

КРОСПЛАТФОРМЕННА ТЕХНОЛОГІЯ ЦИФРОВОЇ ОБРОБКИ ПОТОКОВОГО ВІДЕО НА ОСНОВІ ЛОКАЛЬНИХ ЛІНІЙНИХ ОПЕРАТОРІВ

Національний авіаційний університет

¹E-mail: chindakor@mail.ru

²E-mail: andrey.assaul@gmail.com

Запропоновано інформаційну технологію шейдерної обробки потокового відео в режимі реального часу на основі Adobe AIR та Pixel Bender. Проілюстровано роботу інформаційної технології на прикладі підвищення візуальної якості зображення тексту, що передається через веб-камеру.

Ключові слова: інформаційна технологія, кросплатформенні системи, обробка відео, потокове відео, шейдери.

Постановка проблеми

З розвитком телекомунікаційних технологій все більше набувають популярності відеоконференції.

За даними дослідницького центру Pew Research, на 2010 р. 23 % американських користувачів мережі Internet використовували відео для комунікації [1].

Менеджери компаній, що використовують відеоконференції в повсякденному житті, стверджують, що системи відеоконференцій різко скорочують часові та фінансові витрати на наради, семінари, відрядження та консультації.

Під час організації відеоконференції виникають два основних завдання.

Перше завдання полягає в тому, що канал зв'язку, по якому передається інформація (рис. 1, елемент 1), має володіти високою пропускну здатністю.

Зазвичай для проведення відеоконференцій використовують лінії зі смугою пропускання від 128 до 512 кбіт/с для ISDN відеоконференцій і від 1 до 1,5 Мбіт/с для IP-мереж [2].

Друге завдання – забезпечення швидкості обробки аудіо- і відеопотоку, тобто кодування переданих та декодування отримуваних даних. У відеоконференціях використовують спеціальні та досить ефективні алгоритми стиснення потоку в десятки (сотні) разів. Передаються не самі аудіо- та відеосигнали, а тільки їх найважливіші параметри, які дозволяють відновлювати сигнал на приймальному кінці з прийнятною якістю. Якщо сторона-приймач не встигає обробляти потік, то відбуваються втрати кадрів, збої в мовному каналі

тощо. Тому актуальним є поліпшення візуальної якості відтвореного відеосигналу в режимі реального часу на стороні клієнта (рис. 1, елемент 2).

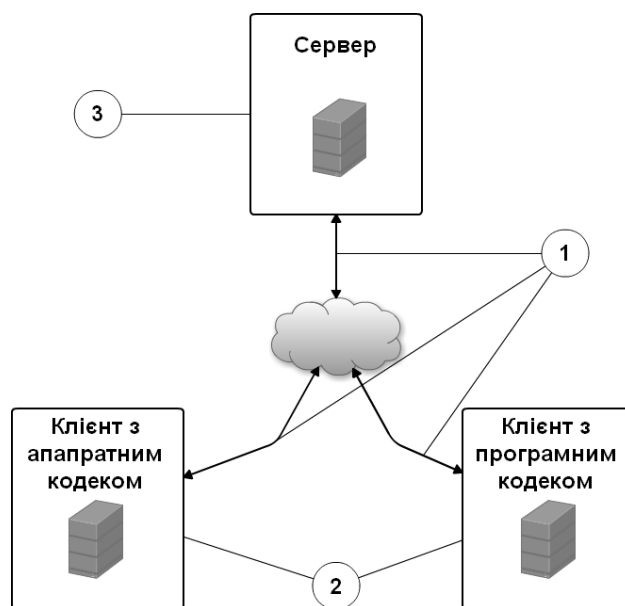


Рис. 1. Організація відеоконференції:

- 1 – канали зв'язку;
- 2 – клієнтські станції;
- 3 – серверна станція

Аналіз досліджень

На ринку систем відеоконференцій є велика кількість як суто програмних рішень, так і програмно-апаратних комплексів для реалізації терміналів відеоконференцій.

Під програмним рішенням розуміють реалізацію процесів відео- і аудіокодування та декодування за допомогою програмного забезпечення, що використовує лише центральний процесор комп'ютера.

Програмно-апаратні рішення ґрунтуються на використанні потужних спеціалізованих процесорів кодування/декодування.

Вартість програмних реалізацій менша, ніж апаратних, і вони можуть бути встановлені майже на будь-якому сучасному персональному комп'ютері, що має звукову плату і володіє можливістю відеозахоплення (підключена відеокамера та драйвер Video for Windows). Типовими представниками є Microsoft NetMeeting та Skype [3; 4].

Незважаючи на зовнішню привабливість, необхідної якості відео з високою деталізацією в програмних рішеннях досягти не вдається. Це зумовлено тим, що кодування відеопотоку накладає високі вимоги на обчислювальні ресурси системи.

Відповідно до прийнятого в рамках рекомендацій H.320 та H.323 стандартом кодування відео H.261 [5] необхідно забезпечити стиснення в реальному часі вихідного сигналу з коефіцієнтом від 100 до 1000 залежно від розміру та характеру кадру. Навіть збільшення потужностей процесорів не в змозі забезпечити якісне кодування і декодування сигналу відеоконференції на персональному комп'ютері. Для реалізації цих функцій у своїх програмних продуктах фірми-розробники змушені встановлювати певні обмеження для процесу кодування: використовувати низьку частоту кадрів, спрощені апроксимуючі алгоритми перетворення відео, що ведуть до зменшення розміру зображення, зниження чіткості і погіршення кольору, наприклад, Skype.

Поява подібних артефактів спричинена високим ступенем усереднення значень кольорових складових графічних даних, що лежить в основі таких алгоритмів. Крім цього більшість сучасних програмних систем відеокommунікацій є програмами з закритим вихідним кодом, що значно зменшує можливості адаптації процесу кодування та обробки відео до конкретних завдань.

Під час фіксації аналогового зображення (кадру відеоряду) відбувається така ситуація. Для розподілу інтенсивності $P(x)$ зображення об'єкта, що формується в площині x його реєстрації, використовують вираз [6]

$$P(x) = \int O(y)H(x-y)dy, \quad (1)$$

де $O(y)$ – розподіл інтенсивності відбиття від об'єкта світлового опроміювання в картинній площині y ;

$H(x-y)$ – розподіл інтенсивності в зображенні осьового точкового джерела (імпульсний

відклик або інакше – функція розсіювання точки системи).

Через технічні властивості систем реєстрації результатом згортки (1) у кожній точці растру кадру буде значення, усереднене в області дискретизації, зокрема:

$$p_{ij} = \tilde{p}_{ij} + \varepsilon_{ij}, \quad i, j \in \mathbf{Z}, \quad (2)$$

де p_{ij} – значення інтенсивності, що розташовано на перетині i -го стовпчика та j -го рядка растру (не зменшуючи загальності, по будь-якій з кольорових складових);

\tilde{p}_{ij} – середнє значення функції $P(x)$ в області, що визначається місцеположенням (i, j) -ї точки растру;

ε_{ij} – випадкова вада, стосовно якої можна припускати будь-який розподіл, проте за замовченням вважають, що відбувається гауссів розподіл.

Виходячи з природи таких даних доцільно проводити обробку декодованого відео математичними операторами, які є близькими до інтерполяційних у середньому, оскільки такі оператори характеризуються стійкістю оцінки наближення за виразом (2).

Крім того, слід ураховувати потребу швидкої обробки, тобто забезпечення мінімуму обчислювальних операцій. Таким вимогам відповідають локальні поліноміальні сплайни на основі B -сплайнів, близькі до інтерполяційних у середньому [7; 8].

Обчислювальний аспект їх застосування при обробці двовимірних послідовностей, у т.ч. і цифрових сигналів, досліджено в роботах [9–12].

Нехай маємо потокове відео, яке отримано на клієнтській системі без спеціальних апаратних пристроїв прискорення декодування з допомогою спрощеного апроксимуючого алгоритму.

Мета роботи – провести обробку відео для контрастування, низько- та високочастотної фільтрації, підвищення візуальної якості відеозображення в режимі реального часу або з мінімальним запізненням.

Цифрова обробка потокового відео

У рамках архітектури RIA (Rich Internet Application) [13] запропоновано інформаційну технологію обробки цифрового потокового відео реального часу:

- математичні методи обробки графічних даних локальними лінійними операторами [9–11];
- програмна реалізація у вигляді автоматизованої системи з відкритим кодом – Spline Video Player (SVP).

Середовищем виконання SVP було обрано Adobe AIR Runtime. Основними перевагами Adobe AIR є:

- кросплатформенність (AIR доступна для платформ: Windows XP, Vista, 7; Mac OS X; Android; QNX(BlackBerry Tablet OS); iOS [14]);
- широкі можливості в області обробки цифрової графіки на графічному пристрої [15].

Розроблена програма SVP дозволяє обробку потокового відео для поліпшення візуальної якості на основі лінійних операторів цифрової фільтрації. Під фільтром розуміємо лінійний оператор $A(p^{i,j})$ дискретної згортки, що призводить до зміни (фільтрації) форми та образу кольорових складових растру кожного окремого кадру:

$$A(p^{i,j}) = \sum_{ii=i-r_i}^{i+r_i} \sum_{jj=j-r_j}^{j+r_j} \gamma_{ii-ij-j} P_{ii,jj};$$

$$i \in [-0.5k_i; 0.5k_i];$$

$$j \in [-0.5k_j; 0.5k_j], \tag{3}$$

де $\gamma_{ii-ij-j}$ – елемент маски фільтра;

k_i, k_j – розміри кадру зображення.

Розмір маски фільтра: $(2r_i + 1) \times (2r_j + 1)$.

Для отримання значення обробленого пікселя локальні лінійні оператори типу (3) застосовуються до всіх кольорових складових пікселів растру. За рахунок лінійності операторів досягається швидкодія при обробці даних на процесорі відеоадаптера. Чим меншою є деталізація зображення, тим більшою має бути маска локального лінійного оператора для досягнення високого поліпшення візуальної якості зображення. Але в разі використання операторів з більшою маскою збільшується і кількість простіших математичних операцій, які необхідно виконати процесору для обробки відеокадру. Тому вибір оптимального оператора для обробки відео покладено на користувача системи.

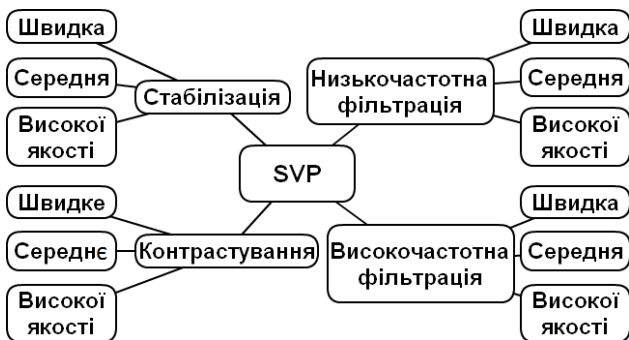


Рис. 2. Математичний функціонал SVP

Загальний перелік функціоналу SVP показано на схемі (рис. 2). Найпростіші маски фільтрів реалізовано в системі:

- низькочастотного

$$\gamma_L = \frac{1}{36} \begin{pmatrix} 1 & 4 & 1 \\ 4 & 16 & 4 \\ 1 & 4 & 1 \end{pmatrix};$$

- високочастотного

$$\gamma_H = \frac{1}{36} \begin{pmatrix} -1 & -4 & -1 \\ -4 & 20 & -4 \\ -1 & -4 & -1 \end{pmatrix};$$

- контрастного

$$\gamma_K = \frac{1}{196} \begin{pmatrix} 1 & -6 & 24 & -6 & 1 \\ -6 & 36 & -144 & 36 & -6 \\ 24 & -144 & 576 & -144 & 24 \\ -6 & 36 & -144 & 36 & -6 \\ 1 & -6 & 24 & -6 & 1 \end{pmatrix};$$

- цифрового стабілізатора

$$\gamma_C = \frac{1}{3136} \begin{pmatrix} 1 & 8 & -74 & 8 & 1 \\ 8 & 64 & -592 & 64 & 8 \\ -74 & -592 & 5476 & -592 & -74 \\ 8 & 64 & -592 & 64 & 8 \\ 1 & 8 & -74 & 8 & 1 \end{pmatrix}.$$

Маски інших фільтрів, що використовуються в системі, наведено в роботах [9–11].

Обробка відеоданих у системі SVP виконується програмними фільтрами (шейдерами) Pixel Bender [15] з використанням зазначеного математичного апарата. Етапи обробки потокового відео показано на рис. 3.

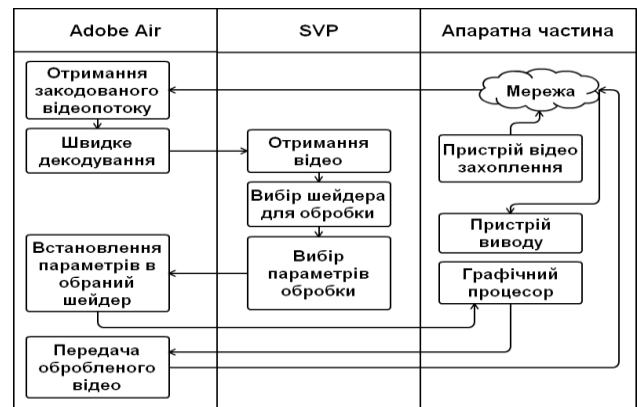


Рис. 3. Обробка відео системою SVP

На етапі «Вибір шейдера для обробки» SVP завантажує скомпільований шейдер Pixel Bender, який реалізує попиксельну обробку відеокadrів локальним лінійним оператором та накладає на об'єкт відео відповідну фільтрацію.

На відміну від аналогічної технології створення шейдерів HLSL [16], яка реалізована в DirectX починаючи з версії 9.0, Pixel Bender використовує API роботи з відеоадаптером Adobe Air, що не залежить від платформи, на якій встановлена SVP.

Фільтрація такими шейдерами виконується безпосередньо на процесорі відеоадаптера.

Шейдери Pixel Bender підтримують роботу не тільки з RGB потоком, а й з альфа-каналом, чим забезпечується можливість контролю «прозорості» зображення, що обробляється.

Незважаючи на пряму роботу з графічним адаптером, шейдери Pixel Bender повністю керовані Air-програмами, оскільки Air Runtime дозволяє додавати параметри виконання шейдера перед відправкою його на обробку відеокартою [14]. Це дозволяє додати у спроектовану систему апарат із контролю за областю обробки відео.

Користувач також має змогу в режимі реального часу задавати математичні фільтри (або їх послідовність), які дозволяють у реальному часі обробляти відеопотік, що надійшов до системи.

Для ілюстрації роботи SVP при обробці поточкового відео було проведено експеримент з передачі зображення (роздрукований текст). Пристроєм захоплення відео обрано веб-камеру A4Tech PK-710MJ [17]. Відстань до зображення надпису «VIDEO» – 13 см (рис. 4).

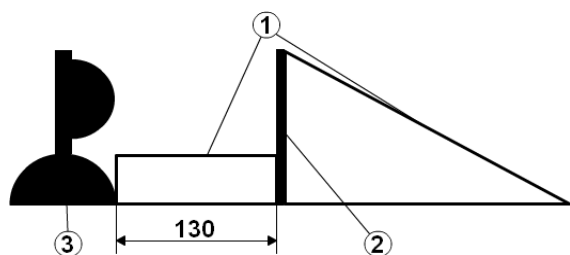


Рис. 4. Робота SVP:
1 – платформа-фіксатор;
2 – зображення;
3 – веб-камера

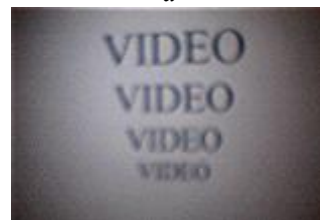
Для більш повної ілюстрації було обрано два варіанти надпису: звичайної щільності літер (рис. 5) та з підвищеною щільністю (рис. 6).

Експеримент проводився у два етапи.

На першому етапі на зображення накладався найсильніший стабілізатор (рис. 5, б; 6, б), на другому – до вже накладеного стабілізатора додавався більш слабкий (рис. 5, в; 6, в).



а



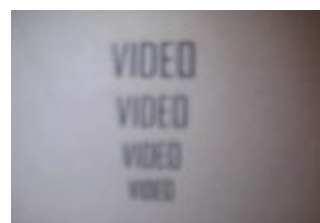
б



в

Рис. 5. Результати обробки зображення тексту Times New Roman:

а – початкове;
б – стабілізоване;
в – двічі стабілізоване



а



б



в

Рис. 6. Результати обробки зображення тексту Arial Narrow:

а – початкове;
б – стабілізоване;
в – двічі стабілізоване

При подвійному накладанні стабілізаторів підвищується чіткість, але починають з'являтися артефакти зображення, що зумовлено «запиранням» кольорових складових деяких пікселів до граничного значення.

Висновки

Запропоновано кросплатформенну інформаційну технологію обробки потокового відеоряду у складі шейдерних підпрограм, що здійснюють обробку зображення на основі алгоритмів з використанням локальних лінійних операторів та прикладного програмного продукту, що здійснює керування процесом обробки, яка дозволяє поліпшувати візуальну якість потокового відео в режимі реального часу, не змінюючи стандартні протоколи передачі відео по мережі. Роботу програмного продукту проілюстровано за допомогою експерименту з передачі зображення тексту.

Надалі планується проводити дослідження в напрямку зменшення Internet-трафіку шляхом попередньої обробки відеоряду низькочастотними операторами та аналіз використання цієї технології на стороні сервера для обробки відео тривимірними локальними лінійними операторами.

Література

1. *Lee Rainie*. Video calling and video chat: Research / Lee Rainie, Kathryn Zickuhr. – Washington, D.C.: Pew Research Center's Internet & American Life Project, 2010. – 12 p.
2. *Филимонов А.* Построение мультисервисных сетей Ethernet / Александр Филимонов. – СПб.: БХВ-Петербург, 2007. – 592 с.
3. *Интернет* ресурс: www.microsoft.com.
4. *Интернет* ресурс: www.skype.com.
5. *Интернет* ресурс: www.itu.int/rec/T-REC-H.261.
6. *Новейшие методы обработки изображений* / под ред. А.А. Потапова. – М.: ФИЗМАТЛИТ, 2008. – 496 с.
7. *Лигун А.А.* Асимптотические методы восстановления кривых / А.А. Лигун, А.А. Шумейко. – К.: ИМ НАН Украины, 1996. – 358 с.
8. *Приставка П.О.* Поліноміальні сплайни при обробці даних: моногр. / П.О. Приставка. – Д.: Вид-во Дніпропетр. ун-ту, 2004. – 236 с.
9. *Приставка П.О.* Обчислювальні аспекти застосування поліноміальних сплайнів при побудові фільтрів / П.О. Приставка // Актуальні проблеми автоматизації та інформаційних технологій: зб. наук. пр. – Д.: Вид-во Дніпропетр. ун-ту, 2006. – С. 3–14.
10. *Приставка П.О.* Побудова контрастних фільтрів за використанням поліноміальних сплайнів / П.О. Приставка // Актуальні проблеми автоматизації та інформаційних технологій: зб. наук. пр. – Д.: Вид-во Дніпропетр. ун-ту, 2007. – С. 15–22.
11. *Приставка П.О.* Застосування комбінованих фільтрів на основі поліноміальних сплайнів при обробці растрових зображень / П.О. Приставка. – Вісник НАУ. – 2008. – № 4. – С. 104–107.
12. *Приставка П.О.* Лінійні комбінації B-сплайнів, близькі до інтерполяційних у середньому, в задачі моделювання аналогових сигналів / П.О. Приставка // Актуальні проблеми автоматизації та інформаційних технологій : зб. наук. пр. – Д.: Вид-во Дніпропетр. ун-ту. – 2011. – Т. 15. – С. 4–17.
13. *Craig James*. Accessible Rich Internet Applications (WAI-ARIA) 1.0: W3C Candidate Recommendation / James Craig, Michael Cooper. – Available from Internet: <<http://www.w3.org>, 18 January 2011>.
14. *Интернет* ресурс: www.adobe.com/ru/products/air/tech-specs.html.
15. *Електронна специфікація технології: Pixel Bender Reference*. Adobe Systems Incorporated – 2009.
16. *Sebastien St-Laurent*. The COMPLETE Effect and HLSL: Guide. – Redmond: Paradoxal Press, 2005. – 39 p.
17. *Интернет* ресурс: www.a4tech.ua.

Стаття надійшла до редакції 31.01.2012.