

Рисунок 3 – Рослина / Дерево L-системи

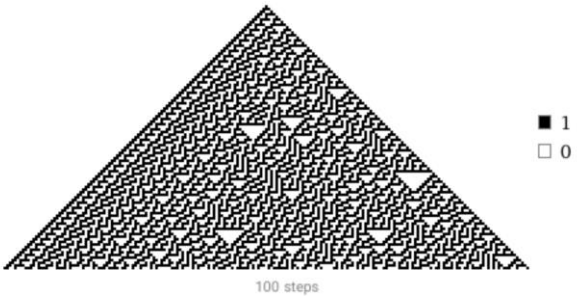

Рисунок 4 – Піраміда Rule 30

Висновки. Проведене порівняння ілюструє, що фрактальні структури можуть формуватися на основі різних математичних принципів: геометричних, ймовірнісних чи алгоритмічних. Результати моделювання свідчать про доцільність застосування методу MRCM у задачах, що вимагають точного відтворення форм, тоді як L-системи демонструють значний потенціал для імітації процесів біологічного росту. Перспективним напрямом подальших досліджень може бути розгляд гібридних підходів, зокрема для можливої процедурної генерації тривимірних (3D) об’єктів.

Список використаних джерел

1. Fractal Generative Models / T. Li et al. New York: arXiv, 2025. 13 p. (Препринт. Cornell University; 2502.17437 (cs)). URL: https://arxiv.org/abs/2502.17437 (дата звернення: 25.11.2025).

2. Fractal generation and analysis using modified fixed-point iteration / A. Tassaddiq et al. *AIMS Mathematics.* 2025. Vol. 10, № 4. P. 9462–9492.

3. A Comparative Study of Fractal Models Applied to Artificial and Natural Data / G. Silva et al. *Fractal and Fractional.* 2025. Vol. 9, № 2. P. 87. DOI: https://doi.org/10.3390/fractalfract9020087 (дата звернення: 25.11.2025).

4. SirXemic. IFS Fractal Generator: інтерактивний генератор фракталів. 2025. URL: https://github.com/SirXemic/jquery-fractals (дата звернення: 03.12.2025).

УДК: 004.738.5

Костюшин М. В., здобувач 2 курсу
ОС «Магістр» спеціальності
F3 Комп’ютерні науки,
Ніколюк П. К., д-р фіз.-мат. наук, професор
кафедри інформаційних технологій

ПОРІВНЯННЯ ПІДХОДІВ ДО ЗЛИТТЯ ДАНИХ У СИСТЕМАХ МУЛЬТИМОДАЛЬНОГО РОЗПІЗНАВАННЯ ОБ’ЄКТІВ

Донецький національний університет імені Василя Стуса, м. Вінниця

Вступ. Сучасні системи штучного інтелекту дедалі частіше працюють не з одним видом даних, а з кількома. Найпоширенішими є зображення, аудіо, текс-

тові описи, сенсорні сигнали. Об'єднання цих джерел інформації дає змогу глибше інтерпретувати складні сцени, підвищувати точність прийняття рішень і забезпечувати стійкість у варіативних умовах. В основі таких систем лежать методи ф'южну, тобто злиття ознак або проміжних представлень, що надходять із різних каналів. Попри значний прогрес у розвитку мультимодальних моделей, оптимальний спосіб поєднання даних залишається відкритим питанням, оскільки різні підходи по-різному впливають на продуктивність, узагальнювальну здатність і обчислювальну складність.

Метою цієї роботи є виконання оглядово-аналітичного порівняння підходів до ф'южну в мультимодальному розпізнаванні об'єктів та виявлення їх переваг, недоліків і сфер доцільного застосування.

Основний текст. Потреба у мультимодальних системах зростає у сферах автономного транспорту, медичної діагностики, робототехніки, відеоаналітики та людино-комп'ютерної взаємодії. Одномодальні рішення вже не забезпечують достатньої надійності, особливо в умовах шуму, часткової втрати інформації чи неоднозначності вхідних сигналів. Ф'южн дає змогу компенсувати недоліки кожної модальності та отримувати синергетичний ефект. Проте наявність кількох принципово різних підходів – раннього, пізнього, гібридного та attention-орієнтованого ф'южну створює потребу в аналітичному розумінні їх переваг і недоліків. Систематизація цих методів є важливою для коректного вибору архітектури мультимодальних систем у наукових і практичних застосуваннях.

У науковій літературі останніх років простежується активний розвиток трансформерних мультимодальних архітектур, де ф'южн реалізується переважно завдяки механізмам cross-attention та contextual alignment. Моделі на кшталт ViLT, CLIP, Flamingo та нові великі мультимодальні моделі (GPT-4o, Gemini, Claude 3.5) демонструють, що багаторівневий ф'южн дає змогу збільшити узгодженість між модальностями та покращувати семантичне розуміння.

Класичні підходи – ранній та пізній ф'южн широко описані в роботах саме з комп'ютерного зору та обробки аудіо, де зазначається, що ранній ф'южн добре працює з низькорівневими ознаками, тоді як пізній ф'южн забезпечує більшу гнучкість, але може втрачати міжмодальні залежності. У сучасних роботах з'являються також гібридні моделі, що комбінують кілька рівнів інтеграції. Однак єдиного універсального підходу не існує, і вибір методу ф'южну залежить від типу задачі, структури даних та вимог до системи [1].

Методи ф'южну зазвичай поділяють на три ключові групи: ранній (early fusion), пізній (late fusion) та гібридний (hybrid/mid-level fusion). Окремо виділяють сучасні моделі ф'южну на основі механізмів уваги (attention), що стали визначальними у трансформерних архітектурах [2].

Ранній ф'южн передбачає інтеграцію модальностей на рівні вхідних даних або низькорівневих ознак, наприклад, шляхом конкатенації каналів (RGB+D) чи сумісного подання сирих сенсорних сигналів. Такий підхід сприяє формуванню спільного латентного простору, у якому зберігаються міжмодальні кореляції й доступні комплексні залежності між модальностями. Основними перевагами є максимальне збереження інформації та раннє виявлення комплементарних ознак; основними недоліками – вимога точної синхронізації й нормалізації модальностей, чутливість до шуму та потенційне домінування нестабільної модальності.

Ранній ф’южн зазвичай ефективний у задачах, де модальності структурно узгоджені та мають близьку просторову розгортку.

Пізній ф’южн виконує інтеграцію на рівні рішень окремих одномодальних моделей через агрегування ймовірностей, голосування або метакласифікатор. Це забезпечує модульність архітектури, можливість незалежного тренування кожної модальності й спрощену інтеграцію pre-trained компонентів. Недоліком є втрата детальних міжмодальних залежностей, оскільки кореляції між ознаками не враховуються на спільному рівні представлень. Метод доцільний у сценаріях із різнорідними даними або в системах, де важливо зберегти автономність і замінність окремих модулів.

Гібридні підходи комбінують раннє злиття ознак із пізнім узгодженням рішень або проміжних представлень. Інтеграція може виконуватися на кількох абстрактних рівнях, наприклад, у просторі фіч-мапів, а фінальне рішення формується шляхом додаткового агрегування модальних прогнозів. Перевагами є кращий баланс між структурною гнучкістю та здатністю моделі виявляти складні кореляції; недоліками – підвищені обчислювальні витрати й складність оптимізації. Такі підходи демонструють високу ефективність у задачах, де потрібна точна взаємодія модальностей за збереження модульності.

Attention-орієнтований ф’южн та трансформерні підходи

Attention-механізми, зокрема cross-attention, дають змогу моделі динамічно зважувати внесок кожної модальності й вибірково підсилювати релевантні ознаки, що істотно підвищує гнучкість і точність мультимодальної інтеграції. Такі підходи забезпечують багаторівневе узгодження модальностей, підтримують масштабування на великі набори даних та формують основу сучасних мультимодальних трансформерів. Обмеженнями є висока обчислювальна вартість, необхідність великих датасетів і складність інтерпретації міжмодальних взаємодій. Attention-ф’южн є стандартом у візуально-мовних та мультимодальних LLM [3].

Підсумовуючи дані: ранній ф’южн оптимальний для структурно узгоджених модальностей із синхронізованою просторовою організацією; пізній – для систем, де критична модульність і незалежність потоків; гібридний – для задач, що вимагають балансу між точністю та гнучкістю; attention-ф’южн – для сценаріїв, де необхідно моделювати складні, контекстно-залежні кореляції на різних рівнях представлення.

Висновки. Отже, оптимальний вибір методу залежить від структури даних, вимог до точності, ресурсних обмежень і характеру взаємодії між модальностями.

Список використаних джерел

1. Baltrušaitis T., Ahuja C., Morency L.-P. Multimodal Machine Learning: A Survey and Taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2018. URL: https://www.researchgate.net/publication/317185818_Multimodal_Machine_LearningA_Survey_and_Taxonomy (дата звернення 22.11.2025).

2. Multimodal fusion for multimedia analysis: A survey / P. Atrey, M. Hossain, A. El Saddik, M. Kankanhalli. *Multimedia Systems.* 2010. URL: https://www.researchgate.net/publication/220460981_Multimodal_fusion_for_multimedia_analysis_A_survey (дата звернення: 22.11.2025).

3. Ahmad J., Del Bue A. mmFUSION: Multimodal Fusion for 3D Object Detection. 2020. URL: https://arxiv.org/abs/2311.04058 (дата звернення 22.11.2025).

УДК: 519.674:517.537.3]:004

Замотайло В. В., здобувач 2 курсу спеціальності 122 Комп'ютерні науки,
Фриз І. В., канд. фіз.-мат. наук, старший викладач кафедри інформаційних технологій

ПРОГРАМНА РЕАЛІЗАЦІЯ АПРОКСИМАЦІЇ ПЕРІОДИЧНИХ СИГНАЛІВ РЯДАМИ ФУР'Є В СИСТЕМАХ ОБРОБКИ ДАНИХ

Донецький національний університет імені Василя Стуса, м. Вінниця

Вступ. У сучасних інформаційних системах важливим завданням є обробка, стиснення та передача даних, які часто мають періодичну природу, наприклад, аудіосигнали, зображення, телеметрія. Для програмної реалізації під час обчислення значень та аналізу таких функцій використовуються методи, що спираються на розклад складного сигналу на простіші складники. Одним з інструментів, який дає змогу це робити, є гармонічний аналіз на основі рядів Фур'є.

Актуальність дослідження зумовлена необхідністю пошуку балансу між точністю відтворення сигналу і обчислювальними витратами процесора. Очевидно, що програмна генерація складних форм хвилі (наприклад, прямокутної або пилкоподібної) в цифрових синтезаторах та системах зв'язку виконується шляхом сумування лише обмеженої кількості гармонік. Отже, метою роботи є програмне моделювання та візуальний аналіз ефективності апроксимації розривного періодичного сигналу (меандру) за допомогою часткових сум ряду Фур'є.

Основний текст. Відомо, що періодична функція $f(x)$ з періодом $2l$, що задовольняє умови Діріхле, може бути розкладена у тригонометричний ряд Фур'є:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

де a_n, b_n – коефіцієнти Фур'є, обчислення яких залежить від інтервалу, періоду функції, парності і непарності [1].

Для дослідження обрано один із класичних прикладів із цифрової електроніки – прямокутний сигнал, який описує ідеальні тактові імпульси в комп'ютерних системах, визначений функцією $f(x) = sign(\sin x)$ з періодом 2π, що визначена на $[0,2\pi]$, [2]. Його розклад у ряд Фур'є містить лише непарні синусні гармоніки:

$$f(x) = \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\sin\big((2k-1)x\big)}{2k-1} = \frac{4}{\pi}\left(\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots\right).$$