

УДК 004.93:681.883.41

*Шумейко К.П., аспірант,
Національний університет «Одеська політехніка»*

ВИКОРИСТАННЯ FUSION-МОДЕЛІ ДЛЯ АНАЛІЗУ МУЛЬТИМОДАЛЬНИХ ДАНИХ

Вступ. В умовах розвитку сучасних інтелектуальних систем особливої актуальності набуває задача спільного аналізу даних, що надходять із різних джерел. На практиці такими джерелами можуть бути зображення, часові ряди, сигнали датчиків, текстові описи та інші типи інформації, які відрізняються за структурою, розмірністю та фізичною природою. Використання лише одного каналу спостереження не завжди забезпечує отримання повного й стійкого уявлення про об'єкт дослідження, оскільки окреме джерело даних може містити шуми, пропуски або обмежений набір інформативних ознак. У зв'язку з цим значний інтерес становить застосування fusion-моделей, призначених для об'єднання різнорідних даних у межах єдиної аналітичної системи [1], [2].

З методичної точки зору fusion-моделі можуть реалізовуватися на різних рівнях обробки інформації. Найпоширенішими є об'єднання на рівні вихідних даних, на рівні ознак і на рівні рішень. Вибір конкретної схеми залежить від структури вхідної інформації, ступеня узгодженості використовуваних джерел і вимог прикладної задачі. Загалом fusion-підхід дає змогу враховувати взаємодоповнювальні властивості різних каналів спостереження і тим самим підвищувати точність, стійкість та інформативність підсумкового аналізу [1], [2].

Основна частина. Одним із перспективних напрямів застосування fusion-моделей є розпізнавання об'єктів у підводному середовищі. У межах цієї задачі найбільший практичний інтерес становить спільне використання відеоданих і даних сонара, оскільки зазначені джерела формують різний, але взаємодоповнювальний опис однієї й тієї самої сцени. Відеокамера забезпечує отримання детального візуального представлення об'єкта, що дає змогу аналізувати його форму, контури, текстурні особливості та, за достатньої якості зображення, колірні характеристики. Водночас ефективність оптичного каналу в підводних умовах істотно знижується внаслідок поглинання і розсіювання світла, зменшення контрасту, кольірних спотворень та обмеженої дальності видимості [3], [4].

На відміну від відеоканалу, сонар зберігає працездатність в умовах низької прозорості води, слабкого освітлення та на значних відстанях. Проте сонаРНі дані, як правило, характеризуються меншою візуальною деталізацією, підвищеним рівнем шумів і обмеженою виразністю ознак, безпосередньо пов'язаних із зовнішнім виглядом об'єкта. Таким чином,

використання лише одного із зазначених джерел не дає змоги забезпечити стійке розпізнавання в усьому діапазоні умов спостереження. Об'єднання відеоданих і даних сонара в межах fusion-моделі є обґрунтованим, оскільки дає змогу компенсувати обмеження кожної з модальностей і сформуванати більш повний опис підводної сцени [3], [5].

З точки зору побудови системи розпізнавання об'єднання відеоінформації та сонарних даних може здійснюватися на кількох рівнях. На рівні вихідних даних можливе спільне використання синхронізованих кадрів після виконання калібрування та просторового вирівнювання сенсорів. На рівні ознак доцільним є спільне використання візуальних характеристик зображення та акустичних структурних ознак, що відображають форму, положення й дистанційні особливості об'єкта. На рівні рішень можуть об'єднуватися результати незалежних моделей, які функціонують окремо для кожного каналу спостереження. Подібна багаторівнева організація дає змогу обирати найефективнішу стратегію fusion залежно від доступності узгоджених наборів даних, обчислювальних обмежень і вимог до точності розпізнавання [1], [2].

Для задач підводного розпізнавання об'єктів особливе значення має здатність fusion-моделі адаптуватися до зміни умов спостереження. У ситуаціях, коли якість відеоданих залишається достатньою, система може більшою мірою спиратися на детальні візуальні ознаки. У разі погіршення прозорості води, зниження освітленості або появи виражених оптичних спотворень зростає значущість сонарного каналу, який дає змогу зберігати інформацію про наявність об'єкта, його просторове положення та загальну структуру. Отже, fusion-підхід забезпечує більш стійке функціонування системи розпізнавання порівняно з використанням лише оптичних або лише акустичних даних [3], [5].

Водночас практична реалізація fusion-моделей для підводного розпізнавання пов'язана з низкою труднощів. До них належать відмінність фізичної природи відеоданих і даних сонара, відсутність прямої піксельної відповідності між ними, необхідність часової синхронізації, різниця просторової роздільної здатності, наявність шумів, а також обмеженість відкритих узгоджених наборів даних, придатних для навчання і тестування двоканальних моделей [3], [5], [6]. У зв'язку з цим найважливішими етапами побудови подібних систем є калібрування сенсорів, просторово-часове узгодження модальностей, формування спільного простору ознак і розроблення механізмів адаптивного врахування внеску кожного джерела даних у підсумкове рішення.

Таким чином, застосування fusion-моделі для об'єднання відеоданих і даних сонара слід розглядати як перспективний напрям підвищення ефективності підводного розпізнавання об'єктів. Використання цього підходу дає змогу перейти від ізольованого аналізу окремих каналів спостереження до формування більш цілісного та стійкого уявлення про підводну сцену. Найперспективнішими є архітектури, що забезпечують спільне навчання ознак, міжмодальну взаємодію та адаптивний перерозподіл значущості вхідних даних залежно від умов спостереження [3], [5].

Висновки. Використання fusion-моделей є перспективним підходом до аналізу різнорідних даних, що забезпечує підвищення точності та стійкості інтелектуальних систем за рахунок об'єднання інформації з кількох джерел. У задачах підводного розпізнавання об'єктів особливу практичну значущість має спільна обробка відеоданих і даних сонара. Відеоканал дає змогу отримувати детальний семантичний опис сцени, тоді як сонар зберігає працездатність в умовах поганій видимості та на значних відстанях. Їх об'єднання в межах fusion-моделі дає змогу компенсувати обмеження кожного окремого каналу та підвищити надійність розпізнавання об'єктів у складних умовах підводного середовища.

Перспективи подальших досліджень пов'язані з удосконаленням методів калібрування й узгодження модальностей, розширенням узгоджених наборів даних, а також розробленням адаптивних fusion-архітектур, здатних ефективно функціонувати в режимі реального часу на борту підводних робототехнічних платформ.

Список використаних джерел:

1. Hall D. L., Llinas J. An Introduction to Multisensor Data Fusion. Proceedings of the IEEE. 1997. Vol. 85, No 1. P. 6–23.
2. Castanedo F. A Review of Data Fusion Techniques. The Scientific World Journal. 2013. Vol. 2013. Article 704504.
3. Kim H. G., Seo J., Kim S. et al. Underwater Optical-Sonar Image Fusion Systems. Sensors. 2022. Vol. 22, No 21. Article 8445.
4. Er M. J., Wang Y., Zhu L. et al. Research Challenges, Recent Advances, and Popular Datasets in Deep Learning-Based Underwater Marine Object Detection: A Review. Journal of Marine Science and Engineering. 2023. Vol. 11, No 2. Article 327.
5. Li Y., Wang B., Sun J., Wu X., Li Y. RGB-Sonar Tracking Benchmark and Spatial Cross-Attention Transformer Tracker. arXiv. 2024. Article 2406.07189.
6. Negahdaripour S., Sekkati H., Pirsiavash H. Opti-acoustic Stereo Imaging: On System Calibration and 3-D Target Reconstruction. IEEE Transactions on Image Processing. 2009. Vol. 18, No 6. P. 1203–1214.