

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

БОГУЧАРСЬКИЙ СЕРГІЙ ІВАНОВИЧ

УДК 004.932.2:004.93'14

**МЕТОДИ І МОДЕЛІ ТОЛЕРАНТНОЇ КЛАСТЕРИЗАЦІЇ
У КОЛЕКЦІЯХ ЗОБРАЖЕНЬ**

05.13.23 – системи та засоби штучного інтелекту

Автореферат
дисертації на здобуття наукового ступеня
кандидата технічних наук

Харків – 2016

Дисертацією є рукопис.

Робота виконана в Харківському національному університеті радіоелектроніки Міністерства освіти і науки України.

Науковий керівник доктор технічних наук, професор
Машталір Володимир Петрович,
Харківський національний університет
радіоелектроніки,
професор кафедри інформатики.

Офіційні опоненти: доктор технічних наук, професор
Пелешко Дмитро Дмитрович,
Національний університет «Львівська політехніка»,
завідувач кафедри інформаційних техно-
логій поліграфічної справи;

кандидат технічних наук, доцент
Гороховатський Олексій Володимирович,
Харківський національний економічний
університет імені Семена Кузнеця,
доцент кафедри інформатики та
комп'ютерної техніки

Захист відбудеться “13” квітня 2016 р. о 15.00 годині на засіданні спеціалізованої вченої ради Д 64.052.01 у Харківському національному університеті радіоелектроніки за адресою: 61166, м. Харків, пр. Науки, 14.

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

Автореферат розісланий “3” березня 2016 р.

Учений секретар
спеціалізованої вченої ради

О.А. Винокурова

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. В цей час і в досить віддаленій перспективі розвиток методів зберігання, обробки, передачі, відображення та інтерпретації динамічної та статичної візуальної інформації залишається одним з важливих шляхів пошуку компромісу між інформаційними потребами та безперервно зростаючими обсягами відеоданих. Під реалізацією інформаційних потреб розуміється вміння в умовах суттєвої невизначеності видобувати знання з відео. Специфікою вирішення задач «розуміння» зображень є необхідність урахування адекватності, достовірності, повноти, надлишковості, об'єктивності, доступності та актуальності інформації, яка видобувається, що пояснюється семантичним розривом між ознаковими описаннями та прикладним трактуванням результатів.

Сегментація зображень (пошук у полі зору однорідних у рамках заданої концепції сприйняття областей, що корелюють з об'єктами сцен) та сегментація відеопотоків (виділення послідовностей відеокадрів, пов'язаних у часі однорідністю зміста) являють собою базові інструменти побудови основних примітивів семантичних конструкцій, що фактично є базисом інтелектуального аналізу візуальної інформації.

Накопичений арсенал методів просторової та часової сегментації (класифікації грануляції) забезпечує ефективну з точки зору предметної області, побудову розбиття поля зору або відеоряду на основі урахування характеристик візуальної інформації, агрегованих ознак, просторових та часових відносин між носіями різних областей (об'єктів) зображень та/або їх фрагментів. Однак для інтелектуального аналізу візуальної інформації, наприклад, CBVIR (Content Based Video Information Retrieval) досить часто необхідно вміти оперувати покриттями даних. З одного боку, це пояснюється природою відеоданих (високий рівень апріорної та поточної невизначеності, розмитість границь областей в полі зору або між сюжетами, тощо), з другої, – механізмами використання відновлених класів, наприклад, при стратифікованому загубленні або підвищенні ступеню деталізації границь областей або навіть деякі області слід одночасно відносити до різних кластерів. Така толерантна кластеризація в якості структуризації відеоданих створює умови для широкого діапазону методів інтелектуального аналізу, а в системах CBVIR забезпечує достатні показники точності та повноти інформаційного пошуку з запитом *'ad exemplum'*. Принциповим моментом ефективного функціонування систем CBVIR є необхідність on-line структуризації візуальної інформації для забезпечення багатоаспектної індексації даних, а у кінцевому підсумку – валідного пошуку сюжетів із запитом у вигляді фрагментів відео, ключових кадрів або різних ознакових запитань.

Аналіз стану та тенденцій розвитку моделей та методів сегментації візуальної інформації дозволяє стверджувати, що одним з перспективних напрямків є фрагментна кластеризація у послідовному режимі обробки інформації.

Дослідження спеціалізованих моделей і методів кластеризації динамічної візуальної інформації веде до висновку, що нейромережеві технології, що реалізують парадигми навчання та самонавчання, можуть виступати в якості адек-

ватного інструментарія для кластеризації відео у сигнальному або ознаковому просторах у послідовному режимі обробки інформації з урахуванням ресурсоемності процесів. Основний вплив на розвиток моделей та методів кластеризації та обробки візуальної інформації забезпечили українські та закордонні вчені С.Г. Антошук, Є. В. Бодянський, Р.А. Воробель, В.М. Крилов, Д.Д. Пелешко, Є.П. Путятін, М.І. Шлезінгер, В.К.Р. Horn, X Liu, S. Mann, O. Marques, K.S. Narendra, D.T. Pham, T. Masters, J.S. Zirilli та ін. Не дивлячись на багаточисельні дослідження у цьому напрямку, зростання обсягів відеоданих з різними властивостями у різних проблемно-орієнтованих областях веде до необхідності створення нових швидкодіючих та надійних засобів кластеризації, що з використанням єдиного інструментарія забезпечать толерантну сегментацію зображень у просторі та відеопотоків у часі.

Зв'язок роботи з науковими програмами, планами, темами. Дисертаційна робота виконана на кафедрі інформатики Харківського національного університету радіоелектроніки у рамках держбюджетних науково-дослідних робіт: «Моделі і методи аналізу та інтерпретації динамічних процесів» (№ ДР 0113U000357, виконавець), «Нейро-фаззі системи для поточної кластеризації та класифікації послідовностей даних за умов їх викривленості відсутніми та аномальними спостереженнями» (№ ДР 0113U000361, виконавець), що виконувалися у відповідності з наказом Міністерства освіти і науки України за результатами конкурсного відбору науково-дослідних робіт. В рамках вказаних робіт здобувачем отримані моделі та методи фрагментної кластеризації растрової інформації у просторі та часі для сегментації зображень та відео потоків як у вигляді розбиття, так і покриття поля зору та часової послідовності сюжетів.

Мета і завдання досліджень. Метою дисертаційної роботи є розробка моделей та методів фрагментної кластеризації візуальної інформації у просторі зображень для сегментації у часі і просторі відеопотоків та відеокадрів для підвищення ефективності систем метричного пошуку та розпізнавання динамічної візуальної інформації у базах даних із запитом за зразком.

Для досягнення цієї мети була обґрунтована необхідність вирішення таких задач:

- аналіз трендів розвитку методів кластеризації растрової інформації для виявлення перспективних засобів, що впливають на ефективність багатоаспектного індексування та інформаційного пошуку у CBVIR;
- синтез методів фрагментної сегментації зображень з нечіткими границями між сегментами довільної форми шляхом побудови матричних модифікацій методів кластеризації;
- модифікація методів ієрархічної агломеративної кластеризації зображень у матричній формі, що призначені для обробки інформації, яка міститься у великих базах даних за умов класів, що перебиваються;
- розробка метода аналізу текстур матричної просторово-часової динамічної моделі послідовності зображень на основі напівконтрольованого матричного векторного квантування;
- дослідження специфіки використання розроблених моделей та методів

для сегментації зображень та відеопослідовностей, створення та впровадження дослідницьких та спеціалізованих програмних засобів.

Об'єкт дослідження – процес аналізу та обробки статичної та динамічної візуальної інформації в інтелектуальних системах комп'ютерного зору.

Предмет дослідження – методи просторової та часової сегментації відеопослідовностей для метричної індексації та пошуку у відеоколекціях.

Методи дослідження – при розробці та дослідженні методів та моделей сегментації відеоданих були використані основні положення теорій розпізнавання образів, штучних нейронних мереж та ідентифікації систем, а також елементи математичної статистики при проведенні та аналізі результатів експериментальних досліджень.

Наукова новизна отриманих результатів. Наукова новизна дисертації полягає у постановці та вирішенні задачі фрагментної толерантної просторової та часової сегментації візуальної інформації для індексації зображень та відеопослідовностей в задачах інформаційного пошуку:

– вперше запропоновано метод напівконтрольованого матричного векторного квантування для аналізу текстур – сегментації послідовності зображень, що дозволяє оброблювати відеодані як у режимі навчання з вчителем, так і самонавчання у послідовному режимі, що дозволяє підвищити швидкодію процесу обробки відеоінформації.

– одержали подальший розвиток методи кластеризації, що засновані на щільності розподілу даних, які оперують на відміну від відомих фрагментами растрової інформації, що дозволяє скоротити кількість оцінюваних параметрів, забезпечує можливість функціонування за умов класів довільної форми, що перекриваються, для обробки зображень та відео потоків, що містяться у великих базах даних;

– одержали подальший розвиток методи кластеризації даних при невідомій кількості кластерів, що за рахунок відмови від векторизації-девекторизації фрагментів зображень дозволяють підвищити швидкодію процесів сегментації растрових даних у послідовному режимі з автоматичним визначенням раціональної кількості толерантних сегментів, що формується;

Практичне значення отриманих результатів. Практична значущість полягає в розробці методів та моделей сегментації уніфікованим інструментарієм, однаково придатним для обробки зображень (фрагментним скануванням поля зору) та відеопотоків (аналізі у часі покриття фрагментами відеокадрів). Результати можуть бути використані для задач структуризації візуальної інформації з метою забезпечення багатоаспектного індексування при інформаційному пошуку у системах CBVIR із запитом «по зразку».

Результати дисертаційних досліджень використані в інституті «ХАРКІВПРОЕКТ» (акт від 12.10.2015 року), Проектному та науково-дослідному інституті «ХАРКІВСЬКИЙ ПРОМБУДНДПРОЕКТ» (акт від 15.10.2015 року). В актах відзначається, що при оцінці дефектності об'єктів, детальному аналізі станів об'єктів на стадіях первинного та робочого проектів є необхідність архівного пошуку прецедентів або власне об'єктів аналізу, інформація про

які зберігається у вигляді фотографій або відеозйомки. У цих випадках запропоновані методи просторово-часової сегментації (структуризації) забезпечують можливість контекстного пошуку з запитом «за зразком», коли наголос робиться саме на носії зображень об'єктів аналізу. Розроблені здобувачем засоби успішно використані при дефектації об'єктів на основі візуальної інформації про них.

Теоретичні та практичні результати роботи знайшли використання у навчальному процесі Харківського національного університету радіоелектроніки при підготовці спеціалістів та магістрів за спеціальністю «Інформатика» (акт від 07.10.2015 року).

Особистий внесок здобувача. Всі положення, що виносяться на захист, отримані здобувачем особисто. У роботах, написаних у співавторстві, дисертанту належить постановка задач, вибір та обґрунтування їх рішення, а саме: в [2, 13,] запропоновані матричні модифікації ітераційного ієрархічного балансового методу кластеризації та методу на основі скороченої репрезентативної вибірки; у [3] обґрунтована та вивчена текстурна кластеризація на основі векторного квантування; у [4] формалізовані та досліджені методи фрагментної кластеризації з використанням рекурентної оптимізації; у [5] введена фрагментна модифікація метода кластеризації з випадковим пошуком в області локального сусідства; у [6, 7, 10] досліджені матричні модифікації метода кластеризації J -середніх; у [8, 12] обґрунтована фрагментна кластеризація з використанням щільності розподілу інформації; у [9] запропонована формалізація задачі толерантної сегментації відеоданих; у [14] запропонований метод сегментації відеоряду; у [15] запропонована толерантна фрагментна кластеризація на основі метода X -середніх. Роботи [1, 11] опубліковані без співавторів.

Апробація результатів дисертації. Основні результати роботи доповідалися, обговорювалися та були схвалені на науково-технічних конференціях та семінарах: VII-й міжнародній школі-семінарі «Теорія прийняття рішень» (Ужгород, 29 вересня – 4 жовтня 2014 року); International workshop and networking event 'Advances in data science' (Poland, Holny Mejera, 6-8 May, 2015); III-й міжнародній науково-практичній конференції «Обчислювальний інтелект (результати, проблеми, перспективи)» (Київ-Черкаси, 12-15 травня 2015 року); міжнародній науковій конференції «Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту» (Залізний Порт, 25-28 травня 2015 року); міжнародній науково-технічній конференції «Електротехнічні та комп'ютерні системи: теорія та практика» (Одеса, 20-24 липня 2015 року); 25th National scientific symposium with international participation 'Metrology and metrology assurance' (Bulgaria, Sozopol, 7-11 September, 2015); X International scientific and technical conference 'Computer science and information technologies' (Lviv, 14-17 September, 2015).

Публікації. Основні положення дисертації відображені у 15 роботах, серед яких 8 статей у фахових виданнях України з технічних наук (з них – 4 статті у міжнародних наукометричних базах); 7 публікацій у працях та матеріалах міжнародних конференцій, симпозіумів та семінарів.

Структура й обсяг дисертації. Дисертація складається зі вступу, п'яти розділів, висновків, списку використаних джерел та додатка. Повний обсяг дисертації становить 150 сторінок; робота містить 20 рисунків (з них 3 на окремих сторінках); список використаних джерел, що включає 133 найменування та займає 15 сторінок; додатка на 4 сторінках.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі обґрунтовано актуальність розробки нових методів сегментації зображень – пошуку у полі зору однорідних у рамках заданої концепції сприйняття областей, що корелюють із значущими об'єктами сцен та сегментації відеопотоків – виділення послідовності відеокадрів, пов'язаних у часі однорідністю змісту, розглянуто сучасний стан проблеми, визначені мета, об'єкт, предмет і методи дослідження, наведені задачі, що розв'язуються у дисертаційній роботі, зв'язок з науковими темами, наведено наукову новизну та практичне значення отриманих результатів, перелік публікацій за темою роботи, надано інформацію про особистий внесок автора.

У першому розділі в результаті ретроспективного аналізу встановлено, що сегментація зображень та відеопотоків – одна з ключових задач комп'ютерного зору. Сегментація представляється як процес побудови розбиття або покриття, що розділяють візуальну інформацію на складові частини з метою видобування об'єктів інтересу та їх сукупності. В силу того, що широка варіативність умов реєстрації того ж самого об'єкта (різні показники освітленості та метеорологічної дальності видимості, змінення взаємного розташування та орієнтації сенсора та об'єкта, часткові закриття, складні складові фону переднього та заднього планів, тощо), а також те, що із змістовної точки зору множина можливих спостережень, що відрізняються різноманітними семантичними та морфологічними властивостями, практично нескінченна, сегментація була і є актуальною задачею.

При аналізі візуальної інформації для основних типів інтерпретації даних (синтаксичній, семантичній, якісній та кількісній) виділяється деяка передоснова – результати сегментації зображень та відеопослідовностей. Для того, щоб конкретизувати або абстрагувати структурні уявлення та забезпечити трактування для конкретної предметної області, необхідно вирішити задачу ліквідації семантичного конфлікту між ознаками, що видобуваються з візуальної інформації, та тематичною інтерпретацією на високому рівні. Компроміс між недостатньою та надлишковою сегментацією дозволяє підвищити ефективність інтеграційної та автоматичної обробки візуальної інформації, для чого необхідно розширити арсенал методів сегментації зображень та відеопотоків.

В результаті проведеного аналізу сформульовані такі висновки. Для усунення семантичного конфлікту між результатами обробки зображень на рівні традиційних ознакових описань та інтерпретацією візуальної інформації на високому рівні необхідно на середньому рівні (наприклад, на основі «просторових» представлень візуальної інформації) *ad hoc* розвивати методи сегментації

як у полі зору, так і у часі при обробці відеопослідовностей. Тому перспективною є модифікація широкого діапазону прогресивних методів кластеризації шляхом їх адаптації до фрагментної обробки двовимірних полів, що у кінцевому підсумку забезпечить створення уніфікованого інструментарія для толерантної сегментації візуальної інформації.

У другому розділі введено модифікації методів кластеризації, що призначені для вирішення задач сегментації зображень та відеоданих, що зберігаються у великих базах даних. Основною відмінністю модифікацій є відмова від операцій векторизації-девекторизації, що дозволяє підвищити швидкодію процесу обробки інформації, можливість формування сегментів довільної форми за умов, коли дані викривлені викидами різної природи, формування кластерів-сегментів з розмитими нечіткими границями з урахуванням просторових кореляційних зв'язків між окремими пікселями.

Зазвичай, у відомих алгоритмах кластеризації припускається, що дані, які подаються на обробку, задані у вигляді багатовимірних векторів, що формують або традиційну таблицю даних, або багатовимірний векторний ряд. Однак, при обробці візуальної інформації, дані доцільно представляти фрагментною (матричною) формою. У такій формі може бути задано як зображення в цілому, так і деякий його фрагмент.

У розділі запропонована модифікація методу CLARANS (Clustering Large Applications based on RANdomised Search), що орієнтована на обробку даних, які надходять у вигляді $(m \times n)$ – матриць.

$$x(k) = \{x_{i_1 i_2}(k)\}, x(k) \in \mathbb{R}^{n \times m}, i_1 = 1, 2, \dots, m, i_2 = 1, 2, \dots, n, \quad (1)$$

де $k = 1, 2, \dots, N$ – номер зображення у колекції або номер фрагменту з деякого покриття (розбиття) або номер відеокадра (або його фрагменту) у відео потоці, що аналізується.

При обробці одного зображення отримуємо традиційну (просторову) сегментацію, при кластеризації відеопослідовності – сегментацію відеоряда у часі. При цьому припускається, що кількість кластерів p задана, а кожен з них описується на основі центроїда $C(l), l = 1, 2, \dots, p$, що також задається у вигляді матриці. Введена модифікація дозволяє обробку даних у матричній формі у ситуаціях з кластерами, що перекриваються, тобто реалізує процес синтезу покриттів.

У рамках перспективного підходу до кластеризації на основі щільності розподілу даних, який дозволяє формувати кластери довільної форми за умов, коли дані «збурені» завадами різної природи, запропонована модифікація методу DBSCAN (Density Based Spatial Clustering of Applications with Noise), що оперує матричними спостереженнями та дозволяє оцінити рівень належності кожного спостереження до кожного з кластерів-сегментів, що формуються. При цьому під кластерами розуміються області з високою щільністю розподілу даних. Ці області розділені областями з низькою щільністю і саме у них концентруються збурення-викиди. Таким чином, у процесі кластеризації «вирощуються» області з високою щільністю розподілу та формуються кластери довільної

форми з відокремленням збурень та шумів.

Для найбільш обґрунтованого методу, що заснований на щільності, DENCLUE (DENSity-based CLUstEring) та створений для обробки великих масивів даних, у розділі введена матрична модифікація.

При модифікації DENCLUE враховані такі припущення:

– вплив кожного вектора-образу на сусідні спостереження описується за допомогою функції впливу, зазвичай ядерної, що описує взаємний зв'язок всіх спостережень у деякому околі конкретного образу;

– загальна щільність розподілу даних у n -вимірному просторі ознак формально може бути описана як сума функцій впливу кожного спостереження;

– кластери визначаються як околиці атракторів щільності (D -атракторів), що є за суттю локальними максимумами загальної функції щільності розподілу даних.

Встановлено, що для деякої довільної точки у просторі ознак y її вплив на образ x може бути описаний за допомогою функції впливу

$$f^y(x) = f(x, y),$$

при цьому в якості такої функції використаний гаусіан

$$f(x, y) = \exp\left(-\frac{D^2(x, y)}{2\sigma^2}\right),$$

де σ – параметр ширини ядерної функції,

$D(x, y)$ – відстань, зазвичай евклідова, між точками x та y .

Тоді для множини спостережень $X = \{x(1), \dots, x(k), \dots, x(N)\}$ загальна функція щільності може бути записана у вигляді

$$f^x(x) = \sum_{k=1}^N f(x, x(k)).$$

Приймаючи замість стандартної векторної евклідової норми її сферичний матричний аналог $D_S^2(x, y) = Sp(x - y)(x - y)^T$, введено матричну функцію впливу

$$f_S(x, y) = \exp\left(-\frac{D_S^2(x, y)}{2\sigma^2}\right) = \exp\left(-\frac{Sp(x - y)(x - y)^T}{2\sigma^2}\right) \quad (2)$$

та матричну функцію щільності

$$f_S^x(x) = \sum_{k=1}^N f_S(x, x(k)).$$

Встановлено, що довільна $(m \times n)$ матриця-образ x притягується до матричного D -атрактору x^* , якщо послідовність ітерацій для матричної функції впливу (2)

$$x^i = x^{i-1} + \eta \frac{\sum_{k=1}^N (x(k) - x^{i-1}) f_S(x^{i-1}, x(k))}{\left(Sp \left(\sum_{k=1}^N (x(k) - x^{i-1}) f_S(x^{i-1}, x(k)) \right) \left(\sum_{k=1}^N (x(k) - x^{i-1}) f_S(x^{i-1}, x(k)) \right)^T \right)^{\frac{1}{2}}},$$

збігається до x^* .

Для оцінки рівня належності $x(k)$ до всіх кластерів (фактично рівня толерантності) використано співвідношення

$$\mu(x(k), C_l) = \frac{(Sp(x(k) - C(l))(x(k) - C(l))^T)^{-\frac{1}{2}}}{\sum_{q=1}^p (Sp(x(k) - C(q))(x(k) - C(q))^T)^{-\frac{1}{2}}}.$$

Запропонована модифікація формує кластери довільної форми при високому рівні збурень.

Третій розділ присвячений модифікації методів кластеризації для вирішення задач сегментації як статичних зображень, так і потоків відео за умов невизначеності про кількість можливих сегментів, їх форму, «розмитість» границь. за умов невідомої кількості класів, особливо при необхідності керування деталізацією розбиття або покриття, досліджені адаптовані для задач обробки зображень модифікації методів X -середніх та J -середніх. Для класів довільної форми введено модифікації процедур BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) та CURE (Clustering Using Representatives), що відрізняються як формою вихідної інформації, так і скороченням обсягу необхідних обчислень.

Фрагментну модифікацію метода X -середніх можна описати наступним чином.

Нехай за умов представлення даних (1) деяким чином одержано p кластерів з набором центроїдів $C = \{C(1), C(2), \dots, C(p)\} \subset \mathbb{R}^{m \times n}$. Для оцінки якості такого розбиття адаптовано критерій Цалінського-Харабаша (що не використовує статистичні передумови) у формі

$$CH(p) = \frac{\frac{1}{p-1} Sp S_B(p)}{\frac{1}{N-p} Sp S_w(p)},$$

де $S_B(p) = \sum_{l=1}^p N_l (C(l) - \bar{C})(C(l) - \bar{C})^T$ – матриця міжкластерного розсію-

вання,

$$S_w(p) = \sum_{l=1}^p \sum_{k=1}^N \mu(x(k), C(l)) (x(k) - C(l))(x(k) - C(l))^T \quad - \quad \text{матриця}$$

внутрішнього кластерного розсіювання,

$$C(l) = \frac{1}{N_l} \sum_{x(k) \in Cl_l} x(k) = \frac{\sum_{k=1}^N \mu(x(k), C(l)) x(k)}{\sum_{k=1}^N \mu(x(k), C(l))} - \text{центроїд кластера } Cl_l,$$

$$\bar{C} = \frac{1}{N_l} \sum_{l=1}^p N_l C(l) - \text{матричний центр важкості масиву } X.$$

Запропонована модифікація методу X -середніх для обробки зображень на основі K -середніх та критерія Цалінського-Харабаша реалізується у вигляді послідовності:

- задання $p = p_{min}$, вирішення задачі за допомогою фрагментної модифікації k -середніх та розрахунок $CH(p_{min})$;
- розщеплення обраного кластера, вирішення задачі за допомогою k -середніх та розрахунок $CH(p_{min} + 1)$;
- при $CH(p + 1) \leq CH(p)$ покладається, що p є найкраща оцінка кількості кластерів у масиві X .

Вибір кластера проводиться на основі ознакових описань, у найпростішому випадку – це кластер з максимальним внутрішнім розсіюванням.

В основу підходу, коли кластери, що формуються в процесі обробки інформації перекриваються, покладено матричну модифікацію методу нечітких C -середніх (FCM), пов'язану з мінімізацією цільової функції самонавчання

$$E(x(k), C(l)) = \sum_{k=1}^N \sum_{l=1}^p \mu^\beta(x(k), C(l)) Sp(x(k) - C(l))(x(k) - C(l))^T$$

при додаткових обмеженнях

$$\sum_{l=1}^p \mu(x(k), C(l)) = 1; 0 < \sum_{l=1}^p \mu(x(k), C(l)) < N, l = 1, 2, \dots, p.$$

Тут $0 \leq \mu(x(k), C(l)) \leq 1$ – рівні належності матричного спостереження $x(k)$ до кластеру Cl_l , $\beta \geq 0$ – параметр, фазифікації, що задає рівень нечіткості (розмитості) границь між сусідніми сегментами.

Вирішення задачі квадратичного програмування, пов'язаної з цільовою функцією (3) при обмеженнях (4), (5) веде до результату

$$\begin{cases} \mu(x(k), C(l)) = \frac{(Sp(x(k) - C(l))(x(k) - C(l))^T)^{(1-\beta)^{-1}}}{\sum_{l=1}^p (Sp(x(k) - C(l))(x(k) - C(l))^T)^{(1-\beta)^{-1}}}, \\ C(l) = \frac{\sum_{k=1}^N \mu^\beta(x(k), C(l)) x(k)}{\sum_{k=1}^N \mu^\beta(x(k), C(l))}. \end{cases}$$

Особливістю введеної модифікації є можливість обробки матричних сиг-

налів за відсутності інформації про статистичні характеристики цих сигналів та кількості кластерів, яка автоматично визначається в процесі аналізу вихідного масиву даних.

Процедурою кластеризації, що має глобальні властивості, є метод *J-means*. Сутність методу полягає в тому, що при «застряганні» у локальному екстремумі в його околі відбуваються «стрибки», що виводять процедуру з цього околу в область критерія більш «глибоких» екстремумів. Основою евристикою при цьому є процес переміщення центроїдів у незайняті точки та знаходження на цій основі розбиття з покращеними значеннями цільової функції. Для підвищення швидкодії пошук відбувається в околах не всіх центроїдів $C(l)$, $l = 1, 2, \dots, p$, а лише у одному кластері, для вибору якого може бути використано або максимальне внутрішнє розсіювання

$$\overline{D}^2 = \frac{1}{N_l} \sum_{x(k) \in Cl_l} Sp(x(k) - C(l))(x(k) - C(l))^T,$$

або максимальна міжкластерна відстань

$$\overline{D}^2 = \frac{1}{P} \sum_{x(k) \in Cl_l} Sp(C(q) - C(l))(C(q) - C(l))^T,$$

остільки найбільш віддалений від інших кластер максимального діаметру має більші шанси містити більш глибокий екстремум.

Для толерантної сегментації у розділі запропонована матрична модифікація метода нечітких *J*-середніх. В основі запропонованого підходу полягає матрична модифікація метода нечітких *C*-середніх, забезпечена механізмом випадкових «стрибків», що дозволяє відшукувати глобальний екстремум цільової функції.

Четвертий розділ присвячений розв'язанню задач класифікації та кластеризації текстур у послідовностях зображень за умов, коли кожний об'єкт аналізу може містити одночасно декілька текстур. В основі запропонованого підходу полягає матрична просторово-часова модель спеціального типу, а також техніка комбінованого векторного квантування з використанням напівконтрольованого навчання.

В основу покладена просторово-часова авторегресійна модель (STAR), що включає передісторію не лише конкретного пікселя, але й його оточення, що задано у формі деякого фрагмента. STAR – модель, що описує просторово-часову еволюцію текстур, повністю задається вектором своїх коефіцієнтів, який надходить на систему класифікації – штучну нейронну мережу.

У розгляд введено матричну модель авторегресії довільного порядку h вигляду

$$x(k) = \sum_{d=1}^h A_d x(k-d) C_d + \xi(k) = A(x(k-1) : x(k-2) : \dots : x(k-h)) C + \xi(k) = \tilde{A} x(k-1) C + \xi(k),$$

$$\text{де } A = (A_1 : A_2 : \dots : A_h),$$

$$C = \begin{pmatrix} C_1 \\ \dots \\ C_2 \\ \dots \\ \vdots \\ \dots \\ C_h \end{pmatrix}, \quad \tilde{x}(k-1) = \begin{pmatrix} x(k-1) & & \mathbf{0} \\ & x(k-2) & \\ \mathbf{0} & & x(k-h) \end{pmatrix} - (hm \times hn) - \text{матриця.}$$

Для визначення $(m \times hm)$ та $(hn \times n)$ -матриць коефіцієнтів A та C використаний адаптивний градієнтний алгоритм ідентифікації матричних моделей

$$\left\{ \begin{array}{l} A(k+1) = A(k) + (SpV_A(k+1)C^T(k)\tilde{x}^T(k)\tilde{x}(k)C(k)V_A^T(k+1)) \\ (SpV_A(k+1)C^T(k)\tilde{x}^T(k)\tilde{x}(k)C(k)C^T(k)\tilde{x}^T(k)\tilde{x}(k)C(k)V_A^T(k+1))^T \\ V_A(k+1)C(k)\tilde{x}^T(k), \\ C(k+1) = C(k) + (SpV_C^T(k+1)A(k+1)\tilde{x}(k)\tilde{x}^T(k)A^T(k+1)V_C(k+1)) \\ (SpA(k+1)\tilde{x}(k)\tilde{x}^T(k)A^T(k+1)V_C(k+1)V_C^T(k+1)A(k+1)\tilde{x}(k) \\ \tilde{x}^T(k+1)\tilde{x}^T(k)A^T(k+1)V_C(k+1), \end{array} \right.$$

$$\text{де } V_A(k+1) = x(k+1) - A(k)\tilde{x}(k)C(k),$$

$$V_C(k+1) = x(k+1) - A(k+1)\tilde{x}(k)C(k).$$

Об'єднуючи матриці коефіцієнтів $A(k+1)$, $C(k+1)$, отримуємо $(m+hn) \times (hm+n)$ блочну матрицю $B(k+1)$:

$$B(k+1) = \begin{pmatrix} A(k+1) & \vdots & \mathbf{0} \\ \dots\dots\dots & \vdots & \dots\dots\dots \\ \mathbf{0} & \vdots & C(k+1) \end{pmatrix},$$

що містить в собі повну інформацію про матричну просторово-часову авторегресійну модель. Саме ця матриця надходить на вхід матричної LVQ-системи.

Оцінка рівнів належності у випадку проведення нечіткої класифікації має вигляд

$$\mu(B(k+1), w_q^L(k+1)) = \frac{(Sp((B(k+1) - w_q^L(k+1))(B(k+1) - w_q^L(k+1))^T))^{-\frac{1}{2}}}{\sum_{q=1}^p (Sp((B(k+1) - w_q^L(k+1))(B(k+1) - w_q^L(k+1))^T))^{-\frac{1}{2}}},$$

$$\text{де } w_q^L(k+1) = \begin{cases} w_q^{L*}(k) + \eta(k+1)(B(k+1) - w_q^{L*}(k)), \\ \text{якщо } B(k+1) \text{ та } w_{q_1q_2}^{L*}(k) \text{ належать одному класу,} \\ w_q^{L*}(k) - \eta(k+1)(B(k+1) - w_q^{L*}(k)), \\ \text{якщо } B(k+1) \text{ та } w_q^{L*}(k) \text{ належать різним класам,} \\ w_q^L(k) \text{ для нейронів, що не перемогли у момент } k+1. \end{cases}$$

Коли частина даних навчальної вибірки є розміченою, а про інші образи відсутня інформація про їх належність до того або іншого класу, використано напівконтрольоване навчання, реалізоване на базі архітектур нейронних мереж Т. Кохонена. Процес обробки інформації в SOM полягає в тому, що на матричний вхід послідовно надходять елементи навчальної вибірки $x(k)$, $k=1,2,\dots, MN/(mn)$, попередньо пронормовані на сферичну матричну норму $\sqrt{Sp(x(k)x^T(k))}$ цих сигналів. Кожний рецептор нульового шару пов'язаний з усіма нейронами шару Кохонена, кожен з яких має mn налаштованих синаптичних ваг, що задаються у вигляді $(m \times n)$ -матриць $w_{q_1q_2}^s$. Скалярний вихідний сигнал кожного нейрона обчислюється за допомогою виразу

$$y_{q_1q_2}^s(k) = Sp(w_{q_1q_2}^T(k-1)x(k)) = E_w^T(w_{q_1q_2}^s(k-1) \odot x(k))E_n,$$

де $(k-1)$ індекс часу матриці синаптичних ваг $w_{q_1q_2}^s(k-1)$, отриманої за допомогою вибірки образів $x(1), x(2), \dots, x(k-1)$;

E_m и E_n – одиничні $(m \times 1)$ и $(n \times 1)$ вектори-стовпці,

\odot – символ прямого матричного добутку.

Процес самонавчання нейронної мережі стартує з початкових значень синаптичних ваг, що обираються досить довільним чином, після чого кожна матриця ваг нормується на $\sqrt{Sp(w_{q_1q_2}^s(0)w_{q_1q_2}^{sT}(0))}$. Далі на вхід мережі надходить перша матриця-образ $x(1)$ та обчислюються вихідні сигнали всіх нейронів шару Кохонена

$y_{q_1q_2}^s(1) = Sp(w_{q_1q_2}^{sT}(0)x(1)) = E_w^T(w_{q_1q_2}^s(0) \odot x(1))E_n$. Наступним етапом є визначення нейрона-переможця $w_{j_1j_2}^{s*}$ такого, що для $q_1 = \overline{1, l}$, $q_2 = \overline{1, p}$

$$Sp((x(1) - w_{q_1q_2}^{s*}(0))(x(1) - w_{q_1q_2}^{s*}(0))^T) < Sp((x(1) - w_{q_1q_2}^s(0))(x(1) - w_{q_1q_2}^s(0))^T),$$

при цьому нейрон-переможець має найбільше в порівнянні з усіма іншими значення вихідного сигналу: $y_{q_1q_2}^{s*}(1) > y_{q_1q_2}^s(1) \forall q_1$ и q_2 .

Налаштування синаптичних ваг нейрона-переможця для $(k+1)$ -го кроку відбувається згідно із правилом «переможець отримує все» у формі

$$w_{q_1 q_2}^{s*}(k+1) = w_{q_1 q_2}^{s*}(k) + \eta(k+1)(x(k+1) - w_{q_1 q_2}^{s*}(k)).$$

Тут $\eta(k+1)$ – крок налаштування, що обирається емпірично так, щоб він монотонно зменшувався у процесі навчання.

На основі техніки векторного квантування, що полягає у компактному представленні великих масивів інформації, яка задана у вигляді $(m \times n)$ -матриць $x(k)$ у формі обмеженого набору центрів (прототипів) $w_{q_1 q_2}^L$, $q_1 = 1, 2, \dots, l$; $q_2 = 1, 2, \dots, p$, реалізується процес контрольованого навчання LVQ.

Для кожного попередньо пронормованого образа-матриці $x(k+1)$ ($Sp(x(k+1)x^T(k+1)) = 1$) визначається нейрон-переможець, синаптичні ваги $w_{q_1 q_2}^{L*}(k)$ якого відповідають центру конкретного кластера. Інакше кажучи, переможцем, як і у SOM, є нейрон з мінімальною відстанню до вхідної матриці такий, що для $q_1 = \overline{1, l}$, $q_2 = \overline{1, p}$

$$Sp((x(k+1) - w_{q_1 q_2}^{L*}(k))(x(k+1) - w_{q_1 q_2}^{L*}(k))^T) < Sp((x(k+1) - w_{q_1 q_2}^L(k))(x(k+1) - w_{q_1 q_2}^L(k))^T).$$

Оскільки тут навчання є контрольованим, належність матриці $x(k+1)$ до конкретної області простору X є відомою. Це дозволяє записати матричну модифікацію процесу навчання LVQ1

$$w_{q_1 q_2}^{L*}(k+1) = \begin{cases} w_{q_1 q_2}^{L*}(k) + \eta(k+1)(x(k+1) - w_{q_1 q_2}^{L*}(k)), & \text{якщо } x(k+1) \text{ та } w_{q_1 q_2}^{L*}(k) \\ & \text{належать одному класу,} \\ w_{q_1 q_2}^{L*}(k) - \eta(k+1)(x(k+1) - w_{q_1 q_2}^{L*}(k)), & \text{якщо } x(k+1) \text{ та } w_{q_1 q_2}^{L*}(k) \\ & \text{належать різним класам,} \\ w_{q_1 q_2}^L(k) & \text{для нейронів, що не перемогли у момент } k+1. \end{cases}$$

Для того, щоб реалізувати процес напівконтрольованого навчання, коли частина вхідних даних є розміченою, а частина – ні, доцільно організувати спільну одночасну обробку інформації SOM та LVQ. Для вирішення задачі доціль-

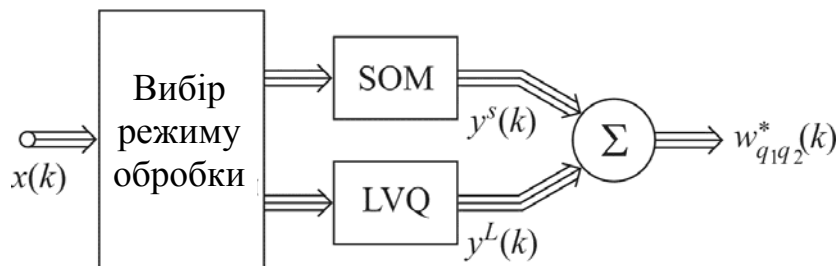


Рисунок 1 – Структурна схема напівконтрольованого навчання

но паралельне поєднання SOM та LVQ так, як це показано на рис. 1.

П'ятий розділ присвячений експериментальному аналізу запропонованих методів пошуку у полі зору однорідних областей, що корелюють із значущими об'єктами сцен, виділення послідовностей відеокадрів, пов'язаних у часі однорідністю змісту.

Експерименти проводилися з відеорядами (представленими медіаконтейнерами AVI), що склалися з 750 кадрів, тобто тривалістю 30 секунд кожне відео. Для валідного порівняння результатів роздільна здатність фіксувалася на рівні 632×488 . Кожен з кадрів кожного відеоряду представлявся у форматі TIFF без стиснення з метою усунення впливу можливих втрат за рахунок згладжування кольорів. При сегментації використана зважена по каналах RGB евклідова метрика, а для порівняння результатів сегментації – нормована метрика

$$\rho(X, Y) = \frac{\sum_{\alpha} \sum_{\beta} \text{card}([x]_{\alpha} \Delta [y]_{\beta}) \text{card}([x]_{\alpha} \cap [y]_{\beta})}{1 + \sum_{\alpha} \sum_{\beta} \text{card}([x]_{\alpha} \Delta [y]_{\beta}) \text{card}([x]_{\alpha} \cap [y]_{\beta})},$$

де $X = \{[x]_{\alpha}\}$, $Y = \{[y]_{\beta}\}$ – результати сегментації (з різними параметрами алгоритму, наприклад, розмірами фрагменту) одного або двох зображень (X або Y може бути ground-truth сегментацією), $[x]_{\alpha} \Delta [y]_{\beta}$ – симетрична різниця. При толерантній сегментації функція належності засновувалась на геометричній ймовірності перетинів, а в результаті усувалася пороговим відсіканням.

Крім окремих відеокадрів просторовій сегментації піддаються зображення ground-truth колекції університету Берклі з роздільною здатністю зображень 484×321 пікселів. При необхідності для отримання точних розбиттів поля зору фрагментами проводилося масштабне перетворення.

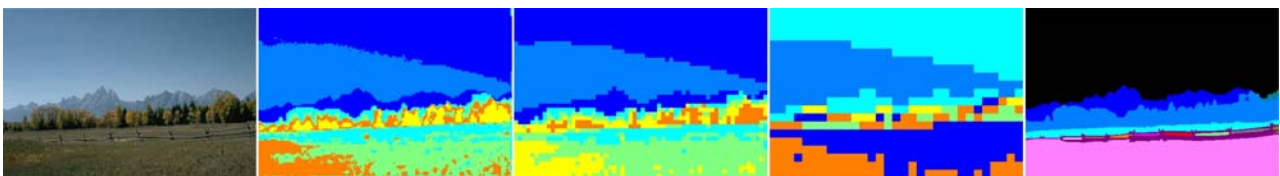


Рисунок 2 – Приклади фрагментної сегментації зображень

На рис. 2 наведені: вихідне зображення, сегментація з фрагментами 3×3 , 8×8 , 16×16 елементів, інтерактивна сегментація. Рис. 3 ілюструє фрагментну кластеризацію у часі: одним кольором означені одні і ті ж кластери, пов'язані із зміною сюжету.

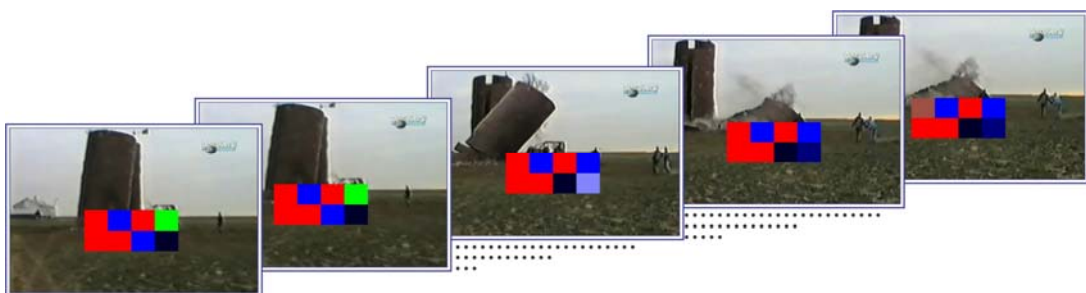


Рисунок 3 – Приклади фрагментної сегментації відеопотоків

У додатку наведені акти впровадження теоретичних та практичних результатів дисертаційних досліджень.

ВИСНОВКИ

В дисертаційній роботі для розвитку інструментарія засобів штучного інтелекту в галузі аналізу та інтерпретації візуальної інформації вирішена актуальна науково-практична задача розробки методів просторової та часової фрагментної сегментації зображень та відеопотоків для підвищення ефективності багатоаспектної індексації і власне систем пошуку та розпізнавання візуальної інформації у великих базах даних з запитом по зразку. Проведені дослідження дозволяють зробити такі висновки.

1. На основі ретроспективного аналізу трендів розвитку методів кластеризації та методів сегментації зображень встановлено, що не дивлячись на значний арсенал існуючих на практиці ефективних методів та підходів, залишається гостра потреба у *ad hoc* розробці методів для фрагментної сегментації, включаючи толерантну, тобто з перекриттям класів, у часі та просторі відеопотоків та відеокадрів.

2. Запропонована матрична модифікація методу кластеризації CLARANS, що призначена для обробки наборів відеоспостережень, що зберігаються у великих базах даних. Особливістю запропонованої модифікації є відмова від векторизації-девекторизації фрагментів зображень, що поряд з покращенням робастності до аномальних викидів довільної природи та можливості перетину класів, дозволяє підвищити швидкодію процесів сегментації растрових даних.

3. Запропонована матрична модифікація методу кластеризації DBSCAN, що орієнтована на вирішення задачі сегментації зображень з нечіткими границями між сегментами довільної форми та відрізняється скороченою кількістю оцінюваних параметрів, що дозволяє підвищити швидкодію обробки растрової інформації.

4. Отримав подальший розвиток метод кластеризації, заснований на використанні розподілів даних та рекурентній оптимізації, що дозволяє формувати сегменти довільної форми при високому рівні завад та відрізняється швидкодією за рахунок використання матричної функції щільності, яка базується на сферичній нормі.

5. Введена матрична модифікація методу кластеризації X -середніх, що відрізняється підвищеною швидкодією за рахунок скорочення кількості точок аналізуємих сегментів та використання модифікованого критерія якості, що дозволяє проводити обробку інформації у послідовному режимі на відміну від традиційних критеріїв. Для роботи за умов класів, що перекриваються, введено модифікацію матричного методу нечітких C -середніх, що дозволяє безпосередньо у процесі обробки автоматично визначати кількість сегментів.

6. Введені матричні модифікації методів ієрархічної агломеративної кластеризації зображень, методів кластеризації J -середніх та нечітких J -середніх, що дозволяють забезпечити досягнення більш глибокого екстремуму цільової

функції якості сегментації зображень, а також підвищити швидкодію процесу обробки інформації за рахунок організації процесів стрибків лише в одному кластері максимального об'єму.

7. Вирішена задача аналізу текстур послідовності зображень, в основі якого полягає техніка комбінованого векторного квантування. Вперше запропонована матрична просторово-часова динамічна модель послідовності зображень та метод її налаштування на основі матричного нечіткого векторного квантування, що дозволило підвищити швидкодію процесу кластеризації текстур.

8. Вперше запропонований метод напівконтрольованого матричного векторного квантування для аналізу текстур – сегментації послідовності зображень, що дозволяє обробляти відеодані як в режимі навчання з вчителем, так і самонавчання в послідовному режимі, що дозволяє підвищити швидкодію процесу обробки відеоінформації.

9. Розроблені методи сегментації зображень та відеопослідовностей для вирішення задач керування великими колекціями візуальної інформації та контекстного пошуку з запитом за зразком реалізовані програмно та використані у розробках інститутів «ХАРКІВПРОЕКТ», «ХАРКІВСЬКИЙ ПРОМБУДНДІПРОЕКТ» та навчальному процесі ХНУРЕ, що підтверджено відповідними актами.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

1. Богучарский, С.И. Обработка изображений на основе комбинированного векторного квантования с использованием полуконтролируемого обучения / С.И. Богучарский // Вісник Національного технічного університету «ХПИ». Збірник наук. праць. Серія: математичне моделювання в техніці та технологіях. – Харків: НТУ «ХПИ». – 2014. – № 6 (1049). – С. 20-29. (Входить до міжнародної наукометричної бази Ulrich's Periodicals Directory.)

2. Богучарский, С.И. Иерархическая агломеративная кластеризация изображений в больших базах данных / С.И. Богучарский, А.Г. Каграманян, С.В. Машталир // Збірник наукових праць ХУПС. – Харків: ХУПС. – 2014. – № 8(124). – С. 93-97.

3. Богучарский, С. И. Анализ текстур в последовательности изображений на основе векторного квантования // С.И. Богучарский, С.В. Машталир // Радиоэлектроника, информатика, управление. – 2014. – №2(31). – С. 94-99. (Входить до міжнародних наукометричних баз Ulrich's Periodicals Directory, Bielefeld Academic Search Engine (BASE), Index Copernicus, Directory of Open Access Journals (DOAJ), EBSCO, WorldCat, та ін.)

4. Богучарский, С.И. Кластеризация коллекций изображений в больших базах данных на основе рекуррентной оптимизации / С.И. Богучарский, С.В. Машталир // Біоніка інтелекту. – 2014. – №1(82) – С. 43-46.

5. Bogucharskiy S. On matrix modification of CLARANS clustering method in large video surveillance databases / S. Bogucharskiy, V. Mashtalir / Вісник Національного університету «Львівська політехніка». Комп'ютерні науки та інформаційні технології. – Львів: Видавництво Львівської політехніки, 2014. – № 800.

– С. 211-216. (Входить до міжнародної наукометричної бази INSPEC.)

6. Bogucharskiy S.I. Image segmentation with fuzzy J-means method // S.I. Bogucharskiy, A.G. Kagramanyan, O.D. Mikhnova // Системні технології. Рег. міжвуз. збірник наук. праць. – Дніпропетровськ: НметАУ, ІВК «Системні технології». – 2014. – Вип. 6 (95). – С.27-34.

7. Богучарский, С.И. Модификация метода J-средних в задачах фрагментной сегментации изображений / С.И. Богучарский, А.Г. Каграманян, С.В. Машталир // Біоніка інтелекту. – 2015. – №1(84) – С. 19-22.

8. Богучарский, С.И. Сегментация изображений в больших базах данных с использованием плотности распределения информации / С.И. Богучарский, В.П. Машталир // Электротехнические и компьютерные системы. – 2015. – 14(90). – С. 119-123. (Входить до міжнародних наукометричних баз Ulrich's Periodicals Directory, Index Copernicus.)

9. Богучарский, С.И. Формализация задачи сегментации видеоданных / С.И. Богучарский, С.В. Машталир // VII Міжнародна школа-семінар «Теорія прийняття рішень»: праці школи-семінару, Ужгород, 29 вересня – 4 жовтня 2014 р. – Ужгород, УжНУ. – 2014. – С. 43-44.

10. Fuzzy J-Means image segmentation / S. Bogucharskiy, V. Mashtalir, O. Mikhnova // Advances in Data Science: Proc. International Workshop and Networking Event, Poland, Holny Mejera, 6-8 May 2015. – Bialystok: BUT, 2015. – P. 15-16.

11. Богучарский, С.И. Обработка изображений с использованием комбинированного векторного квантования / С.И. Богучарский // Обчислювальний інтелект (результати, проблеми, перспективи): праці. III міжнар. наук.-практ. конф., Київ-Черкаси, 12-15 травня 2015 р. – Київ: КНУ, 2015. – С. 55-56.

12. Богучарский, С.И. Фаззи-кластеризация изображений с учетом плотности распределения / С. И. Богучарский С.В. Машталир, Е. Д. Михнова // Интеллектуальные системы принятия решений и проблемы вычислительного интеллекта: сб. научн. трудов междунар. научн. конф., Железный Порт, 25-28 мая 2015 г. – Херсон: ХНТУ, 2015. – С. 251-253.

13. Богучарский, С.И. Иерархическая агломеративная кластеризация в базах данных мультимедиа / С.И. Богучарский, С.В. Машталир // Междунар. науч.-техн. конф. «Электротехнические и компьютерные системы: теория и практика», Одесса, 20-24 июля 2015 г.: Электротехнические и компьютерные системы: специальный выпуск. – 2015. – 19(95). – С. 239-242.

14. Visual means for rotor heat imbalance identification / I. Kalimanova, S. Bogucharskiy, Y. Megel, O. Mikhnova // Metrology and Metrology Assurance: proc. of the 25th nat. scient. symp. with intern. participation, Bulgaria, Sozopol, 7-11 September 2015. – Sofia: TU, 2015. – P. 217-220.

15. Bogucharskiy, S. Image segmentation via X-means under overlapping classes / S. Bogucharskiy, V. Mashtalir // Computer Science and Information Technologies 2015: proc.of the X intern. scient. and techn. conf., Lviv, 14-17 September 2015. – Lviv: LPPH, 2015. – P. 45-47. (Входить до міжнародної наукометричної бази IEEE.)

АНОТАЦІЯ

Богучарський С.І. Методи і моделі толерантної кластеризації у колекціях зображень. – Рукопис.

Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту. – Харківський національний університет радіоелектроніки, Міністерство освіти і науки України, Харків, 2016.

Дисертація присвячена розвитку моделей і методів сегментації візуальної інформації у просторі зображень на основі фрагментної обробки.

Введені матричні модифікації методів кластеризації CLARANS для обробки масивів відеоспостережень у великих базах відеоданих, DBSCAN для сегментації зображень з нечіткими границями між областями довільної форми. Отримані фрагментні методи, засновані на використанні розподілів даних та рекурентній оптимізації, що дозволяють формувати сегменти довільної форми при високому рівні завад. Запропонована матрична модифікація методу кластеризації X -середніх, а також модифікації J -середніх і нечітких J -середніх, що дають можливість забезпечити досягнення більш глибокого екстремуму прийнятої цільової функції якості сегментації. Розроблений метод напівконтрольованого матричного векторного квантування для аналізу текстур – сегментації послідовності зображень, що дозволяє обробляти відеодані як у режимі навчання з вчителем, так і самонавчання у послідовному режимі. Обговорені результати експериментальних досліджень.

Ключові слова: кластеризація, сегментація, зображення, відеодані, фрагментна обробка.

АННОТАЦИЯ

Богучарский С.И. Методы и модели толерантной кластеризации в коллекциях изображений. – Рукопись.

Диссертация на соискание ученой степени кандидата технических наук по специальности 05.13.23 – системы и средства искусственного интеллекта. – Харьковский национальный университет радиоэлектроники, Министерство образования и науки Украины, Харьков, 2016.

Совершенствование и диверсификация программно-аппаратных инструментов, доступных и корпоративным, и индивидуальным пользователям, возрастание объемов визуальной информации в цифровой форме, рост интенсивности ее использования подчеркивают необходимость развития методов сегментации видеоинформации и в пространстве (на отдельных видеокадрах), и во времени (поиск сюжетов видео). Организация данных (грануляция, кластеризация, факторизация) есть суть сегментации – поиска областей носителя изображения или множества изображений на основе заданной или формируемой в процессе обработки модели сходства.

Для устранения семантического конфликта между результатами обработ-

ки изображений на уровне традиционных признаков описаний и интерпретацией визуальной информации на высоком уровне необходимо на среднем уровне (в частности, на основе «пространственных» представлений визуальной информации) ad hoc развивать методы сегментации как в поле зрения, так и во времени при обработке видеопоследовательностей.

Установлено, что определенную перспективу представляет развитие и модификация широкого диапазона прогрессивных методов кластеризации путем их адаптации к фрагментной обработке двумерных полей, что в конечном итоге обеспечивает создание инструментария для толерантной сегментации визуальной информации.

Предложена матричная модификация метода кластеризации CLARANS, предназначенная для обработки массивов видеонаблюдений, хранящихся в больших базах видеоданных. Особенностью введенной модификации является отказ от операций векторизации-девекторизации фрагментов изображений, что позволяет повысить быстродействие процессов сегментации растровых данных, увеличить робастность к аномальным выбросам различной природы, практически всегда присутствующих в реальных видеосигналах. Синтезированная модификация обеспечивает возможность работы в условиях перекрывающихся классов, фактически осуществляя толерантную кластеризацию видеоданных, что в конечном итоге создает предпосылки для повышения эффективности интеллектуального анализа визуальной информации. Предложена матричная модификация метода кластеризации DBSCAN, ориентированная на решение задачи сегментации изображений с нечеткими границами между сегментами произвольной формы и отличающаяся сокращенным количеством оцениваемых параметров, что позволяет повысить быстродействие обработки растровой информации. Получил дальнейшее развитие метод кластеризации, основанный на использовании распределений данных и рекуррентной оптимизации, позволяющий формировать сегменты произвольной формы при высоком уровне помех и отличающийся быстродействием за счет использования матричной функции плотности, базирующейся на матричной сферической норме.

Получена матричная модификация метода кластеризации X -средних, предназначенная для решения задач сегментации визуальной информации и отличающаяся повышенным быстродействием за счет сокращения числа точек анализируемых сегментов и использования модифицированного критерия качества, позволяющая производить обработку информации в последовательном режиме в отличие от традиционных статистических критериев. Для работы в условиях перекрывающихся классов введена модификация матричного метода нечетких C -средних, позволяющая непосредственно в процессе обработки автоматически определять число формируемых сегментов.

Введены матричные модификации методов иерархической агломеративной кластеризации изображений, предназначенные для обработки информации, содержащейся в больших базах данных, и позволяющие сократить объем вычислений, что обеспечивает ускорение обработки в условиях, когда продуцируемые сегменты имеют произвольную форму.

Введены матричные модификации методов кластеризации J -средних и нечетких J -средних, позволяющие обеспечить достижение более глубокого экстремума принятой целевой функции качества сегментации изображений, а также повысить быстродействие процесса обработки информации за счет организации процессов скачков только в одном кластере максимального объема.

Решена задача анализа текстур в последовательности изображений, в основе предложенного метода лежит техника комбинированного векторного квантования. Впервые предложена матричная пространственно-временная динамическая модель последовательности изображений и метод ее настройки на основе матричного нечеткого векторного квантования, что позволило повысить быстродействие процесса классификации и кластеризации текстур.

Впервые предложен метод полуконтролируемого матричного векторного квантования для анализа текстур – сегментации последовательности изображений, позволяющий обрабатывать видеоданные как в режиме обучения с учителем, так и самообучения в последовательном режиме, что также позволяет повысить быстродействие процесс обработки видеoinформации.

Обсуждены результаты экспериментальных исследований предложенных методов фрагментной сегментации и специфика их практического применения. Эксперименты проводились как с отдельными изображениями для выделения их «пространственного» содержания, так и видеопотоками с целью их парсинга, т.е. выделения кластеров, однородных во времени.

ABSTRACT

Bogucharskiy S.I. Methods and models of tolerance clustering in image collection. – Manuscript.

Thesis for the candidate degree of technical sciences in the speciality 05.13.23 – Artificial Intelligence Systems and Means. – Kharkiv National University of Radio Electronics, Ministry Education and Science of Ukraine, Kharkiv, 2016.

The thesis is devoted development of models and methods of visual information segmentation in image space on windows processing basis.

Matrix modifications of clustering methods, viz CLARANS for video streams processing in very large video data bases and DBSCAN for image segmentation with fuzzy boundaries between arbitrary shape regions, are introduced. Window methods based on data distributions and recurrent optimisation, that allow to form clusters with any shape at high noise level, have been proposed. Matrix modification of X -means and also modifications of J -means and fuzzy J -means, which provide possibilities to ensure reaching of deeper extremum of the accepted goal function of segmentation quality, is offered. The method of semicontrolled matrix vector quantization is developed for textures analysis (segmentations of image sequences), allowing to process video data in sequential mode both learning with the teacher and selflearning. Results of experimental investigations are discussed.

Keywords: clustering, segmentation, image, video data, window processing.