

Д.В. Магеровський (Національний університет «Львівська політехніка»),
В.В. Сенник, канд. техн. наук, доцент, *Т.В. Магеровська*, канд. фіз.-мат. наук, доцент
(Львівський державний університет внутрішніх справ)

АНАЛІЗ АЛГОРИТМІВ ТА ОГЛЯД ПЕРСПЕКТИВНИХ ТЕХНОЛОГІЙ, ПРИДАТНИХ ДЛЯ РОЗРОБКИ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ПРОЦЕСУ ПОКРАЩЕННЯ ВІДЕОЗОБРАЖЕННЯ В РЕЖИМІ ON-LINE

Проведено аналіз перспективних технологій (алгоритмів) обробки медіаданих, на основі яких можливі подальші розробки програмного забезпечення для покращення якості відеозображень безпосередньо під час проведення процесу відеоспостереження. Актуальність цієї проблеми, на сьогодні, пов'язана з тим, що, зокрема, правоохоронні органи, маючи значні апаратно-технічні можливості, через певну сукупність таких факторів, як недостатність освітлення, недостатня роздільна здатність камер відеоспостереження, погані погодні умови тощо, не дають змоги чітко встановити окремі деталі правопорушень, а також не дозволяють за біометричними даними встановити особу причетну до таких подій. Розвиток, удосконалення розглянутих у даній публікації алгоритмів покращення якості зображення, у тому числі і в режимі реального часу, дасть можливість розробити та впровадити у практичну діяльність правоохоронних структур відповідне програмне забезпечення, що в свою чергу сприятиме підвищенню кількості розкритих злочинів (правопорушень).

Ключові слова: алгоритм обробки відеозображення, відеоспостереження, машинне навчання

D.V. Magerovskyi, V.V. Senyk, T.V. Magerovska

ALGORITHM ANALYSIS AND OVERVIEW OF PERSPECTIVE TECHNOLOGIES, AVAILABLE FOR ONLINE IMAGING IMPROVEMENT SOFTWARE DEVELOPMENT

In this article was conducted the perspective media data processing technologies (algorithms) analysis, on which base there are possible further software development for video stream quality improvement directly in the process of video surveillance. Actuality of this problem is related to that fact, that having hardware possibilities, law enforcement cannot establish crimes details and/or persons, related to that crime, because of such factors as lack of lighting, cameras low resolution, bad weather conditions, etc. Development and improvement of considered in the article algorithms of imaging quality improvement, including real time calculating, may give a possibility to develop and implement an appropriate software to law enforcements activity, which will cause the increase of disclosed crimes (offenses).

Key words: video surveillance processing algorithm, machine learning.

Постановка проблеми. З останні роки в Україні розпочався стрімкий розвиток технологій відеокommунікації. Найбільшим прогресом, безсумнівно, є поява 3G-інтернету, що дає змогу кожному з нас, фактично мати, постійний доступ до інформації будь-якого типу у будь-якому місці. Не стоять осторонь у використанні сучасних досягнень у галузі відеотехнологій і правоохоронні органи. Частка розкритих злочинів з їх використанням постійно зростає. Однак, попри високу роздільну здатність сучасних відеокамер, вони часто не можуть забезпечити належної якості зображення через низку таких умов, як низький рівень освітленості, недостатньо досконалі алгоритми фокусування тощо. Сукупність усіх цих факторів

призводить до нечітких (часто розмитих) відеозображень, що, у свою чергу, не дозволяє отримати вичерпну інформацію про подію, чітко встановити деталі правопорушень та розпізнати їх виконавців. Враховуючи, зокрема, і той факт, що в Україні є велика кількість застарілих відеокamer спостереження, а також те, що такого роду камери часто розташовані на вулицях, де погодні умови заважають якості зйомки, виникає потреба застосування сучасних алгоритмів покращення відеозображення.

Аналіз останніх досліджень та публікацій. Науковці усього світу давно працюють над розробками розробки, які б дали змогу підвищити якість зйомки безпосередньо під час процесу зйомки (інфрачервоне освітлення, активне шумозаглушення тощо). Розробники ж програмного забезпечення відносно недавно почали активно розробляти алгоритми, які б давали можливість покращення зображення після проведення зйомки, наприклад [1, 2].

Виділення складових проблем, які підлягають розв'язанню. Відсутність активних розробок, які б давали можливість покращення зображення у процесі проведення зйомки пов'язане, насамперед, з тим, що потужність обчислювальних пристроїв, які випускалися промисловістю масово, не була достатньою. Однак, на сьогодні день ситуація докорінно змінилася. Із врахуванням цього актуальність питання розробки алгоритмів та програмного забезпечення, які б покращували якість відео в режимі реально часу лише зростає.

Постановка завдання. Через масовість достатньо потужних пристроїв, з'явилася велика кількість ентузіастів та професіоналів, що задалися ідеєю розробки алгоритмів для підвищення якості фото або потокового відео за фактом його запису у пам'ять обчислюваного пристрою в режимі реального часу. У цій роботі проведено огляд перспективних технологій обробки медіа даних, описано принцип їх роботи, а також наведено ідеї що до використання такого роду алгоритмів у професійній діяльності.

Виклад основного матеріалу. Попри те, що стаття присвячена аналізу технологій покращення якості потокового відео в режимі реального часу, розгляд хотілося б розпочати з технології *Neural Enhance*, оскільки ця технологія має свою перспективу розвитку. Наприкінці жовтня 2016 року, Олександр Шампандард представив свій алгоритм підвищення якості фотографій. Написаний мовою Python, алгоритм містить всього 300 стрічок коду, проте демонструє дуже відчутні результати. Автор демонструє приклад покращення якості зображення з допомогою рис. 1 [3].

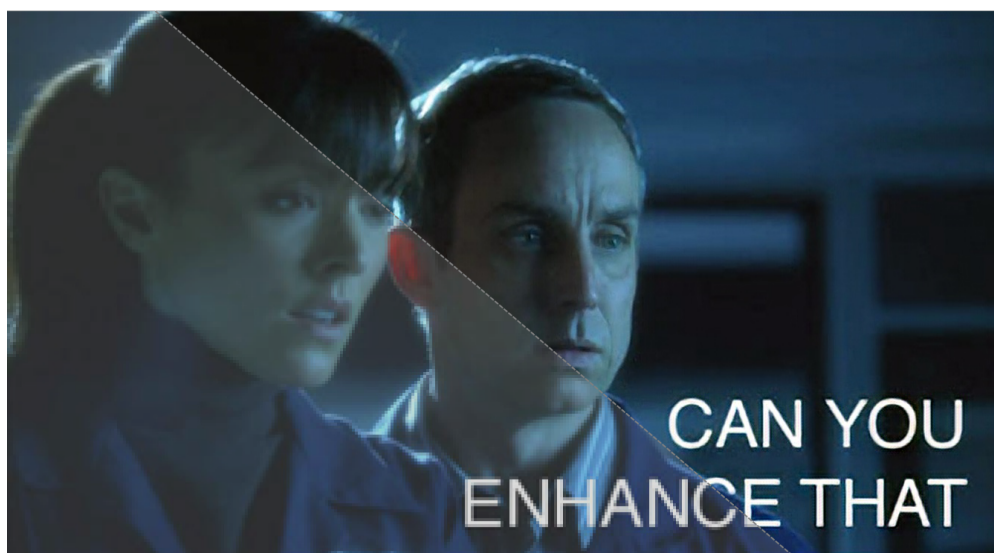


Рисунок 1 – Приклад роботи алгоритму Шампандарда

Алгоритм використовує навчену штучну нейронну мережу. При цьому, автору вдалося уникнути втрат перцептрона, використовуючи оптимізацію, описану в роботі науковців Стендфорського університету [4]. Суть оптимізації полягає у навчанні перцептрона з використанням стохастичного градієнтного спуску для мінімізації зваженої комбінованої функцій втрат:

$$W^* = \arg \min_W E_{x,\{y_i\}} [\sum_{i=1} \lambda_i l_i(f_W(x), y_i)] \quad (1)$$

де λ_i – коефіцієнт ваг, l_i – скалярне значення функції, $f_W(x)$ – функція мапінгу, x – вхідне зображення, y_i – цільове зображення, W – параметризуючі ваги, $E_{x,\{y_i\}}$ – функція втрат.

Мережа трансформації зображення – це глибока надточна нейронна мережа, параметризована з допомогою ваг W . Вона трансформує вхідні зображення x у вихідні зображення y з допомогою механізму мапінгу $y = f_W(x)$. Кожна функція втрат обраховує скалярне значення $l_i(y, y_i)$, визначаючи різницю між вихідним зображенням y та цільовим зображенням y_i . Мережа трансформації зображень навчається використовуючи спадний стохастичний градієнт для мінімізації зваженої комбінації функцій втрат.

Автор провів велику роботу над оптимізацією алгоритму. Так операції конволюції/деконволюції [5] використовують невеликий об'єм пам'яті. Навчання нейронної мережі у даному випадку відбувається за формулою:

$$l(W_{1:L}, b_{1:L}) = \frac{1}{\gamma^2 H W} \sum_{x=1}^{\gamma H} \sum_{y=1}^{\gamma W} \left(I_{x,y}^{HR} - f_{x,y}^L(I^{LR}) \right)^2, \quad (2)$$

де HR – зображення з великою роздільною здатністю, LR – з малою роздільною здатністю, H – висота зображення (пікс), W – ширина зображення (пікс), $l(W_{1:L}, b_{1:L})$ – функція розрахунку середньоквадратичної помилки розташування пікселів навхрест (MSE), γ – коефіцієнт зберігання (0; 1).

На рис. 2 приведено приклад запропонованої ефективної багатопшарової нейронної мережі (ESPCN) з двома шарами для вилучення ознак карти та міжпексельний шар, що агрегує карти функцій з середовища з низьким розширенням та будує зображення з високим розширенням в 1 крок [5].

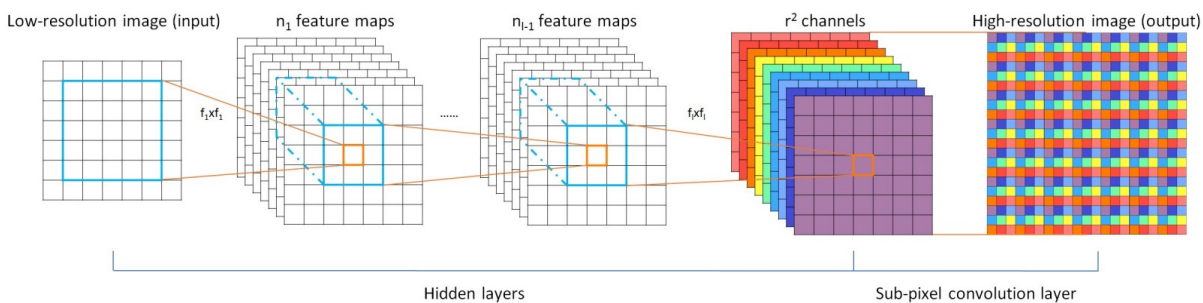


Рисунок 2 – Приклад запропонованої нейронної мережі

Для розрахунку й мінімізації втрат вмісту зображення, що реставрується, автор використав 19-шарову мережу, описану Симоньяном/Цісерманом (мережу VGG) [6]:

$$l_{VGG/i,j}^{SR} = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} \left(\phi_{i,j} \left(G_{\theta_G} (I^{HR}) \right)_{x,y} - \phi_{i,j} \left(G_{\theta_G} (I^{LR}) \right)_{x,y} \right)^2, \quad (3)$$

де $W_{i,j} H_{i,j}$ – розширення карт функції, $\phi_{i,j}$ – карта функції, $G_{\theta_G} (I^{LR})$ – зображення, що реконструюється, $G_{\theta_G} (I^{HR})$ – зображення, в яке буде реконструйовано, I^{LR} – еталонне зображення.

Програма може працювати у двох режимах: рендеринг з допомогою процесора або графічної карти. Автор стверджує, що для рендерингу Full-HD зображення за допомогою процесора буде затрачено близько 40 секунд, за допомогою графічної карти – 5 секунд. Таким чином, на сьогодні про застосування цього алгоритму для покращення зображення в ре-

жимі реального часу мова не йде. Проте, він створений непрофесіоналом й демонструє потенціал технології покращення якості медіаданих. Програма розповсюджується за ліцензією вільного програмного забезпечення, що дає змогу вносити свої зміни в код, та розвивати проект у майбутньому з допомогою інших розробників.

Наступною хотілося б розглянути програму *Magic Pony*. Про стартап з назвою *Magic Pony* не було відомо, поки мережа Twitter не придбала технологію за 150 мільйонів доларів США [7]. Стартап, заснований випускниками Імперського коледжу Лондона, використовує машинне навчання [8], щоб створювати високоякісні відео із зернистих кадрів, й, ймовірно, буде інтегрований у сервіси відтворення потокового відео Twitter, такі як Periscope і Vine. Запатентована технологія зобов'язана покращити якість фото або відео, у тому числі, в умовах поганої освітленості.

Наразі, *Magic Pony* тестувалася в умовах трансляції відеоігор при малій потоковій швидкості, що вело до втрати різкості зображення. Результатом тестування були прийнятні результати з підвищення різкості кадрів. Так фокус-група сприймала ігровий процес без «роздратування», текстова інформація була читабельною. Результат обробки статичного зображення командою розробників *Magic Pony* показано на рис. 3 [9].

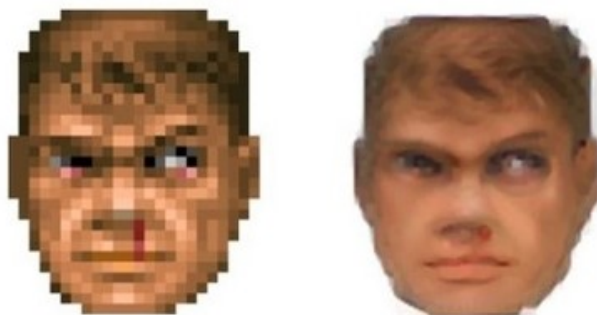


Рисунок 3 – Результат роботи Magic Pony над статичним зображенням

Можливості для застосування цієї технології колосальні. Перш за все, використання технології для трансляції потокових відео зменшить використання трафіку, що використовується при їх ідентичній вихідній якості. Таким чином, наприклад, власники мобільних пристроїв матимуть значно менше обмежень під час користування 3G Інтернетом. Технологія застосовна всюди, де використовується потокове відео. Військові зможуть отримувати якісніші знімки ландшафту за допомогою аеророзвідки на великих висотах або роботою із безпілотними літальними засобами. Камери відеоспостереження стануть кориснішими для правоохоронних структур, оскільки підвищиться відсоток розкритих правопорушень завдяки тому, що вдалося проаналізувати ті записи, які через низьку якість не можна було проаналізувати до цього.

Одним з базових алгоритмів підвищення якості зображень є використання *карт Кохонена*, які самоорганізуються за допомогою неймереж. Карта Кохонена складається з набору нейронів, кількість яких задається аналітиком. Кожен нейрон, у свою чергу, описується двома векторами: а) вектор ваги нейрона t (у загальному випадку, розмірність вектора збігається з розмірністю векторів вхідних даних); б) вектор координат p (у загальному випадку, вектор описує точку в просторі, в якій розташований нейрон). Зазвичай нейрони мають на площині в вершинах регулярної решітки з квадратними або шестикутними осередками.

Існує кілька способів початкової ініціалізації карти, шляхом присвоєння всім векторам ваг нейронів, значень одним з таких способів:

1. Задання всіх координат випадковими числами.
2. Привласнення вектору ваги значення випадкового спостереження з вхідних даних.

3. Вибір векторів ваги з лінійного простору, натягнутого на головні компоненти набору вхідних даних.

Варто зауважити, що кожен нейрон після ініціалізації стає нерухомий на карті, тобто вектор p не змінюється протягом усього навчання.

Навчання відбувається за ітераціями.

Нехай t – номер ітерації. Припустимо, що ініціалізація – це ітерація з номером 0. Далі, виконуються такі операції:

1. Обирається випадковий вектор $x(t)$ з набору вхідних значень.

2. Знаходяться відстані до всіх векторів-ваг усіх нейронів карти.

Для цієї операції обирається одне із значень (в загальному випадку – середньоквадратичне відхилення). Ведеться пошук найближчого значення до вхідного $x(t)$:

$$d(x(t), m_e(t)) \leq d(x(t), m_i(t)), \quad (4)$$

де $m_e(t)$ – вектор ваги нейрона-переможця $M_e(t)$, $m_i(t)$ – вектор ваги i -го вузла на карті, $d(x(t), m_i(t))$ – міра відхилення.

У випадку кількох найближчих нейронів, один з них обирається випадковим чином.

3. Визначення ступеня сусідства нейронів й зміна ваг нейронів на карті:

3.1. Вибирається міра сусідства, функція $h_{ei}(t)$ (зазвичай у якості цієї функції використовується функція Гауса):

$$h_{ei}(t) = \alpha(t) \exp\left(-\frac{\|r_e - r_i\|^2}{2\sigma^2(t)}\right), \quad (5)$$

де $0 < \alpha(t) < 1$ – навчальний співмножник, що монотонно спадає з кожною наступною ітерацією; r_e, r_i – координати вузлів $M_e(t)$ і $M_i(t)$ на карті, $\sigma(t)$ – співмножник, що зменшує кількість сусідів з ітераціями, монотонно спадає.

Параметри α, σ та характер їх спадання задаються аналітиком.

3.2. Обчислення помилки карти.

Вимірюється вектор ваг за формулою:

$$m_i(t) = m_i(t-1) + h_{ei}(t)(x(t) - m_i(t-1)). \quad (6)$$

4. Вибір умови зупинки.

Для визначення критерію зупинки у більшості випадків використовується помилка карти, наприклад, як середня арифметична відстань між спостереженнями і векторами ваги відповідних їм ВМУ.

$$\frac{1}{N} \sum_{i=1}^N \|x_i - m_e\|, \quad (7)$$

де N – кількість елементів набору вхідних даних.

Можна виділити такі елементи для моделювання системи відновлення мультимедійних даних:

x – вектор спостережуваних даних. Під час відновлення мультимедійної інформації це може бути набір пікселів (семплів, набір блоків стиснення зображень), згрупованих за певним правилом;

m – вектор ваги нейрона. За визначенням, повинен бути однієї розмірності з x ;

$h_{ei}(t)$ – міра сусідства нейронів. Деяка функція, яка повертає відстань між нейронами і залежить від номера ітерації t . За певних умов, може визначати топологію поверхні, на яку нанесена карта;

$(x(t), m_i(t))$ – відхилення, що показує, наскільки вектор $x(t)$ не схожий на вектор $m_i(t)$.

Таким чином, за присутності даних елементів в моделі можна покращувати якість зображення або відновлювати пошкоджені ділянки з допомогою карт Кохонена [10].

Висновок. Таким чином, розвиток та удосконалення, розглянутих у цій публікації, алгоритмів покращення якості зображення, у тому числі і в режимі реального часу, дасть можливість розробити та впровадити у практичну діяльність правоохоронних структур відповідне програмне забезпечення, що в свою чергу сприятиме підвищенню кількості розкритих злочинів (правопорушень) завдяки можливості не лише встановлювати окремі деталі таких подій, а й розпізнавати осіб-виконавців через on-line аналіз їх біометричних даних.

Список літератури:

1. Як поліпшити якість відео [Електронний ресурс]. – Режим доступу : <http://uk.wondershare.com/video/improve-video-quality.html>
2. Качество видео и его улучшение – программа Neat Video [Електронний ресурс]. – Режим доступу : <http://vide0.org.ua/kachestvo-video-i-ego-uluchshenie-programma-neat-video/>
3. Neural Enhance [Електронний ресурс]. – Режим доступу : <https://github.com/alexjc/neural-enhance>
4. J. Johnson, A. Alahi, L. Fei-Fei. Perceptual Losses for Real-Time Style Transfer and Super-Resolution [Електронний ресурс]. – Режим доступу : <https://arxiv.org/pdf/1603.08155v1.pdf>.
5. W. Shi , J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, Z. Wang. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network [Електронний ресурс]. – Режим доступу : <https://arxiv.org/pdf/1609.05158v2.pdf>
6. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [Електронний ресурс]. – Режим доступу : <https://arxiv.org/pdf/1609.04802v2.pdf>
7. Madhumita Murgia. Twitter pays \$150m for London AI startup Magic Pony [Електронний ресурс]. – Режим доступу : <http://www.telegraph.co.uk/technology/2016/06/20/twitter-buys-london-startup-magic-pony-for-close-to-150m/>
8. Evolution of machine learning [Електронний ресурс]. – Режим доступу : http://www.sas.com/en_us/insights/analytics/machine-learning.html.
9. Will Knight. Artificial Intelligence Can Now Design Realistic Video and Game Imagery [Електронний ресурс]. – Режим доступу : <https://www.technologyreview.com/s/601258/artificial-intelligence-can-now-design-realistic-video-and-game-imagery/>
10. T. Kohonen Self-organizing Maps – 3 ed. – Berlin, Heidelberg, New York, Barcelona, Hong Kong, London, Milan, Paris, Singapore, Tokyo, Springer, 2001.

References:

1. Yak polipshyty yakist video [Electronic resource] / Access mode: <http://uk.wondershare.com/video/improve-video-quality.html>.
2. Kachestvo video i ego uluchshenie – programma Neat Video [Electronic resource] / Access mode: <http://vide0.org.ua/kachestvo-video-i-ego-uluchshenie-programma-neat-video/>
3. Neural Enhance [Електронний ресурс]. – Режим доступу : <https://github.com/alexjc/neural-enhance>
4. J. Johnson, A. Alahi, L. Fei-Fei. Perceptual Losses for Real-Time Style Transfer and Super-Resolution [Electronic resource] / Access mode: <https://arxiv.org/pdf/1603.08155v1.pdf>.
5. W. Shi , J. Caballero, F. Huszar, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, Z. Wang. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network [Electronic resource] / Access mode: <https://arxiv.org/pdf/1609.05158v2.pdf>
6. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [Electronic resource] / Access mode: <https://arxiv.org/pdf/1609.04802v2.pdf>

7. Madhumita Murgia. Twitter pays \$150m for London AI startup Magic Pony [Electronic resource] / Access mode: <http://www.telegraph.co.uk/technology/2016/06/20/twitter-buys-london-startup-magic-pony-for-close-to-150m/>

8. Evolution of machine learning [Електронний ресурс]. – Режим доступу : http://www.sas.com/en_us/insights/analytics/machine-learning.html.

9. Will Knight. Artificial Intelligence Can Now Design Realistic Video and Game Imagery [Electronic resource] / Access mode: <https://www.technologyreview.com/s/601258/artificial-intelligence-can-now-design-realistic-video-and-game-imagery/>

10. T. Kohonen Self-organizing Maps – 3 ed. – Berlin, Heidelberg, New York, Barcelona, Hong Kong, London, Milan, Paris, Singapore, Tokio, Springer, 2001.

