

АНАЛІЗ ТА КЛАСИФІКАЦІЯ ОСНОВНИХ МЕТОДІВ РОЗПІЗНАВАННЯ ОБРАЗІВ НА ПЛОЩИНІ ПРОЕКЦІЇ

© Басюк Т., Пушко Я., 2015

Проаналізовано основні методи розпізнавання образів у інформаційних системах. Розглянуто проблеми у сфері розпізнавання і методи їх вирішення. Наведено критерії для здійснення порівняльного аналізу та сформульовано висновки щодо можливості використання описаних методів для пошуку зображення лазера на площині проекції. В результаті подальшого дослідження здійснено системний аналіз предметної області згідно з методологією Rational Unified Process, що дало змогу сформулювати розширений набір прецедентів. Проведено моделювання з використанням засобів мови UML, що сприяло специфікації та документуванню артефактів, які необхідні для розроблення програмної системи.

Ключові слова: розпізнавання образів, класифікація методів розпізнавання, евристичні методи, математичні методи, лінгвістичні методи.

The paper analyzes the main methods of pattern recognition in information systems. The existing problems in the recognition and methods for their solution are studied. Criteria for comparative analysis are provided and conclusions about the possibility of using the described methods for searching the laser image on the plane projection are formulated. As a result of further investigation carried out a systematic analysis of the subject area on the methodology Rational Unified Process, which made it possible to form an expanded set of precedents. The simulation with the use of language UML, which contributed to the specification and documentation of artifacts that are needed in the development of a software system.

Key words: pattern recognition, classification of recognition methods, heuristic methods, mathematical methods, linguistic methods.

Вступ

Сьогодні діяльність будь-якої галузі промисловості чи науки не можна уявити собі без проекційних засобів, які незамінні у процесі відображення значного масиву інформації для широкого загалу. Інтенсивність впровадження цих засобів сприяє розвитку відповідних технологій для їх реалізації. Найперспективніше застосування технологій: DLP та LCOS [1]. DLP (англ. Digital Light Processing) – технологія, яка широко використовується у проекторах, основою якої є електромеханічна система, що створює зображення за допомогою мікроскопічних дзеркал, розташованих у вигляді матриці на напівпровідниковому кристалі DMD (Digital Micromirror Device) [2]. LCOS (англ. Liquid Crystal on Silicon) використовує спеціальні LCoS-кристали, що складаються з рідкокристалічної матриці та поляризатора. Під дією електричних імпульсів рідкі кристали перебувають або у відкритому, або у закритому стані, тим самим пропускаючи або відбиваючи світло від підкладки LCoS-кристала [3].

Особливістю цих технологій є висока якість проекції та можливість інтерактивного керування процесом її опрацювання. Проте недоліком є відсутність мобільності та висока ціна, зокрема, проектор NEC VE281, побудований на основі технології DLP, коштує 7 540 грн, що є значним недоліком для освітніх закладів, які в сучасному інформаційному суспільстві прагнуть забезпечити якісний виклад інформації. Як вихід з цієї ситуації пропонується використання зв'язку: комп'ютер –

веб-камера – лазерна указка. Цей ланцюжок характеризується і значною мобільністю, і відсутністю додаткових засобів для реалізації інтерактивної взаємодії доповідача з проекційним зображенням. Проте його реалізація вимагає розв’язання множини задач, таких як визначення типу лазера, його розпізнавання на площині, організація взаємодії тощо.

Зв’язок висвітленої проблеми із важливими науковими та практичними завданнями

Проектування системи розпізнавання лазера на площині проекції є важливим завданням у межах розвитку як IT-галузі, так і промисловості загалом, оскільки немає жодного підприємства чи суб’єкта господарювання, в арсеналі якого не застосовується проекційне обладнання, що сприяє необхідності розвитку відповідного математико-алгоритмічного апарату для його реалізації.

Внаслідок неординарності поставленої задачі немає загального її розв’язання для різних типів лазера та проекції, тому з певними допущеннями можна застосовувати як методи розпізнавання зображень, так і технології штучного інтелекту. Побудова нових методів опрацювання сприятиме розв’язанню задачі розпізнавання та забезпечить істотне здешевлення інтерактивного обладнання.

Аналіз останніх досліджень та публікацій

Створення систем підвищення інтерактивності взаємодії для забезпечення як промислового, так і навчального процесу набуває все більшої популярності, оскільки представлення значних потоків інформації вимагає застосування як нових методів, так і засобів для їх коректного відображення [4].

У роботах [5–6] описано можливі підходи до розв’язання описаної задачі, проте за всієї різноманітності праць у жодному з наведених досліджень автори не проаналізували основні методи розпізнавання зображень на площині проекції з метою їх застосування для вирішення поставленого науково-практичного завдання.

Основні завдання дослідження та їх значення

Основними завданнями дослідження є: аналіз методів розпізнавання зображень та оцінка їх ефективності під час побудови системи розпізнавання зображень на площині проекції. Відповідно до вказаних завдань дослідження, для їх вирішення необхідно сформулювати критерії для порівняння методів розпізнавання та вибору найоптимальнішого, що можна застосувати для розв’язання досліджуваної задачі; проаналізувати етапи розпізнавання зображень та сформулювати основні концепції побудови прикладної системи розпізнавання образів на площині.

Розв’язання зазначених задач дасть змогу спроектувати прикладну систему, яка надасть засоби інтерактивної взаємодії з елементами проекції з мінімальною кількістю як апаратних, так і програмних ресурсів.

Основні результати досліджень

Методи розпізнавання образів використовуються в різноманітних сферах життя людини: у медицині, металургії, банківській справі, біометрії тощо. У кожній галузі розпізнавання виконує певну роль. Це може бути як контроль якості, наприклад, у банкоматі під час ідентифікації купюри, так і процеси автентифікації користувачів. Для кожної галузі застосовуються свої методи, оскільки саме вони дають змогу отримати найкращі результати за заданих вхідних даних. Для розв’язання досліджуваної задачі можна застосувати різні класи методів: евристичні, лінгвістичні та математичні [6].

Якщо опис об’єктів розпізнавання здійснюється за допомогою підобразів і їх співвідношень, то для конструювання системи використовують лінгвістичний або синтаксичний підхід з використанням принципу загальності властивостей. Основне припущення, яке робиться у цьому методі, ґрунтується на тому, що образи, які належать тому самому класу, володіють низкою загальних властивостей або ознак, які відображають їх подібність. У випадку представлення неklasифікованого образу вибирається набір ознак, які порівнюються із наявними в системі розпізнавання. У разі використання цього методу основне завдання полягає у виділенні загальних властивостей образів, причому їхня належність до цього класу відома [5].

Найпростішим методом, що реалізується з використанням цього підходу, є розпізнавання образів на основі їх порівняння з еталонами. Класифікація ґрунтується на раніше вибраному критерію зіставлення подібності. Інакше кажучи, якщо вхідний образ краще відповідає еталону i -го класу, то він класифікується як його представник. Такий підхід використовується для розпізнавання друкованих літер та банківських чеків. Недоліком цього підходу є складність вибору еталона, який найбільше підходить до кожного класу образів, та встановлення необхідного критерію відповідності. Зазначені труднощі особливо істотні, коли образи належать до одного класу, типовим прикладом якого є розпізнавання рукописних літер.

В основу математичного підходу покладено правила класифікації, які формуються і виводяться в межах визначеного математичного формалізму за допомогою принципів загальності властивостей і кластеризації. Побудова системи розпізнавання, яка основана на реалізації цих принципів, визначається просторовим розміщенням окремих кластерів. Якщо кластери, що відповідають різним класам, рознесені один від одного, то застосовують прості схеми розпізнавання, наприклад, класифікацію за принципом мінімальної відстані [6].

В основу евристичного підходу покладено поняття інтуїції та досвіду: використовуються принципи перелічення членів класу та загальні властивості. Зазвичай системи, побудовані на основі таких методів, містять набір специфічних процедур, розроблених для конкретних задач розпізнавання. Структура і якість евристичної системи великою мірою визначається досвідом розробників. Найчастіше цей клас методів використовується у задачах біометрії [7].

Наведений аналіз свідчить, що сьогодні існує множина методів розпізнавання, які застосовуються в різних галузях та ґрунтуються на різних критеріях. Можливість їх використання для розв'язання цієї задачі актуалізує процедуру їх детального аналізу та класифікації.

З метою проведення аналізу та здійснення класифікації пропонується застосовувати такі критерії оцінювання: кількість зразків для збереження, розмірність ознакового простору, ризик втрати, кількість образів, що розпізнаються, можливість безпосереднього використання, швидкодія роботи. Результати порівняння наведено в таблиці.

Порівняння методів розпізнавання

Методи	Критерії порівняння					
	A	B	C	D	E	F
Метод масок	Один ($a=1$)	Мінімальна	Середній	Залежить від кількості збережених класів образів	Так	повільно ($f>2$)
Метод фрагментів	Один ($a=1$)	Середня	Мінімальний	Залежить від кількості збережених класів образів	Так	Швидко ($f<1$)
Метод проекції	Один ($a=1$)	Велика	Великий	Залежить від кількості збережених класів образів	Так	Середньо ($1<f<2$)
Метод аналізу штрихових елементів	Декілька ($1<a<10$)	Велика	Мінімальний	Формується інженером із знань	Ні	Повільно ($f>2$)
Метод синтезу системи	Декілька ($1<a<10$)	Середня	Середній	Формується інженером із знань	Ні	Середньо ($1<f<2$)
Метод потенціальних функцій	Декілька ($1<a<10$)	Середня	Середній	Невелика кількість ($d<=3$)	Ні	Середньо ($1<f<2$)
Метод кластерів	Багато ($a>10$)	Велика	Середній	Формується інженером із знань	Так	Швидко ($f<1$)

Кількість зразків для збереження (A) – цей критерій вказує кількість навчальних зразків, що містяться в системі, та відповідну кількість порівнянь, які необхідно здійснювати. Що стосується цієї задачі, то він є критичним, оскільки, з одного боку, під час збереження всіх навчальних зразків зменшується ризик нерозпізнавання зображення, а, з іншого, потребує великих обсягів пам'яті та обчислювальних ресурсів.

Розмірність ознакового простору (B) здебільшого є мінімальним значенням, оскільки сприяє зменшенню кількості необхідних вимірів, спрощується процес обчислення, збільшується статична стійкість результатів розпізнавання. Разом з тим зменшення розмірності ознакового простору збільшує ризик втрати зображення. З огляду на це, визначення цього критерію є компромісом, який можна розділити на дві частини: формування початкового ознакового простору та його мінімізація. Початковий ознаковий простір зазвичай формується на основі досвіду та інтуїції (основна мета, яка ставиться – мінімізація ризику втрат), а мінімізація розмірності реалізується із застосуванням формальних методів та відповідних алгоритмів розпізнавання.

Ризик втрати (C) – критерій, за яким формується найінформативніший ознаковий простір. Кількість навчальних зразків має бути такою, щоб, за можливістю, мінімізувати ризик втрати зображення. Крім того, цей критерій формують обчислювальні затрати на ліквідацію помилок розпізнавання та виміри ознак розпізнаваних об'єктів.

Кількість образів, що розпізнаються (D), завжди скінченна і не може бути меншою від двох значень. Іноді розглядаються випадки, коли такий образ лише один, проте цей випадок є виродженим, оскільки всі реалізації належать до того самого класу. Перелік образів, що розпізнаються, може надаватися системі ззовні (інженером із навчання) або формуватись в автономному режимі.

Можливість безпосереднього використання (E) – цей критерій вказує на необхідність попередньої обробки зображення з використанням досліджуваного методу під час фільмування зображення відеокамерою.

Швидкість роботи (F) – важливий критерій, що безпосередньо впливає на роботу всієї системи. Швидкі методи можуть призводити до втрати образу розпізнавання, тим самим призводячи до помилок роботи всієї системи, а повільні, навпаки, унеможливають «комфортну» роботу користувача. Оскільки робота проводиться в реальному масштабі часу, то необхідно забезпечити компроміс між швидкістю роботи та потрібною якістю розпізнавання.

Як показує аналіз, наведені методи не можуть гарантувати виконання всіх критеріїв під час розпізнавання зображення лазера на площині. Отже, частковим варіантом використання може бути застосування методів фрагментів чи масок, які утворюють множину методів евристичного аналізу.

Системний аналіз предметної галузі

Що стосується реалізації, то з метою успішної побудови будь-якої складної інформаційної системи необхідно виконати ґрунтовний системний аналіз предметної області, що дає змогу визначити мету і призначення системи, зрозуміти структуру і механізми функціонування. Аналіз проводиться згідно з методологією Rational UnifiedProcess [8]. Найінформативнішим при цьому є розгорнутий опис прецедентів:

- *зацікавлені особи прецеденту та їхні вимоги*: прикладний програміст – потребує наявності відповідного програмного та апаратного забезпечення; адміністратор – вимагає наявності простого механізму встановлення та налаштування; користувач – хоче отримати просту у використанні систему з повною функціональністю;
- *основний актор прецеденту* – користувач, що використовує розроблену систему, щоб збільшити інформативність доповіді;
- *передумови прецеденту* – успішність налаштування програмної системи;
- *основний успішний сценарій* – налаштування системи; ідентифікація вказівника лазера; керування процесом відображення даних; завершення роботи;

- спеціальні системні вимоги – забезпечення роботи в реальному часі; створення множини фільтрів з метою покращення розпізнаваного зображення; керування вказівником лазера;
- додаткові пристрої – веб-камера, проектор, лазерна указка (лазер);
- розширення основного сценарію або альтернативні потоки – полягають у можливій появі помилок в процесі роботи системи, а саме неможливості ідентифікації користувачького вказівника. Альтернативні потоки:
- фон площини або проекції збігається з кольором керуючого вказівника (система виводить повідомлення про помилку ідентифікації користувачького вказівника і переходить в стан очікування; адміністратор змінює фон проекції; програмна система відновлює роботу та переходить у звичний режим);
- на площині є декілька вказівників (система виводить повідомлення про помилку ідентифікації, продовжуючи відстежувати попередній вказівник; адміністрування ситуації та продовження функціонування).

Наступним етапом розроблення стало проведення моделювання з використанням засобів мови UML з метою специфікації та документування всіх артефактів, які необхідні для розроблення програмної системи [9]. Розроблено такі діаграми:

- **варіантів використання (use case)** – акторів цієї діаграми подано в розгорнутому описі прецедентів (рис.1). Акторами в цій діаграмі виступають адміністратор, прикладний програміст та користувач. Користувач – актор, який здійснює керування комп'ютером та володіє обмеженим функціоналом системи: зміною слайдів, збільшенням частини екрана тощо. Адміністратор – актор, що налаштовує систему, вибирає робочу поверхню проекції, змінює набір фільтрів. Прикладний програміст – змінює внутрішню структуру системи шляхом модифікації алгоритмів роботи.

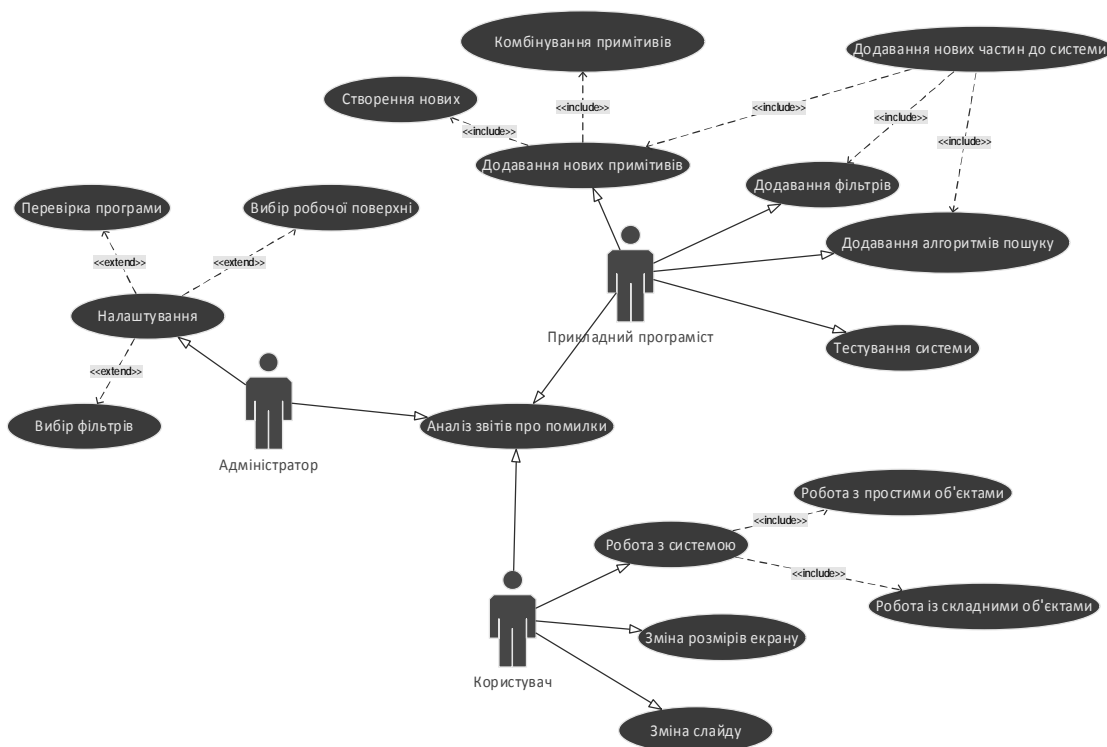


Рис.1. Діаграма варіантів використання

- **класів (class)** – використовується для представлення статичної структури моделі системи в термінології класів об'єктно-орієнтованого програмування (рис. 2). Основні класи: камера

(Camera Connect), журнал (Logger), фільтр (Filters), розпізнавання (Recognition), позиція курсора (CursorPos), зв'язок з операційною системою (OSconnect), графічний інтерфейс (GUI). Клас *камера* містить відомості щодо засобу спостереження, способів та типу підключення, швидкості роботи тощо. *Журнал* є особливим класом, оскільки здійснює не лише протоколювання роботи системи, але й реалізує функції інформування користувача про можливі помилки запису/читання.

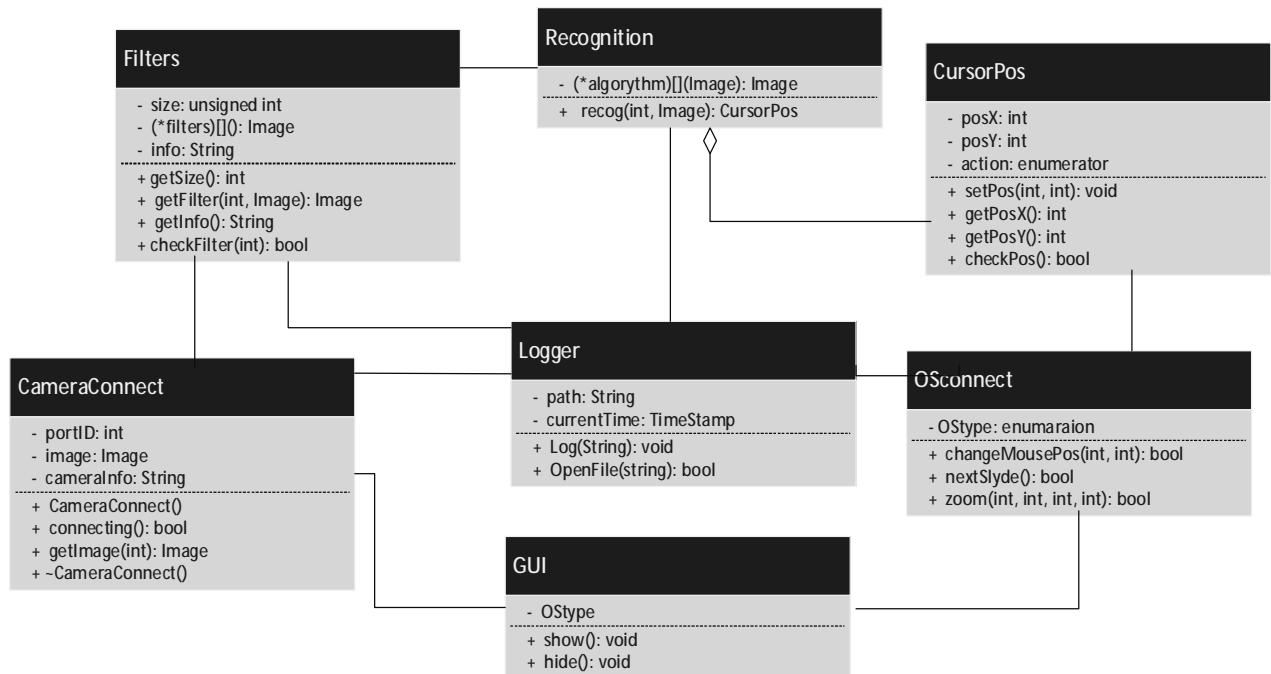


Рис. 2. Діаграма класів

Клас *фільтр* є одним з основних класів, оскільки містить методи, які застосовуються для покращення розпізнавання зображення. Клас *розпізнавання* містить інформацію про масив вказівників на функції, що реалізують процедуру розпізнавання зображення. В класі *позиція курсора* зберігаються дані щодо поточного розташування вказівника та перелік дій, які він передбачає. Клас *зв'язок з операційною системою* реалізує інтерфейси доступу до API функцій, які забезпечують необхідний функціонал. Графічний інтерфейс містить методи побудови графічного оформлення програмного продукту;

- **послідовності (sequence)** – використовуються для відображення в часі взаємодії об'єктів та особливостей передавання і приймання повідомлень. Наведені об'єкти (рис. 3) обмінюються повідомленнями згідно з потоками, представленими на діаграмі. Після закінчення роботи адміністратор завершує роботу системи.
- **активності (activity)** – являє собою статичне представлення дій системи під час аналізу активності користувача на площині проекції (рис. 4).

Наведена діаграма ілюструє можливі дії користувача під час роботи з системою. А саме: на початку роботи система вибирає необхідну множину фільтрів та застосовує певний алгоритм розпізнавання. Далі відбувається власне робота із системою (переміщення курсора, масштабування зображення чи зміна слайда). Після завершення роботи відповідні відомості надсилаються засобам операційної системи та відбувається термінація програмного засобу

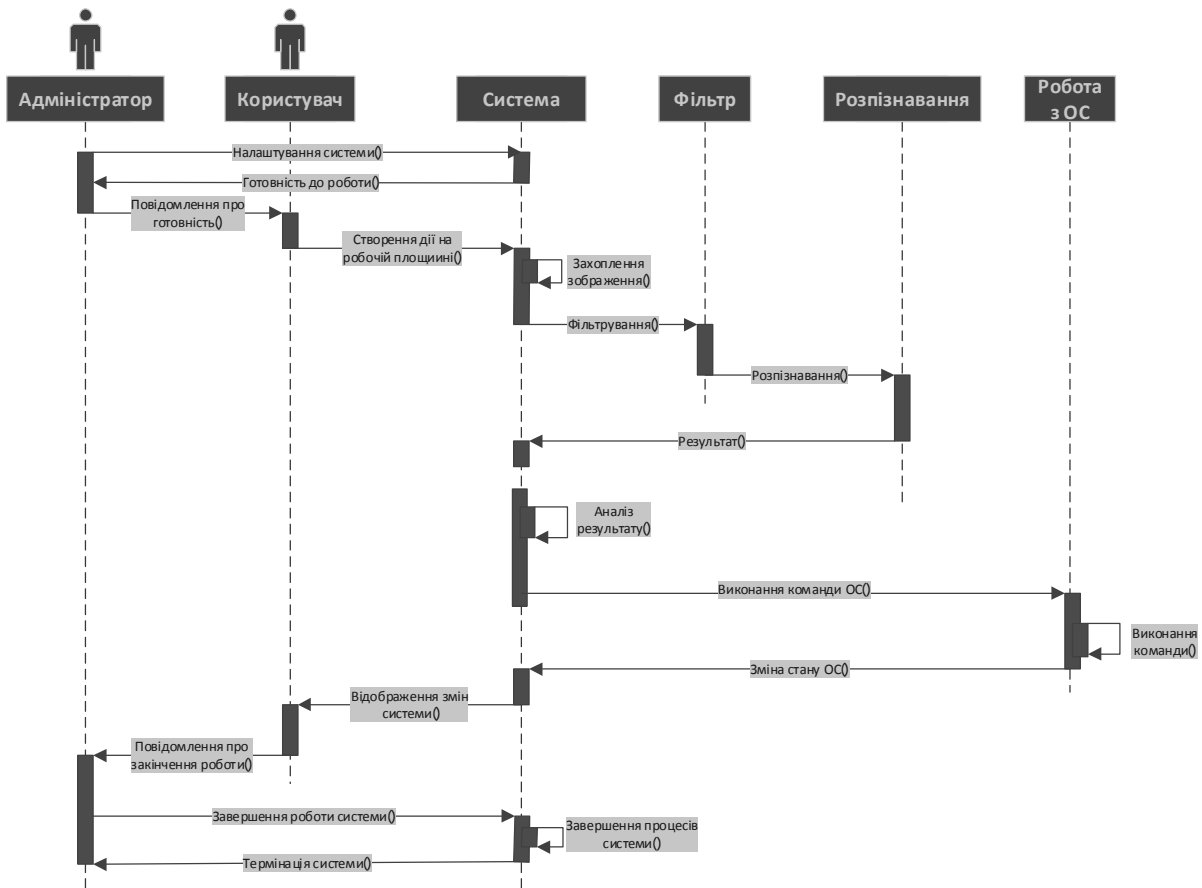


Рис. 3. Діаграма послідовності

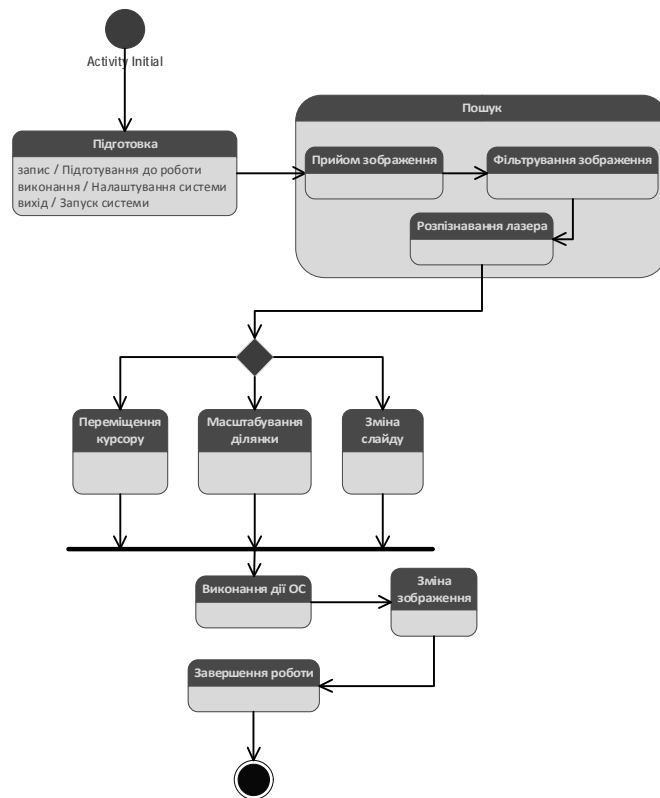


Рис. 4. Діаграма активності

Висновки

В результаті проведеного дослідження розглянуто актуальні проблеми у сфері розпізнавання образів та описано основні підходи, що застосовуються для їх вирішення. На основі здійсненого аналізу розроблено критерії, які застосовано для проведення порівняльного аналізу методів розпізнавання лазера на площині проекції. Наведена класифікація надала засоби з підбору методів, які застосовані в процесі конструювання прототипу функції пошуку зображення лазера на площині. Подальші дослідження стосуватимуться проектування модулів системи відповідно до створеного методологічного підґрунтя та їх тестування на відповідність поставленим вимогам.

1. Чепмен Н. Цифровые технологии мультимедиа / Н. Чепмен, Д. Чепмен. – М.: Вильямс, 2006. – 624 с.
2. How DLP Technology Works [Electronic Resource] – Mode of access: <http://www.dlp.com/technology/how-dlp-works/default.aspx>. – 21.01.2015. – The name of the screen title
3. LCoS Projectors [Electronic Resource] – Mode of access: <http://www.projectorreviews.com/projector-categories/lcos-projectors>. – 21.01.2015y. – The name of the screen title
4. Мерков А. Б. Распознавание образов. Введение в методы статистического обучения / А. Б. Мерков. – М.: Эдиториал, 2011. – 256 с.
5. Форсайт Д. А. Компьютерное зрение. Современный подход / Д. А. Форсайт, Д. Понс – М.: Вильямс, 2004. – 928 с.
6. Bishop C. M. Neural Networks for Pattern Recognition / C. M. Bishop – Oxford University Press, 1996. – 504 p.
7. Блог “Розпізнавання образів” [Електронний ресурс] – Режим доступу: <http://patterns-recognition.blogspot.com/> – 12.11.2014y. – Назва з титулу екрана.
8. Крачтен Ф. Введение в Rational Unified Process / Ф. Крачтен. – СПб.: Вильямс, 2002. – 240 с.
9. Рамбо Д. UML 2.0. Объектно-ориентированное моделирование и разработка / Дж. Рамбо, М. Блаха. – СПб.: Питер, 2007. – 544 с.