

УДК 621.391.3

І. В. МИРОНІВ,

Чернівецький національний університет імені Юрія Федьковича

КОНКУРУЮЧІ КЛІТИННІ АВТОМАТИ В ЗАДАЧАХ РОЗПІЗНАВАННЯ СИМВОЛІВ ТЕКСТУ

Розглянуто особливості розробки нової інформаційної технології розпізнавання символів тексту, в основу якої покладено новий тип клітинних автоматів — конкуруючі клітинні автомати. Переваги цієї технології — простота правил взаємодії, можливість легкого розпаралелювання процесу розпізнавання, а також розпізнавання спотворених і частково накладених символів. Здійснено порівняння якості роботи та швидкодії з відповідними характеристиками комерційної системи ABBYY FineReader.

Ключові слова: конкуруючий клітинний автомат; рухомий клітинний автомат; автомат Мура; граф переходів.

ВСТУП

Постановка задачі. Найважливішим чинником високоякісного розпізнавання тексту виступає розробка алгоритму розпізнавання тексту, який у поєднанні з попередньою обробкою зображення (очищення зображення від шуму та його поділ на зображення символів) забезпечує необхідний результат.

Аналіз літературних джерел

Останніми роками розробникам доводиться дедалі частіше розв'язувати задачу розпізнавання образів доволі складної природи. Досягнуті в цій сфері успіхи пов'язані з іменами таких учених, як І. С. Місунов [1], котрий запропонував підхід векторного подання інформації в задачі пошуку та класифікації символів; О. А. Поліщук [2], котрий здійснив аналіз і синтез систем розпізнавання рукописних математичних текстів; І. А. Імад [3], котрий розробив методи обробки зображення текстів арабської абетки на основі ознакових методів.

Клітинні автомати (КА) дотепер не так часто залучалися до задач розпізнавання. Чи не єдиним поглибленим дослідженням у цій галузі є праця, основну частину якої присвячено вивченню характеристик КА в процесах розпізнавання тексту. Автор використовує послідовності різних КА для виділення характерних ознак символів тексту: петель, перетинів, положень кінців. У [4] досліджено можливості КА щодо розпізнавання рукописних символів. Основними недоліками таких підходів вважається їх громіздкість та необхідність навчання.

Інші праці, пов'язані із залученням КА до таких задач або дуже застаріли, або стосуються використання КА в супутніх щодо розпізнавання задачах. Так, у [5] запропоновано новий алгоритм розпізнавання з JPEG зображення водяних знаків на основі клітинних автоматів, а в [6] побудовано спеціалізовані клітинно-автоматні структури для аналізу зображення контуру.

Водночас усі згадані дослідження показали свою неефективність при розпізнаванні рукописних та частково деформованих символів, а також символів, що частково накладаються один на одиний.

Невирішені питання

За результатами аналізу наукових джерел і наукових методів, спрямованих на досягнення поставленої мети, сформульовано актуальне наукове завдання, сутність якого полягає в створенні математичної моделі рухомих конкуруючих клітинних автоматів, які давали б змогу розпізнавати спотворені символи та символи, які частково накладаються один на одиний, і побудувати інформаційну технологію, на базі якої реалізувати програмний продукт і порівняти його показники розпізнавання з відомими аналогами.

Мета та завдання дослідження

Ідеться про розробку програмного забезпечення для дослідження задачі розпізнавання символів тексту.

Для досягнення цієї мети розв'язуються такі наукові задачі:

- побудова математичної моделі конкуруючого рухомого клітинного автомата;
- розробка інформаційної технології на базі математичної моделі;
- реалізація інформаційної технології у програмному забезпеченні.

Математична модель рухомих конкуруючих
клітинних автоматів та алгоритм її застосування

Множину клітинних автоматів, які використовуються в статті, можна записати у вигляді

$$\Sigma = \{\sigma_i\}_{i=1}^N, \quad (1)$$

де N — потужність алфавіту.

Кожний автомат має свою множину U станів і мітку ξ (колір, положення в алфавіті тощо) та залежить від дискретного часу:

$$\sigma_i = \sigma_i(U, t, \xi). \quad (2)$$

Тут $U = \{u_k\}_{k=1}^K$; K — кількість станів поточного автомата. Перехід автомата з поточного стану k в наступний $k + 1$ можна описати так:

$$(u_k, t, \xi) \xrightarrow{t \rightarrow t+1} \sigma_i(u_{k+1}, t+1, \xi). \quad (3)$$

Переходом КА в новий стан керує функція переходу ϕ , тобто

$$\sigma_i(u_k, t, \xi) = \phi(\sigma_i(u_k, t, \xi)). \quad (4)$$

Рухомий КА пересувається з поточного стану в наступний, використовуючи правила, що формуються функцією переходів (графи переходів наведено далі).

Припустимо, що ми подали зображення символів, які необхідно розпізнати, у вигляді сукупності станів, аналогічних станам КА. Тоді $\Omega = \{\omega_p\}_{p=1}^P$ — множина цих станів. Кількість P станів символа невідома.

Потрапивши на символ, автомат буде рухатися по його станах ω_p , кількість яких P , взагалі кажучи, не дорівнює кількості станів поточного автомата K , $P \neq K$.

Переходи будуть виконуватися КА за станами символа:

$$\sigma_i(\omega_p, t, \xi) \xrightarrow{t \rightarrow t+1} \sigma_i(\omega_{p+1}, t+1, \xi), \quad (5)$$

причому тоді і тільки тоді, коли $\omega_p = u_k \in U$ та $\omega_{p+1} = u_{k+1} \in U$, тобто коли стани символа збігаються з аналогічними станами КА. Якщо ж $\omega_p, \omega_{p+1} \neq u_k, u_{k+1} \in U$, то такий КА не має дозволених переходів і вилучається з клітинно-автоматного поля.

Якщо КА може реалізувати всі свої стани, задані функцією переходів на поточному символі, тобто якщо $\forall u_k \in U \exists \omega_p \in \Omega$, то будемо вважати, що цей КА «вдало» описує поточний символ.

Якщо ж $\forall u_k \in U \bar{\exists} \omega_p \in \Omega$, тобто хоча б для одного стану КА не існує аналогічного стану символа, то такий автомат вилучається з КА поля.

Таких КА, що реалізують усі свої можливі стани на поточному символі, може бути кілька. Для того щоб вибрати з них той, який точно відповідає цьому символу, використовується механізм конкуренції, котрий полягає ось у чому. Нехай на одному символі рухаються три КА:

$$\sigma_i \rightarrow U = \{u_k\}_{k=1}^K, \quad \sigma_j \rightarrow V = \{v_l\}_{l=1}^L, \quad \sigma_h \rightarrow W = \{w_s\}_{s=1}^S. \quad (6)$$

Кожний із них реалізує на ньому всі свої стани. Будемо вважати, що конкуренцію «виграє» той КА, кількість станів якого найбільша, тобто ми маємо знайти

$$\max(K, L, S). \quad (7)$$

Нехай $\max(K, L, S) = S$. Тоді КА $\sigma_h = \sigma_h(S, t, \eta)$ буде вважатися таким, що найбільш «вдало» описує поточний символ. Зчитуючи його мітку η , ми дізнаємося, який символ розпізнали.

Сам алгоритм розпізнавання ґрунтується на побудові КА та його графа переходів (рис. 1), за станами якого він має змогу рухатися.

Відповідно до математичної моделі символа, які розпізнаються, також подаються у вигляді множини станів, аналогічних станам графа переходів КА (рис. 2).

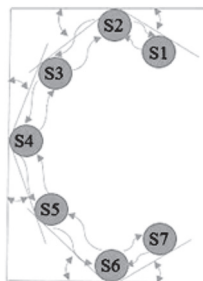
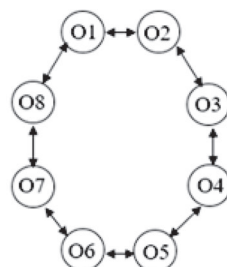


Рис. 1. Приклад графа переходів імовірнісних автоматів Мура для символа С



[O1,O2,O3,O4,O5,O6,O7,O8]

Рис. 2. Подання символа у вигляді множини станів, аналогічних станам графа переходів КА

Зображення тексту, який необхідно розпізнати, вважається КА полем, на якому у випадковому порядку розташовуються всі типи КА. Таким чином, ми отримуємо багат шарову структуру сукупності КА, які починають рухатися по сусідніх станах КА поля. При цьому рух має задовольняти такі принципи.

1. Якщо КА не потрапив на символ, він не може почати свій рух і вилучається з КА поля.
2. Якщо КА не реалізує на символі стани, визначені його графом переходів, він також вилучається з КА поля.
3. Якщо КА реалізує на символі всі свої стани, визначені його графом переходів, він породжує новий КА даного типу.

Із метою оптимізації та зменшення обчислювальних ресурсів нами було розроблено операцію редукції КА поля. Процес редукції полягає в тому, що КА поле генерується лише на символах тексту, а вільний простір зображення не використовується. Для цього застосовано алгоритм розкладання Карунена—Лоева [7], поданий на рис. 3.

Як бачимо, символ поділяється на 18 рівних частин проти годинникової стрілки, через кожні 20 градусів, після чого аналізується множина чорних пікселів у даній області та зводиться до вигляду масиву станів для КА.

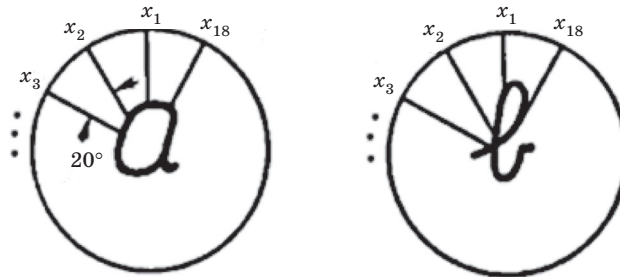


Рис. 3. Приклад розкладання Карунена—Лоева для символів *a*, *b*

У результаті отримуємо багат шарову структуру КА, де рух відбувається лише в межах виділеного символу (рис. 4, *a*).

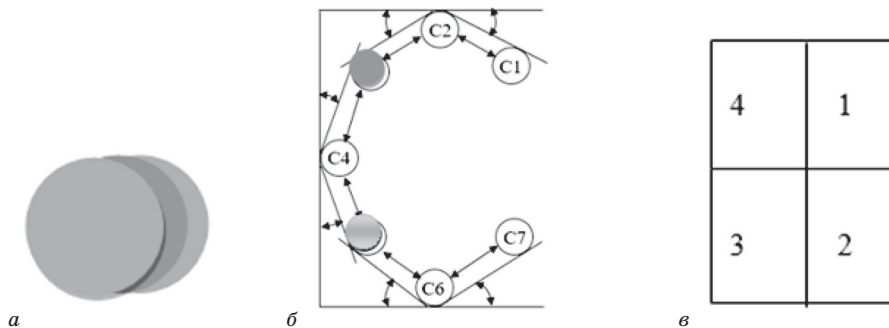


Рис. 4. Приклад процесу розпізнавання:

a — багат шаровість КА; *b* — розміщення КА на колекції станів символу; *c* — позначення чверті розбиття символу

Як випливає з рис. 4, *b*, на КА поле символу *C* потрапляють КА двох типів: перший відповідає символу *C*, а другий йому не відповідає. При русі по КА полю згідно з формулами (3)–(5) аналізується досяжність актуальних станів, кути між сусідніми станами (які не повинні відрізнятись від дозволеного більш ніж на 5 градусів), що вимірюються по-різному в різних квадрантах (рис. 4, *c*): у першому квадранті вони вимірюються від верхньої та правої сторін, у другому — від правої та нижньої і т. д.

У результаті такого аналізу КА, який задовольняє всі свої стани, породжує КА даного типу. У протилежному випадку він вилучається з КА поля.

Відскановане зображення необхідно підготувати для генерування КА поля. Процес підготовки полягає в розбитті зображення тексту на зображення рядків та символів за допомогою аналізу груп чорних пікселів. Використання алгоритму Карунена—Лоева і побудова КА поля дозволяють нам позбутися залежності від таких властивостей, як товщина ліній та розмір символу.

Ілюстрацію цього подано на рис. 5, де показано приведення символів *e* та *l* до станів КА поля.

Як випливає з математичної моделі та графів переходів, КА, які реалізують всі свої стани на КА поля одного символу, може бути кілька: *I* та *H*, *T*, *L*, *C* та *O* тощо. У таких випадках використовується розроблений нами алгоритм конкуренції, який описано формулами (6) та (7). Якщо КА символу *I* рухається по станах символу *H* і реалізує *S* своїх станів, а КА символу *H* при цьому реалізує *K* своїх станів, причому $K > S$, то КА символу *I* вилучається з КА поля. У цьому випадку конкуренцію виграв КА символу *H*, і вважається, що він буде результатом розпізнавання.

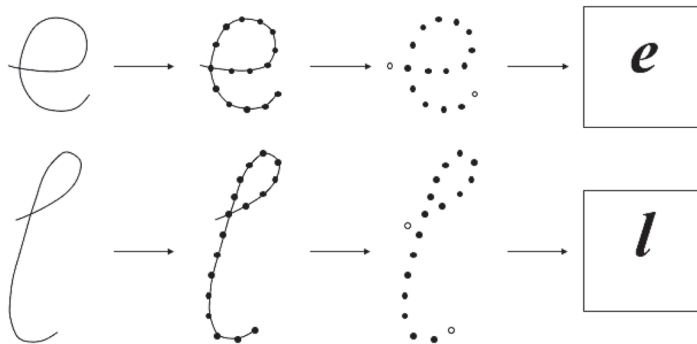


Рис. 5. Приклад процесу приведення зображення символів е та l до станів КА поля

Інформаційна технологія розпізнавання символів тексту на основі конкуруючих клітинних автоматів

Для реалізації запропонованої математичної моделі та алгоритму рухомих конкуруючих клітинних автоматів було використано інформаційну технологію розпізнавання символів тексту, структуру якої ілюструють рис. 6 і 7.

Блоки 1–4 (див. рис. 6) призначено для попередньої обробки зображення, розбиття на рядки та символи. Залежно від способу реалізації вони можуть бути різні.

Блок 5 (див. рис. 6) генерування КА поля використовує алгоритм Карунена—Лоева та графи переходів відповідних символів.

Після створення КА поля запускається процес розпізнавання символів, структуру якого унаочнює рис. 7.

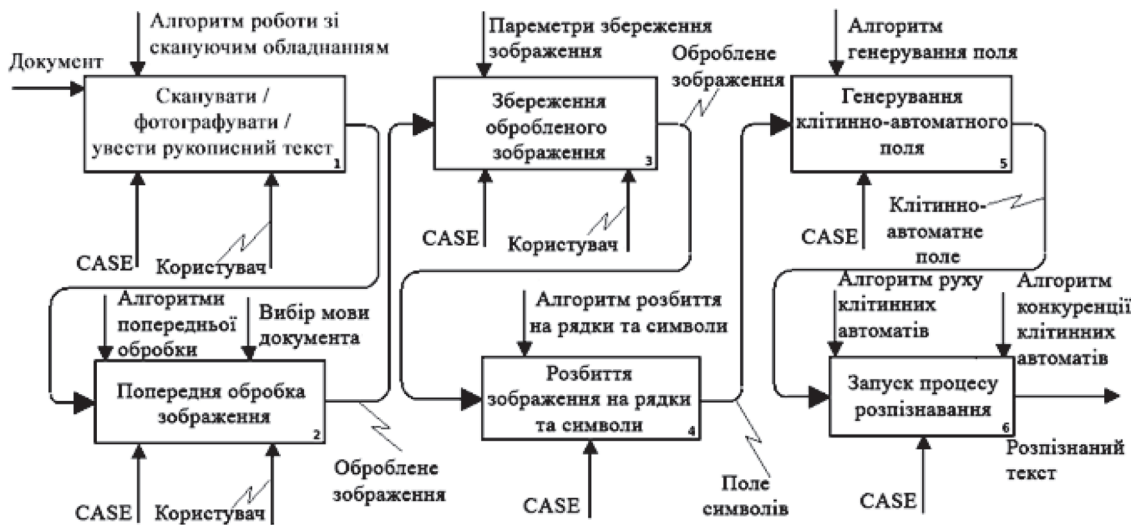


Рис. 6. Загальна схема розробленої інформаційної технології розпізнавання символів на основі конкуруючих КА

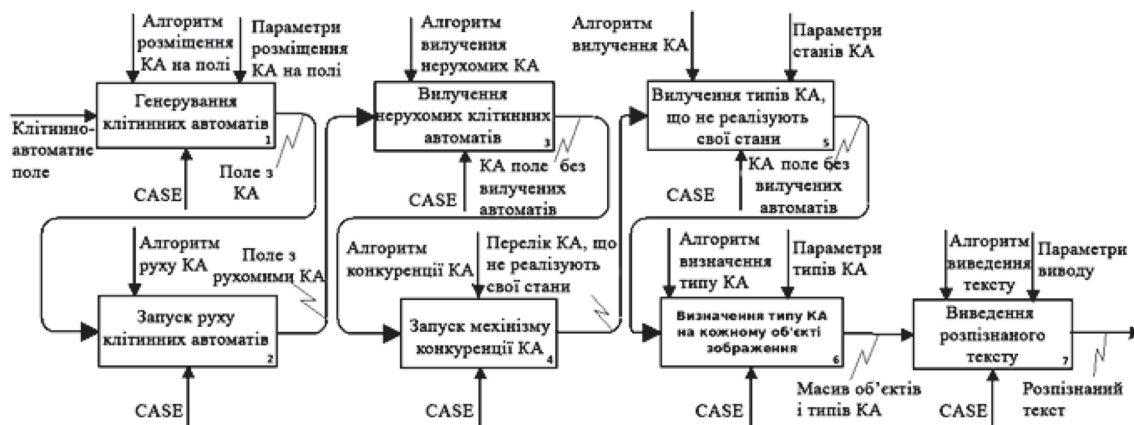


Рис. 7. Етапи процесу розпізнавання символів на основі клітинних автоматів

Процес розпізнавання починається з генерування КА, що відповідає за кожний символ мови, якою написано оригінал. Клітинні апарати генеруються випадково, рівномірно по всьому полю із зображеннями літер. Далі запускається механізм руху КА по зображеннях символів (блок 2).

Під час руху КА працює алгоритм вилучення нерухомих автоматів (блок 3), після чого запускається механізм конкуренції, результатом якого є вилучення типів КА, що не реалізують усіх своїх станів (блоки 4–5). У разі неможливості визначення КА, що виграв конкуренцію, ІТ залишає для користувача можливість вибору правильного символу.

Тепер залишається визначити унікальний параметр (наприклад, колір) тих КА, які залишилися на зображеннях літер (блок 6), та вивести розпізнаний текст (блок 7).

Перевагами такої інформаційної технології є те, що запропоновані алгоритми здатні виконувати розпізнавання спотворених символів і таких, що частково накладаються один на один, а також рукописного тексту «друкованими» літерами. Окрім того, як це буде показано далі, зазначена технологія демонструє високу ймовірність розпізнавання в усіх розглядуваних випадках.

Розробка програмного забезпечення та апробація результатів

Для визначення показників якості та швидкості запропонованої інформаційної технології було реалізовано програмні продукти (десктопну та мобільну під ОС Android) і проведено порівняння якості та швидкості розпізнавання з відомими аналогами.

Спрощену діаграму класів, яка реалізує роботу з КА, подано на рис. 8.

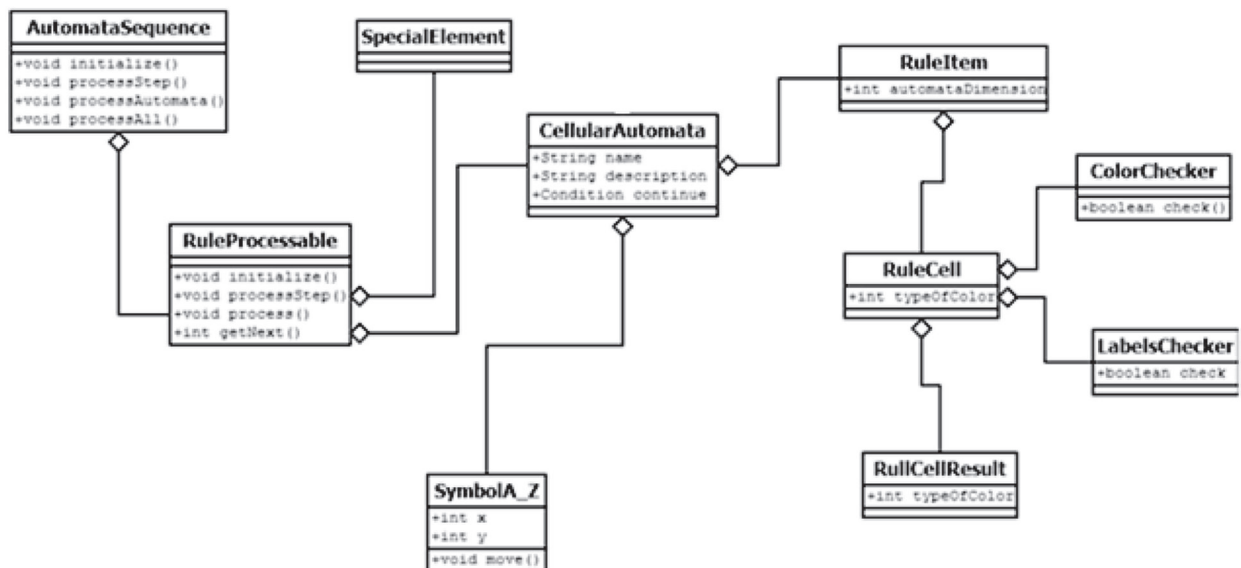


Рис. 8. Діаграма класів, яка реалізує роботу з клітинними автоматами

Система складається з двох базових класів: CellularAutomata та AutomataSequence, які описують роботу КА та їх послідовностей для запуску механізму конкуренції. У клітинно-автоматну взаємодію включено такі елементи, як поділ зображення на окремі символи, перевірка умов у графі переходу тощо. Стани кожного автомата, який відповідає певній літері алфавіту, та правила переходів по графу реалізовано в класах RuleItem, RuleCell і SymbolA_Z. Для спрощення діаграми описи всіх літер показано в одному класі. Ці класи відповідають за реалізацію руху КА та описують граф переходів. Індексом типу автомата, який дозволяє його однозначно ідентифікувати, виступає кольорова мітка. Зчитавши цю мітку, ми можемо визначити розпізнану літеру. Цей процес охоплюють класи RuleCellResult, LabelsChecker, ColorChecker. Результатом їхньої роботи буде розпізнаний за допомогою конкуруючих клітинних автоматів текст.

Взаємодія з апаратним забезпеченням та виведення розпізнаного тексту виконується за допомогою стандартних функцій API Windows та Android. Обробка зображень, отриманих зі скануючого обладнання, виконувалася за допомогою бібліотек з відкритим кодом OpenCV (для ОС Android).

Із використанням розробленого програмного забезпечення було проведено тестування якості розпізнавання символів тексту, набраного різними шрифтами, рукописних символів та символів, які частково накладаються. Також було проведено порівняння якості розпізнавання розробленого програмного продукту з комерційним програмним забезпеченням FineReader.

Для генерування вхідних зображень використовувалися генератори тексту.

При розпізнаванні 1000 символів шрифтом Times New Roman та Arial програма показала 98% розпізнаних символів. Символи, які не було розпізнано, — це символи *r* і *n*, які йдуть у тексті один за одним. Результатом помилкового розпізнавання став символ *m*. Така помилка постійно трапляється й у FineReader.

При розпізнаванні 100 рукописних символів програма показала 84%, оскільки рукописні символи *e* та *l* розпізнаються залежно від правильності їх написання.

Якщо символи мають спільні лінії, вони також погано розпізнаються, але в цьому разі вони погано розпізнаються й людиною. Якщо ж символи накладаються частково, то результат розпізнавання становить 56%.

Частково деформовані символи програма успішно розпізнала у 68% випадків. Слід зауважити, що успішність розпізнавання залежить від ступеня деформації, і якщо цей ступінь більший за 30%, існує велика ймовірність помилкового розпізнавання символа.

Результати порівняння розпізнавання символів програмним забезпеченням (від нормального їх вигляду до відповідного ступеня спотворення), а також символів, що частково накладаються, із результатами розпізнавання комерційного продукту FineReader наведено на рис. 9, де вісь *X* показує ступінь спотворення, а вісь *Y* — відсоток розпізнавання. У разі розпізнавання символів, що частково накладаються, вісь *X* показує ступінь накладання, а вісь *Y* — відсоток розпізнавання.

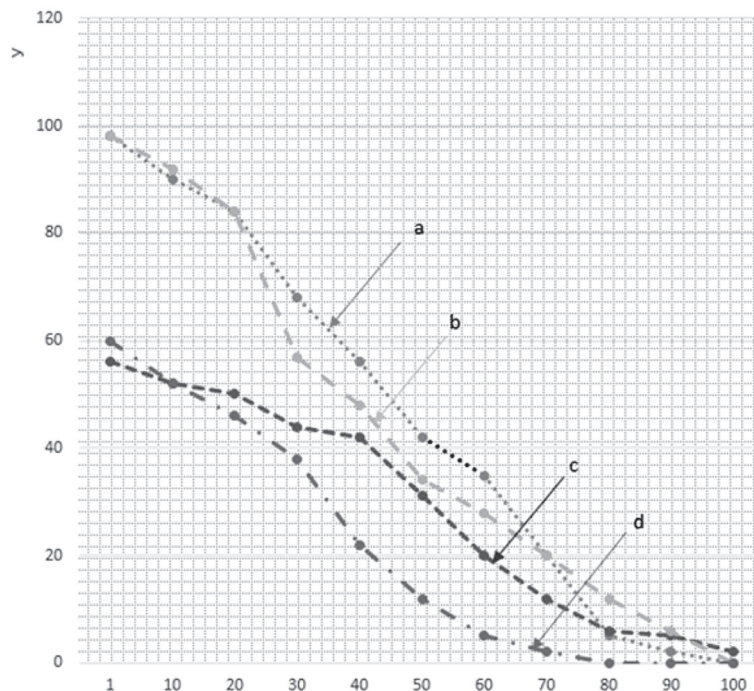


Рис. 9. Графік якості результату розпізнавання інформаційною системою та комерційним продуктом FineReader

Порівняння розроблених програмних версій показало, що мобільна версія не поступається десктопній за якістю розпізнавання, але поступається щодо швидкості (хоча цей показник залежить від характеристик ПК), а також абсолютно не посткомерційному продукту FineReader.

ВИСНОВКИ

У результаті проведення досліджень щодо можливості застосування клітинних автоматів до задачі розпізнавання тексту було розроблено математичну модель та алгоритм, в основу якого покладено рухомі конкуруючі клітинні автомати, котрі дозволяють розпізнавати не лише символи в наявних системних шрифтах, а й такі, що мають певний ступінь деформованості чи накладаються один на один. Показано, що реалізована у програмному продукті інформаційна технологія, створена на базі алгоритму та математичної моделі, продемонструвала високу швидкодію та ефективність розпізнавання деформованих символів порівняно з відомими системами. Досягнуто таких результатів.

1. При розпізнаванні 1000 символів шрифтом Times New Roman та Arial програма показала 98% розпізнаних символів.

2. При розпізнаванні 100 рукописних символів програма показала 84%, оскільки рукописні символи *e* та *l* розпізнаються залежно від правильності їх написання.

3. Якщо символи мають спільні лінії, то вони також погано розпізнаються, але в цьому разі їх погано розпізнає й людина. Якщо ж символи накладаються частково, то результат розпізнавання становить 56%.

4. Частково деформовані символи програма успішно розпізнала у 68% випадків. Слід зауважити, що успішність розпізнавання залежить від ступеня деформації, і якщо цей ступінь вищий за 30%, існує велика ймовірність помилкового розпізнавання символа.

Список використаної літератури

1. Мисуно, И. С. Разработка и исследование векторных представлений информации для задач поиска и классификации: дис. ... канд. техн. наук; спец. 05.13.23 - Системы и средства искусственного интеллекта / И. С. Мисуно; НАН Украины, МОН Украины, Междунар. науч.-учеб. центр информац. технологий и систем, 2006.— 167 с.
2. Поліщук, О. А. Аналіз та синтез систем розпізнавання рукописних математичних текстів : дис. ... канд. техн. наук; спец. 01.05.04 - Системний аналіз і теорія оптимальних рішень / О. А. Поліщук; Київ. нац. ун-т ім. Тараса Шевченка, 2006.— 135 с.
3. Імад, І. А. Методи опрацювання зображень текстів арабської абетки: автореф. дис. ... на здобуття наук. ступеня кандид. техн. наук; спец. 05.13.06 - Інформаційні технології / І. А. Імад.— Нац. ун-т «Львівська політехніка».— Львів, 2008.— 18 с.
4. Oliveira, C. C. An Approach to Searching for Two-Dimensional Cellular Automata for Recognition of Handwritten Digits / C. C. Oliveira, P. P. Oliveira // Lecture Notes In Artificial Intelligence. Berlin; Heidelberg: Springer-Verlag, 2008.— Vol. 5317.— P. 462–471.
5. Суясов, Д. И. Выделение структурных признаков изображений символов на основе клеточных автоматов с метками / Д. И. Суясов // Программное и аппаратное обеспечение.— 2010.— № 4.— С. 39–45.
6. Wu, H. A new JPEG image watermarking algorithm based on cellular automata / H. Wu, J. Zhou, X. Gong, Y. Wen, B. Li // Journal of Information & Computational Science.— 2011.— V. 8, Issue 12.— P. 2431–2439.
7. Фу, К. Последовательные методы в распознавании образов и обучении машин / К. Фу; пер. с англ., под. ред. Л. А. Мееровича, Я. З. Цыпкина.— М.: Наука, 1971.— 256 с.

Рецензент: доктор техн. наук, доцент В. В. Онищенко, Державний університет телекомунікацій, Київ.

И. В. Миронов

КОНКУРИРУЮЩИЕ КЛЕТОЧНЫЕ АВТОМАТЫ В ЗАДАЧАХ РАСПОЗНАВАНИЯ СИМВОЛОВ ТЕКСТА

Рассмотрены особенности разработки новой информационной технологии распознавания символов текста, основанной на новом типе клеточных автоматов — конкурирующих клеточных автоматов. Преимуществами этой технологии являются простота правил взаимодействия, возможность легкого распараллеливания процесса распознавания, возможность распознавания искаженных и частично наложенных символов. Проведено сравнение качества работы и быстродействия с соответствующими характеристиками коммерческой системы ABBYY FineReader.

Ключевые слова: конкурирующий клеточный автомат; движущийся клеточный автомат; автомат Мура; граф переходов.

Ivan Myroniv

COMPETING CELLULAR AUTOMATA IN THE TASKS OF RECOGNIZING THE TEXT CHARACTERS

The research on the possibility of using cellular automata in solving the problem of text character recognition is carried out. For this purpose, the concept of competing cellular automata has been introduced and a mathematical model of their functioning and interaction has been developed. To implement the proposed algorithms a simulation program has been developed, which allowed to evaluate the effectiveness of cellular automata algorithms, conduct experiments on the recognition of the characters of the English alphabet and demonstrate a number of advantages over the other methods.

The paper proposes to use a set of cellular automata based on the Mealy and Moore transition graphs for each type of cellular automaton, in order to determine endings, intersections, and cycles in characters. The work simulation program is observed.

Among the benefits of this technology are simple rules of interaction, the ability to easily parallel the recognition process, the ability to recognize distorted and partially overlapped characters and symbols. The comparison of recognition quality and performance with the commercial system ABBYY FineReader has demonstrated high performance indicators of the developed recognition technology.

Keywords: competing cellular automaton; mobile cellular automaton; Moore machine; transition graph.