

# Хвилькове та Карунена-Лоева перетворення зображень у реальному часі

Мусатенко Ю. С., Курашов В. Н.

Лабораторія оптичної та мікрохвильової обробки інформації та теорії середовищ,  
радіофізичний факультет, Київський національний університет імені Тараса Шевченка  
Тел. (044) 266-0483; E-mail: ysm@mail.univ.kiev.ua

В роботі представлений розроблений авторами програмний пакет для здійснення хвилькового (wavelet)<sup>1</sup> та Карунена-Лоева(КЛ) перетворень зображень в реальному часі на персональному комп'ютері. При розробці пакету максимальна увага приділялась отриманню високої швидкодії. Для забезпечення високих швидкісних показників реалізацію перетворень було здійснено в цілих числах з використанням технології MMX. Розроблений пакет призначений для забезпечення роботи в реальному часі системи стиснення наборів зображень, розробленої авторами раніше.

## ВСТУП

Ортогональні перетворення широко використовуються в системах стиснення зображень. Найбільш широко використовуваним, мабуть, є блочне косинусне перетворення, на основі якого побудований популярні стандарти стиснення JPEG та MPEG-2. Проте, на сьогоднішній день, було запропоновано більш ефективні системи стиснення на основі хвилькового перетворення [1-2] та перетворення КЛ [3] (в тому числі авторами даної роботи [4]). Як правило, базовим алгоритмом для даних систем стиснення є алгоритм EZW[4] і хвилькове перетворення. Багато інших систем стиснення використовують модифікації даного алгоритму з деякими покращеннями. Одним з основних недоліків таких систем є їх вища складність реалізації та низька швидкодія, що не дозволяє їм конкурувати з відомими стандартами в задачах реального часу таких як стиснення та декомпресія відео. Основою нашої існуючої системи стиснення є схема перетворення на основі перетворення КЛ вздовж осі часу та двовимірного хвилькового перетворення в площині зображень з подальшим застосуванням оригінальної схеми квантування. Для ефективної роботи даного методу необхідно є ефективна реалізація кожної з

складових. В даній роботі представлений пакет, який з високою швидкістю реалізує частину системи стиснення, яка відповідає перетворенню КЛ та хвильковому перетворенню.

Для хвилькового перетворення використовується відомий фільтр 9/7, який є одним з найкращих з точки зору використання в задачах стиснення з втратами.

Робота має таку структуру. В першому розділі коротко розглянуто перетворення КЛ, хвилькове перетворення та використовувані алгоритми їх здійснення. В ньому ж представлено схему розвинення для якої необхідно забезпечити високу швидкодію. В розділі 2 розглянуто реалізацію згаданих перетворень в цілих числах з використанням технології MMX. В третьому розділі представлено результати тестування швидкодії та точності відновлення. Наприкінці приведені висновки з роботи.

## 1. ПЕРЕТВОРЕННЯ СИГНАЛІВ ТА ЗОБРАЖЕНЬ

### 1.1 Перетворення Карунена-Лоева

Зміст перетворення КЛ полягає в мінімізації середньоквадратичної похибки  $\varepsilon^2$  для  $K$ -вимірної проекції сигналу  $\{U \in R^d\}$  для всіх  $K$ . Іншими словами, ми маємо знайти ортогональний базис  $\{\phi_k^0\}_0^d$  такий, що задовольняє умові:

$$\varepsilon^2 = \left\langle \left\| U_n - \sum_{k=1}^K a_n^k \phi_k^0 \right\|_n^2 \right\rangle = \min. \quad (1)$$

Тут  $\|\cdot\|$  позначає  $L^2$  норму, а  $\langle * \rangle$  - усереднення по генеральній сукупності. Функції  $\phi_n^0$ , що задовольняють умові (1.1), утворюють базис КЛ для заданого ансамблю  $\{U\}$ ; а набір коефіцієнтів  $\{a_n^k\}$ ,  $k=1,2,\dots,d$ , є коефіцієнтами розкладу сигналів ансамблю в базис. Варіаційна задача (1) дає рівняння на власні значення, котре відоме, як теорема КЛ [3]:

<sup>1</sup> Термін "хвилькове перетворення" є перекладом з англійської виразу "wavelet transform", даний переклад автори вважають найбільш вдалим.

$$G\phi_k^0 = \lambda_k \phi_k^0, \quad (2)$$

де  $G$  – кореляційна матриця вибірки  $\{U_n\}$ , а  $\lambda_k$  – власні числа матриці  $G$ . Оскільки кореляційна матриця після перетворення КЛ є діагональною, то координати  $\{\mu_n^k\}$  векторів  $U_n$  в просторі, породженому базисними векторами  $\{\phi_k^0\}_0^K$ , є некорельованими випадковими величинами.

Основною проблемою, пов'язаною з перетворенням КЛ, є громіздкість чисельного розв'язку рівняння (2), яке має складність  $O(Nd^2 + d^3)$ ;  $N$  – обсяг навчаючої вибірки.

## 1.1 Хвилькове перетворення

Останнім часом в системах обробки інформації широко використовується хвилькове перетворення. Воно забезпечує т.зв. мультирозрізнявальну декомпозицію, подає ансамбль через набір промасштабованих "копій" і є потужним засобом для аналізу та стиснення зображень, виділення ознак та розпізнавання образів. Однією з важливих переваг цього перетворення є низька складність, вона лінійно зростає з кількістю відліків. Нижче ми розглянемо основні принципи хвилькового перетворення [1-2].

Поняття мультирозрізнявального аналізу дозволило відшукати метод побудови хвилькового ортонормованого базису для простору сигналів скінченної енергії  $L_2(R)$ . У 1989 році Маллат [2] представив цей математичний метод, основна ідея якого полягає в такому. Ми маємо послідовність підпросторів  $V_j \subset L_2(R)$ ,  $V_j \subset V_{j-1}$ ,  $j \in Z$  і функцію  $\phi(x)$ , таку, що послідовність функцій  $\{2^{-j/2} \phi(2^{-j}x - n)\}$ ,  $n \in Z$ , утворює ортонормований базис у  $V_j$ . Традиційно  $W_j$  позначає ортогональне доповнення  $V_j$  в  $V_{j-1}$ , так що  $V_{j-1} = V_j \oplus W_j$ , де символ  $\oplus$  позначає суму підпросторів. Підпростори  $V_j$  обрано таким чином, що підпростори  $W_i$  є ортогональні між собою.  $W_j$  містить деталі  $j$ -го рівня аналізованого сигналу, тому, що він отримується як різниця підпросторів  $V_{j-1}$  і  $V_j$ , які містять апроксимації сигналу  $j-1$ -го і  $j$ -го рівня відповідно. Можна показати, що можливо створити материнську функцію  $\psi(x)$ , таку, що послідовність функцій  $\{2^{-j/2} \psi(2^{-j}x - n)\}$ ,  $n \in Z$ , утворює ортонормований базис у  $W_j$ . Звідси очевидно, що набір функцій  $\{2^{-j/2} \psi(2^{-j}x - n)\}$ ,  $n \in Z$ , утворює ортонормований базис у  $L_2(R)$ . Апроксимація і деталі сигналу отримуються його ортогональним проектуванням у апроксимуючий і різницевий підпростори. Інформація, котра міститься у підпросторах  $V_j$  та  $W_j$

може розглядатись як апроксимація та деталі  $j$ -го рівня сигналу  $f(x)$  відповідно. Проекції сигналу в  $j$ -тий апроксимуючий та різницевий підпростори є:

$$c_n^j = \int_{-\infty}^{+\infty} f(x) \cdot 2^{-j/2} \phi(2^{-j}x - n) \cdot dx; \quad (3)$$

$$d_n^j = \int_{-\infty}^{+\infty} f(x) \cdot 2^{-j/2} \psi(2^{-j}x - n) \cdot dx.$$

В дійсності коефіцієнти подання сигналу, що відповідають підпросторам  $V_j$  та  $W_j$  обчислюються з попереднього рівня апроксимації  $V_{j-1}$  за допомогою пари фільтрів  $h$  і  $g$  за алгоритмом Маллата:

$$c_n^j = \sum_k \tilde{h}(2n - k) \cdot c_k^{j-1},$$

$$d_n^j = \sum_k \tilde{g}(2n - k) \cdot c_k^{j-1}. \quad (4)$$

Відновлення сигналу відбувається з використанням майже тих самих фільтрів і зворотнього ходу алгоритму:

$$c_n^{j-1} = \sum_k h(n - 2k) \cdot c_k^j + \sum_k g(n - 2k) \cdot d_k^j. \quad (5)$$

Тут відгуки пари фільтрів традиційно позначаються символами  $h$  та  $g$  для відновлюючої частини і  $\tilde{h}(n)$   $\tilde{g}(n)$  для аналізуючої частини. Вони між собою пов'язані простими рівностями.

$$\tilde{h}(n) = h(-n),$$

$$\tilde{g}(n) = g(-n). \quad (6)$$

Якщо оператори  $\hat{H}$  і  $\hat{G}$  позначають згортку сигналу з низькочастотним та високочастотним відгуками  $h$  та  $g$  відповідно, то декомпозиція сигналу з  $(j-1)$ -го рівня в  $j$ -й рівень здійснюється застосуванням операторів  $\hat{H}$  і  $\hat{G}$  до коефіцієнтів  $(j-1)$ -го рівня. В результаті цього процесу ми отримуємо бінарне дерево для обчислення коефіцієнтів розвинення в набір підпросторів, котрий визначається хвильковим мультирозрізнявальним аналізом. Це дерево зображено на Рис.1. На кожному кроці ми здійснюємо розкладання коефіцієнтів розвинення сигналу в дві частотні смуги за допомогою фільтрів  $\hat{H}$  і  $\hat{G}$  та рекурсивно продовжуємо цю фільтрацію для низькочастотної частини розвинення сигналу. Коефіцієнти розкладання по хвильковому базису лишатимуться у блоках, котрі виділені на рисунку. Відновлення здійснюється за подібною схемою, тільки в зворотній бік.

Крім ортогональних хвилькових фільтрів часто користуються неортогональними, серед яких найпоширенішими є т.зв. біортогональні фільтри. До таких фільтрів відноситься і використовуваний нами відомий фільтр 9/7. Особливістю цих фільтрів є те, що не зважаючи на відсутність ортогональності вони дозволяють повне відновлення, крім того на відміну від ортогональних фільтрів вони можуть бути симетричними для довжин фільтра  $> 2$ . Повне

відновлення можливе за рахунок того, що фільтрація вперед є проектуванням в деякий афінний базис, а зворотня фільтрація являє собою відновлення за допомогою взаємного базису. В зв'язку з цим фільтри  $h$  та  $g$  пов'язані з  $\tilde{h}(n)$  та  $\tilde{g}(n)$  зовсім інакшим чином, але фільтрація проводиться так само. Свою назву (біортогональні) ці фільтри отримали через те, що попарно ортогональними є фільтри  $h$  та  $\tilde{g}(n)$  і  $g$  та  $\tilde{h}(n)$ . В багатьох роботах відмічено вищі показники біортогональних фільтрів поряд з ортогональними в задачах стиснення зображень.

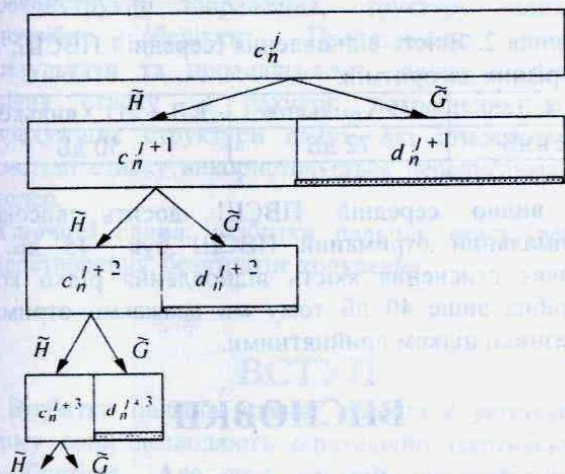


Рис. 1 Дерево хвилькових розвинень.

$A_{2^{-3}}^d f$	$D_{2^{-3}}^1 f$	$D_{2^{-2}}^1 f$	$D_{2^{-1}}^1 f$
$D_{2^{-3}}^2 f$	$D_{2^{-3}}^3 f$		
$D_{2^{-2}}^2 f$	$D_{2^{-2}}^3 f$		
$D_{2^{-1}}^2 f$	$D_{2^{-1}}^3 f$		

Рис. 2 Схема двовимірного хвилькового перетворення.

При застосуванні до зображень часто користуються простим сепарабельним узагальненням одновимірного хвилькового перетворення на двовимірний випадок. Як можна показати, схема перетворення, тоді має вигляд представлений на Рис. 2. При цьому оператор  $D_{2^{-j}}^k f, k=1,2,3$  відповідає відшукуванню деталей  $j$ -го рівня з  $k$ -ю

орієнтацією, а оператор  $A_{2^{-j}}^d f$  є оператором побудови апроксимації сигналу. В одновимірному випадку є два оператори, а в двовимірному – чотири, які можна подати як добутки одновимірних.

Для задач стиснення наборів зображень нами проводиться перетворення фактично тривимірного набору даних. При цьому здійснюється перетворення КЛ по вісі часу, і двовимірне хвилькове перетворення в площині зображень. При цьому схема перетворення виглядає так як на Рис. 3.

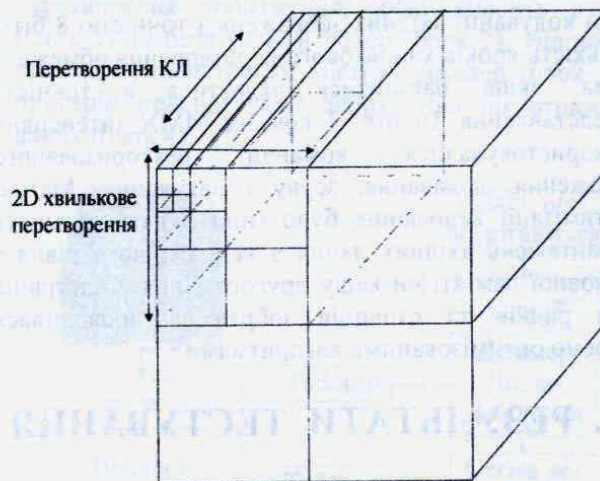


Рис. 3. Схема розрізнення при застосуванні перетворення КЛ по часовій осі і 2D хвилькового перетворення в площині зображень.

Далі нами розглядатиметься яким чином забезпечити здійснення представленого тривимірного перетворення в реальному часі.

## 2. РЕАЛІЗАЦІЯ ПЕРЕТВОРЕНЬ

Побудована ефективна реалізація описаного тривимірного перетворення містила такі складові.

1. У всіх найважчих ділянках використовувалися цілі числа з точністю 16 біт.
2. Найважчі обчислення здійснювалися з використанням команд MMX.
3. Для перетворення КЛ і хвилькового перетворення проводилася ґрунтовна оптимізація стратегії кешування.

Типовим ансамблем, для якого необхідно здійснювати перетворення є 16 зображень 768x768 з глибиною сірого 8 бітів. Зрозуміло, що оптимізація проводилася саме для подібних конфігурацій. Найважчими ділянками, які вимагали суттєвих оптимізацій були такі:

- Обчислення кореляційної матриці часових послідовностей пікселів. (Векторів часових послідовностей стільки, скільки пікселів у зображенні, тобто дуже багато).
- Отримання коефіцієнтів перетворення КЛ для кожного вектора часової послідовності пікселів.

Таблиця 1. Порівняльна таблиця швидкодії алгоритмів перетворень реалізованих з командами MMX та з плаваючою комою для ансамблю 16 зображень 768x768 на ПК Pentium III 1GHz.

Перетворення	Для алгоритму з командами MMX			Для алгоритму з плаваюч. комою			Пришвидшення, разів
	Час всього, мс	Час на одне зображ., мс	Тактів на піксель	Час всього, мс	Час на одне зображ., мс	Тактів на піксель	
КЛ	336	21	35	20480	1280	2170	62
2D Хвилькове	672	42	70	27200	1700	2940	42
КЛ+Хвильк.	1008	63	105	47680	2980	5110	48

- Фільтрація за виразами (4)-(5) окремо для рядків та стовпців зображень.

При кодуванні вхідних зображень з точністю 8 бітів кількість кроків хвилькового перетворення обмежена 4-ма, якщо намагатися вкласти в точність представлення 16 біт. З команд MMX інтенсивно використовувалися команди векторизованого множення, додавання, зсуву і пакування. Метою оптимізації кешування було мінімізувати кількість завантажень вхідних даних в кеш першого рівня з основної пам'яті чи кешу другого рівня. Фільтрація для рядків та стовпців зображень проводилася окремо оптимізованими алгоритмами.

### 3. РЕЗУЛЬТАТИ ТЕСТУВАННЯ

Тестування проводилося з метою порівняння швидкодії з існуючими реалізаціями в числах з плаваючою комою та визначення точності відновлення. Для оцінки швидкодії використовувався ансамбль 16 зображень 768x768. Проводилося порівняння розробленого алгоритму реального часу з таким самим алгоритмом реалізованим в числах з плаваючою комою на ПК Pentium III 1GHz. Крім цього проводилося порівняння для складових – окремо для перетворення КЛ та для 2D хвилькового перетворення. Результати порівняння представлені в Таблиці 1. Видно, що проведені оптимізації дозволили досягнути великого виграшу в швидкості, а саме, в 40-60 разів. Як видно час обробки в перерахунку на одне зображення є 63мс, що дозволяє обробляти 15 зображень за секунду, а для зображень 512x512 вдасться обробляти 30 зображень на секунду. Подібні показники дозволяють використовувати даний алгоритм в системах реального часу. Зрозуміло, що підвищення частоти процесора, та перехід від команд MMX до SSE2 дасть ще додаткове підвищення швидкодії.

Оскільки перетворення здійснювалися в цілих числах з обмеженою точністю, то проводилося тестування якості відновлення. Якість оцінювалася на основі Пікового Відношення Сигнал Шум(ПВСШ) в децибелах, яке визначається так:

$$\text{ПВСШ} = 20 \log(255/\text{MSE}) \quad (7)$$

де 255 - максимальна амплітуда значень пікселів в зображенні, а MSE - величина середньоквадратичного відхилення пікселів відновленого зображення від

оригіналу. Як тестовий використовувався набір зображень Mobile.

Таблиця 2. Якість відновлення (середній ПВСШ, дБ) для різних алгоритмів.

КЛ	2D Хвилькове	КЛ + 2D Хвилькове
61 дБ	72 дБ	50 дБ

Як видно середній ПВСШ досить високий. Мінімальний отриманий ПВСШ був ~45 дБ. В задачах стиснення якість відновлення рідко коли потрібна вище 40 дБ тому ми вважаємо отримані показники цілком прийнятними.

### ВИСНОВКИ

В роботі представлений розроблений авторами програмний пакет для здійснення хвилькового та КЛ перетворень зображень в реальному часі на персональному комп'ютері з використанням команд MMX. Досягнутий виграш порівняно з традиційною реалізацією є 48 разів. Отримана швидкодія тривимірного (КЛ x 2D Хвилькове) перетворення є 15 кадрів на секунду для ансамблю 16 зображень 768x768 на ПК Pentium III 1GHz. Крім використання на ПК розроблений алгоритм чи його складові можуть бути легко перенесені для використання на сигнальному процесорі типу TMS320C64x, що дозволить його використання в окремих пристроях без ПК.

### ЛІТЕРАТУРА

1. Daubechies I. Ten lectures on wavelets. — Philadelphia, Society for Industrial and Applied Mathematics, 1992. — 358с.
2. Mallat S.G. A Theory for multiresolution signal decomposition: the wavelet representation // IEEE Trans. Pattern Anal. Machine Intell. — 1989. — v.31. — P. 674-693.
3. Фрэнкс Л. Теория сигналов. — М.: Сов. радио, 1974. — 344с.
4. О.М. Soloveyko, Musatenko Yu.S., V.N. Kurashov, V. A. Dubikovskiy Video compression based on enhanced EZW scheme and Karhunen-Loeve transform. — SPIE Proceedings. — v.4041. 2000. — pp.180-191.