

**Міністерство освіти і науки, молоді та спорту України  
Харківський національний університет радіоелектроніки**

**ЩЕРБІНІН КОСТЯНТИН СЕРГІЙОВИЧ**

**УДК 004.932.2:004.93'14**

**МЕТОДИ ТА МОДЕЛІ СЕГМЕНТАЦІЇ ВІДЕО  
ДЛЯ СИСТЕМ МЕТРИЧНОГО ПОШУКУ**

**05.13.23 – системи та засоби штучного інтелекту**

**Автореферат**

**дисертації на здобуття наукового ступеня**

**кандидата технічних наук**

**Харків – 2011**

Дисертацією є рукопис.

Робота виконана в Харківському національному університеті радіоелектроніки Міністерства освіти і науки, молоді та спорту України.

**Науковий керівник** доктор технічних наук, професор  
**Машталір Володимир Петрович,**  
Харківський національний університет  
радіоелектроніки, декан факультету  
комп'ютерних наук

**Офіційні опоненти:** доктор технічних наук, професор  
**Антощук Світлана Григорівна,**  
Одеський національний політехнічний університет,  
завідуюча кафедрою інформаційних систем

доктор технічних наук, професор  
**Кучеренко Євгеній Іванович,**  
Харківський національний університет радіоелектроніки, професор кафедри штучного інтелекту

Захист відбудеться «\_\_\_»\_\_\_\_\_ 20\_\_ р. о \_\_\_\_\_ годині на засіданні спеціалізованої вченої ради Д 64.052.01 у Харківському національному університеті радіоелектроніки за адресою: 61166, м. Харків, пр. Леніна, 14.

З дисертацією можна ознайомитись у бібліотеці Харківського національного університету радіоелектроніки за адресою: 61166, м. Харків, пр. Леніна, 14.

Автореферат розісланий «\_\_\_»\_\_\_\_\_ 2011 р.

Вчений секретар  
спеціалізованої вченої ради

С.Ф. Чалий

## ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

**Актуальність теми.** Диверсифікація використання мультимедіа і гіпермедіа в різноманітних видах людської діяльності акцентує увагу розробників на вдосконаленні засобів реєстрації, обробки, аналізу, зберігання та пошуку у великих колекціях візуальної інформації. Тому, перш за все, постають завдання семантичної обробки відеоданих. Набутий досвід у зоні цифрової обробки й автоматичної предметно-орієнтованої інтерпретації статичних зображень, з одного боку, зумовлює базис створення та впровадження перспективних моделей і методів обробки відео, з іншого, є певною обструкцією, оскільки облік динамічних властивостей послідовності відеокадрів потребує синтезу принципово нових моделей з урахуванням зміни інформації в часі. Певна річ, що розумне поєднання обох підходів створює передумови для виникнення нових ефективних методів усунення семантичного конфлікту між результатами обробки алгоритмами низького рівня і перцептуальним трактуванням цих результатів. Валідний пошук відеоданих за змістовними критеріями, зокрема із запитом 'ad exemplum', забезпечується розвитком методів і моделей індексації відеоданих, синтезом мір і метрик для порівняння подій, сюжетів, сценаріїв, реалізацією релевантного зворотного зв'язку з метою ітераційного уточнення інформаційних потреб Користувача.

Міждисциплінарна багатofакторна інтеграція методів розпізнавання образів, комп'ютерного зору, керування базами даних, аналізу багатовимірних часових рядів, штучного інтелекту в цілому — шлях створення подібних систем, які зазвичай називаються CBVIR (Content Based Video Information Retrieval). Основне значення для розвитку методів обробки візуальної інформації, що визначають інструментальні засоби CBVIR, мають наукові розробки українських та зарубіжних вчених, серед яких найбільш впливовий внесок зробили С.Г. Антошук, А.М. Ахметшин, Є.В. Бодянський, Р.А. Воробель, В.Н. Крилов, Є.П. Путятін, М.І. Шлезінгер, В.К.Р. Horn, S. Mann, O. Marques, D.C. Montgomery, M. Petkovic, P. Zezula та ін.

Аналіз стану і тенденцій розвитку методів пошуку візуальної інформації із запитом «за зразком» надає змогу стверджувати, що, незважаючи на численні дослідження в цьому напрямі, зростання обсягів накопичуваних відеоданих і інтенсифікації їх використання потребує створення нових швидкодіючих засобів валідного пошуку. Доцільним та розумним компромісом у цьому разі можуть бути системи метричного пошуку, що забезпечують таку парадигму, при якій спочатку здійснюється пошук завідомо непотрібних відеосюжетів і лише серед тих, що залишилися, виконується детальний пошук. Певну перспективу має проведення теоретико-експериментальних досліджень з розробки моделей і методів забезпечення підвищеної швидкодії метричного пошуку на основі двоетапної сегментації відео: у часі — для отримання однорідних (у широ-

кому сенсі) відеосегментів та в просторі (у полі зору) — для отримання багатовимірних часових рядів, аналіз яких і створює передумови для індексації відео.

Таким чином, вивчення методів сегментації багатовимірних часових рядів, індукованих сегментацією окремих відеокадрів, а також виявлення необхідних і достатніх умов, які дозволяють істотно прискорити пошук, є актуальним завданням, що забезпечує можливість створення ефективних CBVIR.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційна робота виконана на кафедрі інформатики Харківського національного університету радіоелектроніки в рамках держбюджетних науково-дослідних робіт: «Дослідження і розробка методів аналізу зображень в умовах складних перетворень» (№ ДР 0103U001572, виконавець), «Розробка моделей і методів факторизації інформації в умовах багатозначності», (№ ДР 0104U003432, виконавець), «Інтелектуальний аналіз і обробка даних у реальному часі на основі засобів обчислювального інтелекту» (№ ДР 0104U003432, виконавець), «Синтез методів обробки інформації в умовах невизначеності на основі самонавчання і м'яких обчислень» (№ ДР 0107U003028, виконавець), які здійснювалися відповідно до наказів Міністерства освіти і науки, молоді та спорту України за результатами конкурсного відбору науково-дослідних робіт. У рамках зазначених робіт здобувачем отримані методи сегментації багатовимірних часових рядів в умовах невизначеності, моделі підвищення швидкодії пошуку через унеможливлення повного перебору.

**Мета і задачі дослідження.** Метою дисертаційної роботи є розробка моделей і методів сегментації відеоданих, обробки, аналізу й індексації результатів сегментації в часі та просторі для підвищення ефективності систем метричного пошуку і розпізнавання динамічної візуальної інформації в базах даних із запитом за зразком.

Досягнення поставленої мети потребувало вирішення таких задач:

- аналіз відомих методів обробки зображень і відеопослідовностей, трендів їх розвитку для виявлення інструментальних засобів CBVIR, які найбільше впливають на ефективність в аспектах складності обчислювальних моделей і обліку семантики;
- синтез методів адаптивного виявлення в умовах апіорної невизначеності змін властивостей багатовимірних часових рядів, індукованих ознаками окремих відеокадрів, для пошуку в послідовному режимі семантично однорідних сюжетів;
- обґрунтування і дослідження необхідних і достатніх умов елімінації відеосюжетів під час метричного пошуку без трудомісткого порівняння відеосегментів зі зразком-запитом;
- дослідження специфіки використання моделей і методів, які розроблялися, для пошуку відеопослідовностей у базах даних, створення і впровадження дослідних і спеціалізованих програмних засобів.

*Об'єкт дослідження* — процеси сприйняття і контекстна інтерпретація динамічних просторових сцен.

*Предмет дослідження* — методи сегментації динамічної візуальної інформації для метричної індексації і пошуку у відеоколекціях.

*Методи дослідження* — розробка і дослідження моделей сегментації і індексації відеоданих потребували використання основних положень математичного апарату розпізнавання образів, штучних нейронних мереж, теорії множин, алгебри, елементів математичної статистики.

**Наукова новизна отриманих результатів.** Наукова новизна дисертації полягає у постановці і розв'язанні задачі сегментації та індексації відеопослідовностей для швидкого пошуку у великих відеоколекціях із запитом за зразком:

– *уперше запропоновані методи* виявлення зміни властивостей багатовимірних часових рядів на основі багатовимірної модифікації процедури експоненціального згладжування у формі адаптивного контролю характеристик, що створює передумови для виявлення відеосегментів у процесі послідовного аналізу динамічної візуальної інформації;

– *набули подальшого розвитку методи* метричного пошуку, які, на відміну від відомих, враховують у загальному вигляді необхідні і достатні умови елімінації відеосегментів без порівняння із запитом, що при довільній кількості опорних точок у фазовому просторі забезпечує суттєве підвищення швидкодії;

– *удосконалено методи* усунення семантичного конфлікту між низькорівневою обробкою відеоданих і цілями перцептивного сприйняття, які, на відміну від відомих, враховують «просторовий зміст» у вигляді результатів сегментації окремих кадрів відеопослідовностей, що створює достатні підстави для валідного пошуку відеосегментів у ряді предметно-орієнтованих зон.

**Практичне значення отриманих результатів.** Розроблені методи і моделі сегментації відеоданих, знайдені умови елімінації відеосегментів, отримано уявлення «просторового змісту» відеокадрів і відеоряду в цілому можуть використовуватися для вирішення широкого кола завдань пошуку й інтерпретації візуальної інформації. Використання цих результатів дослідження забезпечило стійку автоматичну сегментацію відеоданих і зображень, що дозволило здійснити реалізацію у вигляді підсистем CBVIR з метричним пошуком, при якому відеосегменти порівнюються з ключовими кадрами. Синтезовані методи підтвердили свою ефективність при використанні їх у системах медичної діагностики, зокрема, в процесі пошуку в ендоскопічних відеоархівах Інституту загальної та невідкладної хірургії АМН України (акт про впровадження від 19.05.2011 р.) та під час вирішення пошукових завдань у відеоархівах коронарної ангіографії Центральної клінічної лікарні «Укрзалізниця» (акт про впровадження від 29.03.2011 р.). Наукові положення, що викладені в дисертації, використовувалися в навчальному процесі в Харківському національному універси-

теті радіоелектроніки (акт про впровадження від 11.04.2011 р.). Результати дисертаційних досліджень реалізовані у вигляді комплексів програмних засобів для баз відеоданих, один з яких сертифікований системою УкрСЕПРО.

**Особистий внесок здобувача.** Всі положення, що виносяться на захист, отримані здобувачем особисто. У роботах, написаних у співавторстві, авторові належать постановка задач, вибір та обґрунтування методів їх вирішення, а саме: [2, 3, 8, 14] запропоновані і досліджені методи і моделі сегментації багатовимірних часових рядів, що відповідають відеопослідовностям; [4, 6, 11] отримані, обґрунтовані та досліджені умови вилучення з аналізу об'єктів під час метричного пошуку без обчислення відстані до них; [1, 9, 5, 12] формалізовані і досліджені методи сегментації відеокادрів, що забезпечують синтез багатовимірних часових рядів; [7, 10, 13] конкретизовані і вивчені ознакові простори, які являють собою «просторовий зміст» візуальної інформації.

**Апробація результатів дисертації.** Основні результати роботи доповідалися, обговорювалися і були схвалені на науково-технічних конференціях і семінарах: 14-th International Conference in Central Europe for Computer Graphics, Visualization and Computer Vision (Plzen, Czech Republic, January 30 – February 3, 2006), 4-th International Workshop Adaptive Multimedia Retrieval (Geneva, Switzerland, July 27–28, 2006); Fifth International Conference on Information Research and Applications (Varna, Bulgaria, June 26–30, 2007); міжнародній конференції «Інтелектуальні системи прийняття рішень і прикладні аспекти інформаційних технологій» (Євпаторія, 14–18 травня 2007, Євпаторія, 19–23 травня 2008); 15-th Zittau East-West Fuzzy Colloquium (Zittau, Germany, September 17–19, 2008); міжнародній науковій конференції «Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту» (Євпаторія, 18–22 травня 2009, Євпаторія, 17–21 травня 2010).

**Публікації.** Основні положення дисертації відображені в 14 роботах, серед яких: 4 статті у виданнях, що входять до переліку ВАК України, 2 статті в зарубіжних журналах, 8 публікацій у працях і матеріалах міжнародних конференцій, з них 4 — за кордоном.

**Структура дисертації.** Дисертація складається зі вступу, чотирьох розділів, висновків, додатка та списку використаних джерел. Повний обсяг дисертації становлять 154 сторінки; обсяг основного тексту — 130 сторінок; робота містить 34 рисунки (з них 3 — на окремих сторінках); 5 таблиць (з них 4 — на 7 окремих сторінках); додаток — на 4 сторінках; список використаних джерел містить 136 найменувань — 14 сторінок.

## ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** обґрунтовано актуальність теми дисертаційної роботи, сформульовано мету та задачі дослідження, розкрито наукову та практичну цінність

отриманих результатів. Наведено відомості про публікації та апробацію роботи.

**У першому розділі** проаналізовано стан проблеми пошуку і розпізнавання зображень у базах даних і розглянуто підходи, спрямовані на поліпшення якості та зменшення витрат часу на пошук.

З'ясовано, що диверсифікація мультимедіа довела необхідність розвитку методів організації та впорядкування сховищ візуальної інформації з метою створення ефективних методів пошуку за змістовними критеріями спочатку зображень, а потім, як природний розвиток, відеоданих. Основна парадигма такого пошуку (за асоціацією, за категоріями, цільового) — синтез автоматичного опису для індексації і структурування вмісту баз відеоданих власне зі змісту візуальної інформації. Під «змістом» розуміється деякий набір ознак, добутих із відеопотоку або його частин. У такому разі при задаванні запиту у вигляді зразка пошук виконується в ознаковому просторі, зазвичай, із заданим порогом подібності. Однак виникає семантичний конфлікт, який потребує усунення, між низькорівневими моделями відеосегментів в ознакових просторах і рівнем моделювання сприйняття візуальної інформації людиною в заданому контексті. Проміжним етапом між оцінюванням подібності ознакових описів і аналізом високорівневих концепцій сприйняття відео є «просторовий» зміст відеокадрів, який наданий у вигляді розбиття поля зору. Таке уявлення індукує багатовимірні часові ряди, які з достатнім ступенем точності моделюють відеопотоки.

Ретроспективний аналіз існуючих систем пошуку засвідчив, що за відсутності достовірних метаданих інтелектуальний аналіз візуального (в нашому випадку просторового) змісту запиту-зразка в більшості випадків є найефективнішим. У цьому сенсі методи метричної індексації, основані на аксіомі нерівності трикутника, дозволяють скорочувати кількість операцій порівняння на етапі реалізації запиту. Для цього потрібна подвійна сегментація: у часі — для виділення однорідних послідовностей відеокадрів і в полі зору — для побудови зон, що корелюються з об'єктами сцен, які спостерігаються. У розділі проаналізовано потенційні підходи до побудови ознакових просторів.

Доведено, що до нині для обробки часових рядів склався ряд відносно незалежних ефективних підходів, серед яких, у першу чергу, необхідно відзначити статистичний (кореляційний, спектральний, регресійний аналізи), адаптивний (експоненціальне згладжування; аналіз слідкувального сигналу) і оснований на методах класифікації (кластеризація, сегментація). І якщо для одновимірних рядів розроблені ефективні алгоритми аналізу, то результати, пов'язані з багатовимірними нестационарними послідовностями, потребують відповідного розвитку в плані побудови розбиття тимчасових рядів на семантично гомогенні сегменти і виявлення моментів переходу від одного сегмента до іншого.

Аналіз поточного стану методів і моделей семантичного пошуку свідчить, що перспективним підходом є реалізація пошуку за зразком з двоетапною (у часі і в полі зору) сегментацією відео. Існуючі методи сегментації відео доці-

льно розвивати відповідно до створення процедур виявлення зміни властивостей багатовимірних часових рядів в умовах невизначеності. Одним з перспективних підходів до скорочення часових витрат пошуку із запитами ‘ad exemplum’ є вивчення необхідних і достатніх умов, які дозволяють на етапі пошуку вилучати з розгляду цілі сімейства невідповідних до запиту відеосегментів без трудомісткого обчислення відстаней до них.

У розділі сформульовані цілі та задачі дисертаційних досліджень.

У **другому розділі** розвивається послідовний підхід до сегментації, при якому прийняття рішень проводиться покадрово в on-line режимі завдяки аналізу багатовимірних нестационарних часових рядів з апіорі невідомими характеристиками.

Нехай відеопослідовність задана в деякому ознаковому просторі багатовимірним тимчасовим рядом  $x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T$ , де  $k = \overline{1, N}$  — дискретний час. Сегмент ряду  $S(a, b) = \{a \leq k \leq b, a, b \in \mathbb{N}\}$  — це статистично однорідна послідовність  $x(a), x(a+1), \dots, x(b)$ , а завдання часової сегментації — це пошук  $c$  непересічних сегментів  $S^c = \{S_i(a_i, b_i), 1 \leq i \leq c\}$ , таких, де  $a_1 = 1, b_c = N, a_i = b_{i-1} + 1$ , тобто необхідно знайти  $c$  інтервалів  $S_1 < S_a < \dots < S_c$  та їх межі  $a_i, b_i$ .

Для послідовного аналізу відеопотоку введена модель

$$x(k) = BX(k) + \xi(k), \quad (1)$$

де  $B = (B_0 : B_1 : \dots : B_p) - (n \times (pn + 1))$  — матриця;

$X(k) = (1, x^T(k-1), \dots, x^T(k-p))^T$  — вектор розмірності  $((pn + 1) \times 1)$ ;

$B_0 = \{b_{0i}\} - (n \times 1)$  — вектор невідомих середніх значень;

$B_l = \{b_{lij}\} - (n \times n)$  — матриці параметрів, що підлягають визначенню;

$l = \overline{1, p}$ ;  $p$  — порядок моделі;

$\xi(k)$  — випадкові збурення з обмеженим другим моментом.

Матриця апіорі невідомих параметрів  $B$  містить практично всю необхідну інформацію про властивості відеоданих.

Ряду (1) відповідає модель, що настроюється

$$\hat{x}(k) = B(k-1)X(k), \quad (2)$$

матриця параметрів якої уточнюється під час аналізу кожного  $k$ -го відеокадру завдяки мінімізації заданого критерію — деякої функції відмінності обчислених  $\hat{x}(k)$  і реальних  $x(k)$ .

Синтезована модель (2) повинна бути працездатною і в режимі прогнозування, а порушення прогнозованих властивостей — ознака виникнення розладів, які аналізуються за допомогою рекурентних процедур навчання виду



$$\begin{cases} B(k) = B(k-1) + \gamma(k)e(k)X^T(k); \\ e(k) = x(k) - \hat{x}(k) = x(k) - B(k-1)X(k), \end{cases}$$

де  $\gamma(k)$  — скалярний або матричний коефіцієнт посилення алгоритму, що визначає його властивості і залежить від прийнятого критерію ідентифікації;  
 $e(k)$  — векторна помилка оцінювання.

Зазвичай, використовуються алгоритми, пов'язані з критерієм мінімуму суми квадратів помилок, зокрема однокрокова процедура

$$B(k) = B(k-1) + \frac{e(k)X^T(k)}{X^T(k)X(k)}, \quad (3)$$

яка є узагальненням алгоритму Качмажа на векторно-матричну модель (2). Володіючи високою швидкістю, процедура (3) не має фільтруючих властивостей, а тому нездатна розрізнити зміни в сигналі і вплив стохастичного компонента  $\xi(k)$ . Використання традиційного експоненціально зваженого (з коефіцієнтом  $\beta$ ) рекурентного методу найменших квадратів ускладнюється недостатньою зумовленістю інформаційної матриці

$$\sum_{u=1}^k \beta^{k-u} X(u)X^T(u),$$

яка породжується високим рівнем кореляції між компонентами  $x_i(k)$ . Ця проблема може бути вирішена таким чином: замість операції обертання виваженої інформаційної матриці, яка основана на формулі Шермана-Моррісона, використовується операція псевдообертання Гревіля. Але отриманий таким чином алгоритм вельми громіздкий з обчислювальної точки зору, що особливо позначається при великих  $n$ .

Запропоновано використовувати багатовимірну модифікацію алгоритму експоненціально зваженої стохастичної апроксимації у вигляді

$$\begin{cases} B(k) = B(k-1) + \frac{e(k)X^T(k)}{\beta r(k-1) + \|X(k)\|^2}; \\ r(k) = \beta r(k-1) + \|X(k)\|^2. \end{cases}$$

Число параметрів моделі (2), яка настроюється, становить  $n(pn+1)$ , що може викликати певні труднощі при великих  $n$  і реальних частотах обробки відеокадрів (особливо високого розрішення).

Зручним математичним апаратом для розв'язання задачі виявлення зміни властивостей одновимірних стохастичних послідовностей є експоненціальне згладжування. Аналізується внутрішня структура стаціонарної багатовимірної часо-

вої послідовності за допомогою її кореляційної матриці виду

$$R(k, \tau) = \frac{1}{k} \sum_{u=1}^k (x(u) - \bar{x})(x(u - \tau) - \bar{x})^T, \tau = 0, 1, 2, \dots, \tau_{\max},$$

яка містить інформацію про автокореляційні і взаємно кореляційні функції всіх компонентів  $x_i(k)$ .

Для виявлення змін властивостей уведена кореляційна матриця, що є експоненціально згладженою

$$S_R(k, \tau) = \alpha(x(k) - S(k))(x(k - \tau) - S(k))^T,$$

а контроль проводиться за допомогою нерівності

$$Sp(S_R(k, \tau)S_R^T(k, \tau)) - Sp(S_R(k - 1, \tau)S_R^T(k - 1, \tau)) \geq \gamma_R,$$

де  $Sp(\circ)$  — слід матриці,

$Sp(S_R S_R^T)$  — квадрат сферичної норми матриці  $S_R$ .

Показано, що на основі методології експоненціального згладжування можна в реальному часі забезпечити контроль за змінами всіх характеристик багатовимірних часових рядів.

Важливою проблемою в процесі аналізу великих масивів (як за обсягом, так і за розмірністю) відеоданих, заданих у формі часових рядів, є задача їх стиску з метою виділення латентних факторів, які визначають внутрішню структуру контрольованого сигналу, що має на меті зробити вихідний часовий ряд простішим щодо інтерпретації виявлення змін властивостей.

Одним з найефективніших підходів стосовно вирішення цього завдання є апарат факторного аналізу, в межах якого найбільшого поширення набув метод головних компонент. Вихідною інформацією для аналізу є  $k \times n$  матриця спостережень

$$X = \begin{pmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ x_1(2) & x_2(2) & \dots & x_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(u) & x_2(u) & \dots & x_n(u) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(k) & x_2(k) & \dots & x_n(k) \end{pmatrix},$$

яка створена масивом із  $k$   $n$ -мірних векторів спостережень  $x(u) = (x_1(u), x_2(u), \dots, x_n(u))^T$  і її кореляційна  $(n \times n)$  матриця виду

$$R(k) = \frac{1}{k} \sum_{u=1}^k x^c(u)x^{cT}(u),$$

де  $x^c(u) = x(u) - \bar{x}$  — центровані щодо середнього вихідні дані.

Доцільним є використання рекурентних on-line процедур для знаходження власних векторів матриці  $R(k)$  завдяки послідовній обробці спостережень багатовимірного часового ряду  $x(1), x(2), \dots, x(k), x(k+1) \dots$  без обчислення самої кореляційної матриці.

Для попередньо центрованих даних адаптивний алгоритм знаходження власних значень і векторів має такий вигляд

$$\begin{cases} w^1(k+1) = w^1(k) + \eta(k+1)(x(k+1) - y(k)w^1(k))y(k+1), \\ y(k+1) = x^T(k+1)w^1(k), w^1(0) \neq 0, y'(1) = x^T(1)w^1(0), \end{cases}$$

де  $\eta(k+1)$  — параметр кроку настроювання, який обирається досить малим для забезпечення стійкої роботи алгоритму.

Цей алгоритм забезпечує нормування вектора  $w^1(k)$ , тобто  $\|w^1(k)\|^2 = 1$ , тоді як вектор  $w'(k)$  є власним вектором матриці  $R(k)$ , що відповідає максимальному власному значенню, а вихідний сигнал  $y(k)$  характеризується максимально можливою дисперсією, тобто містить максимум інформації про багатовимірний вхідний сигнал  $x(k)$ .

Далі вихідний сигнал  $x(k)$  піддається експоненціальному згладжуванню, що фільтрує шумові компоненти  $\xi(k)$ , а виявлення змін проводиться за допомогою одновимірного слідкувального сигналу.

Таким чином, запропоновані три методи щодо виявлення змін властивостей багатовимірних часових рядів, індукованих відеоданими в певному ознаковому просторі. Специфіка ознак, що змінюються в часі, тут не враховується, тому отримані результати в цілому справедливі для аналізу довільних часових рядів.

Перший з розглянутих методів, оснований на використанні моделей, які налагоджуються, є найбільше обґрунтований з математичної точки зору. Однак для побудови адекватної математичної моделі, що здатна ефективно виявляти розлади, які виникли, можуть знадобитися великі за обсягом навчальні вибірки, але це не завжди можливо, наприклад, у разі частішої зміни короточасних сюжетів. Другий метод базується на використанні багатовимірного експоненціального згладжування і є найпростішим з обчислювальної точки зору. При цьому для своєї роботи він не потребує значних обсягів інформації і дозволяє ефективно виявляти розладнання, що виникають, у формі різких стрибків ознакових описів. Водночас йому властива деяка інерційність, яка призводить до запізнювання в процесі виявлення, що, втім, може компенсуватися використанням процедур регулювання параметра згладжування за допомогою слідкувального сигналу, але, на жаль, це ускладнюватиме обчислювальну модель. Нарешті, третій метод щодо виявлення зміни властивостей багатовимірних часових рядів за до-

помогою аналізу головних компонент (перетворення Карунена-Лосва) дозволяє використовувати весь арсенал існуючих методів виявлення розладнань одномірних сигналів. Однак його застосування ускладнюється тим, що алгоритм навчання компресора не є оптимальною за швидкістю процедурою, що в деяких випадках призводитиме до «затягування» процесу навчання. Таким чином, в умовах апріорної невизначеності при сегментації відеопотоків доцільно паралельне використання всіх запропонованих методів.

**Третій розділ** присвячений дослідженню із загальних позицій умов, які забезпечують вилучення з розгляду об'єктів без обчислення відстаней до них при пошуку в довільному метричному просторі.

Нехай у деякому метричному просторі  $\Phi$   $K \subseteq \Phi$  — кінцева конфігурація точок з матрицею відстаней, що обчислюється прелімінарно

$$\mathcal{P} = (\rho(x_i, x_j))_{i,j=1}^n, \quad (4)$$

де  $\rho(x_i, x_j)$  — відстань між точками  $x_i, x_j$ .

Для запиту  $y \in \Phi$  відома відстань до довільної опорної точки  $x^* \in K$ :  $\delta = \rho(y, x^*)$ . Введемо поріг  $\Delta$ , який характеризує ступінь подібності (рівень нерозрізненості) точок метричного простору  $\Phi$  із запитом  $y$ .

Виникнення задачі елімінації об'єктів зумовлене трудомісткістю обчислення відстаней  $\rho(y, x_k)$ , що при великих потужностях конфігурацій  $K \subseteq \Phi$  може призводити до неприпустимих часових витрат. Результатом пошуку є найближча в заданій метриці точка або кілька точок, упорядкованих за подібністю в межах заданого порогу. Іншими словами, результатом  $(\Delta, k)$ -пошуку за запитом  $y \in \Phi$  є елементи множини  $K^m = \{x_{i_1}, x_{i_2}, \dots, x_{i_m}\} \subseteq K$ ,  $0 < m \leq k \leq n$ , для яких

$$\begin{aligned} \forall x_{i_j} \in K^m, \forall x \in K \setminus K^m, \forall y \in \Phi \quad \rho(y, x_{i_j}) \leq \Delta, \Delta \geq 0, \\ \rho(y, x_{i_j}) \leq \rho(y, x), \rho(y, x_{i_j}) \leq \rho(y, x_{i_{j+1}}), j = \overline{1, m-1}. \end{aligned}$$

Визначена зона достатності  $S$  — це зона з  $\mathbb{R}^n$  і є такою, де існує функція  $\varphi: K \times K \rightarrow \mathbb{R}^n$ , тобто  $\varphi(x_1, x_2) = (\alpha_1, \alpha_2, \dots, \alpha_n)$ , і якщо  $\varphi(x_1, x_2) \in S$ , то цього достатньо для знаходження  $\min \{\rho(y, x_1), \rho(y, x_2)\}$ . Відповідно, зона необхідності  $N \subseteq \mathbb{R}^n$  — це зона, для якої існує  $\psi: K \times K \rightarrow \mathbb{R}^n$ , і якщо можна вказати правильну нерівність з набору  $\{\rho(y, x_1) \geq \rho(y, x_2), \rho(y, x_2) \geq \rho(y, x_1)\}$ , то  $\psi(x_1, x_2) \in N$ .

Установлено, що зони області достатності  $S$  для довільної пари точок  $x_1, x_2 \in K$  дозволяє знайти найближчу точку до запиту  $y$ .

Зони  $S$  і  $N$  розбиваються на дві підзони  $S', S''$  ( $N', N''$ ). Зона  $S'$  відповідає ситуації, коли  $\rho(y, x_1) \geq \rho(y, x_2)$ , а  $S''$  —  $\rho(y, x_2) \geq \rho(y, x_1)$ , тобто  $S = S' \cup S''$  і  $S' \cap S'' = \emptyset$  за винятком  $\rho(y, x_1) = \rho(y, x_2)$ . Аналогічна властивість виконується і для  $N$ . Показано, що в процесі метричного пошуку на основі зазначених альтернатив зона достатності збігається із зоною необхідності.

Побудовано зони достатності при одній і двох опорних точках, що відповідає більшості алгоритмів, які використовуються на практиці. Під опорними точками маються на увазі об'єкти, до яких обчислені відстані до запиту. Наведено геометричні інтерпретації достатніх умов.

Розглянуто загальний випадок, коли процедура визначення найближчої точки конфігурації відбувається з використанням опорних точок  $\{x_i^*\}_{i=1}^n$ . Ця ситуація фактично відповідає попередній кластеризації ансамблю відеоданих, а опорними точками є центри кластерів. Це означає, що конфігурація  $K$  складається з  $n+2$  точок (до опорних додаються дві порівнювані  $x_1, x_2$ ), а з урахуванням запиту  $y$  цей випадок схематично зображено на рис. 1.

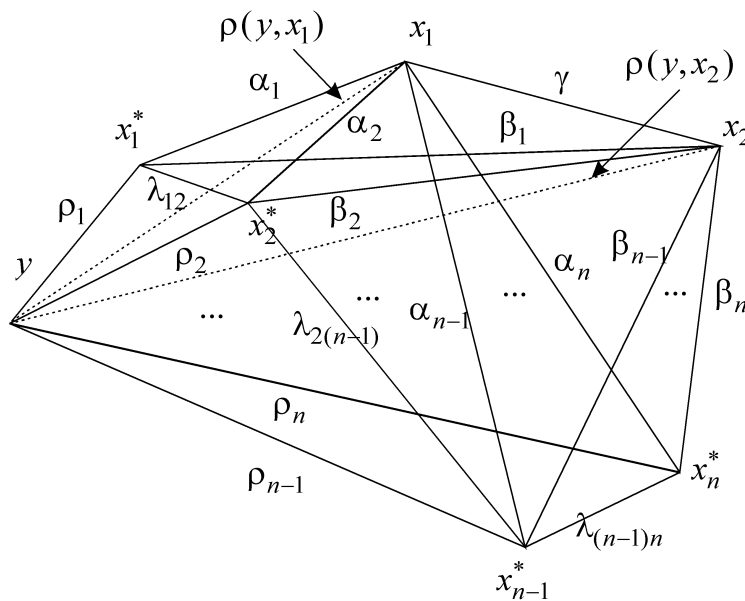


Рис. 1. Використання  $n$  опорних точок

Дані структуруються таким чином: порівнювані відстані позначаються  $\rho(y, x_1)$  і  $\rho(y, x_2)$ , далі  $\rho(x_1, x_2) = \gamma$ ;  $\rho(x_i^*, x_1) = \alpha_i$ ,  $\rho(x_i^*, x_2) = \beta_i$ ;  $\rho(y, x_i^*) = \rho_i$ ;  $\rho(x_i^*, x_j^*) = \lambda_{ij}$ ,  $i = \overline{1, n}$ ,  $i \neq j$ ,  $\lambda_{ij} = \lambda_{ji}$ . У цих позначеннях для конфігурації  $K$  підматриця відстаней  $\mathcal{P}$  розмірності  $(n+2) \times (n+2)$  має вигляд

$$P = \begin{pmatrix} 0 & \gamma & \alpha_1 & \dots & \alpha_i & \dots & \alpha_n \\ \gamma & 0 & \beta_1 & \dots & \beta_i & \dots & \beta_n \\ \alpha_1 & \beta_1 & 0 & \dots & \lambda_{1i} & \dots & \lambda_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_i & \beta_i & \lambda_{i1} & \dots & 0 & \dots & \lambda_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \alpha_n & \beta_n & \lambda_{n1} & \dots & \lambda_{ni} & \dots & 0 \end{pmatrix} \begin{matrix} x_1 \\ x_2 \\ x_1^* \\ \dots \\ x_i^* \\ \dots \\ x_n^* \end{matrix} .$$

$$x_1 \quad x_2 \quad x_1^* \quad \dots \quad x_i^* \quad \dots \quad x_n^*$$

Отже,  $\psi(x_1, x_2) \in \mathbb{R}^r$  ( $r = \frac{1}{2}(n+2)(n+1)$ ), тобто зона достатності повинна будуватися в  $\mathbb{R}^r$ .

Доведено, якщо елементи матриці відстаней (4) відповідають системам умов:

$$\begin{cases} \alpha_i + \beta_i \geq \gamma; \\ \alpha_i + \gamma \geq \beta_i; \quad (i = \overline{1, n}); \\ \gamma + \beta_i \geq \alpha_i, \end{cases} \quad \begin{cases} |\rho_i + \alpha_i| \leq |\rho_j - \beta_j|; \\ |\rho_j + \beta_j| \leq |\rho_i - \alpha_i|, \end{cases} \quad (i, j = \overline{1, n});$$

$$\begin{cases} \alpha_i + \alpha_j \geq \lambda_{ij}; \\ \alpha_i + \lambda_{ij} \geq \alpha_j; \\ \lambda_{ij} + \alpha_j \geq \alpha_i, \end{cases} \quad \begin{cases} \beta_i + \beta_j \geq \lambda_{ij}; \\ \beta_i + \lambda_{ij} \geq \beta_j; \\ \lambda_{ij} + \beta_j \geq \beta_i, \end{cases} \quad \{ |\rho_i - \rho_j| \leq \lambda_{ij} \leq \rho_i + \rho_j, \quad (i, j = \overline{1, n}, i \neq j);$$

$$\begin{cases} \lambda_{ij} + \lambda_{il} \geq \lambda_{jl}; \\ \lambda_{ij} + \lambda_{jl} \geq \lambda_{il}; \quad i, j, l = \overline{1, n}, i \neq j, i \neq l, l \neq j, \\ \lambda_{jl} + \lambda_{il} \geq \lambda_{ij}, \end{cases}$$

належать  $\Delta$ -оточенню запиту  $y$ , тобто відповідають обмеженням  $\{\rho_i + \alpha_i < \Delta; \{\rho_i + \beta_i < \Delta$ , де  $0 < \rho_i < \Delta$ , то цього достатньо для вибору найближчої точки конфігурації  $K$  до запиту  $y$  без обчислення відстані до неї.

Вивчено в загальному вигляді зони достатності і необхідності за умов виключення з розгляду об'єктів без обчислення відстаней до них в задачах метричного пошуку. Встановлено, що до задач метричного пошуку відеосюжетів із запитами за зразком з позицій трудомісткості обчислень доцільне послідовне використання однієї опорної точки фазового простору, однак (зокрема, при поданні відеосегмента ознаками ключового кадру) застосування двох і більше опорних точок у випадках існування «доброї» кластеризації є доцільнішим.

**Четвертий розділ** присвячений імітаційному моделюванню систем по-

шуку відео і, насамперед, запропонованих методів сегментації відеоданих.

Оскільки генерація багатовимірних часових рядів здійснюється на базі сегментації окремих відеокадрів відеопослідовностей, детально проаналізовані ознаки форми. Установлено набір геометричних ознак, що характеризують деякі зони, а в сукупності описують подання відеокадрів результатами просторової сегментації. Такий підхід дав змогу трактувати поняття «зміст» зображення взаємозв'язками зон та їх позиціонуванням у полі зору.

Здійснено класифікацію ознак та вибір характеристик форми, які є основою експериментального аналізу та практичної реалізації дисертаційних досліджень. Використання «просторового змісту» візуальної інформації дозволило зменшити семантичний конфлікт між низькорівневою обробкою відеоданих і цілями тематичної інтерпретації.

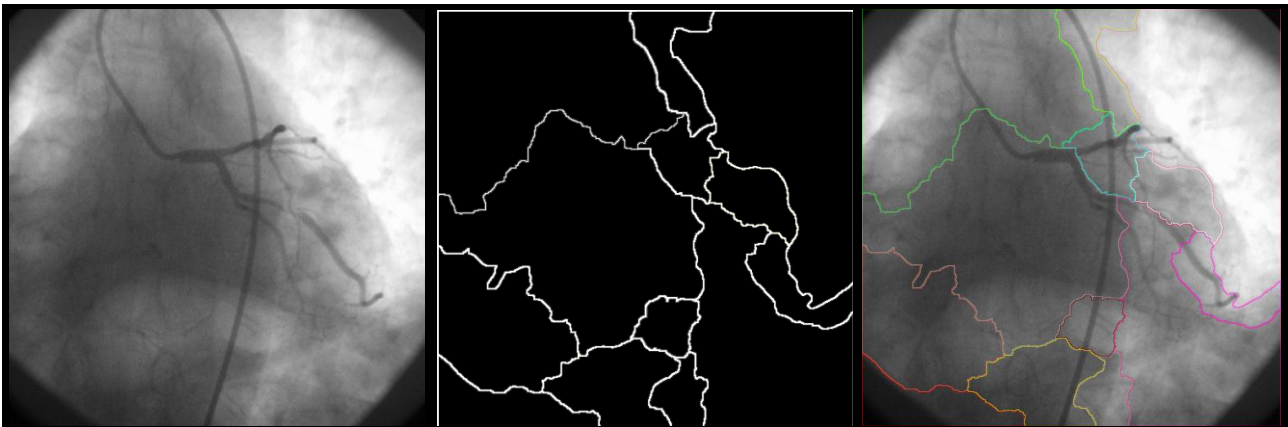


Рис. 2. Приклад сегментації відеокадру з ангіографічної колекції

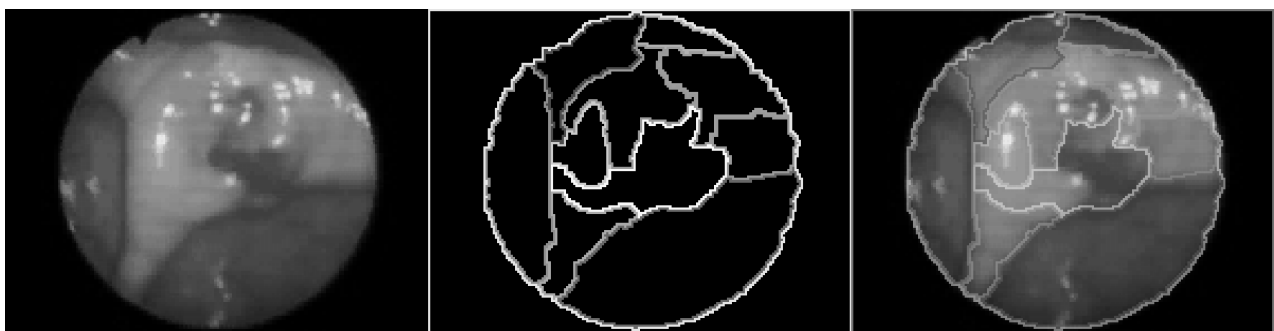


Рис. 3. Приклад сегментації відеокадру з ендоскопічної колекції

Під час використання відеоданих медичної тематики (при пошуку у відеоархівах коронарної ангіографії і ендоскопічних відеонаборах з патологією стравоходу та дванадцятипалої кишки) проаналізовані можливості щодо використання селективних ознак, а також запропонованих підходів з аналізу зміни сцен у кольорових і напівтонових відеоданих. На рис. 2, 3 наведені приклади просторової сегментації зображень із відеопослідовностей, ознаки зон яких і є джерелом багатовимірних часових рядів. У процесі моделювання пошуку як

ключового кадру використовувалося зображення, яке найменшим чином відрізнялося за результатами сегментації від усіх інших відеокадрів, що належать до даного сегмента.

Як цілі експериментальних досліджень були висунуті задачі вивчення специфіки сегментації в часі і просторі, а саме: впливу деталізації просторової сегментації на якість виявлення семантичних змін відеоданих, вибору ознакових описів форми зон для продукування багатовимірних часових рядів, визначення необхідної частоти обробки відеокадрів.

Для експериментів використовувалися однакові за часом (550 кадрів) відеодані з рівною розв'язністю. Параметри чутливості кольоровотекстурної сегментації підбиралися інтерактивно. Експериментально встановлено, що раціональна розмірність ознакового простору повинна бути близько 10. Проте збільшення кількості параметрів призводить до непотрібних часових витрат, а зменшення в цілому збільшує кількість помилок детектування (із семантичної точки зору) змін властивостей багатовимірних часових рядів. З огляду на нерівномірність впливу ознак форми і власне зон необхідне нормування ознак. Моделювання дозволило дійти висновку, що при «просіюванні» відеокадрів (більше 5 пропущених) виникає можливість пропуску різких змін у багатовимірних часових рядах.

На рис. 4, як приклад, наведено результати сегментації ендоскопічного відео на базі переходу від багатовимірного тимчасового ряду до одновимірного. Якщо при такому підході миттєві зміни властивостей тимчасового ряду легко детектуються, то плавні зміни просторового змісту можна відстежити тільки з використанням багатовимірного ознакового простору.

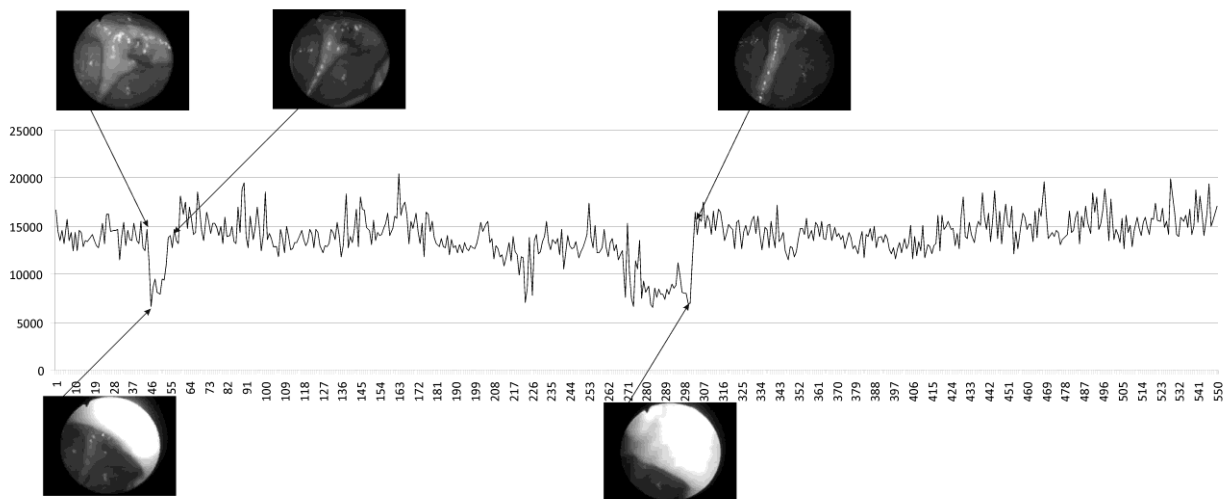


Рис. 4. Приклад тимчасової сегментації на основі методу головних компонент

У розділі розглянуто питання практичної реалізації запропонованих методів і специфіка впровадження результатів дисертаційних досліджень.



## ВИСНОВКИ

У дисертаційній роботі для систем гіпермедіа, зокрема для аналізу й інтерпретації відеопотоків, запропоновано нове рішення актуальної задачі сегментації відеоданих, обробки, аналізу та індексації результатів сегментації в часі і просторі для підвищення ефективності систем метричного пошуку і розпізнавання динамічної візуальної інформації в базах даних із запитом за зразком. Під час проведення теоретико-експериментальних дисертаційних досліджень була досягнута мета та вирішені всі завдання, що дозволило дійти основних наукових висновків і отримати відповідні результати.

1. З'ясовано, що для великих відеоколекцій перспективним напрямом досліджень є організація семантичного пошуку із запитом за зразком. Перехід до змісту відеоданих доцільно здійснювати подвійною сегментацією: у часі — це пошук змін подій або сюжетів, у просторі (у полі зору) — це пошук зон, корелювальних з об'єктами сцен.

2. Синтезовано метод адаптивного виявлення зміни властивостей багатовимірних часових рядів в умовах апріорної невизначеності на основі моделі, яка настроюється і навчання якої проводиться послідовно в процесі перегляду відео. Введена багатовимірна модифікація алгоритму експоненціально зваженої стохастичної апроксимації, що має як фільтруючі, так і слідкувальні властивості з автоматичним регулюванням параметра згладжування.

3. Запропоновано, обґрунтовано та досліджено методи сегментації відеопослідовностей на основі багатовимірної модифікації процедури експоненціального згладжування, а також на базі попередньої компресії за допомогою адаптивної модифікації аналізу головних компонент, після чого можуть ефективно використовуватися відомі адаптивні процедури виявлення розладнань в одновимірному випадку.

4. У загальному вигляді вивчені зони достатності і необхідності для метричного пошуку за змістовними критеріями з парадигмою видалення із розгляду об'єктів без обчислення відстаней до них, що створює передумови для раціонального щодо побудови обчислювальних моделей, вибору ознакового простору і синтезу швидкодіючих процедур пошуку.

5. Отримані достатні умови формування критеріїв елімінації об'єктів в процесі пошуку в довільному метричному просторі з однією, двома і  $k$  опорними точками. При цьому відстані від запиту обчислюються тільки до цих точок, вибір яких є довільним, тобто вони можуть бути центрами деякої категоризації або кластеризації, і тому як підсумок забезпечується підвищення швидкодії пошукових процедур у цілому.

6. Досліджені набори ознак, які характеризують форму зон, продукованих сегментацією окремих відеокадрів і в сукупності відображаючих «просторовий зміст» зображень, а в результаті формуючих багатовимірні тимчасові ряди,

аналіз яких і забезпечує пошук подій, сюжетів і сценаріїв у вихідному відеопотоці.

7. Завдяки експериментальним дослідженням доведена валідність використання теоретичних результатів сегментації відеоданих при організації пошуку відеосегментів за ключовими зображеннями на основі їх просторового змісту. Результати теоретико-експериментальних досліджень реалізовані й упроваджені у вигляді прикладних і дослідницьких програмних комплексів, що використовуються при пошуку у великих відеоархівах (ендоскопічних і коронарної ангіографії), один з яких сертифікований УкрСЕПРО.

## СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Егорова Е.А. Интеллектуальная обработка результатов сегментации синтезированных изображений / Е.А. Егорова, А.Н. Чупиков, К.С. Щербинин // Прикладная радиоэлектроника. — 2006. — Т. 5, № 3. — С. 408–411.
2. Робастные алгоритмы самообучения карты Кохонена в задачах обработки изображений / В.Я. Винарский, С.В. Машталир, Е.С. Сакало, К.С. Щербинин // Бионика интеллекта. — 2009. — №1 (70). — С. 86–89.
3. Машталир С.В. Поиск изменения сцен в видеоданных на базе анализа результатов сегментации / С.В. Машталир, С.В. Постульга, К.С. Щербинин // Бионика интеллекта. — 2010. — №1 (72). — С. 65–69.
4. Машталир С.В. Области достаточности и необходимости при решении задач альтернативного поиска видеоданных / С.В. Машталир, К.С. Щербинин // Прикладная радиоэлектроника. — 2010. — Т. 9, № 4. — С. 580–583.
5. Image quotient set transforms in segmentation problems / D. Kinoshenko, S. Mashtalir, K. Shcherbinin, E. Yegorova // International Journal Information Technologies & Knowledge. — 2008. — Vol. 2, No 4. — P. 372–378.
6. Regions of sufficiency for metrical data retrieval / V. Mashtalir, K. Shcherbinin, V. Shlyakhov, E. Yegorova // International Journal Information Technologies & Knowledge. — 2011. — Vol. 5, No 1. — P. 31–45.
7. Mashtalir S. Internal and external salient points under affine transformations. Comparative study / S. Mashtalir, E. Yegorova, K. Shcherbinin // Proc. of 14-th International Conference in Central Europe for Computer Graphics, Visualization and Computer Vision (Plzen, Czech Republic, January 30 – February 3, 2006). — Plzen: Science Press, 2006. — P. 156–161.
8. Image retrieval with segmentation-based query / A. Chupikov, D. Kinoshenko, V. Mashtalir, K. Shcherbinin // Adaptive multimedia retrieval: user, context, and feedback / S. Marchand-Maillet et al. (eds.). — 4-th International Workshop Adaptive Multimedia Retrieval (Geneva, Switzerland, July 27–28, 2006). — Berlin–Heidelberg: Springer-Verlag. — Lecture Notes in Computer Science. — 2007. — Vol. 4398. — P. 208–222.

9. Image partition transforms for faithful segmentation search / D. Kinoshenko, S. Mashtalir, K. Shcherbinin, E. Yegorova // Proc. of Fifth International Conference on Information Research and Applications (Varna, Bulgaria, June 26–30, 2007). — Sofia: FOI-COMMERCE, 2007. — Vol. II. — P. 385–389.
10. Мультипороговая сегментация с использованием признаков формы / А.Г. Каграманян, Д.К. Киношенко, А.Н. Чупиков, К.С. Щербинин // Материалы международной конференции «Интеллектуальные системы принятия решений и прикладные аспекты информационных технологий» (Евпатория, 14–18 мая 2007). — Херсон: ПП Вышемирский В.С., 2007. — Т. 3. — С. 143–146.
11. Метрическая индексация для быстрого поиска изображений в базах данных / Е.А. Егорова, Д.К. Киношенко, А.Г. Каграманян, К.С. Щербинин // Интеллектуальні системи прийняття рішень і проблеми обчислювального інтелекту. Т. 3, Ч. 1. Теоретичні і прикладні аспекти систем прийняття рішень. — (Євпаторія, 19–23 травня 2008). — Херсон: ПП Вишемирський В.С., 2008. — С. 117–120.
12. Mashtalir V. Adjacent classes transformation under multithresholding image fuzzy-segmentation / V. Mashtalir, K. Shcherbinin, V. Shlyakhov // Proc. of 15-th Zittau East-West Fuzzy Colloquium (Zittau, Germany, September 17–19, 2008). — Zittau: HS Zittau/Görlitz, 2008. — P. 151–154.
13. Машталир С.В. Методы контекстного сжатия информации / С.В. Машталир, С.В. Постульга, К.С. Щербинин // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта: Сборник трудов междунар. науч. конф. (Евпатория, 18–22 мая 2009). — Херсон: ХНТУ, 2009. — Т. 2. — С. 377–379.
14. Машталир С.В. Анализ результатов сегментации видеоданных / С.В. Машталир, С.В. Постульга, К.С. Щербинин // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта: Матер. междунар. науч. конф. (Евпатория, 17–21 мая 2010). — Херсон: ХНТУ, 2010. — С. 431–435.

## АНОТАЦІЯ

Щербінін К.С. Методи і моделі сегментації відео для систем метричного пошуку. — Рукопис.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 — системи та засоби штучного інтелекту. — Харківський національний університет радіоелектроніки, Харків, 2011.

Дисертацію присвячено розвитку методів і моделей пошуку відеоданих по змістовним критеріям. Для систем CBVIR (Content Based Video Information Retrieval) із запитом ‘ad exemplum’ вирішена задача сегментації багатовимірних часових рядів в умовах апріорної невизначеності. Запропоновані методи, завдяки яким забезпечується послідовний (покадровий) аналіз відео з метою отримання однорідних відеосегментів, що характеризуються ключовими кадрами, які й використовуються в якості пошукового образу та пошукового запиту.

Для підвищення швидкодії метричного пошуку в загальному вигляді знайдені умови, що дозволяють вилучати з розгляду відеосегменти або їх сімейства без трудомісткого обчислення відстані від них до запиту. Умови конкретизовані для однієї, двох і довільного числа опорних точок метричного простору, який використовується.

Експериментальні дослідження підтвердили ефективність запропонованих методів і моделей. Методи подвійної сегментації (відео — в часі, зображень — у полі зору) покладені в основу побудови пошукових систем обслуговування великих відеоархівів (ендоскопічних і коронарної ангіографії).

Ключові слова: змістовний пошук, багатовимірні тимчасові ряди, сегментація, ознаки форми.

## АННОТАЦИЯ

Щербинин К.С. Методы и модели сегментации видео для систем метрического поиска. — Рукопись.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.23 — системы и средства искусственного интеллекта. — Харьковский национальный университет радиоэлектроники, Харьков, 2011.

Глобальная аккумуляция данных в плохоструктурированных видеоархивах индуцирует целый класс сложных задач, связанных с установлением семантического соответствия поискового запроса и поискового образа в условиях априорной неопределенности и ограничений на время принятия решений и вычислительных ресурсов. Диссертация посвящена разработке моделей и методов сегментации многомерных временных рядов, индуцированных видеоданными, и поиску условий для повышения быстродействия поиска видеопоследовательностей в коллекциях динамической визуальной информации по содержательным критериям.

Исследовано состояние проблемы поиска видео в базах данных. В усло-

виях постоянно возрастающих объемов регистрируемых видеоданных необходим синтез новых, удовлетворяющих повышающимся требованиям по валидности и быстродействию методов и моделей поиска видеоданных. Анализ текущего состояния методов и моделей поиска визуальной информации по содержанию показал, что перспективным и уже в настоящее время практически важным подходом является реализация поиска по образцу с двухэтапной (во времени и в поле зрения) сегментацией видео. Отмечено, что существующие методы сегментации видео целесообразно развивать в плане создания процедур для обнаружения изменения свойств многомерных временных рядов в условиях неопределенности. Установлено, что одним из перспективных подходов к сокращению временных затрат поиска с запросами 'ad exemplum' является изучение необходимых и достаточных условий, которые позволяют на этапе поиска исключать из рассмотрения целые семейства «непохожих» на запрос видеосегментов без трудоемкого вычисления расстояний до них.

В диссертационной работе проведены исследования по сегментации в условиях априорной неопределенности многомерных временных рядов, которые индуцируются признаками результатов сегментации отдельных видеокадров. Предложены методы адаптивного обнаружения изменения свойств многомерных временных рядов на основе: настраиваемой с помощью специализированных адаптивных алгоритмов идентификации с конечной памятью модели; многомерной модификации процедуры экспоненциального сглаживания с адаптивной процедурой контроля характеристик, позволяющей обнаруживать возникающие разладки в on-line режиме; на базе предварительной компрессии с помощью адаптивной модификации метода анализа главных компонент с переходом к одномерному сигналу с помощью преобразования Карунена-Лоева, после чего может использоваться весь арсенал известных адаптивных процедур обнаружения разладок. Предложенные процедуры адаптивного обнаружения изменения свойств многомерных временных рядов отличаются вычислительной простотой, достаточно высоким быстродействием и позволяют в последовательном режиме (по мере поступления) решать задачу сегментации видео для поиска сюжетов на основе динамических признаков описаний. Установлено, что в условиях априорной неопределенности при сегментации видеопотоков целесообразно параллельное использование предложенных подходов, так как границы между сегментами могут быть и четкими, и малоконтрастными, и размытыми, по сути, являющимися самостоятельными сюжетами.

Для задач метрического поиска в общем виде изучены необходимые и достаточные условия для исключения из рассмотрения объектов без вычисления расстояний до них, что создает предпосылки для рационального, с точки зрения построения вычислительных моделей, выбора характеристик фазового пространства и синтеза собственно алгоритмов поиска. Необходимые и достаточные условия конкретизированы для использования с одной, двумя или произвольным числом опорных точек, которые могут выбираться на основе предварительной метрической обработки видеопоследовательностей в базах данных, в

частности, кластеризации наборов видео.

Проанализированы вопросы формирования признаков пространств, отражающих «пространственное» содержание видеок кадров. Путем имитационного моделирования исследовано влияние специфики формирования многомерных временных рядов на качество поиска однородных сегментов. Экспериментальные исследования проводились с учетом специфики видеоданных в предметно-ориентированной области, а именно — с видеопотоками, представляющими собой эндоскопические данные и информацию о коронарной ангиографии.

Основные компоненты разработанных программных комплексов легли в основу внедрения результатов диссертационного исследования. Путем экспериментальных исследований доказана валидность использования теоретических результатов сегментации видеоданных при организации поиска видеосегментов по ключевым изображениям на основе их пространственного содержания. Результаты теоретико-экспериментальных исследований реализованы и внедрены в виде прикладных и исследовательских программных комплексов, используемых при поиске в больших видеоархивах (эндоскопических и коронарной ангиографии), один из которых сертифицирован УкрСЕПРО.

Ключевые слова: содержательный поиск, многомерные временные ряды, сегментация, признаки формы.

## ABSTRACT

Shcherbinin K.S. Methods and models of video segmentation for metric search systems. — Manuscript.

The thesis for the candidate degree in technical sciences on the specialty 05.13.23 — Artificial Intelligence Systems and Means. — Kharkiv National University of Radio Electronics, Kharkiv, 2011.

The thesis devoted to developing of methods and models of search of the video by content criteria. For systems CBVIR (Content Based Video Information Retrieval) with a query ‘ad exemplum’ the problem of multidimensional time series segmentation under a priori uncertainty is solved. To detect homogeneous video segments that characterised by key frame which are used as a search pattern and a search query, methods ensuring the sequential (frame-by-frame) analysis of video are offered.

For metric search speedup in a general case the conditions, allowing to eliminate videosegments or even their sets without a time-consuming evaluation of distance from them to a query have been proposed. The conditions are rendered concrete for one, two and arbitrary number of pivot points of exploited metric space.

Experimental researches have confirmed efficiency of the offered methods and models. Methods of double segmentation (video — in time, images — in field of view) are taken as a basis of search engines construction to maintain largescale videoarchives (endoscopic and coronary angiography).

Keywords: content-based search, multidimensional time series, segmentation, shape features.

Відповідальний за випуск проф. Бодянський Є.В.

Підп. до друку 03.11.11. Формат  $60 \times 84^{1/16}$ . Спосіб друку – ризографія.  
Умов. друк. арк. 1,2. Облік. вид. арк. 1,0. Тираж 100 прим. Зам. №

---

Україна, 61166 Харків, просп. Леніна, 14, ХНУРЕ.

---

Віддруковано в навчально-науковому видавничо-поліграфічному центрі  
ХНУРЕ.

Україна, 61166 Харків, просп. Леніна, 14.