

ВИКОРИСТАННЯ ІНДЕКСУ СТРУКТУРНОЇ ПОДІБНОСТІ ДЛЯ ЗНАХОДЖЕННЯ КОЕФІЦІЄНТІВ ІНТЕРПОЛЯЦІЙНОГО ФІЛЬТРА

На даний час поширеною задачею є відображення зображень із низькою роздільною здатністю на дисплеях з високою роздільною здатністю. Саме тому виникає необхідність проведення просторового розширення зображень в режимі «реального часу». Для вирішення поставленого завдання пропонується новий алгоритм отримання коефіцієнтів інтерполяційного фільтра на основі адаптивного кодування з динамічним діапазоном та наводяться результати його досліджень.

The displaying of lower resolution video sequences on higher resolution displays has become a very common task. Thus, there is a strong demand for real-time, high quality video interpolation. To address this task the new method for coefficients training of adaptive coding with dynamic ranges (ADRC) based method for video interpolation is described and evaluated.

1. ВСТУП

В останній час праці багатьох вчених присвячені розгляду низки методів та алгоритмів для просторового розширення зображень на основі адаптивного кодування блоку зображення з динамічним діапазоном [2,10,16]. Дані методи у процесі навчання характеризуються високою швидкістю та простотою обчислення їх інтерполяційних коефіцієнтів. Згадані коефіцієнти оптимізуються відносно метрики оцінювання якості зображення. В переважній більшості в ролі такої метрики виступає найменша середньоквадратична помилка (НСП, англ. Mean Square Error – MSE). Проте використання даної метрики слабо корелюється із існуючими суб'єктивними оцінками якості [2,8]. Заміна метрики MSE на метрику, результати якої мають кращу кореляцію із суб'єктивними оцінками, дозволяє отримати вихідні інтерполювані зображення суттєво вищої якості.

2. ІНТЕРПОЛЯЦІЙНИЙ ФІЛЬТР НА ОСНОВІ АДАПТИВНОГО КОДУВАННЯ З ДИНАМІЧНИМ ДІАПАЗОНОМ

Одним з найбільш відомих інтерполяційних фільтрів є фільтр побудований на основі адаптивного кодування з динамічним діапазоном

¹ Національний університет «Львівська політехніка»

– АКДД (англ. Adaptive Dynamic Range Coding – ADRC) [1, 7, 10], вихідний сигнал якого залежить від вхідних даних. Моментні коефіцієнти фільтра протягом інтерполяції залежать від вмісту локального блоку зображення.

На рис. 1. наведена узагальнена блок-схема інтерполювання зображення з використанням фільтра на основі АКДД. На вхід схеми подається зображення формату стандартної роздільної здатності (англ. Standard-definition SD). Інтерполяція проводиться в межах блоку 3×3 пікселів. У кожному блоці виконується адаптивне кодування і знаходиться бітова маска блоку. На основі отриманої маски з пошукової таблиці вибираються коефіцієнти інтерполяційного фільтра і обчислюється результуюче значення.

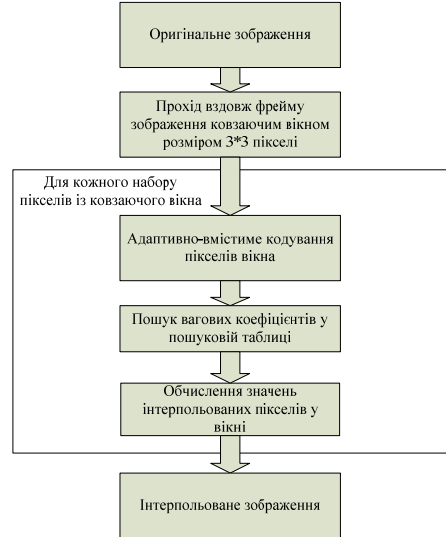


Рис. 1. Схема алгоритму адаптивного кодування з динамічним діапазоном

Нехай F_{HD} – яскравість оригінальних пікселів формату високої роздільної здатності (англ. High-definition– HD), тоді F_{HI} – значення яскравості отримане в результаті інтерполювання, яке є зваженою сумою 9-ти пікселів формату SD інтерполяційного вікна і обчислюється згідно виразу:

$$F_{HI} = \sum_{l=0}^8 w_c(k) \cdot F_{SD}(k), \quad (1)$$

де $w_c(k)$ – вагові коефіцієнти класу c .

Для обчислення коефіцієнтів фільтра необхідно проводити процес навчання. Коефіцієнти інтерполяційного фільтра знаходяться шляхом мінімізації значення MSE. Результуючі коефіцієнти зберігаються в пошуковій таблиці (англ. Look Up Table – LUT). Використання критерію MSE є доволі складним, оскільки великим є число класів. Однак навчання проводиться лише один раз.

3. ІНДЕКС СТРУКТУРНОЇ ПОДІБНОСТІ ЗОБРАЖЕННЯ

В процесі знаходження коефіцієнтів фільтра на основі АКДД здійснюється мінімізація сумарної квадратичної помилки між блоками оригінального та інтерпольованого зображення. При такому підході не враховується структурна інформація блоку. У випадку врахування такої інформації інтерпольовані зображення матимуть кращу суб'єктивну оцінку якості.

У працях [12, 15] розглядається новий підхід до оцінки якості зображень. При цьому показано, що людська зорова система (ЛЗС) є надзвичайно заплутаною, високонелінійною системою і її основною функцією є виділення структурної інформації із поля спостереження. Вимірювання структурних спотворень дозволяє проводити оцінку спотворення зображення, яке сприйняте спостерігачем. Саме спотворення зображення розглядається як структурна інформаційна втрата. Вимірювання структурних інформаційних втрат ґрунтується на припущенні, що ЛЗС функціонує таким самим чином, оскільки “структурна інформація” представляє такі атрибути, які відображають структуру об'єктів в місці дії і є незалежними від середнього сигналу якості і контрасту зображення. Це приводить до підходу оцінки якості зображення, який відокремлює вимірювання сигналу за яскравістю, контрастністю і структурними спотвореннями. У працях [11, 13] запропоновано простий алгоритм індексування рівності зображення.

Нехай x і y два додатні сигнали, які були зрівноважені один з одним (наприклад, два фрагменти витягнуті з одного просторового розташування з порівнюваних зображень). Для даних сигналів μ_x – середнє значення x , μ_y – середнє значення y , σ_x – дисперсія x , σ_y – дисперсія y , σ_{xy} – коваріація x і y відповідно. Середнє і стандартне відхилення (квадратний корінь з дисперсії) сигналу вважаються наближеними оцінками яскравості і контрасту сигналу.

Коваріація (нормалізована дисперсією) може вважатися показником, наскільки один сигнал нелінійно змінений відносно іншого.

Порівняльні вимірювання яскравості, контрасту і структурної інформації визначаються відповідно наступними характеристиками [13]:

$$\begin{aligned} l(x, y) &= \frac{2\mu_x\mu_y}{\mu_x^2 + \mu_y^2} \\ c(x, y) &= \frac{2\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2} \\ s(x, y) &= \frac{\sigma_{xy}}{\sigma_x\sigma_y}. \end{aligned} \quad (2)$$

Вищенаведені умови – концептуально незалежні у розумінні, що перші дві залежать лише від сигналів яскравості і контрасту двох порівнюваних зображень. Зміни сигналів яскравості або контрасту будь-якого зображення не впливають на третю умову. Характеристика $s(x, y)$ не використовує пряме описове представлення структур зображення, проте вона відображає схожість між структурами двох зображень. Вона рівна одиниці лише при умові рівності структур порівнюваних зображень (структурна інформація розглядається як така, що відмінна від інформації про яскравість і контраст).

Структурна подібність $S(x, y)$ між сигналами x і y визначається згідно [13]:

$$S(x, y) = l(x, y) \cdot c(x, y) \cdot s(x, y) = \frac{4\mu_x\mu_y\sigma_{xy}}{(\mu_x^2 + \mu_y^2)(\sigma_x^2 + \sigma_y^2)} \quad (3)$$

Якщо два сигнали представлені дискретно, статичні характеристики можуть бути оцінені наступним чином:

$$\mu_x = \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, \quad \mu_y = \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (4)$$

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2, \quad \sigma_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2 \quad (5)$$

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}), \quad (6)$$

де N – кількість пікселів, для яких проводиться оцінка даних параметрів.

При обчисленні індекса згідно виразу (3), коли $(\mu_x^2 + \mu_y^2)$ або $(\sigma_x^2 + \sigma_y^2)$ близькі до нуля, результуюче вимірювання є нестійким. Даний ефект часто спостерігається при роботі з плоскими областями зображення. Для уникнення даної ситуації використовують новий підхід, який отримав назву індекса структурної подібності (Structural SIMilarity (SSIM)) між сигналами x і y [13]:

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (7)$$

Дві константи C_1 і C_2 обчислюються згідно наступних виразів:

$$C_1 = (K_1 L)^2, \quad C_2 = (K_2 L)^2, \quad (8)$$

де L динамічний діапазон значень піксела (для 8-ми бітного зображення у тонах сірого $L = 255$); K_1 та K_2 дві константи. Значення констант пропонується наступним [13]: $K_1 = 0,01$ і $K_2 = 0,03$.

4. ЗНАХОДЖЕННЯ КОЕФІЦІЄНТІВ ІНТЕРПОЛЯЦІЙНОГО ФІЛЬТРА НА ОСНОВІ ОПТИМІЗАЦІЇ ЗНАЧЕННЯ ІНДЕКСУ СТРУКТУРНОЇ ПОДІБНОСТІ

Врахування структурної інформації під час знаходження коефіцієнтів інтерполяційного фільтра дозволяє отримати інтерпольоване зображення вищої якості.

Пропонований алгоритм змінює навчальний процес отримання коефіцієнтів фільтра. Існуючий метод отримання коефіцієнтів на основі оптимізації значень квадратичної похибки замінюється на оптимізацію значень індексу структурної подібності. При такому підході зростає кількість та складність обчислень, а також об'єм використаної пам'яті. Проте процес знаходження коефіцієнтів фільтра відбувається одноразово і не впливає на роботу алгоритму з реальними даними. Схема отримання результуючого зображення в режимі «реального часу» залишається незмінною.

Нехай F_{HD} – яскравість оригінальних пікселів формату високої роздільної здатності (англ. High-definition– HD), тоді F_{HI} – значення яскравості отримане в результаті інтерпольування. Припускаючи, що клас c в загальному містить N зразків під час навчального процесу,

значення індекса структурної подібності для всіх пікселів, що належать даному класу обчислюється згідно виразу:

$$SSIM = \frac{(2\mu_{F_{HI}}\mu_{F_{HD}} + C_1)(2\sigma_{F_{HI}F_{HD}} + C_2)}{(\mu_{F_{HI}}^2 + \mu_{F_{HD}}^2 + C_1)(\sigma_{F_{HI}}^2 + \sigma_{F_{HD}}^2 + C_2)} \quad (9)$$

де:

$$\mu_{F_{HI}} = \overline{F_{HI}} = \frac{1}{N} \sum_{i=1}^N F_{HIi}, \quad \mu_{F_{HD}} = \overline{F_{HD}} = \frac{1}{N} \sum_{i=1}^N F_{HDi} \quad (10)$$

$$\sigma_{F_{HI}}^2 = \frac{1}{N-1} \sum_{i=1}^N (F_{HIi} - \overline{F_{HI}})^2, \quad \sigma_{F_{HD}}^2 = \frac{1}{N-1} \sum_{i=1}^N (F_{HDi} - \overline{F_{HD}})^2 \quad (11)$$

$$\sigma_{F_{HI}F_{HD}} = \frac{1}{N-1} \sum_{i=1}^N (F_{HIi} - \overline{F_{HI}})(F_{HDi} - \overline{F_{HD}}) \quad (12)$$

Дані обчислення повторюються для всіх класів протягом процесу навчання. Для знаходження оптимальних коефіцієнтів фільтра перша похідна $\frac{\partial(SSIM)}{\partial w_c(k)}$ відносно кожного коефіцієнта $w_c(k)$ прирівнюється до нуля:

$$\begin{cases} \frac{\partial(SSIM)}{\partial w_c(1)} = 0; \\ \dots; \\ \frac{\partial(SSIM)}{\partial w_c(n)} = 0; \end{cases} \quad (13)$$

Ввівши змінні $A_1, A_2, A_3, A_4, B_{1,k}, B_{2,k}$ та $B_{3,k}$, які відповідно дорівнюють:

$$A_1 = \frac{1}{(\mu_{F_{HI}}^2 + \mu_{F_{HD}}^2 + C_1)}, \quad (14)$$

$$A_2 = \frac{1}{(\sigma_{F_{HI}}^2 + \sigma_{F_{HD}}^2 + C_2)}, \quad (15)$$

$$A_3 = (2\mu_{F_{HI}}\mu_{F_{HD}} + C_1), \quad (16)$$

$$A_4 = (2\sigma_{F_{HI}F_{HD}} + C_2), \quad (17)$$

$$B_{1,k} = \frac{\partial(\sigma_{F_{HI}F_{HD}})}{\partial w_c(n)}, \quad (18)$$

$$B_{2,k} = \frac{\partial(\sigma_{F_{HI}}^2)}{\partial w_c(k)}, \quad (19)$$

$$B_{3,k} = \frac{\partial(\mu_{F_{HI}})}{\partial w_c(n)}, \quad (20)$$

рівняння (13) представимо наступним чином:

$$\left\{ \begin{array}{l} \frac{\partial(SSIM)}{\partial w_c(1)} = B_{3,1} \cdot 2\mu_{F_{HD}} A_4 A_1 A_2 + 2B_{1,1} A_3 A_1 A_2 - \\ \quad - 2\mu_{F_{HI}} B_{3,1} A_3 A_4 A_1^2 A_2 - B_{2,1} A_3 A_4 A_1 A_2^2 \\ \dots \\ \frac{\partial(SSIM)}{\partial w_c(n)} = B_{3,n} \cdot 2\mu_{F_{HD}} A_4 A_1 A_2 + 2B_{1,n} A_3 A_1 A_2 - \\ \quad - 2\mu_{F_{HI}} B_{3,n} A_3 A_4 A_1^2 A_2 - B_{2,n} A_3 A_4 A_1 A_2^2 \end{array} \right. \quad (21)$$

Отримана система рівнянь є нелінійною. Для її розв'язання, тобто знаходження вагових коефіцієнтів інтерполяційного фільтра $w_c(k)$, використаний метод Ньютона, який забезпечує швидку збіжність і не вимагає значної кількості ітерацій для знаходження коренів рівняння.

5. РЕЗУЛЬТАТИ ПРОВЕДЕНИХ ДОСЛІДЖЕНЬ

Для об'єктивного оцінювання запропонованого алгоритму обчислювались значення MSE та SSIM. Кадри результуючих інтерпольованих зображень запропонованим алгоритмом наведено на Рис. 2 Як показують результати досліджень (Рис. 2.), при використанні запропонованого підходу значення метрики MSE погіршується, а значення метрики SSIM покращується. Це зумовлено врахуванням структурної інформації під час обчислення коефіцієнтів інтерполяційного фільтра. Така заміна дозволяє отримати інтерпольовані зображення з меншою кількістю блочних спотворень на границях зображень (Рис. 2.). Всі зображення для проведення процесу навчання взяті з бази даних Експертної групи з питань якості відеозображень (англ. Video Quality Experts Group – VQEG), яка є доступною за посиланням [4].



Рис.2. Кадри інтерпольованої послідовності *Mobile and Calendar* з використанням коефіцієнтів фільтра оптимізованих відносно а) метрики MSE ($MSE = 616.46$, $SSIM = 0.642$) б) метрики SSIM ($MSE = 635.31$, $SSIM = 0.661$).

6. ВИСНОВКИ

Наведений алгоритм інтерполяції зображень на основі методу адаптивного кодування вмістимого блоку зображення з динамічним діапазоном забезпечує достатню швидкість однак його якісні характеристики вимагають суттєвого покращення, скільки його коефіцієнти, які знаходяться в процесі навчання оптимізуються відносно метрики MSE. Даний підхід враховує лише абсолютні значення пікселів протягом навчання, ігноруючи структурною інформацією. Методи оцінки якості, що базуються на використанні структурної інформації, краще корелюють із суб'єктивними оцінюваннями якості. На основі цього в роботі запропонований новий алгоритм отримання коефіцієнтів інтерполяційного фільтра з використанням метрики структурної подібності зображення (SSIM). Такий алгоритм вимагає значно більшого числа обчислень в процесі навчання, а також збереження значень усіх пікселів навчальних відеозображень у оперативній пам'яті. Враховуючи, що процес навчання відбувається лише один раз, дані вимоги не змінюють характеристики процесу інтерполяції в режимі «реального часу». Отримані інтерпольовані зображення суттєво вищої якості у порівнянні з зображеннями отриманими при використанні коефіцієнтів фільтра, оптимізованих відносно значення метрики MSE.

1. Верес З.Є., Наконечний А.Й. Класифікація методів просторового розширення зображення // *Методи та прилади контролю якості*. – 2009. – №22. – С. 76-80. 2. Верес З.Є., Наконечний А.Й. Методи оцінки якості зображень та шляхи їх вдосконалення // *Комп'ютерні технології друкарства*. – 2008. – №20. –

C. 69-81. 3. Bovik A.C. *The Essential Guide to Image Processing* //Academic Press. – 2009. – P. 672. 4. *Final report from the video quality experts group on the validation of objective models of video quality assessment, phase II* //VQEG. – 2003. 5. Inglis A. F. and Luther A. C. *Video engineering, 2nd Edition* //McGraw-Hill. – 1996. – P. 350. 6. Jack K. *Video Demystified: a Handbook for the Digital Engineer, 5rd Edition* // Newnes. – 2007. – P. 944. 7. Kondo T., Fujiwara T., Okumura Y., and Node Y. *Picture conversion apparatus, picture conversion method, learning apparatus and learning method* //US-patent 6,323,905, 2001. 8. Poynton C. A., *A Technical Introduction to Digital Video* //John Wiley & Sons. – 1996. – P. 352. 9. Schultz R. R. and Stevenson R. L. *A bayesian approach to image expansion for improved definition* //IEEE Transactions on Image Processing. – 1994. – 3, № 3. – 3. 233-242. 10. Shao L. *An Overview and Performance Evaluation of Least Mean Square Trained Filters* // Technical Note PR-TN 2007/00230, Philips Research Europe. – 2007. – P. 35. 11. Wang Z. and Bovik A. C. *A universal image quality index* //IEEE Signal Processing. – 2002. – 9, №3. – P. 81–84. 12. Wang Z., Bovik A. C. and Lu L. *Why is image quality assessment so difficult?* //Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Proc. – 2002. – № 4. – P. 3313-3316 13. Wang Z., Bovik A. C., Sheikh H. R. and Simoncelli E.P. *Image Quality Assessment: From Error Visibility to Structural Similarity* //IEEE Transactions on Image Processing. – 2004. – 13, № 4. – P. 600–612. 14. Wang Z., Lu L. and Bovik A. C. *Video quality assessment based on structural distortion measurement* //Signal processing: Image communication. – 2004. – 19, № 2. P. 121-132. 15. Wang Z. *Rate Scalable Foveated Image and Video Communications* //PhD thesis, Dept. of ECE, The University of Texas at Austin. – 2001. 16. Zhao M., et al. *Content adaptive image de-blocking* //Proc. IEEE International Symposium on Consumer Electronics. – 2004.