

## ПОВТОРЮВАНІСТЬ ЕЛЕМЕНТІВ ЦИФРОВИХ ЗОБРАЖЕНЬ ІЗ ШУМАМИ

### REPETITIONS OF ELEMENTS OF NOISY DIGITAL IMAGES

Олексій Кравчук, Іван Довгань, Олег Кушнір

*кафедра оптоелектроніки та інформаційних технологій, Львівський національний університет імені Івана Франка, м. Львів вул. Тарнавського 107, 79017, Україна*

В останні десятиліття спостерігаємо бурхливий розвиток галузей, пов'язаних зі штучним інтелектом, зокрема комп'ютерного зору, розпізнавання образів і опрацювання природної мови [1, 2]. Раніше в праці [3] було запропоновано аналогії між текстами та цифровими зображеннями, а також низку підходів до аналізу зображень, запозичених зі статистичної лінгвістики та заснованих на вивченні явища спалахів (intermittance) і довгосяжних кореляцій. Було показано, що такі підходи дають змогу кількісно відрізнити інформативні зображення від шумів, за аналогією до задачі розрізнення природних текстів від випадкових символічних послідовностей [3].

Дане дослідження продовжує цей напрям і спрямоване на аналіз характеристики повторюваності в цифрових зображеннях залежно від відносної кількості білого шуму на цих зображеннях, а також на виявлення можливих кореляцій між цими величинами. Характеристику повторюваності  $v(t)$  у статистичній лінгвістиці було вперше введено Гольхером (див. працю [4]) як функцію поточного дискретного «часу» (позиції)  $t$  символу в тексті. Вона дорівнює «словникові»  $V_0$  завершених повторень символічних або лексичних  $n$ -грам будь-яких довжин  $n$ , поділеному на поточне значення позиції  $t$ .

Оскільки в розрахунках характеристики повторюваності текст розглядають як одновимірну послідовність символів або слів, то двовимірну структуру зображення слід перетворити на одновимірний часовий ряд за рецептами [3]. Характеристику повторюваності для текстів означають на лінгвістичних рівнях символів або слів, а тому в разі цифрових зображень її розраховують для рівнів яскравості пікселів зображення («слів») або відповідних цифр, з яких складаються ці рівні яскравості («символів»).

Шляхом експериментальних досліджень низки цифрових зображень ми виявили, що насичене значення повторюваності  $v_0$  при великих  $t$  є наближено монотонною функцією відсотку білого шуму на зображеннях. Взаємозв'язок повторюваності  $v_0$  у зображеннях і частки білого шуму на них дає змогу використати параметр  $v_0$  як своєрідну метрику шуму і цим запропонувати новий метод встановлення співвідношення між інформативною та шумовою складовою зображень.

### Бібліографія

1. R. C. Gonzalez, R. E. Woods, Digital Image Processing, Pearson (2007), 976 p.
2. A. Clark, C. Fox, S. Lappin, The Handbook of Computational Linguistics and Natural Language Processing, Wiley-Blackwell (2010), 802 p.
3. І. Я. Довгань, О. С. Кушнір, Ю. М. Фургала, Г. І. Паночко, Електроніка та інформ. технол., 17, 3 (2022).
4. O. S. Kushnir, L. B. Ivanitskyi, A. I. Kashuba, M. R. Mostova, V. B. Mykhaylyk, Repetition characteristic for single texts, Proc. 5<sup>th</sup> Int. Conf. on Computational Linguistics and Intelligent Systems (COLINS-2021). Vol. I: Main Conference. Ed. by N. Sharonova et al., CEUR Workshop Proc., vol. 2870, p. 629–641.