

ФІЗИЧНО ОБҐРУНТОВАНЕ КОМПЛЕКСУВАННЯ ВИДИМИХ ТА ІНФРАЧЕРВОНИХ ЗОБРАЖЕНЬ ІЗ ВИКОРИСТАННЯМ МОДЕЛІ TeVNet

С. О. Горбачов^{1,а}, Н. О. Гордійко¹

¹ Навчально-науковий Фізико-технічний інститут

Анотація

У статті розглянуто фізично обґрунтований підхід до комплексування видимих та інфрачервоних зображень із використанням моделі TeVNet. Проведено комп'ютерне моделювання процесу декомпозиції інфрачервоних даних на температурні та відбиті компоненти. Результати оцінено за допомогою безеталонної метрики BRISQUE, що дозволяє якісно визначати інформативність комплексованих зображень.

Ключові слова: інфрачервоне зображення, комплексування зображень, TeVNet, мультиспектральний аналіз, оцінка якості

Вступ

Сучасні технології комп'ютерного зору все частіше використовують багатоспектральні зображення для розв'язання різноманітних задач, таких як безпековий моніторинг, медична діагностика, автономне керування транспортом, дистанційне зондування та метеорологія [1, 2, 3]. Використання кількох спектральних каналів дозволяє компенсувати обмеження від окремих сенсорів та значно підвищити інформативність отриманих зображень.

Особливе місце серед таких каналів посідає інфрачервоний (ІЧ), що є особливо актуальним при роботі в умовах недостатньої видимості, наприклад вночі або за несприятливих погодних умов [4]. ІЧ зображення формуються не тільки за рахунок відбитого світла, але й завдяки власному тепловому випромінюванню об'єктів.

Фізична природа формування ІЧ зображень ускладнює процес їх інтеграції з видимими зображеннями. Традиційні методи комплексування часто нехтують такими фізичними факторами, як емісивність поверхонь, атмосферні впливи або відбите навколишнє випромінювання, що може призводити до неточностей та появи артефактів у результаті злиття [5].

Останнім часом набули популярності нейромережіві підходи до комплексування багатоспектральних зображень, які здатні враховувати ці складні фізичні процеси. Серед них особливе місце займає фізично обґрунтована модель TeVNet, яка виконує декомпозицію ІЧ зображень на три основні фізичні компоненти: температурну карту, карту емісивності та карту відбитого випромінювання.

Метою цієї роботи є дослідження фізично обґрунтованого підходу до комплексування видимих та

інфрачервоних зображень із застосуванням моделі TeVNet і сучасних методів глибокого навчання.

1. Постановка задачі

Традиційне комплексування видимих та інфрачервоних зображень часто нехтує складною фізичною природою ІЧ випромінювання, що може призводити до втрати важливої температурної інформації або появи артефактів при злитті. Проблема полягає в тому, що ІЧ зображення формується як сукупність власного теплового випромінювання об'єкта, відбитого випромінювання навколишнього середовища та атмосферного впливу.

Для коректного об'єднання даних із різних спектральних каналів необхідно спочатку провести фізичну декомпозицію ІЧ сигналу на складові, що відповідають:

- температурному розподілу об'єктів сцени;
- емісивності поверхонь;
- впливу відбитого випромінювання та атмосфери.

У цій роботі акцент зроблено на використанні моделі TeVNet, яка дозволяє розділяти ІЧ зображення на фізично інтерпретовані компоненти. Завдяки цьому комплексування з видимим спектром може здійснюватися більш обґрунтовано, із збереженням температурної достовірності сцени.

Додатково, для оцінювання якості результатів комплексування пропонується застосування безеталонних метрик, зокрема BRISQUE, що дозволяє об'єктивно визначити інформативність і якість злитих зображень без потреби в еталонних даних [6].

^аse1712rega@gmail.com

2. Метод декомпозиції інфрачервоного зображень

2.1. Фізична модель формування інфрачервоного зображення

Інфрачервоні зображення формуються в результаті суми трьох основних компонент [5]:

- власного теплового випромінювання об'єкта;
- відбитого випромінювання від навколишніх об'єктів;
- власного випромінювання атмосфери.

Цей процес описується рівнянням:

$$S_{\lambda} = \tau_{atm} e_{\lambda} B_{\lambda}(T) + \tau_{atm} (1 - e_{\lambda}) B_{\lambda}(T_{env}) + (1 - \tau_{atm}) B_{\lambda}(T_{atm}), \quad (1)$$

де $B_{\lambda}(T)$ — спектральна густина випромінювання згідно закону Планка, e_{λ} — емісивність матеріалу, τ_{atm} — коефіцієнт пропускання атмосфери, T_{env} — температура навколишнього середовища, T_{atm} — температура атмосфери.

2.2. Архітектура моделі TeVNet

Моделі TeVNet [5] була розроблена для фізично обґрунтованої декомпозиції інфрачервоного зображення на три складові:

- температуру об'єкта (T);
- карту емісивності (e);
- карту відбитого випромінювання (V).

TeVNet має структуру згорткової нейронної мережі, яка включає:

- вхідний згортковий шар (розмір фільтра 3×3) для первинної екстракції ознак;
- три згорткових блоки із зростаючим числом фільтрів (32, 64, 128), кожен із активацією ReLU;
- три окремі вихідні гілки для передбачення температури, емісивності та відбитого випромінювання.

2.3. Функція втрат

Функція втрат для навчання TeVNet враховує не лише точність реконструкції сигналу S , але й точність прогнозування фізичних компонент [5]:

$$L_{rec} = \|S - \hat{S}\|_2^2 + \lambda_1 \|T - T_{real}\|_2^2 + \lambda_2 \|e - e_{real}\|_2^2, \quad (2)$$

де λ_1 та λ_2 — вагові коефіцієнти для відповідних складових. В роботі використовували значення $\lambda_1 = 1.0$ та $\lambda_2 = 0.5$ для балансування точності реконструкції та фізичної інтерпретації вихідних даних.

2.4. Переваги підходу

Завдяки використанню фізичних законів теплового випромінювання (закон Планка, закон Стефана–Больцмана), модель TeVNet дозволяє більш коректно реконструювати фізичні властивості сцени. Це

створює підґрунтя для побудови більш інформативного і точного результату комплексування видимого та ІЧ зображень.

3. Експериментальні результати

3.1. Підготовка даних

Для навчання і тестування моделі використовувався датасет LLVIP (Low-Light Visible and Infrared Person Dataset) [7], що містить понад 15000 пар видимих та інфрачервоних зображень, отриманих у різних умовах слабкого освітлення.

Попередня обробка включала нормалізацію, видалення шуму за допомогою гаусівського фільтра, а також геометричні аугментації для підвищення узагальнюючої здатності моделі.

3.2. Навчання моделі TeVNet

Моделі TeVNet навчалась з використанням оптимізатора Adam із початковим коефіцієнтом навчання 1×10^{-4} . Розмір батча складав 2 зображення. Навчання проводилося протягом 5 епох на графічному процесорі NVIDIA RTX 2060.

Під час навчання мінімізувалась функція втрат, яка враховувала точність реконструкції спектральної густини сигналу, а також точність передбачення температурних та емісивних карт.

3.3. Результати комплексування із застосуванням TeVNet

Після декомпозиції інфрачервоного зображення за допомогою TeVNet окремо виділялись температурна карта об'єктів та карта відбитого випромінювання. Подальше комплексування із видимим спектром проводилося шляхом поєднання текстурної інформації із температурною картою.

Результати показали:

- Значне покращення температурної достовірності об'єктів на комплексованому зображенні, особливо для холодних фонів і теплих об'єктів.
- Зменшення кількості артефактів, пов'язаних із некоректним врахуванням відбитого випромінювання.
- Більш чітке відтворення контурів об'єктів навіть у зонах слабкого освітлення.

3.4. Оцінка якості

Застосування метрики BRISQUE до результатів комплексування показало зниження середнього значення оцінки порівняно з вхідними зображеннями. Це свідчить про вищу якість зображень за сприйняттям людиною та кращу передачу фізичних властивостей сцени. Приклади комплексованих зображень з підтвердженням оцінок наведено на рисунках 1(a) та 1(б).

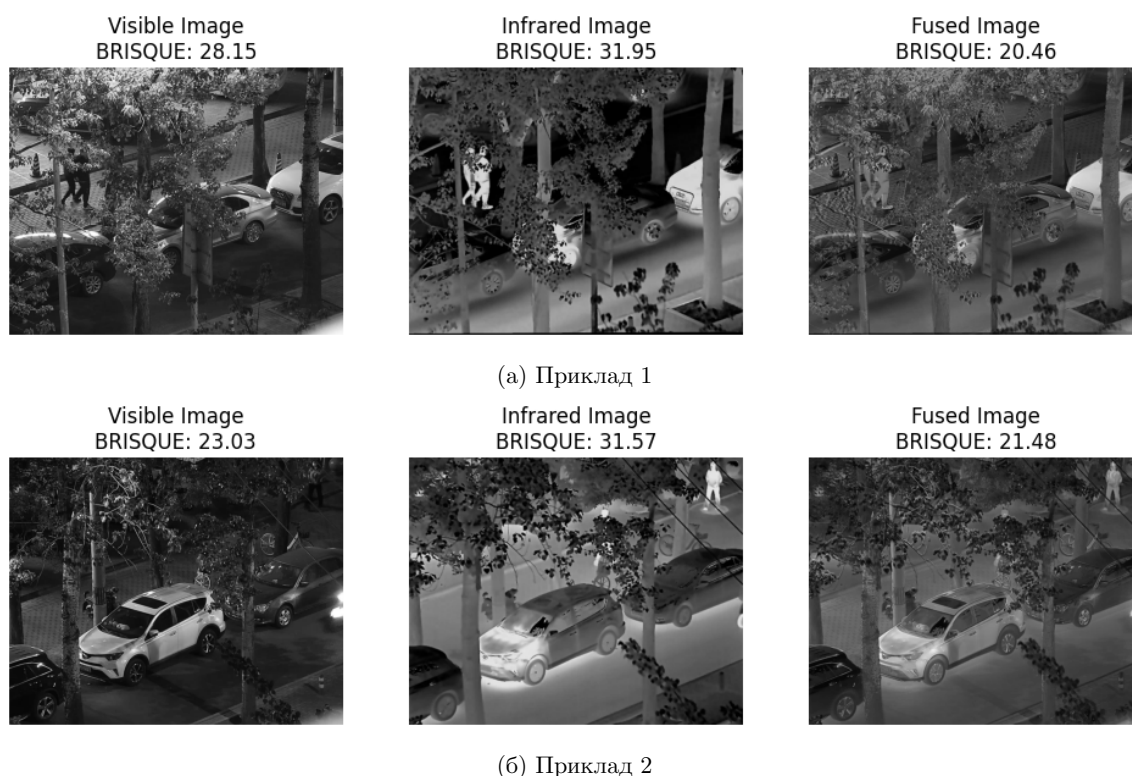


Рис. 1. Результати комплексування видимого та ІЧ-зображень [7]

Висновки

У цій роботі запропоновано фізично обґрунтований підхід до комплексування видимих та інфрачервоних зображень із використанням моделі TeVNet.

Моделю TeVNet дозволяє виконувати декомпозицію інфрачервоного зображення на температурну карту, карту емісивності та карту відбитого випромінювання. Це забезпечує фізично достовірне представлення сцени та значно покращує результати подальшого комплексування зображень.

Проведені експерименти на наборі даних LLVIP показали, що застосування TeVNet перед комплексуванням дозволяє:

- зберегти температурні особливості об'єктів навіть у складних умовах освітлення;
- зменшити кількість артефактів, спричинених відбитим випромінюванням;
- покращити візуальну якість злитих зображень за оцінкою безеталонної метрики BRISQUE.

Результати роботи підтверджують ефективність використання фізично обґрунтованої декомпозиції інфрачервоних зображень для покращення мультиспектрального комплексування. Подальші дослідження можуть бути спрямовані на інтеграцію моделей підвищеної спектральної роздільної здатності та розширення підходу для гіперспектральних даних.

Перелік використаних джерел

1. Infrared and Visible Image Fusion Methods and Applications: A Survey / J. Ma, H. Zhang, Y. Liu, X. Fan // IEEE Access. — 2023. — Т. 11. — С. 12456—12472.
2. Nalepa J. Recent Advances in Multi- and Hyperspectral Image Analysis // Sensors. — 2021. — Т. 21. — С. 6002.
3. NASA SPoRT. Satellite Multispectral Image Processing for Meteorology : tex. звіт. / NASA Short-term Prediction Research ; Transition Center. — 2023. — Technical report.
4. Edge-Preserving Deep Learning for Infrared and Visible Image Fusion / Y. Liu, L. Wang, J. Cheng, H. Li, X. Chen // Sensors. — 2023. — Т. 23. — С. 4182. — DOI: [10.3390/s23094182](https://doi.org/10.3390/s23094182). — URL: <https://www.mdpi.com/1424-8220/23/9/4182>.
5. Authors. TeVNet: A Physically Grounded Infrared Image Decomposition Model // ArXiv preprint. — 2024. — Available: <https://arxiv.org/abs/2407.09299>.
6. Горбачов С. О., Гордійко Н. О. Оцінювання якості зображень систем мультиспектрального моніторингу без порівняння з еталоном // Збірник матеріалів XXII Всеукраїнської науково-практичної конференції студентів, аспірантів та молодих вчених «Теоретичні і прикладні проблеми фізики, математики та інформатики». — Київ, Україна, 13.05.2024—17.05.2024. — С. 19—21.
7. LLVIP: A Low-Light Visible and Infrared Person Dataset for Detection and Recognition / R. Xu, Y. Wang, H. Yu, C. Yuan // arXiv preprint arXiv:2112.02425. — 2021. — URL: <https://arxiv.org/abs/2112.02425>.