

С.М. Єфремов, Т.А. Зайцева

Дніпровський національний університет імені Олеся Гончара

МОДЕЛЮВАННЯ ПОШУКУ НЕВІДОМИХ ЛОКАЛЬНО-ПОВТОРЮВАНИХ ПАТЕРНІВ НА ЗОБРАЖЕННЯХ

В роботі розглядається алгоритм контекстного моделювання обмеженого порядку для знаходження невідомих локально-повторюваних патернів на зображенні. Виконана програмна реалізація алгоритму, та на тестових прикладах проведена оцінка роботи програми.

Ключові слова: патерн, контекстне моделювання обмеженого порядку, комп'ютерний зір, диспропорція зображення, кут спостереження.

S.M. Yefremov, T.A. Zaytseva

Oles Honchar Dnipro National University

MODELLING OF THE SEARCH FOR UNKNOWN LOCALLY REPEATED PATTERNS IN IMAGES

The ability of detecting unknown repeated patterns that appear more than one time in a digital image is a serious challenge. The possibilities of the optimization of pattern recognition using finite-context modelling is described based on the provided image examples. The algorithm doesn't change in regards to the input data and suits for the repeating pattern recognition based on any image. The algorithm idea is taken form the work of Diogo Pratas and Armando J. Pinho [1], which is a continuation of their previous studies aimed at improvement in detection of unknown repeated patterns that occur in a close proximity, allowing to utilize a compression based approach to find exact repetitions. Finite-context modelling is being used to pinpoint the possible locations of the repetitions by exploring the connection between lossless image compression and image complexity. The additional step has been added to account for the viewing angle of the camera in order to get the context of how the objects in image actually look in correspondence to how they look in the image. Mentioned addition helps to improve the precision of input image and, as a consequence, pattern recognition itself. The experimental results show that proposed approach provides increased ability to eliminate false positives.

The computer program has been developed according to the selected method of finite-context modelling calculation. It was used during the image content features description, specifically unknown locally repeating patterns and in the further computer vision research. The software implementation of the algorithm takes into account the capabilities of modern computer technology, Javascript programming language and modern needs of the image content repetitions recognition speed and accuracy. The implementation of computer program logic is shown in the script examples, utilizing modularity and parallel calculation as the strong sides of Javascript, further improving repeating patterns recognition speed.

Keywords: pattern, finite-context modelling, computer vision, image disparity, viewing angle.

С.М. Єфремов, Т.А. Зайцева

Дніпровський національний університет імені Олеся Гончара

МОДЕЛИРОВАНИЕ ПОИСКА НЕИЗВЕСТНЫХ ЛОКАЛЬНО-ПОВТОРЯЮЩИХСЯ ПАТТЕРНОВ НА ИЗОБРАЖЕНИЯХ

В работе рассматривается алгоритм контекстного моделирования ограниченного порядка для нахождения неизвестных локально-повторяющихся паттернов на изображении. Выполнена программная реализация алгоритма и на тестовых примерах продемонстрированы результаты работы программы.

Ключевые слова: паттерн, контекстное моделирование ограниченного порядка, компьютерное зрение, диспропорция изображения, угол наблюдения.

Вступ. Знаходження невідомих повторюваних патернів, що з'являються декілька разів на цифровому зображенні є досить складною задачею. На даний час розроблено ряд методів і алгоритмів аналізу та синтезу зображень, які ґрунтуються на різних теоретичних засадах. Підхід на основі стиснення достатньо ефективним робить знаходження точних повторень образів у змісті зображення [1]. На відміну від підходу з використанням тренованої згорткової нейронної мережі [2], метод для знаходження невідомих повторюваних патернів [1] є найбільш вдалим варіантом для розрахунку та класифікації отриманих результатів. У роботах [1, 3] використовується контекстне моделювання обмеженого порядку для розпізнавання змісту зображення та виявлення повторюваних патернів. Досліджується зв'язок між стисненням зображення та погіршенням якості, а також надається порівняльна оцінка можливості виявлення якісного зображення у залежності від рівню складності змісту зображення. Треба відзначити, що роботи присвячені розпізнаванню повторюваних патернів на фасадах будинків, таких як вікна чи елементи декору та ін. стають все більш популярні [4].

В загальному вигляді задача пошуку патернів сформульована в [5]. Запропоновано низку різних методів виділення ознак зображень для вирішення задачі виявлення повторюваних патернів.

Широко застосовуються перцептивні хеш-алгоритми [6-8], які описують клас функцій для генерації індивідуального відбитка (хеша) для їх подальшого порівняння. Перцептивний хеш являє собою розвиток концепції криптографічних хеш-функцій.

В роботі [1] реалізується та досліджується алгоритм, що дозволяє поліпшувати розрахунок та знаходити невідомі повторювані патерни на зображеннях, позбувшись надлишкових операцій при кодуванні, при побудові контекстної моделі обмеженого порядку. Особливо важливим аспектом точного розпізнавання є кут спостереження, під яким образи містяться на зображенні. Оскільки зміст зображення може бути розташований під певним кутом [9]

(позиція камери щодо будинку тощо), то його врахування є досить важливим фактором для максимально точного розпізнавання, компенсуючим вплив перспективи на геометрію образів.

Постановка задачі. Нехай маємо область розпізнавання на зображенні розміром $W_x \times H_y$, де W_x – ширина зображення у пікселях, H_y – висота зображення у пікселях. Розглянемо побудову поверхонь складності ϕ із використанням контекстної конфігурації:

$$\phi_{i,j} = -\log_2 P(F_{i,j} = f_{i,j} | c_{k,i,j}), \quad (1)$$

де P – обчислювальні ймовірності, $f_{i,j}$ – кожен піксель вхідного зображення f , $c_{k,i,j}$ – двовимірний контекст порядку k , i, j – координати пікселів.

Необхідно покращити результати знаходження невідомих повторюваних патернів, розробити програмне забезпечення, що містить імплементацію методу контекстного моделювання обмеженого порядку для побудови поверхонь складності [1] та дослідити комплексні випадки змісту зображень, для випадків коли кут спостереження впливає на геометрію образів.

Необхідно програмно реалізувати даний метод максимально адаптовано до можливостей мови програмування Javascript з урахуванням сучасних технологій обробки зображень.

Основний матеріал. Основна ідея полягає у використанні контекстного моделювання обмеженого порядку для дослідження накопичених знань, а саме всіх минулих локальних входжень патернів.

Контекстна модель обмеженого порядку накопичує статистичну інформацію. Для кожного результату вона призначає ймовірнісні оцінки символам алфавіту $A = \{s_1, s_2, \dots, s_A\}$, де $|A|$ є розміром A . Оцінки обчислюються з урахуванням обумовлюючого контексту, обчислююмого за кінцевим фіксованим числом $k > 0$ з минулих розрахунків контекстної моделі $x_{n-k+1..n} = x_{n-k+1} \dots x_{n-1} x_n$ порядку k . У разі даних зображень недавнього минулого, зазвичай, посиляються на просторову близькість. Отже $x_{n-k+1..n}$ може посилатись на набір k просторово найближчих один до одного зразків та не обов'язково k останніх оброблених зразків.

Ймовірності $P(X_{n+1} = s | x_{n-k+1..n})$, $\forall_s \in A$, обчислюються за допомогою підрахунку символів, які накопичуються під час обробки зображення. Тут кількість символів накопичується відповідно до вікна фіксованого розміру останніх символів, відкидаючи решту. Таким чином, коли розмір вікна менший, модель досліджує «місцеву» статистику. З іншого боку, коли вікно має попередній розмір, модель досліджує всю статистику зображень. Ймовірності оцінюються за формулами:

$$P(X_{n+1} = s | x_{n-k+1..n}) = \frac{C(s | x_{n-k+1..n}) + \alpha}{C(x_{n-k+1..n}) + \alpha |A|}, \quad (2)$$

$$C(x_{n-k+1..n}) = \sum_{a \in A} C(a | x_{n-k+1..n}). \quad (3)$$

У формулі (2) $C(s|x_{n-k+1..n})$ являє собою кількість разів, коли раніше джерело інформації генерувало символ s , що має контекст $x_{n-k+1..n}$. Це загальна кількість подій, що відбулись до цього часу у зв'язку з контекстом $x_{n-k+1..n}$.

Параметр α забезпечує баланс оцінки максимальної правдоподібності для рівномірного розподілу, не даючи генерувати нульові ймовірності. Можна побачити, що коли загальна кількість подій n є великою, дана оцінка поводиться по суті як статистична оцінка максимальної правдоподібності. Для $\alpha = 1$ – це медіана – відомий оцінювач Лапласа, що мінімізує ризик відносно функції втрати абсолютного відхилення.

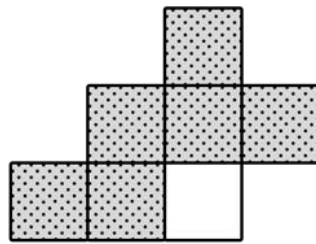


Рис. 5. Контекстний шаблон порядку шість, застосований у контекстній моделі обмеженого порядку

Було застосовано моделювання з обмеженим контекстом для побудови зображень поверхонь складності ϕ . Контекстна конфігурація, зображена на рис. 1. Поверхня складності ϕ – це зображення з такою ж геометрією, що і вхідне зображення f , де кожен піксель $\phi_{i,j}$ (1) відповідає довжині коду, необхідного для кодування $f_{i,j}$, оціненого за моделлю з обмеженим контекстом, де $c_{k,i,j}$ позначає (зазвичай) двовимірний контекст порядку k , а i, j – координати пікселів.

Зображення скануються піксель за пікселем у порядку растрового сканування. Якщо певний візерунок A знайдено вперше, йому буде потрібно надати певну кількість бітів, що буде відповідати певній складності. Коли цей самий шаблон знову зустрінеться, кількість бітів, необхідних для кодування цього входження, буде меншою, оскільки кодер вже побудує внутрішнє представлення шаблону. Використовуючи цей підхід, першим виникненням шаблону можна було б знехтувати. За бажанням цю залежність легко видалити: досить відсканувати зображення в декількох напрямках і зберегти мінімальну складність, виявлену при всіх скануваннях.

Для моделювання з обмеженим контекстом потрібні ресурси пам'яті, які зростають експоненціально із розміром алфавіту. Як правило, навіть зображення сірого рівня використовують алфавіти, що робить ці моделі досить складними для реалізації через потреби у великих ресурсах. Щоб впоратися з цією проблемою, перед обчисленням поверхонь складності ϕ ми виконуємо зменшення кількості інтенсивностей до двадцяти максимум, використовуючи квантування Ллойда-Макса [1].

Результати чисельних експериментів. Досліджено здатність застосованого методу знаходити як точні, так і наближені повторення. На рис.2 представлені пари поверхонь різної ступені складності. У кожній парі зліва – оригінал, справа – побудована поверхня обраним методом. Складність поверхні зображення зростає від темних до світлих областей, тобто, зверху вниз по кожному стовбцю пар. На першому зображенні першого рядка рис. 2 можна побачити, що метод виконує точні повторення, виявляючи два однакові квадрати. У першому стовбці, зверху вниз, кожна з пар зображень мають диспропорції у 0%, 10% та 20% пікселів відповідно. Другий стовбець, зверху вниз, містить пари зображень із диспропорцією у 30%, 40% та 50% пікселів відповідно. Для більшого рівня складності поверхні, природньо, що ефективність методу знижується.

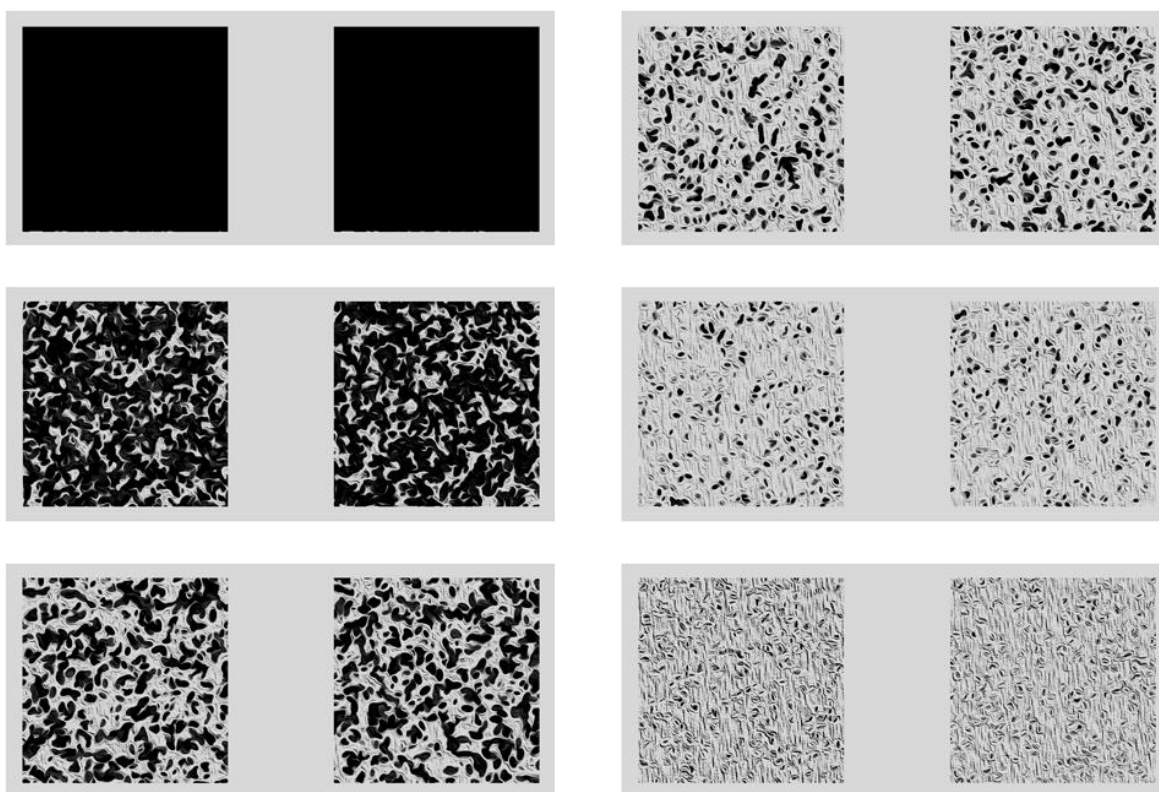


Рис. 6. Складність поверхні зображення, зростаюча від темних до світлих областей

Інколи необхідно знайти невідомі повторювані патерни, наприклад, вікон та елементів декору на будинках для подальшого поліпшення при побудові тривимірної моделі, коли застосовується один і той самий об'єкт (вікно тощо) декілька разів. Даний метод є вдалим для досягнення вищезазначеної мети.

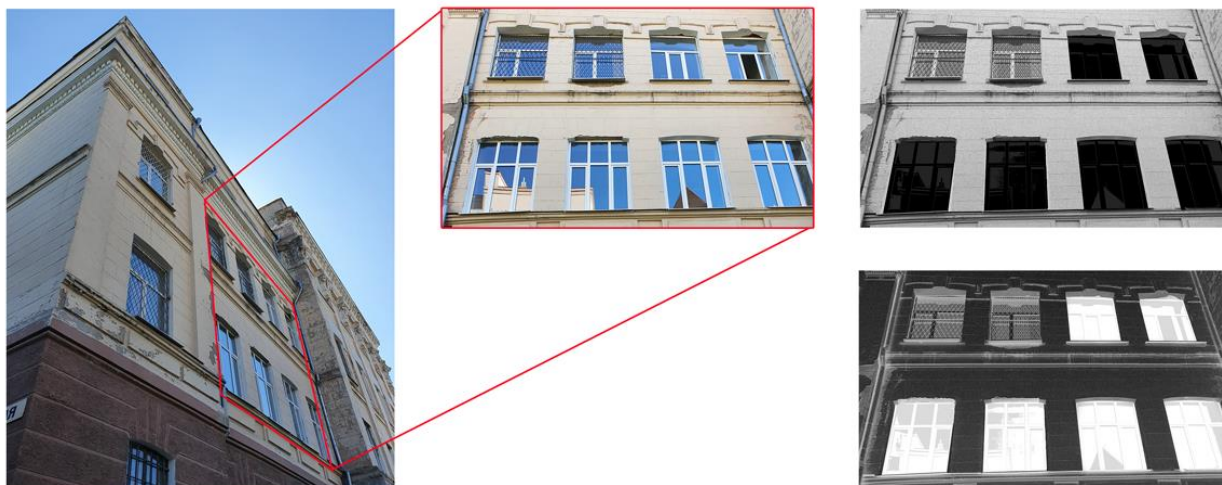


Рис. 7. Фрагмент фасаду корпусу №3 ДНУ з демонстрацією поетапних перетворень.

На рис. 3 продемонстрована ефективність методу для більш складних фігур (в даному випадку – вікна). Відбувається моделювання фрагменту фасаду корпусу №3 Дніпровського національного університету імені Олеся Гончара. Як можна побачити на зображеннях, метод добре працює, як для поверхні в цілому, так і для локальних фрагментів. Однак ракурс оригінальної фотографії, а саме кут зору, під яким відображаються вікна будинку, значно ускладнює задачу, оскільки наявність перспективи створює додаткову диспропорцію між вікнами.

На зображенні ліворуч, що є оригінальною фотографією фасаду корпусу, зробленою під кутом, вказана область, що винесена на центральне зображення, для обробки запропонованим методом. Зображення праворуч демонструють результати комп'ютерного моделювання поверхонь різної складності відповідно.

Для покращення результатів, важливим кроком є виявлення саме кута, під яким зроблено фотографію, а саме під яким камера бачить сцену, у даному випадку, фрагмент фасаду корпусу №3. Маючи можливість виявити кут спостереження камери, наступним кроком є трансформація зображення для відображення дійсної форми об'єктів (вікон), що є необхідним для покращення отриманих результатів, надаючи більш точне представлення образів на зображенні. Це необхідно не лише для покращення початкових даних, але й для виявлення повторюваних патернів, оскільки якщо представлення образів на сцені є неточним (рис. 3), то диспропорція образів на зображенні значно зростає, перевищуючи 20% та значно знизить ефективність даного методу обробки зображень.

Розглянемо інший випадок, що містить ще більш складні образи (рис. 4). Маємо два види об'єктів: лампи та антаблементи колон. Як і на рис. 3, кут спостереження ще більш ускладнює ситуацію. У даному випадку вдалось ідентифікувати лампи, оскільки вони мають форму близьку до сферичної, тому кут спостереження не надав значного впливу на форму та подальшу ідентифікацію даних образів.

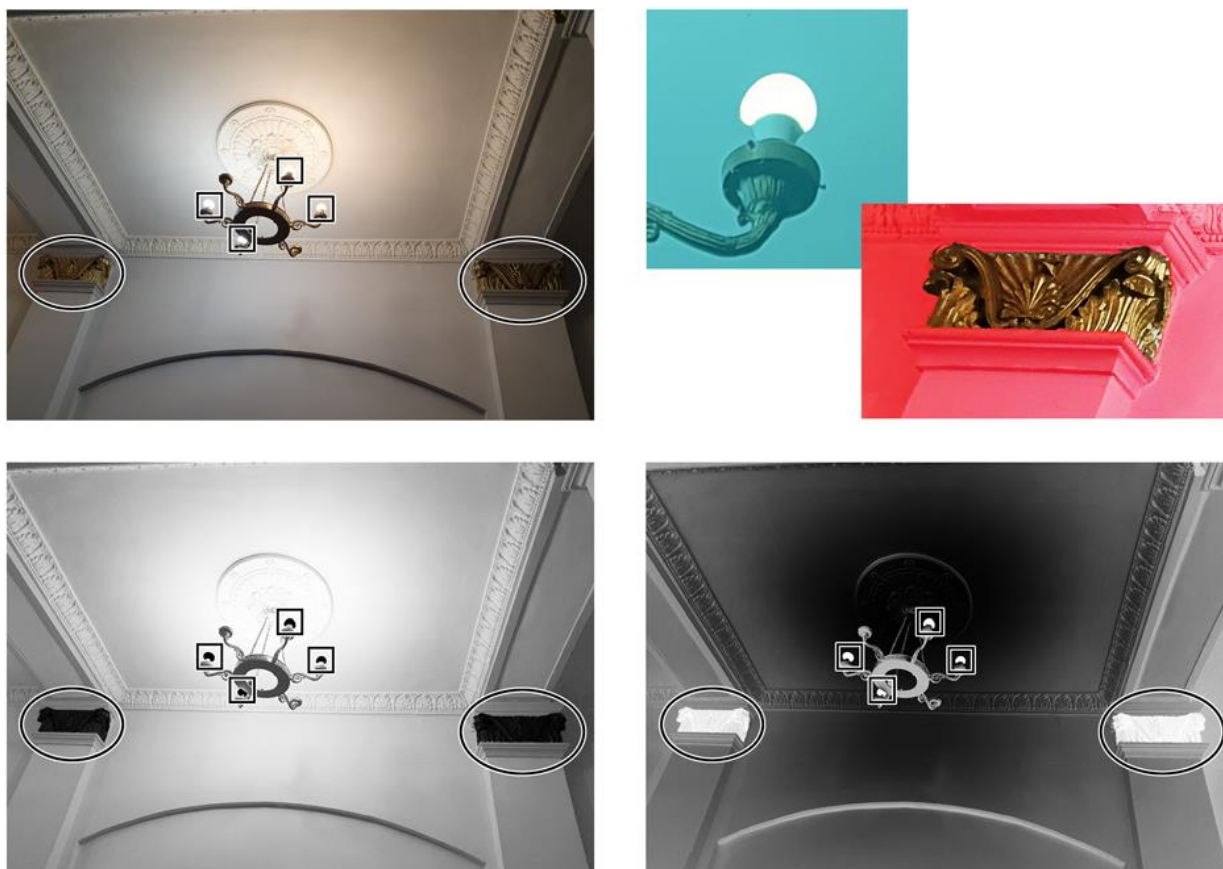


Рис. 8. Фрагмент стелі корпусу №3 ДНУ.

На відміну від ламп, антаблементи не вдалось би ідентифікувати без значного покращення вхідних даних, оскільки диспропорція між зображеннями виявилася достатньо значною та геометрія локальних образів – більш комплексною. Застосовуючи локальні минулі входження (третє зображення), стало можливим досягти необхідного результату, але в цьому випадку з фрагментами сірого кольору на зображенні.

В результаті проведених експериментів, ефективність методу без урахування кута спостереження впала до 4.76% при куті 40° , тоді як з урахуванням кута із подальшою трансформацією склала 83.34% відповідно (рис.5). Таким чином, ефективність модифікованого методу була підвищена на 7.14%, 16.66%, 42.85% та 78.58% при куті спостереження 10° , 20° , 30° та 40° відповідно.

Програмна реалізація методу. Швидкість розрахунків при програмній реалізації залежить не лише від реалізованого алгоритму, але й від оптимального використання операцій, наданих мовою програмування Javascript.

Оскільки алгоритм розрахунку базується на трансформації зображення, вхідне зображення потребує розбиття на двовимірний масив індивідуальних пікселів та подальшого «проходження необхідної кількості разів по даному масиву».

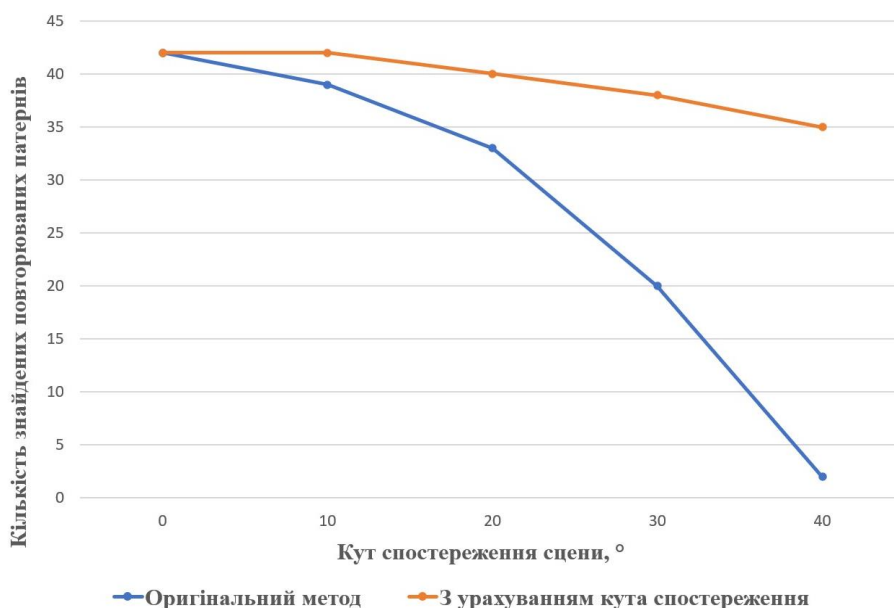


Рис. 9. Порівняння ефективності базового та модифікованого алгоритмів знаходження повторюваних патернів

```
function prepareImage(rawImage) {
  let angle = 0;
  let image = null;

  angle = calculateVisionAngle(rawImage);
  image = normalizeImage(rawImage, angle);

  return image;
}
```

Рис. 10. Функція підготовки зображення (визначення кута спостереження та подальшої трансформації зображення)

Підготуємо зображення, створивши функцію *prepareImage*, що готує зображення до контекстного моделювання обмеженого порядку. Слід зауважити, що розбиття коду на функції (модульність) – є невід’ємним стандартом мови програмування Javascript.

Оскільки «проходження по двомірному масиву» – достатньо складна операція, зменшення кількості «проходжень» є пріоритетом для підвищення ефективності розрахунків. Але визначення кута спостереження та трансформування зображення повинні виконуватись у чіткій послідовності. Тому зменшення кількості «проходжень» є складним, оскільки розрахунок кута спостереження, що виконується функцією *calculateVisionAngle*, та відповідна трансформація зображення, що виконується функцією *normalizeImage*, розраховуються незалежно одне від одного. Підготовка зображення є невід’ємною складовою виявлення повторюваних патернів. Дану задачу й вирішує функція *normalizeImage*, визначаючи кут спостереження та відповідно трансфор-

муючи зображення, нівелюючи перспективу, що значно полегшує початкову задачу знаходження повторюваних патернів на зображенні.

Маючи підготоване зображення, наступним кроком є застосування вищезазначеного методу для знаходження повторюваних патернів на зображенні. Слід зауважити, що завдяки підготовці зображення, а саме трансформації, компенсуючий кут спостереження, а як наслідок, відсоток диспропорції образів значно зменшується.

Висновки. Для програмної реалізації було обрано та програмно втілено знаходження кута спостереження із подальшою трансформацією, що нівелює перспективу та метод знаходження невідомих повторюваних патернів, представлений у роботі [1]. Контекстне моделювання обмеженого порядку надало можливість виконати розрахунки не виходячи за межі традиційних для користувача ресурсів пам'яті, що є важливою складовою процесу моделювання.

Були досліджені різні випадки використання методу, із різними степенями ефективності та також досліджена проблема впливу величини кута спостереження камери на якість моделювання пошуку невідомих локально-повторюваних патернів на зображеннях. Було покращено результати розрахунку повторюваних патернів, враховуючи кут спостереження та відповідну трансформацію вхідного зображення.

Тестування програмного продукту на вище зазначених прикладах продемонструвало, що даний підхід не тільки надає досить високу точність розпізнавання, але й покращує якість побудови контекстної моделі обмеженого порядку та проведення розрахунків.

Програмна реалізація виконана на мові програмування Javascript та задовольняє стандартам відповідного програмного забезпечення, надаючи можливість виконання контекстного моделювання обмеженого порядку у випадку обмежених ресурсів пам'яті.

Бібліографічні посилання

1. **Pratas, D.** On the Detection of Unknown Locally Repeating Patterns in Images [Text]/ Diogo Pratas, Armando J. Pinho // Image Analysis and Recognition (ICIAR 2012). – 2012. – С. 158–165.
2. **Lettry, L.** Repeated Pattern Detection using CNN activations[Text] / Louis Lettry, Michal Perdoch, Kenneth Vanhoey, Luc Van Gool // 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). – 2017. – Vol. 1. – С. 47–55.
3. **Armando, J.P.** On the Representability of Complete Genomes by Multiple Competing Finite-Context (Markov) Models [Text] / Armando J. Pinho, Paulo J. S. G. Ferreira, António J. R. Neves, Carlos A. C. Bastos // PLoS ONE. – 2011. – Vol. 6.
4. **Haider, A.** Window Detection in Facades [Text] / H. Ali, C. Seifert, N. Jindal, L. Paletta, G. Paar // 14th International Conference on Image Analysis and Processing (ICIAP). – 2007. – Vol. 1. – С. 837–842.
5. **Башков, Е.А.** Поиск изображений по содержимому в графических базах данных: моногр. / Е.А. Башков, О.Л. Вовк, Н.С. Костюкова. – Донецк, ГВУЗ «ДонНТУ», 2014. – 120 с.

6. **Рудаков, И.В.** Исследование перцептивных хеш-функций изображений [Текст] / И.В. Рудаков, И.М. Васютович // Журнал «Наука и образование» МГТУ им. Н.Э. Баумана. – 2015. – №8.
7. **Чалая, Л.Э.** Поиск неполных дубликатов в системах анализа цифровых изображений [Текст] / Л.Э. Чалая, П.Ю. Попаденко // Вісник КрНУ імені Михайла Остроградського. – 2014. – № 5(88). – С. 42–47.
8. **Методы и алгоритмы поиска изображений** [Електроний ресурс] / Г.Д. Огневой // II Международная научно-техническая интернет-конференция. – 2014. – Режим доступа: <http://rep.bntu.by/handle/data/12606>.
9. **Schindler, G.** Detecting and matching repeated patterns for automatic geo-tagging in urban environments [Text] / Grant Schindler, Panchapagesan Krishnamurthy, Roberto Lubliner-man, Yanxi Liu, Frank Dellaert // Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008) – 2008.

Надійшла до редколегії 11.09. 2020.