x) \rightarrow Q(x)$. На основі отриманих даних розроблено правила тифлокоментування.

Розроблено програмно-алгоритмічний комплекс який автоматизовано реалізує всі етапи побудови адаптованого відеоконтенту.

У **четвертому розділі** наведено результати практичного впровадження запропонованих методів та засобів адаптації відеоконтенту для осіб з вадами зору. Описано створення адаптованого відеоконтенту та налаштування алгоритму пошуку місць для вставлення тифлокоментарів. Апробовано результати реалізації програмно-алгоритмічного комплексу. Проведено порівняльний аналіз результатів експериментів.

Розроблено програмно-алгоритмічний комплекс “Audio Editor”, у якому використано запропоновані методи та алгоритми, які описані вище. Цей ПАК розв’язує такі задачі: 1) визначення місць у звукоряді відеоконтенту для вставлення тифлокоментарів; 2) начитування тифлокоментарів; 3) вибір з множини начитаних тифлокоментарів релевантного, який відповідає певним параметрам (час, семантика); 4) редагування та нормалізація звукової доріжки; 5) побудова паралельної звукової доріжки, яка складена з начитаних тифлокоментарів; 6) поєднання звукоряду з відеоконтентом. Результатом роботи ПАК є змонтована звукова доріжка, яку можна експортувати або в звуковий файл форматів WAV та MP3, або у формат відеофайлів AVI, MPEG, WMV. За допомогою цього ПАК адаптовано для осіб з вадами зору короткометражний фільм “Доторкнись і побач”, який знаходиться у вільному доступі в мережі Internet (www.youtube.com/watch?v=UBHv5jiPySY).

У табл. 1 наведено часові межі сюжетів кінофільму “Доторкнись і побач” та їхній опис за допомогою тифлокоментарів.

З метою апробації результатів тестування адаптованого відеоконтенту для осіб з вадами зору проведено два незалежні перегляди короткометражного фільму “Доторкнись і побач” для учнів та педагогів комунального закладу Львівської обласної ради (Львівської спеціальної загальноосвітньої школи-інтернату I–III ступенів № 100) та для Львівської обласного товариства сліпих (УТОС). Таким чином охоплено усі вікові групи (від 9-ти до 76-ти років).

Результати експерименту наведено на рис. 6. Ліворуч наведено відносну кількість схвальних відповідей на питання після перегляду фільму “Доторкнись і побач” без тифлокоментарів, а з правої сторони – після перегляду з тифлокоментарями. Були задані такі запитання:

- Чи зрозуміли Ви сам сюжет та його розвиток протягом відеоконтенту?
- Чи не було моментів у відеоконтенті, коли Ви втратили зрозумілість того, що відбувається на екрані (незрозумілі звуки, шуми)?
- Чи задоволені Ви переглядом цього відеоконтенту з тифлокоментарем?

Таблиця 1. Опис окремих сюжетів фільму “Доторкнись і побач” за допомогою тифлокоментарів.

Оригінал фільму	Часовий проміжок	Тифлокоментарі
0:00–1:10 (без будь-яких пояснень грає саундтрек)	0:00–0:08	титр: Доторкнись і побач.
	0:12–0:16	Дощовий вечірній Львів.
	0:17–0:25	У мерехтливих вогнях ліхтарів площі Ринок “пробігає” трамвай.
	0:32–0:40	Чоловік одягнутий у коричневу куртку з капюшоном йде вулицею.
	0:41–0:57	Під пазухою він несе коробку з подарунковою стрічкою.
	0:59–1:07	Необачно переходячи дорогу його збиває автомобіль. До нього вискакує водій.

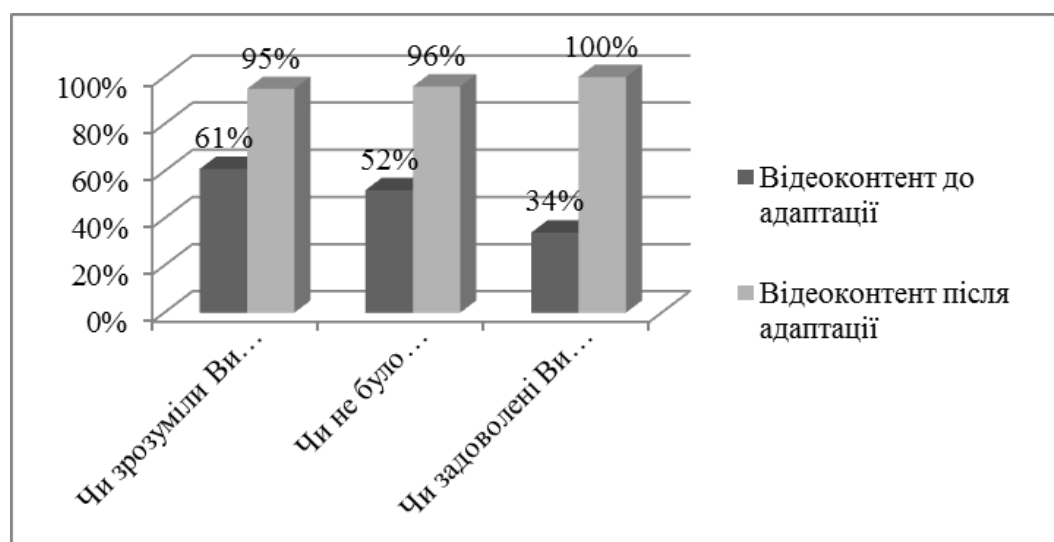


Рис. 6. Графік зіставлення відповідей глядачів

У такий спосіб експериментально доведено доцільність застосування методу тифлокоментування для опису сюжетів відеоконтенту, що дало змогу до 30 % підвищити сприйняття відеоконтенту особам з вадами зору.

ОСНОВНІ РЕЗУЛЬТАТИ ТА ВИСНОВКИ

У дисертаційній роботі розв’язано актуальне наукове завдання комп’ютерного опрацювання та подання відеоконтенту (художніх фільмів, мультиплікації, науково-популярних і навчальних матеріалів) у форматі, доступному для людей з вадами зору. Розроблено математичне та програмне забезпечення тифлокоментування відеоконтенту, що дає змогу створювати відеоконтент, адаптований для осіб з вадами зору.

Основні наукові та практичні результати дисертаційної роботи:

- 1) На основі онтологічного підходу визначено методи адаптації людей з різними типами нозологій до соціальних комунікацій, зокрема, доступу

людей з вадами зору до відеоконтенту, які ґрунтуються на процесі тифлокоментування.

- 2) Розроблено математичне забезпечення подання відеоконтенту для осіб з вадами зору на основі процесу тифлокоментування, яке ґрунтується на спектральному аналізі та координації побудови паралельної звукової доріжки, що дало змогу формалізувати процес адаптування відеоконтенту для осіб з вадами зору.
- 3) На основі теорії асоціативних правил та логіки першого порядку розроблено правила побудови тифлокоментарів, які слугують базою для сценаристів та звукорежисерів, що здійснюють процес адаптування відеоконтенту для потреб осіб з вадами зору.
- 4) Для автоматизації побудови відеоконтенту для осіб з вадами зору розроблено архітектуру та реалізовано програмно-алгоритмічний комплекс тифлокоментування, функціонування якого ґрунтується на пошуку місць у відеоконтенті для вставлення тифлокоментарів, начитуванні у часові межі відповідних тифлокоментарів, редагуванні і накладанні аудіоряду з тифлокоментарем на відео ряд.
- 5) Експериментально доведено доцільність застосування методів, алгоритмів та програмних засобів для опису сюжету відеоконтенту для осіб з вадами зору, що дає змогу до 30 % підвищити ефективність сприйняття відеоконтенту такими глядачами.

Під час виконання дисертаційного дослідження за допомогою розроблених методів і програмних засобів адаптовано та викладено у вільний доступ короткометражний фільм “Доторкнись і побач” (www.youtube.com/watch?v=UBHv5jiPySY).

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Демчук А. Б. Відеоконтент для незрячих: метод тифлокоментування / А. Б. Демчук // *Радіоелектроніка, інформатика, управління : наук. журн. Запор. нац. техн. ун-ту. – Запоріжжя : ЗНТУ, 2014. – № 1(30). – С. 146–149.*
2. Демчук А. Б. Підхід до розв’язування задачі вибору тифлокоментатора опису сюжету для людей з вадами зору / А. Б. Демчук, В. В. Литвин // *Інформаційні системи та мережі : вісн. Нац. ун-ту “Львівська політехніка”. – Львів, 2013. – № 770. – С. 138–142.*
3. Демчук А. Б. Основні прийоми і способи тифлокоментування / А. Б. Демчук // *Інформаційні системи та мережі : вісн. Нац. ун-ту “Львівська політехніка”. – Львів, 2014. – № 783. – С. 329–335.*
4. Демчук А. Б. Тифлокомментирование как метод адаптации видеоконтента для незрячих / А. Б. Демчук // *Новости науки Казахстана : междунар. науч.-техн. журн. – Алматы, 2014. – № 1 (119). – С. 136–143.*
5. Демчук А. Б. Информационные технологии для незрячих и незлышащих людей / О. В. Лозинская, А. Б. Демчук // *Наука и мир : междунар. науч. журн. – Волгоград, 2014. – № 1 (5). – С. 102–104.*

6. Литвин В. В. Метод автоматизованої розбудови та оцінювання якості онтологій баз знань / В. В. Литвин, М. Я. Гопяк, А. Б. Демчук // Автоматизовані системи управління та прилади автоматики: всеукраїнський міжвідомчий науково-технічний збірник. – Харків, 2012. – № 161. – С. 62–69.
7. Демчук А. Б. Метод побудови інтелектуального агента на основі онтології предметної області / А. Б. Демчук, В. В. Литвин, М. М. Войчишен // Інформаційні системи та мережі : вісник Нац ун-ту “Львівська політехніка”. – Львів, 2011. – № 715. – С. 215–225.
8. Демчук А. Б. Метод оцінювання якості онтологій баз знань інтелектуальних систем на основі iso/іес 25012. / А. Б. Демчук, М. Я. Гопяк // Інформаційні системи та мережі : вісн. Нац. ун-ту “Львівська політехніка”. – Львів, 2014. – № 805. – С. 80–86.
9. Демчук А. Б. Освітній відеоконтент для людей з вадами зору. Напрацювання та досвід Нац. ун-ту “Львівська політехніка” / А. Б. Демчук, В. В. Литвин / Львів : Видавництво Львівської політехніки, 2013. – 20 с.
10. Демчук А. Б. Метод вибору тифлокоментарів опису фільмів для людей з вадами зору / А. Б. Демчук, В. В. Литвин // Сучасні проблеми математичного моделювання та обчислювальних методів : матер. Всеукраїнської наукової конференції. – Рівне, 2013. – С. 96.
11. Гопяк М. Я. Використання онтологій баз знань в інтелектуальних системах / М. Я. Гопяк, В. В. Литвин, А. Б. Демчук, Т. І. Черна // Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту : матер. Міжнар. наук. конф. – Херсон : ХНТУ, 2013. – С. 352–353.
12. Демчук А. Б. Тифлокоментування: зародження та перспективи розвитку / А. Б. Демчук, О. В. Лозинська // Наука как основа мирного диалога : матер. Міжнар. конф. – Донецьк, 2014. – С. 17-20.
13. Демчук А. Б. Основні прийоми і способи створення тифлокоментарів / А. Б. Демчук, О. В. Лозинська, М. Я. Гопяк // Інформаційні технології, економіка та право: стан та перспективи розвитку : матер. Міжнар. конф. – Чернівці : ПВНЗ “Буковинський університет”, 2014. – С. 4-5.
14. Демчук А. Б. Основи тифлокоментування / А. Б. Демчук // Наука та сучасність: виклики ХХІ ст. : матер. Міжнар. конф. – Київ, 2014. – С. 27–30.
15. Демчук А. Б. Подання відеоконтенту незрячим / А. Б. Демчук // Проблеми та перспективи розвитку науки у країнах СНД : матер. ХІХ Міжнар. наук.-практ. інтернет-конф. – Переяслав-Хмельницький : Переяслав-Хмельницький державний педагогічний університет ім. Г. Сковороди, 2014. – С. 249–251.
16. Литвин В. В. Загальний підхід до розроблення інтелектуальних систем підтримки прийняття рішень на основі адаптивних онтологій / В. В. Литвин, Р. О. Голощук, А. С. Мельник, А. Б. Демчук, М. М. Войчишен // Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту : матер. Міжнар. наук. конф. – Євпаторія, 2011. – Т. 1. – С. 184–188.
17. Литвин В. В. Використання зважених онтологій в інтелектуальних системах управління військами / В. В. Литвин, А. Б. Демчук, М. Л. Бойчук // Сучасні

- проблеми прикладної математики та інформатики : тези XVIII Всеукр. наук. конф. – Львів : Львів нац. ун-ту ім. І. Франка, 2012. – С. 102.
18. Литвин В. В. Підхід до побудови інтелектуальної системи якості трубопроводів на основі онтології / В. В. Литвин, Н. Б. Шаховська, А. Б. Демчук, М. М. Войчишен // Обчислювальний інтелект : матер. I Міжнар. наук.-техн. конф. – Черкаси, 2011. – С. 305.
 19. Лозинська О. В. Використання дерев синтаксичного розбору для перекладу української жестової мови / О. В. Лозинська, А. Б. Демчук, М. В. Давидов // Інформаційні управляючі системи та технології : матер. Міжнар. наук.-практ. конф. – Одеса, 2013. – С. 321–324.
 20. Демчук А. Б. Багаторівневі інтелектуальні інформаційні системи опрацювання відеоконтенту для осіб з вадами зору / А. Б. Демчук, Р. В. Вовнянка, М. Я. Гопяк // Інформація, комунікація, суспільство – 2014 : матер. III Міжнар. наук. конф. “ІКС-2014” / Нац. ун-т “Львівська політехніка”, Кафедра соціальних комунікацій та інформаційної діяльності, кафедра інформаційних систем та мереж. – Львів : Видавництво Львівської політехніки, 2014. – С. 168–169.

АНОТАЦІЇ

Демчук А. Б. Математичне та програмне забезпечення подання відеоконтенту для осіб з вадами зору. – На правах рукопису.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 01.05.03 – математичне та програмне забезпечення обчислювальних машин і систем. – Національний університет “Львівська політехніка”, Міністерства освіти і науки України, Львів, 2015.

Дисертація присвячена розробленню програмно-алгоритмічного комплексу адаптації відеоконтенту для осіб з вадами зору, зокрема вдосконаленню методу тифлокоментування за рахунок пошуку вільних від діалогів та важливих шумів місць у звукоряді відеоконтенту.

Розроблено математичне забезпечення подання відеоконтенту для осіб з вадами зору на основі використання теорії координат, спектрального аналізу. Побудовано архітектуру програмно-алгоритмічного комплексу адаптації відеоконтенту для осіб з вадами зору. Реалізовано метод пошуку вільних місць у відеоконтенті, які слугують для вставки тифлокоментарів. На основі логіки 1-го порядку та асоціативних правил розроблено правила тифлокоментування для тифлокоментаторів, сценаристів та звукорежисерів.

Розв’язано актуальне наукове завдання подання відеоконтенту у форматі, доступному для сприйняття особам з вадами зору. Наукова новизна одержаних результатів полягає у обґрунтуванні архітектури та практичній реалізації програмно-алгоритмічного комплексу, призначеного для побудови, редагування та адаптування відеоконтенту для осіб з вадами зору. Проведено апробацію результатів використання програмного комплексу.

Експериментально доведено доцільність застосування методу тифлокоментування для опису сюжетів відеоконтенту, що дало змогу до 30 % підвищити сприйняття відеоконтенту особам з вадами зору.

Основні результати роботи використовуються у Львівському обласному товаристві сліпих (УТОС), Львівській спеціальній школі №100 та у навчальному процесі НУ “Львівська політехніка”.

Ключові слова: програмно-алгоритмічний комплекс, відеоконтент, особа з вадою зору, тифлокоментування, тифлокоментар, аудіодескрипція.

Демчук А. Б. Математическое и программное обеспечение представления видеоконтента для лиц с нарушениями зрения. – На правах рукописи.

Диссертация на соискание ученой степени кандидата технических наук по специальности 01.05.03 – математическое и программное обеспечение вычислительных машин и систем. – Национальный университет “Львівська політехніка” Министерства образования и науки Украины, Львов, 2015.

Диссертация посвящена разработке программно-алгоритмического комплекса адаптации видеоконтента для лиц с недостатками зрения, в частности совершенствованию метода тифлокомментирования за счет поиска свободных от диалогов и важных шумов мест в звукоряде видеоконтента.

Разработано математическое обеспечение представления видеоконтента для лиц с недостатками зрения на основе использования теории координации, спектрального анализа. Построено архитектуру программно-алгоритмического комплекса адаптации видеоконтента для лиц с недостатками зрения. Реализован метод поиска свободных мест в видеоконтенте, которые служат для вставки тифлокомментариев. На основе логики 1-го порядка и ассоциативных правил разработаны правила тифлокомментирования для тифлокомментаторов, сценаристов и звукорежиссеров.

Решено актуальное научное задание представления видеоконтента в формате доступном для восприятия лицам с недостатками зрения. Научная новизна исследования заключается в обосновании архитектуры и практической реализации программно-алгоритмического комплекса, предназначенного для построения, редактирования и адаптации видеоконтента для лиц с недостатками зрения.

Экспериментально доказана целесообразность применения метода тифлокомментирования для описания сюжетов видеоконтента, что позволило до 30 % повысить восприятие видеоконтента лицам с недостатками зрения.

Основные результаты работы используются в Львовском областном обществе слепых (УТОС), Львовской специальной школе №100 и в учебном процессе НУ "Львовская политехника".

Ключевые слова: программно-алгоритмический комплекс, видеоконтент, лица с нарушениями зрения, тифлокомментирование, тифлокомментарий, аудіодескрипція.

Demchuk A. B. Mathematical and software submission of video content for people with visually impaired. – Manuscript.

Thesis for Ph.D. degree in technical sciences in specialty 01.05.03 – mathematical and software support of computer machines and systems. – Lviv Polytechnic National University, Ministry of Education and Science of Ukraine, Lviv, 2015.

One of the most urgent problems of people with visual impairments is an almost complete unavailability of video content. To adapt video content to the needs of visually impaired is selected method of audiodescription.

The thesis is devoted to the development of program and algorithmic complex of video content adaptation for people with visual impairments, including improvement of audio description method by searching places free of dialogues and important noises in the video content scale.

The mathematical support of video content representation for visually impaired people based on the theory of coordination for the construction of additional track audiodescription comments and spectral analysis for search available include video to insert audiodescription comments is developed. The architecture of program and algorithmic complex of video content adaptation for people with visual impairments by using the unified modeling language UML is created. The method of searching places free of dialogues and important noises in the video content scale, which serve to insert typhlocomments is realised. The rules of audiodescription comments for typhlocommentators, scriptwriters and sound engineers based on the logic of the 1st order and associative rules is developed. Used object-oriented and sampled programming - to implement complex algorithmic software and adapt video content for people with visual impairments.

The actual scientific problem of video content submission in the format accessible for perception of people with visual impairments is solved. Scientific novelty of the obtained results lies in the argumentation of architecture and practical implementation of program and algorithmic complex, appointed to build, edit and adapt video content for people with visual impairments. The approbation of program complex implementation results is made.

With the purpose test the results of testing video content adapted for people with visual impairments views held two independent short film "Touch and see".

The feasibility of the typhlocomments method for the video content plot description, that allows to increase up to 30 % the perception of video content by people with visual impairments is experimentally proved.

Main results of this work are already used in Lviv Boarding School for Blind Children №100, Lviv Regional Society for the Blind people (UTOS) and in educational process Lviv Polytechnic National University.

Key words: applied programmed algorithmic complex, video content, people with visual impairments, typhlocomments, typhlocommentar, audiodescription.