НАВЧАЛЬНО-МЕТОДИЧНА ЛІТЕРАТУРА

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ

Кафедра: "Комп'ютерні науки"

К О Н С П Е К Т Л Е К Ц І Й

з дисципліни

Обробка сигналів та зображень

для студентів денної форми навчання напряму підготовки 6.170101 – Безпека інформаційних і комунікаційних систем

ТЕРНОПІЛЬ 2015

УДК 519.7

Конспект лекцій з дисципліни “Обробка сигналів та зображень” (для студентів денної форми навчання напряму 6.170101 «Безпека інформаційних і комунікаційних систем») / Укладачі: к.т.н., доцент Фриз М.Є., Стадник М. А. – Тернопіль: ТНТУ, 2015 – 97 с.

Конспект лекцій призначений для виконання освоєння теоретичного матеріалу студентами із дисципліни “ Обробка сигналів та зображень”. Складається з урахуванням модульної системи навчання, рекомендацій до самостійної роботи і індивідуальних завдань, тем лабораторних занять, екзаменаційних питань, типової форми та вимог для комплексної перевірки знань з дисципліни.

Укладачі: М. Є. Фриз, к.т.н., доцент
асистент М. А. Стадник

Відпов. за випуск М. В. Приймак, д.т.н., професор

Рецензент Р. О. Козак, доцент

Затверджено
на засіданні кафедри комп’ютерних наук
Протокол № 2 від “ 09 ” 09 2015 р.

Схвалено та рекомендовано до друку методичною комісією факультету комп’ютерно-інформаційних систем і програмної інженерії Тернопільського національного технічного університету імені Івана Пулюя.
Протокол № 2 від “ 25 ” 09 2015 р.

Вказівки складені з урахуванням матеріалів літературних джерел, названих у списку.

ЗМІСТ

ВСТУП	4
Тема 1. Теоретичні основи цифрової обробки сигналів.	5
Тема 1. Математичні моделі сигналів.	15
Тема 3. Z-перетворення та його властивості.	33
Тема 4. Фур'є аналіз сигналів.	40
Тема 5. Згортка та кореляція, їх обчислення та застосування.	44
Тема 6. Цифрова фільтрація сигналів**.**	51
Тема 7. Основи цифрового представлення зображень.	58
Тема 8. Просторові методи покращення зображень.	68
Тема 9. Частотні методи покращення зображень.	77
Тема 10. Цифрова обробка кольорових зображень.	79
Тема 11. Стиснення зображення.	82
Тема 12. Основи сегментації зображення.	89
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	95

ВСТУП

Мета навчальної дисципліни «Обробка сигналів та зображень» полягає у вивченні основних методів, алгоритмів та засобів цифрової обробки сигналів та зображень в різноманітних системах.

Завдання дисципліни – надати студентам знання та основні поняття з основ теорії цифрової обробки сигналів та зображень, що охоплює відомості про математичні моделі та методи цифрової обробки інформації; ефективні алгоритми перетворення та аналізу сигналів і зображень в лінійних/нелінійних стаціонарних/нестаціонарних системах.

Після вивчення курсу студент повинен знати:

- поняття дискретизації та квантування сигналу;
- технічні особливості та характеристики каналу зв'язку та його складових для організації передачі інформації;
- сучасні алгоритми кодування сигналу;
- основні види цифрових фільтрів, методи їх аналізу і синтезу;
- основні методи статистичної обробки даних;
- спектральний аналіз сигналів;
- принципи формування цифрових зображень;
- просторові та частотні методи покращення зображень.

Після вивчення курсу студент повинен вміти:

- обчислювати параметри статистичних розподілів, будувати гістограми;
- володіти прикладними програмними пакетами математичного аналізу;
- виконувати пряме і зворотне перетворення Фур'є;
- обчислювати і будувати амплітудно-частотні характеристики;
- застосовувати просторові та частотні методи для покращення зображення; розробляти алгоритми та створювати на їх основі програмні модулі.

ЛЕКЦІЯ 1

Тема: Теоретичні основи цифрової обробки сигналів.

План

1. Поняття сигнал у обробці сигналів. Розмірність сигналу.
2. Основні завдання цифрової обробки сигналів та зображень.
3. Класифікація шумів та завад.

Сигнал - це інформаційна функція, що несе повідомлення про фізичні властивості, стан або поводження якої-небудь фізичної системи, об'єкта або середовища, а метою обробки сигналів можна вважати отримання певних інформаційних відомостей, що відображені в цих сигналах (коротко - корисна або цільова інформація) й перетворення цих відомостей у форму, зручну для сприйняття й подальшого використання**.**

Фізичний процес с невід’ємною властивістю об’єкта, пов'язаний з ним вектором змінних стану певної розмірності і є їхнім набором, отриманим для різних значень часу. Результатом спостереження є вимірювання у дискретні проміжки часу деяких фізичних величин, які є відображенням змінних стану процесу.

У природі всі вимірювані фізичні величини представлені аналоговими сигналами. Назву “аналоговий” дали неперервним сигналам, тому що вони “аналогічні фізичним процесам, які відбуваються в дійсності.

Деякі сигнали - це реакції на інші сигнали. Наприклад, відбитий сигнал радара або ультразвукової системи відображення - це модель випромінюваного сигналу.

Промислові системи вимірювання, збирання даних, управління процесом для вироблення відповідних сигналів зворотного зв’язку та безпосереднього керування процесом використовують інформацію, отриману від давачів. Давачі і пов’язані з ними кола застосовують для вимірювання різних фізичних власти-

востей, наприклад, температури, сили, тиску, напруженості, прискорення, інтенсивності тощо.

Давач (sensor) - це пристрій, який перетворює вхідний сигнал або збудження (що відображає фізичну властивість) на електричний вихідний сигнал. Часто вживається також термін “перетворювач”, хоча, строго говорячи, *перетворювач (transducer)* - це конвертор одного типу енергії в іншу.

Давачі, як правило, є елементами великої системи, яка складається із формувачів сигналів та різних аналогових та цифрових кіл їхньої обробки. Щодо формування сигналу розрізняють активні та пасивні сенсори. *Активний* сенсор вимагає наявності зовнішнього джерела збудження. Це, наприклад, термістори, резистивні термометри (RTD - Resistance Temperature Detectors) і тензометри. Через них протікає струм. Щоб знайти значення опору, вимірюється відповідне

значення напруги. Іншим способом с улаштування пристрою у мостову схему. Тут також необхідний зовнішній струм або зовнішня напруга по колу.

Пасивні сенсори генерують свій електричний вихідний сигнал без використання зовнішнього струму або зовнішньої напруги по колу. Прикладом пасивних давачів є термоелементи та фотодіоди, які, незалежно від зовнішніх кіл, створюють термоелектричні напруги або фотодіодний струм, відповідно.

Границю шкали виходів більшості давачів становлять малі зміни напруги, струму або опору, і тому, перед виконанням аналогової або цифрової обробки сигналу, виходи повинні бути відповідно погоджені. Для цього існує цілий ряд кіл, які належать до *кіл формування сигналу.* Основними функціями формування сигналу є підсилення, зсув рівня, гальванічна ізоляція, трансформація опору, лінеаризація, фільтрація.

Більшість виходів сенсорів є нелінійними, тому для одержання коректних вимірювань їх виходи повинні бути лінеризовані. Для виконання цієї функції може бути використана аналогова техніка, хоча поява останнім часом високопродуктивних АЦП дає змогу виконувати ефективнішу і точнішу лінеаризацію - програмну.

Цифрова техніка стає усе популярнішою при обробці виходів з давачів у системах збирання даних, управління процесом та вимірювання. Восьми розрядні мікроконтролери (наприклад, 8051-сумісні) володіють достатньою швидкістю обробки даних. Забезпечивши аналого-цифрове перетворення та можливості програмування мікроконтролера всередині давача, можна реалізувати “інтелектуальний давач” (“smart sensor”), який мас функції самокалібрування, лінеаризації тощо.

Серія виробів MicroConverter от Analog Devices містить вбудовані високопродуктивні мультиплексори, АЦП, ЦАП, з’єднані з flash-пам’яттю ядром мікроконтролера індустріального стандарту 8052, і підтримує декілька стандартів конфігурації послідовного порту. Ці вироби є першими інтегрованими схемами, які можна назвати “інтелектуальними давачами” для систем збирання даних на одному кристалі.

Першочергова причина обробки фізичних сигналів полягає в необхідності отримання інформації, яка міститься в них. Ця інформація наявна в амплітуді сигналу, в частоті чи в спектральному складі, у фазі чи у відносних часових залежностях декількох сигналів.

У деяких випадках бажано переформатувати інформацію, яка міститься в сигналі. Наприклад, зміна формату відбувається при передаванні звукового сигналу в телефонній системі з багатоканальним доступом і частотним розділенням (FDMA). У цьому випадку аналогові методи використовують для створення “стеку” голосових каналів у частотному спектрі для передавання через радіорелейну станцію мікрохвильового діапазону, коаксіальний чи оптоволоконний кабель. У випадку цифрового зв’язку аналогова звукова інформація спочатку перетворюється на цифрову за допомогою аналого-цифрового перетворювача (АЦП). Цифрова інформація, яка представляє індивідуальні звукові канали, мультиплексується в часі (багатоканальний доступ з часовим розділенням, (TDMA) і передається по послідовній цифровій лінії зв’язку.

Ще одна причина обробки сигналів полягає у стисненні смуги частот сигналу (без істотної втрати інформації) з подальшим форматуванням та пере-

даванням інформації на понижених швидкостях, що дає змогу звузити смугу пропускання каналу. Високошвидкісні модеми і системи адаптивної імпульсно-кодової модуляції (ADPCM) широко використовують алгоритми усунення надлишковості даних (стиснення), так само, як і цифрові системи мобільного зв'язку, MPEG рекордери, програвачі та телебачення високої чіткості (HDTV).

У деяких випадках у сигналі, який містить інформацію, є шум, і основною метою є відновлення сигналу. Такі методи, як фільтрація, автокореляція, згортка тощо, часто застосовують для виконання цього завдання і в аналоговій, і в цифровій областях.

Нерідко фізичні аналогові сигнали можуть бути безпосередньо згенеровані (штучно сформовані). Наприклад, музика і мова, які штучно синтезовані. При генерації фізичних аналогових сигналів з використанням тільки цифрових методів, вони основуються на інформації, попередньо отриманій із джерел аналогічних фізичних аналогових сигналів. Розробляючи звукові системи, попередньо задаються певними статистичними властивостями звуків, які генеруються, широко застосовуючи для цього методи цифрової обробки сигналів (ЦОС).

Сигнали можуть бути оброблені з використанням аналогових методів (аналогова обробка сигналів, або ASP), цифрових методів (цифрова обробка сигналів, або DSP) або комбінації аналогових і цифрових методів (комбінована обробка сигналів, або MSP). Основна відмінність DSP від традиційного комп'ютерного аналізу даних полягає у високій швидкості та ефективності виконання складних функцій цифрової обробки, таких як фільтрація, аналіз із використанням швидкого перетворення Фур'є, стиснення даних у реальному масштабі часу.

У випадку комбінованої обробки сигналів система виконує як аналогову, так і цифрову обробку. Така система може бути реалізована у вигляді друкованої плати, гібридної інтегральної схеми або окремого кристала з інтегрованими елементами. АЦП і ЦАП є пристроями комбінованої обробки сигналів, оскільки в кожному з них реалізуються як аналогові, так і цифрові функції.

Неможливо обробити фізичні аналогові сигнали, використовуючи тільки цифрові методи, оскільки всі давачі (мікрофони, термопари, тензорезистори, п’єзоелектричні кристали, головки накопичувана на магнітних дисках тощо) є аналоговими пристроями. Тому деякі види сигналів вимагають наявності схем нормалізації для подальшої обробки сигналів аналоговими або цифровими методами. Схеми нормалізації сигналу на практиці реалізуються за допомогою аналогових процесорів, які виконують такі функції, як підсилення, накопичування, виявлення сигналу на фоні шуму, динамічне стиснення діапазону та фільтрація.

Більша частина обробки аналогового сигналу здійснюється в АЦП, оскільки він розташований найближче до давача. Багатофункційність АЦП може виявлятися у збільшенні частоти дискретизації, розширенні динамічного діапазону, підвищенні роздільної здатності, відсіканні вхідного шуму тощо.

Загальне уявлення про принцип обробки аналогового сигналу можна отримати зі структурної схеми (рис. 1.3).

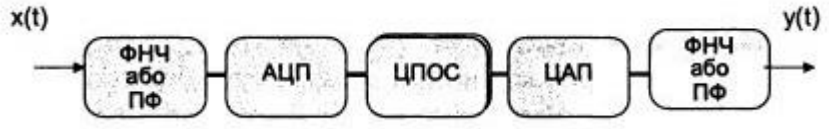


Рисунок 1.1 – Структурна схема цифрової обробки сигналів

Як правило, перед здійсненням реального аналого-цифрового перетворення аналоговий сигнал проходить через кола нормалізації, які виконують функції підсилення, послаблення (атенюації) та фільтрації. Для придушення небажаних сигналів поза смугою пропускання та запобігання накладанню спектрів необхідний ФНЧ/ПФ.

Аналого-цифровий перетворювач неперервно дискретизує сигнал із частотою, яка дорівнює частоті дискретизації - fд, та видає новий відлік цифровому процесору обробки сигналів (ЦПОС) із такою самою частотою. Для забезпечення роботи в режимі реального часу ЦПОС повинен закінчити усі

обчислення в межах інтервалу дискретизації 1/*f*д та передати вихідний відлік на ЦАП до надходження наступного відліку з АЦП.

Цифроаналоговий перетворювач (ЦАП) потрібний, коли дані необхідно перетворити назад у аналоговий сигнал (наприклад, у випадку голосового або звукового додатка). Тоді, після нього, для усунення небажаних ВЧ-гармонік необхідно застосовувати ФНЧ.

У реальних процесах аналого-цифрового та цифро-аналогового перетворення є два ключові етапи: дискретизація за часом та квантування за амплітудою, які визначають роздільну здатність цих операцій.

Найпростішими сигналами мультимедійних систем є одномірні сигнали, як, наприклад, аудіосигнали, виміри яких- небудь параметрів полів (електричних, магнітних, та ін.). Значення одномірних сигналів залежать тільки від однієї незалежної змінної просторових, часових і інших незалежних змінних, наприклад, відеосигнали. Все більше застосування знаходять також багатомірні сигнали, утворені деякою множиною одномірних сигналів.

Багатомірні сигнали можуть мати різне представлення за своїми аргументами, наприклад, дискретним за рівнем і безперервним за часом.

Багатомірний сигнал може розглядатися, як упорядкована сукупність одномірних сигналів. З урахуванням цього, при аналізі й обробці сигналів, багато принципів і практичні методи обробки одномірних сигналів, математичний апарат яких розвинений досить глибоко, поширюються й на багатомірні сигнали. Фізична пририрода сигналів для математичного апарата їхньої обр значення не має.

Разом з тим обробка багатомірних сигналів має свої особливості і може істотно відрізнятися від обробки одномірних сигналів, внаслідок більшого числа ступенів волі. Так, при дискретизації багатомірних сигналів має значення не тільки частотний спектр сигналів, але й форма растра дискретизації. Приклад не дуже корисної особливості - багатомірні поліноми сигнальних функцій, на відміну від одномірних, не розкладаються на прості множники. Що стосується порядку розмірності багатомірних сигналів, то її збільшення вище

двох практично не змінює принципи й методи аналізу даних і позначається, в основному, тільки на ступені громіздкості формул і технічному ускладненні обчислень.

З огляду на ці фактори, при розгляді загальної теорії аналізу, перетворень і обробки сигналів обмежимося, в основному, одно- і двовимірними сигнальними функціями, а як універсальні незалежні змінні (аргументи функцій) будемо використовувати, зазвичай, змінну ***t*** для одномірних сигналів і змінні x,t або ***x,x*** для двовимірних сигналів, безвідносно до їхнього фізичного змісту (простір, час, енергія та ін.).

У загальному випадку сигнали є багатомірними функціями.

Коли говорять про "аналіз" сигналів, то йдеться не лише про суто математичні перетворення, але й про одержання на основі цих перетворень висновків стосовно специфічних особливостей відповідних процесів та об'єктів. Метою аналізу сигналів зазвичай є:

- визначення або оцінка числових параметрів сигналів (енергія, середня потужність, середньо квадратичне значення та ін.);
- розкладання сигналів на елементарні складові для порівняння властивостей різних сигналів;
- порівняння ступеня близькості, "подібності", "спорідненості" різних сигналів, у тому числі з певними кількісними оцінками.

Математичний апарат аналізу сигналів досить широко застосовується на практиці в усіх без винятку галузях науки й техніки.

З поняттям сигналу нерозривно зв’язаний термін ***реєстрації*** сигналів, використання якого є також широким й неоднозначним, як і самого терміна "сигнал". У найбільш загальному розумінні цей термін означає операцію виділення сигналу і його перетворення у форму, зручну для подальшого використання. Так, у процесі інформації щодо фізичних властивостей яких-небудь об'єктів під реєстрацією сигналу розуміють процес виміру фізичних

властивостей об'єкта й перенесення результатів вимірювань на матеріальний носій сигналу або безпосереднє енергетичне перетворення яких-небудь властивостей об'єкта в інформаційні параметри матеріального носія сигналу (як правило електричного). Але так само широко термін реєстрації сигналів використовують і для процесів виділення вже сформованих сигналів, що несуть певну інформацію, із суми інших сигналів (радіозв'язок, аудіо й відеотехніка тощо), і для процесів фіксації сигналів на носіях довгострокової пам'яті, і для багатьох інших процесів, пов'язаних з обробкою сигналів.

Стосовно даної роботи під терміном реєстрації будемо розуміти ***реєстрацію даних*** які проходять через конкретну систему або точку системи й певним чином фіксуються на якому-небудь матеріальному носії або в пам'яті системи. Що стосується процесу одержання інформації за допомогою технічних засобів, які забезпечують дослідним шляхом знаходження співвідношення вимірюваної величини із прийнятої за визначенням зразковою одиницею цієї величини, і подання обмірюваного співвідношення в якій-небудь фізичній або числовій формі інформаційного сигналу, то для цього процесу будемо застосовувати, в основному, термін ***детектування.***

При детектуванні сигналів, що несуть цільову для даного виду вимірів інформацію, у сумі з основним сигналом одночасно реєструються й сигнали, що заважають, - шуми й завади всілякої природи.

До завад належать також перекручування корисних сигналів під впливом різних дестабілізуючих факторів на процеси вимірів. Виділення корисних складових із загальної суми зареєстрованих сигналів або максимальне придушення шумів і завад в інформаційному сигналі при збереженні його корисних складових є однієї з основних задач первинної обробки сигналів (результатів спостережень).

Типи завад розділяють за джерелами їхнього виникнення, за енергетичними спектрами, за характерами впливів на сигнал, за імовірнісними характеристиками і іншими ознаками.

Джерела завад бувають внутрішні й зовнішні.

Внутрішні шуми можуть бути властиві фізичній природі джерел сигналів, як, наприклад, теплові шуми електронних потоків в електричних ланцюгах або дробові ефекти в електронних приладах, або ті, що виникають у вимірювальних пристроях і системах передачі й обробки сигналів під впливом різних дестабілізуючих факторів - температури, підвищеної вологості, нестабільності джерел живлення, впливу механічних вібрацій на гальванічні з'єднання тощо.

Зовнішні джерела шумів бувають штучного й природного походження. До штучних джерел завад належать індустріальні завади - двигуни, перемикачі, генератори сигналів різної форми й природними джерелами завад є блискавки, флукт, магнітних полів, сплески сонячної енергії тощо.

Електричні й магнітні поля різних джерел завад внаслідок наявності індуктивних, ємнісних і резистивних зв'язків створюють на різних ділянках і ланцюгах сигнальних систем паразитні різниці потенціалів і струми, що накладаються на корисні сигнали.

Завади підрозділяються на флуктуаційні, імпульсні й періодичні. Флуктуаційні або шумові завади представляють хаотичний і безладний у часі процес у вигляді нерегулярних випадкових сплесків різної амплітуди. Як правило, флуктуаційні завади розподілені за нормальним законом з нульовим середнім і впливають тільки на сигнали низького рівня.

Імпульсні завади багато в чому схожі на шумові завади й проявляються як у вигляді окремих імпульсів, так і у вигляді послідовності імпульсів, форма й параметри яких мають випадковий характер. Причинами імпульсних завад є різкі стрибки струму й напруги в промислових установках, транспортних засобах, а також природні електричні явища. Розподіл імпульсних перешкод симетричне з довільною щільністю розподілу.

Періодичні завади викликаються періодичними низькочастотними або високочастотними полями ліній електропередач, силових електроустановок та ін. Якщо основна потужність завад зосереджена на окремих ділянках діапазону частот, наприклад, на частоті напруги промислової мережі або кратна цій

частоті, то такі завади називають зосередженими.

Залежно від характеру впливу на сигнал завади розділяють на адитивні й мультиплікативні. Адитивні (тобто ті, які накладаються) завади підсумовуються із сигналом, не залежать від його значень і форми й не змінюють інформативної складової самого сигналу. Мультиплікативні або деформуючі завади можуть змінювати форму інформаційної частини сигналу, залежати від його значень і від певних особливостей у сигналі й таке інше. При відомому характері мультиплікативних завад можлива корекція сигналу на їхній вплив.

Варто зазначити, що поділ сигналів на корисні й ті, що заважають (шумові) є досить умовним. Джерелами сигналів, що заважають, також є певні фізичні процеси, явища або об'єкти. Пририрода сигналів для математичного апарата їхньої обр значення не має.

Разом з тим обробка багатомірних сигналів має свої особливості і може істотно відрізнятися від обробки одномірних сигналів, внаслідок більшого числа ступенів волі. Так, при дискретизації багатомірних сигналів має значення не тільки частотний спектр сигналів, але й форма растра дискретизації. Приклад не дуже корисної особливості - багатомірні поліноми сигнальних функцій, на відміну від одномірних, не розкладаються на прості множники. Що стосується порядку розмірності багатомірних сигналів, то її збільшення вище двох практично не змінює принципи й методи аналізу даних і позначається, в основному, тільки на ступені громіздкості формул і технічному ускладненні обчислень.

З огляду на ці фактори, при розгляді загальної теорії аналізу, перетворень і обробки сигналів обмежимося, в основному, одно- і двовимірними сигнальними функціями, а як універсальні незалежні змінні (аргументи функцій) будемо використовувати, зазвичай, змінну $\boldsymbol{t}$ для одномірних сигналів і змінні x,t або $\boldsymbol{x,x}$ для двовимірних сигналів, безвідносно до їхнього фізичного змісту (простір, час, енергія та ін.).

ЛЕКЦІЯ 2

Тема: Математичні моделі сигналів.

План

1. Математичний опис сигналів.
2. Математичні моделі сигналів. Класифікація сигналів.
3. Дискретизація та квантування сигналів.

Сигнали можуть бути об’єктами теоретичних досліджень і практичного аналізу тільки в тому випадку, якщо зазначено спосіб їхнього математичного опису - математична модель сигналу. Математичний опис дозволяє абстрагуватися від фізичної природи сигналу й матеріальної форми його носія, проводити класифікацію сигналів, виконувати їхнє порівняння, установлювати ступінь тотожності, моделювати системи обробки сигналів. Як правило, опис сигналу задається функціональною залежністю певного інформаційного параметра сигналу від незалежної змінної (аргументу) – ***s(x), y(t)*** і т.п. Функції математичного опису сигналів можуть бути як дійсними, так і комплексними. Вибір математичного апарата опису визначається простотою й зручністю його використання при аналізі й обробці сигналів.

Відзначимо подвійність застосування опису сиг: функціями типу s(t) й т.п. з одного боку, s(t) - це величина, рівна значенню функції в момент часу ***t***. З іншого, - ми позначаємо через s(t) і саму функцію, тобто маємо правило, за яким кожному значенню ***t*** ставиться у відповідність певна величина s. У більшості аналітичних виражень це не обумовлює непорозумінь і за однозначної відповідності значень сигналів їхнім аналітичним вираженням приймається за умовчанням.

Теорія аналізу й обробки фізичних даних базується на математичних моделях відповідних фізичних полів і фізичних процесів, на основі яких

створюються математичні моделі сигналів. Математичні моделі сигналів дають можливість узагальнено, абстрагуючись від фізичної природи, робити висновок про властивості сигналів, передбачувати зміни сигналів в умовах, що змінюються, заміняти фізичне моделювання досліджуваних процесів математичним. За допомогою математичних моделей є можливість описувати властивості сигналів, які є головними, визначальними в досліджуваних процесах, та ігнорувати велику кількість другорядних ознак. Знання математичних моделей сигналів дає можливість класифікувати їх за різними ознаками, що характерні для того або іншого типу моделей. Так, сигнали розділяють на невипадкові й випадкові, залежно від можливості точного передбачення їхніх значень у будь-які моменти часу. Сигнал є невипадковим і називається ***детермінованим,*** якщо математична модель дозволяє здійснювати таке передбачення. Детермінований сигнал задається, як правило, математичною функцією або обчислювальним алгоритмом, а математична модель сигналу може бути представлена у вигляді

$$s=F\,(t,\ z,\omega,\ldots,;\ A,B,C,\ldots),$$

де s - інформативний параметр сигналу; t, z, ω, незалежні аргументи (час, просторова координата, частота та ін.); $A,B,C,\ldots,-$ параметри сигналів. Модель повинна бути, за можливістю, простіш, мінімізована за кількістю незалежних аргументів і адекватна досліджуваному процесу, що в основному визначає результати вимірів.

Математичний опис не може бути всеосяжним та ідеально точним і, по суті, завжди відображає не реальні об’єкти, а їх спрощені або гомоморфні моделі. Моделі можуть задаватися таблицями, графіками, функціональними залежностями, рівняннями станів і переходів з одного стану в інше і т.п. Формалізований опис може вважатися математичною моделлю оригіналу, якщо він дозволяє з певною точністю прогнозувати стан і поводження досліджуваних об'єктів шляхом формальних процедур над їхнім описом.

Невід'ємною частиною будь-якої математичної моделі сигналу є також

область визначення сигналу, що встановлюється інтервалом завдання незалежної змінної. Приклади завдання інтервалу для змінних:

$$a \le x \le b,\ x \in [a,b];\quad a < y \le b,\quad y \in (a,b];$$

$$a < z < b,\quad z \in (a,b).$$

Простір значень незалежної змінної звичайно позначається через індекс R. Наприклад, $R := (-\infty,+\infty)$, $x \in R$.

Крім завдання області визначення сигналу можуть бути також задані види чисельних значень змінних (цілі, раціональні, дійсні, комплексні).

Математичні моделі полів і сигналів на першому етапі обробки й аналізу результатів спостережень повинні дозволяти якоюсь мірою ігнорувати їхню фізичну природу й повертати її в модель тільки на заключному етапі інтерпретації даних. При аналізі фізичних даних використовують два основних підходи до створення математичних моделей сигналів.

Перший підхід оперує з ***детермінованими*** сигналами, значення яких у будь-який момент часу або в довільній точці простору (а також залежно від будь-яких інших аргументів) є аріорно відомими або можуть бути досить точно визначені (обчислені). Для опису невипадкових сигналів використовуються також ***квазідетерміновані*** моделі, у яких значення одного або декількох параметрів апріорно невідомі та вважаються випадковими величинами з малим випадковим компонентом, впливом якого можна зневажати.

Другий підхід припускає випадковий характер сигналів, закон зміни яких у часі (або в просторі) має випадковий характер і які набувають конкретних значень з деякою імовірністю. Модель такого сигналу являє собою опис статистичних характеристик випадкового процесу шляхом завдання законів розподілу ймовірностей, кореляційної функції, спектральної щільності енергії та ін.

Випадковість може бути обумовлена як власною фізичною природою сигналів, що характерно, наприклад, для звукових сигналів, так й імовірнісним характером сигналів, що реєструються, як за часом або місцем їхньої появи, так і за змістом. Із цих позицій випадковий сигнал може розглядатися як

відображення випадкового за своєю природою процесу або фізичних властивостей об'єкта (процесу), які визначаються випадковими параметрами чи складною структурою, результати вимірів у якій важко передбачити.

Між цими двома видами сигналів немає різкої границі. Строго кажучи, детермінованих процесів і детермінованих сигналів, що відповідають їм, у природі не існує. Навіть сигнали, добре відомі на вході в середовище (при зовнішньому впливі на неї), за місцем їхньої реєстрації завжди ускладнені випадковими завадами, впливом дестабілізуючих факторів і апріорно невідомими параметрами й будовою самого середовища. З іншого боку, модель випадкового поля часто апроксимується методом суперпозиції (накладення) сигналів відомої форми. Детерміновані моделі можуть використатися й для вивчення чисто випадкових процесів, якщо рівень корисного сигналу в цьому процесі значно вище рівня статистичних флуктуацій.

На вибір математичної моделі в тому або іншому випадку значно впливає також складність математичного апарату обробки сигналів і сформовані традиції інтерпретації результатів спостережень. Не виключається й зміна моделі, як правило, з ймовірнісної в детерміновану, у процесі нагромадження інформації про досліджуване явище або об'єкт.

Класифікація сигналів здійснюється на підставі істотних ознак відповідних математичних моделей сигналів. Усі сигнали розділяють на дві великі групи: детерміновані й випадкові (рис.2.1).

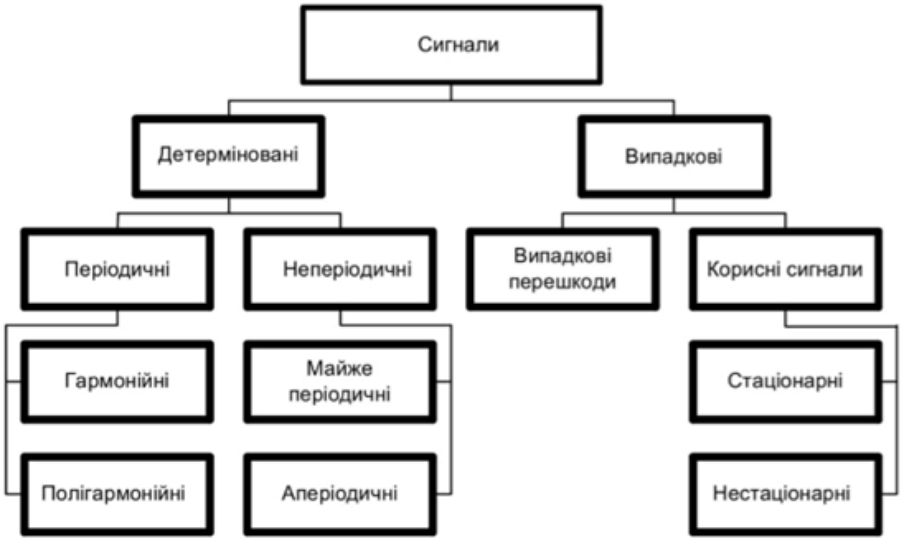


Рисунок 2.1 – Класифікація сигналів

З енергетичних позицій сигнали розділяють на два класи: з обмеженою (кінцевою) енергією та з нескінченною енергією.

Для сигналів з обмеженою енергією (інакше сигналів, що інтегруються з квадратом) повинне виконуватися співвідношення скінченності енергії сигналу, тобто її обмеженості. Як правило, до цього класу сигналів відносять аперіодичні й імпульсні сигнали, що не мають розривів 2-го роду при обмеженій кількості розривів 1-го роду, і особливих точок, що йдуть у нескінченність. Будь-які періодичні, полігармонійні й майже періодичні сигнали, а також сигнали з розривами й особливими точками 2-го роду, є сигналами з нескінченною енергією. Для їхнього аналізу застосовуються спеціальні методи.

Іноді в окремий клас виділяють сигнали кінцевої тривалості, відмінні від нуля тільки на обмеженому інтервалі аргументів (незалежних змінних). Такі сигнали звичайно називають ***фінітними.***

Гармонічні сигнали (синусоїдальні), описуються наступними формулами:

$$s(t) = A \cdot \sin(2\pi f_o t + \phi) = A \cdot \sin(\omega_o t + \phi),$$

$$s(t) = A \cdot \cos(\omega_o t + \varphi),$$

де A, f_o, ω_o, φ, ϕ - постійні величини, які можуть виконувати роль

інформаційних параметрів сигналу, : A - амплітуда сигналу, f_o - циклічна частота в герцах, $\omega_o = 2\pi f_o$ - кутова частота в радіанах, φ і ϕ- початкові фазові кути в радіанах. Період одного коливання $T = 1/f_o = 2\pi/\omega_o$. При $\varphi = \phi - \pi/2$ синусоїдальні і косинусоїдальні функції описують один і той же сигнал. Частотний спектр сигналу представлений амплітудним і початковим фазовим значенням частоти f_o (при t = 0).

Полігармонічні сигнали складають найбільш широко поширену групу періодичних сигналів і описуються сумою гармонійних коливань :

$$s(t) = \sum_{n=0}^{N} A_n \sin(2\pi f_n t + \varphi_n),$$

чи безпосередньо функцією $s(t) = y(t \pm kT_p)$, $k = 1,2,3,...$, де T_p - період одного повного коле-бания сигналу y(t), заданого на одному періоді. Значення $f_p = 1/T_p$ називають фундаментальною частотою коливань. Полігармонічні сигнали є сумою певної постійної складової ($f_o=0$) і довільного (у межі - нескінченного) числа гармонійних складових з довільними значеннями амплітуд A_n і фаз φ_n, з періодами кратними періоду фундаментальної частоти f_p. Іншими словами, на періоді фундаментальної частоти f_p, яка рівна або кратна менше мінімальної частоти гармонік, укладається кратне число періодів усіх гармонік, що і створює періодичність повторення сигналу. Частотний спектр полі гармонічних сигналів дискретний, у зв'язку з чим друге поширене математичне пред-ставление сигналів - у вигляді спектрів (рядів Фур'є).

Як приклад на рис. 4 приведений відрізок періодичної сигнальної функції, яка отримана підсумовуванням постійної складової (частота постійної складової дорівнює 0) і трьох гармонійних коливань з різними значеннями частоти і початкової фази коливань. Математичний опис сигналу задається формулою:

$$s(t) = \sum_{k=0}^{3} A_k \cdot \cos(2 \cdot \pi \cdot f_k \cdot t + \varphi_k),$$

де: A_k = {5, 3, 4, 7} - амплітуда гармонік; f_k = {0, 40, 80, 120} - частота в герцах; φ_k = {0, - 0.4, - 0.6, - 0.8} - початковий фазовий кут коливань в радіанах; k = 0, 1, 2, 3. Фундаментальна частота сигналу 40 Гц.

Рисунок 2.2 – Модель сигналу

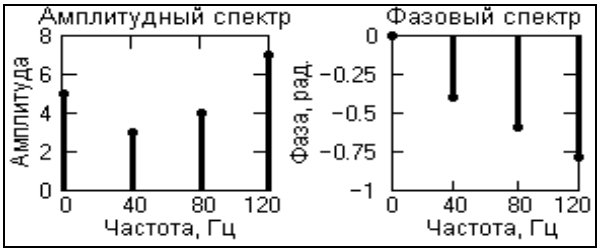


Рисунок 2.3 – Спектр сигналу

Частотне представлення цього сигналу (спектр сигналу) приведене на рис.2.3. Звернемо увагу, що частотне представлення періодичного сигналу s(t), обмеженого по числу гармонік спектру, складає всього вісім відліків і дуже компактно в порівнянні з тимчасовим представленням.

Періодичний сигнал будь-якої довільної форми може бути представлений у вигляді суми гармонійних коливань з частотами, кратними фундаментальній частоті коливань $f_p = 1/T_p$. Для цього досить розкласти один період сигналу в ряд Фур'є по тригонометричних функціях синуса і косинуса з кроком по частоті, рівним фундаментальній частоті коливань $\Delta f = f_p$:

$$s(t) = \sum_{k=0}^{K} (a_k \cos 2\pi k\Delta ft + b_k \sin 2\pi k\Delta ft),$$

$$a_o = (1/T)\int_0^T s(t)\,dt, \quad a_k = (2/T)\int_0^T s(t)\cos 2\pi k\Delta ft\,dt,$$

$$b_k = (2/T)\int_0^T s(t)\sin 2\pi k\Delta ft\,dt.$$

Кількість членів ряду Фур'є K = k_{max} зазвичай обмежується максимальними частотами f_{max} гармонійних складових в сигналах так, щоб $f_{max} < K \cdot f_p$. Проте для сигналів з розривами і скачками має місце $f_{max} \to \infty$, при цьому кількість членів ряду обмежується по допустимій погрішності апроксимації функції s(t).

Одночастотні косинусні і синусні гармоніки можна об'єднати і представити розкладання в компактнішій формі:

$$s(t) = \sum_{k=0}^{K} S_k \cos(2\pi k \Delta f t - \varphi_k),$$

$$S_k = \sqrt{a_k^2 + b_k^2}, \quad \varphi_k = \text{argtg}\,(b_k/a_k).$$

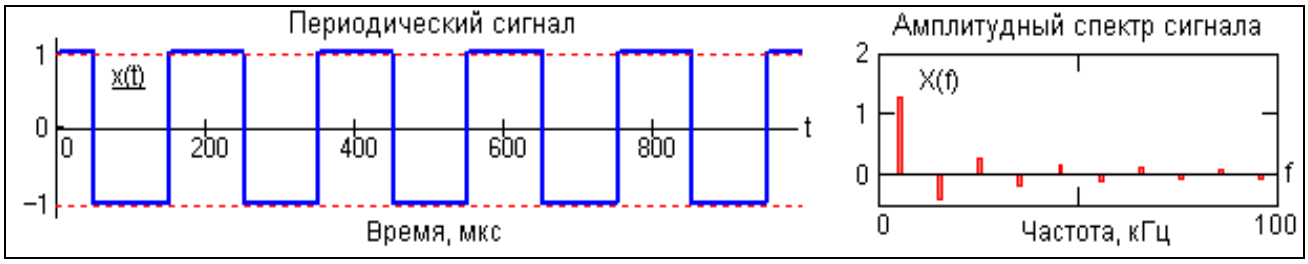


Рисунок 2.4 – Прямокутний періодичний сигнал

Приклад представлення прямокутного періодичного сигналу (меандру) у вигляді амплітудного ряду Фур'є в частотній області наведений на рис.6. Сигнал парний відносно t=0, не має синусних гармонік, усі значення φ_k для цієї моделі сигналу дорівнюють нулю.

Інформаційними параметрами **полігармонічного сигналу** можуть бути як певні особливості форми сигналу (розмах від мінімуму до максимуму, екстремальне відхилення від середнього значення, і тому подібне) так і параметри певних гармонік в цьому сигналі. Так, н-д, для прямокутних імпульсів інформаційними параметрами можуть бути період повторення імпульсів, тривалість імпульсів, шпаруватість імпульсів (відношення періоду до тривалості). При аналізі складних періодичних сигналів інформаційними параметрами можуть також бути:

- Поточне середнє значення за певний час, наприклад, за час періоду:

$(1/T)\int_{t}^{t+T} s(t)\,dt.$

- Постійна складова одного періоду :

$(1/T)\int_{0}^{T} s(t)\,dt.$

- Середнє випрямлене значення:

$(1/T)\int_{0}^{T} |s(t)|\,dt.$

- Середнє квадратичне значення:

$\sqrt{\frac{1}{T}\int_{0}^{T} x(t)^2\,dt}$.

До неперіодичних сигналів відносять: майже періодичні і аперіодичні сигнали. Основним інструментом їх аналізу також є частотне представлення.

Аперіодичні сигнали складають основну групу неперіодичних сигналів і задаються довільними функціями часу. На рис. 7 показаний приклад аперіодичного сигналу, заданого формулою на інтервалі (0, ∞):

$$s(t) = \exp(-a\cdot t) - \exp(-b\cdot t),$$

де a і b - константи, в даному випадку a = 0.15, b = 0.17.

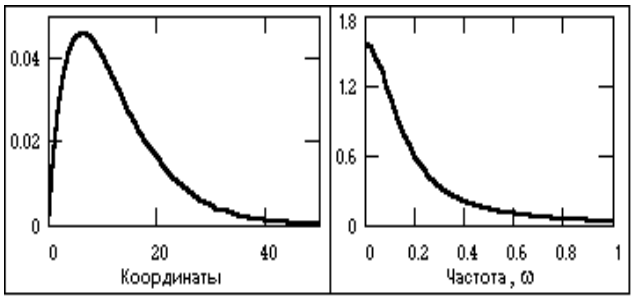


Рисунок 2.5 – Аперіодичний сигнал і модуль спектра

До аперіодичних сигналів відносяться також імпульсні сигнали, які в радіотехніці і в галузях, що широко її використовують, часто розглядають у вигляді окремого класу сигналів. Імпульси є сигналами, як правило, певної і досить простої форми, існуючі в межах кінцевих тимчасових інтервалів. Сигнал, приведений на рис. 2.5, належить до імпульсних.

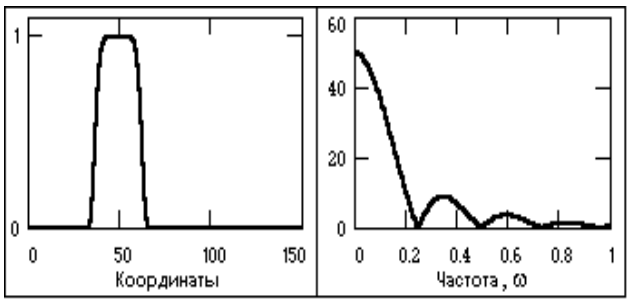


Рис. 2.6 – Імпульсний сигнал і модуль спектра

У класі імпульсних сигналів виділяють підклас радіоімпульсів. Приклад радіоімпульса наведений на рис. 2.7.

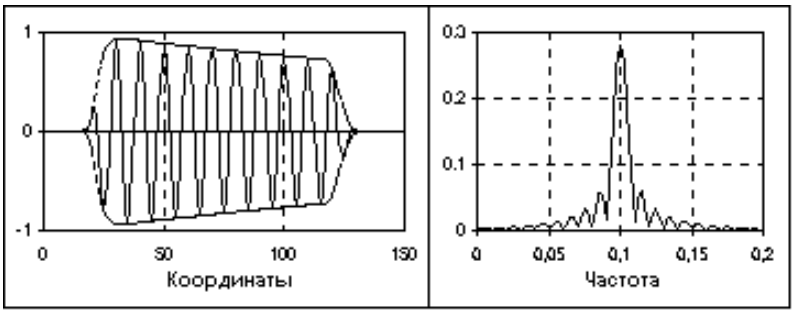


Рисунок 2.7 – Радіоімпульс і модуль його спектра

Рівняння радіоімпульсу має вигляд:

$$s(t) = u(t)\cos(2\pi f_o t+\varphi_o).$$

де $\cos(2\pi f_o t+\varphi_o)$ - гармонійне коливання заповнення радіоімпульсу, u(t) - що огинає радіоімпульсу. Положення головного піку спектру радіоімпульсу на частотній шкалі відповідає частоті заповнення f_o, а його ширина визначається тривалістю радіоімпульсу. Чим більше тривалість радіоімпульсу, тим менше ширина головного частотного піку.

Виділяють наступні типи сигналів, яким відповідають певні форми їхнього математичного опису.

Аналоговий сигнал є безперервною функцією безперервного аргументу, тобто визначений для будь-якого значення аргументів. Джерелами аналогових сигналів, як правило, є фізичні процеси і явища, безперервні в динаміці свого

розвитку в часі, у просторі або за будь-якої іншої незалежної змінної. Сигнал, що реєструється, є подібним (“аналогічним”) процесу, що його породжує. Приклад математичного запису сигналу: y(t)= 4,8 exp([-(t-4)^2/2,8].

Приклад графічного відображення даного сигналу наведений на рис. 5; при цьому як сама функція, так і її аргументи, можуть приймати будь-які значення в межах деяких

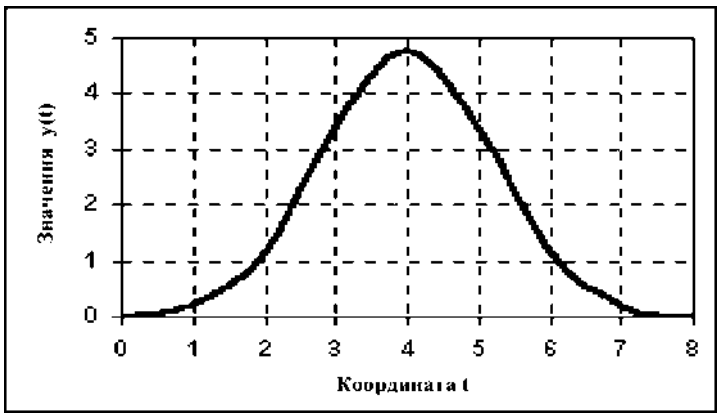


Рисунок 2.8 – Аналоговий сигнал

Під дискретизацією розуміють перетворення неперервного сигналу в дискретний, представлений сукупністю відліків, за якими неперервний сигнал може бути відновлений із заданою точністю. Вважатимемо, що відліки сигналу є відліками часу. Дискретизація за часом (дискретизація) - це процедура зняття миттєвих значень сигналу $x(t)$ через рівні проміжки часу T. На рис. 1.4 показаний процес дискретизації неперервного сигналу $x(t)$.

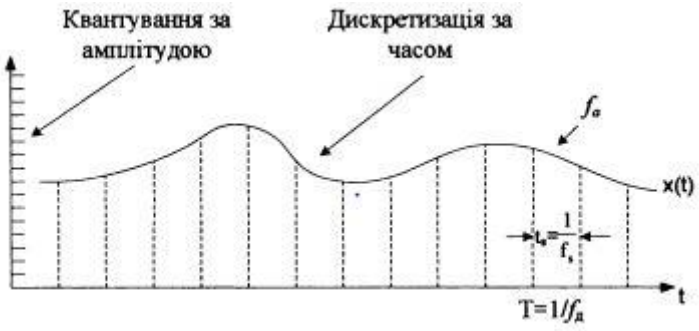


Рисунок 2.9 – Дискретизаця та квантування аналогового сигналу

Миттєві значення x[nT] називаються вибірками, або відліками, час T - періодом (інтервалом, кроком) дискретизації, n вказує порядковий номер відліку. Очевидно, що чим частіше брати відліки, тобто чим менший період дискретизації T, тим точніше послідовність відліків x[nT] відображатиме вихідний сигнал x(t). Період дискретизації T визначає частоту дискретизації fд, яка дорівнює:

$$f_д = 1/T; \;\; T = 1/f_д$$

звідси чим менше T тим вище fд. З іншого боку, чим вища частота дискретизації, тим важче обчислювачу виконувати велику кількість операцій над відліками в темпі їхнього надходження на переробку і тим складнішою повинна бути його будова.

Отже, точність представлення сигналу вимагає збільшення fд, а прагнення зробити обчислювач якомога простішим приводить до бажання її понизити. Якщо брати відліки дуже рідко, то у випадку швидкозмінного сигналу вони не міститимуть інформації, достатньої для його ідентифікації. Швидкість зміни сигналу характеризується верхньою частотою його спектра. Тому існує обмеження на мінімальне значення fд - для повного відновлення неперервного сигналу за його відліками x[nT] необхідно і достатньо, щоб частота дискретизації fд була, принаймі удвічі більшою від найвищої частоти F у спектрі сигналу x(t), який передається, тобто:

$$f_д \geq 2F, \quad T \leq 1/2F$$

Отже, якщо сигнал не обмежений за спектром (верхня частота F прямує до нескінченності), то мінімально допустима величина інтервалу дискретизації прагнутиме до нуля. Звідси можна зробити важливий висновок - без втрати інформації дискретними відліками можуть бути представлені лише обмежені за спектром аналогові сигнали. Якраз тому в системах цифрової обробки сигналів перед виконанням дискретизації сигналу його спектр обмежують до частоти F =

f_B застосуванням фільтра низьких частот, який називається ще антиелайзинговим фільтром.

Явище елайзингу полягає у виникненні спотворень сигналу за рахунок накладання спектра у разі невдалого вибору частоти дискретизації. Дискретизація в часі призводить до появи періодичних копій спектра сигналу. За надто малої частоти дискретизації ці копії перекриваються, що спричиняє спотворення сигналу під час його відновлення. Гармоніки сигналу з частотами, вищими від частоти дискретизації, відображаються в частоти, нижчі від цієї частоти, створюючи завади. Гранична частота дискретизації fд=2F за якої перекриття ще не відбуваються, називається ще частотою Найквіста.

У спектрі будь-якого скінченного сигналу є такі вищі складові, які, починаючи з деякої верхньої частоти f_B, мають незначні амплітуди і тому ними можна знехтувати без істотного спотворення самого сигналу. Значення f_B визначається конкретним типом сигналу та розв'язуваної задачі. Наприклад, для стандартного телефонного сигналу f_B = 3,4 кГц , а мінімальна стандартна частота його дискретизації f_B =8кГц .

Антиелайзинговий фільтр формує аналоговий сигнал зі значно придушеними верхніми частотними складовими у смузі затримки, починаючи з частоти $F = f_B$. Це дає підставу вважати сигнал практично обмеженим за частотою і таким, що не зазнає ефекту накладання при частоті дискретизації, не меншій за 2F. Дискретизація з частотою Найквіста називається граничною дискретизацією. Сигнал, дискретизований із $f_д$ > 2F, називається передискретизованим сигналом. Незважаючи на те, що в цьому випадку одержується надлишкова кількість відліків, інколи така техніка необхідна, особливо під час аналізу сигналів з виділенням якихось ознак.

Якщо крок дискретизації постійний, то дискретизація називається рівномірною, у протилежному випадку - нерівномірною. У разі нерівномірної дискретизації крок "підлаштовується" під швидкість зміни сигналу, збільшуючись на гладких, малоінформативних ділянках. Виникає необхідність

в зберіганні значення інтервалу дискретизації між кожною парою відліків, тому нерівномірну дискретизацію рідко використовують на практиці.

Аналогом дискретизації для цифрового сигналу слугує операція проріджування. Суть її полягає в тому, що за певними правилами можна без втрати інформації “викидати” деякі члени послідовності, тобто проріджувати цифровий сигнал. Операція проріджування називається ще децимацією. Її виконує пристрій, який називається дециматором. Зворотна їй операція відновлення сигналу називається інтерполяцією, її реалізує інтерполятор. Децимація позначається стрілкою, напрямленою вниз - ↓, а інтерполяція - стрілкою, спрямованою до гори - ↑. Поряд із стрілкою стоїть коефіцієнт децимації (інтерполяції"). Наприклад, запис ↓2 означає, що з послідовності вилучається кожний другий відлік; запис - ↑2 означає вставлення нуля між сусідніми відліками. Дециматор й інтерполятор є лінійними системами.

Під квантуванням розуміють перетворення деякої величини з неперервною шкалою значень на величину з дискретною шкалою значень. Воно полягає в заміні будь-якого значення відліку одним із кінцевої множини дозволених значень або рівнів квантування.

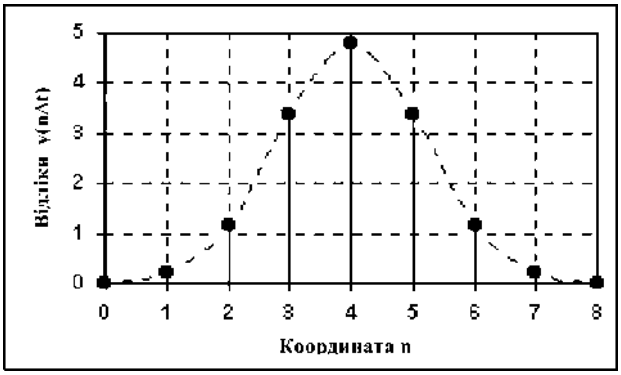


Рисунок 2.10 – Цифровий сигнал

Квантування виконують з метою формування послідовності чисел $x[nT] = x[n]$, поданих у двійковому коді. Для цього весь діапазон змін величини відліків ділиться на деяку кількість дискретних рівнів і кожному відліку за певними правилами присвоюється значення одного з двох найближчих рівнів квантування, між якими розміщений цей відлік. Різниця між сусідніми

квантованими значеннями визначає крок квантування, а різницю між фактичним і квантованим значенням називають помилкою квантування ε. Максимальне значення цієї помилки дорівнює половині кроку квантування. На підставі потрібного значення помилки квантування вибирають кількість рівнів квантування. Що більше рівнів квантування, то менша помилка.

Оскільки цілі числа (рівні квантування) розташовані через однакові інтервали, то говорять про рівномірне квантування. Відомо, що середня помилка рівномірного квантувача випадкового сигналу дорівнює $\Delta^2/12$, де Δ- крок квантувача. Отже, у випадку операції заокруглення середня помилка дорівнює 1/12. За рівномірного квантування випадкового сигналу співвідношення сигнал/шум на виході квантувача приблизно дорівнює 6*N* дБ, де N - кількість бітів, які використовуються дня кодування одного відліку. Наприклад, у 12-розрядного АЦП відношення сигнал/шум може дорівнювати 72 дБ.

Кількість рівнів квантування визначається розрядністю АЦП. Так, АЦП на три розряди може мати лише вісім рівнів квантування (к =2^3), а мінімальне і максимальне значення відліків дорівнюють відповідно 0 ↔ 000 і 7 ↔ 111. Отже, що більша розрядність АЦП, то точніше поданий відлік, однак складнішим і дорожчим є АЦП, який необхідний для розв’язання поставленої задачі.

Помилка квантування на виході ідеального рівномірного квантувача має рівномірний спектр від 0 до частоти, що дорівнює половині частоти дискретизації. Однак система людського слуху (зору) має різну частотну чутливість. Існують методи побудови квантувачів, які враховують аудіовізуальні влас-тивості людини. Спектр помилки квантування при цьому переміщується за межі видимого (чутного) діапазону, тобто здійснюється шейпінг шуму квантування. Такий шейпінг здійснює, наприклад, модулятор сигма-дельта АЦП.

Помилка квантування може призводити до певних проблем. Так, помилка квантування корельована з сигналом, наприклад, музичним, дає відчуття “брудного” звуку під час прослуховування. Для надання помилці

випадковішого характеру до вихідного аналогового сигналу можна додати деяку кількість високочастотного шуму. Цей метод називається дизеризацією, а відповідний квантувач - дизеризованим. Енергія помилки в разі використання цього методу дещо зростає, але помилка стає декорельованою із сигналом, що у деяких випадках покращує якість звучання. Найефективніший, але трудомісткий метод декореляції помилки квантування називається дизеризацією з відніманням.Його суть така: спочатку генерують і запам’ятовують послідовність псевдо- випадкових чисел; потім її пропускають через АЦП і складають з аналоговим сигналом. На виході АЦП запам’ятовані випадкові числа віднімаються від оцифрованого сигналу.

Зменшити помилку квантування можна, збільшивши кількість рівнів квантування. Однак, це не завжди можливо. Разом з тим, у техніці широко застосовують компандери. Це пристрої, які здійснюють стиснення динамічного діапазону аналогового сигналу перед квантувачем (компанлування) і його розширення після квантувача (експанлування) Використання компанлерів лаг змогу зменшити помилки квантування.

Останнім часом під квантуванням часто розуміють процес перетворення сигналів із дискретною шкалою значень також на цифрові сигнали, діапазон значень яких с меншим. Особливо важливе значення таке перетворення мае для стиснення сигналів, тобто зменшення розмірності їх цифрового опису. Під час цього аналоговий сигнал спочатку переитворюється на цифровий за допомогою квантування достатньо великою кількістю рівнів, яке виконує АЦП, а потім цей сигнал "стискається” в системі ЦОС. Стиснення досягається за рахунок двох речей. По-перше, цифрові значення відліків реальних сигналів виявляються корельованими. Так, коефіцієнт кореляції сусідніх відліків мови приблизно дорівнює 0.9. Ця кореляція може застосовуватись, наприклад, у диференційних системах кодування або в системах стиснення з прогнозуванням. По- друге, при стисненні використовуються властивості людського слуху (зору), який подібний на фільтр, що пропускає певні частоти (та типи спотворень) і затримує інші. Завдяки цим властивостям часто спостерігається ефект

розбіжності між суб'єктивно спостережуваною якістю сигналу та об’єктивно виміряною мірою спотворення.

Вибір міри спотворення с одним із найважливіших при проектуванні квантувача Найчастіше використовують середньоквадратичну помилку, яка виражається в зручній математичній формі і багато в чому відображає властивості чуттєвого сприйняття.

Інколи квантування безпосереднього сигналу не настільки ефективне, як квантування його перетвореної версії. Причиною цього є те, що. за належного вибору перетворення, його коефіцієнти можуть виявитися декорельованими. Зазвичай використовують ортогональні перетворення, наприклад, Фур’є, дискретне косинусне, вейвлет-перетворения. Найчастіше кодування з перетворенням застосовують дія стиснення зображень. Так. в основу стандарту JPEG покладено дискретне косинусне перетворення, а в основу стандарту JPEG2000 - вейвлет-перегворення зображення. На завершальній стадії багатьох алгоритмів стиснення використовується ентропійний кодер. Таким кодером найчастіше слугує поєднання кодера довжин серій з кодером Хаффмана або арифметичним кодером.

Під час побудови алгоритму стиснення існують дві можливості: аналізувати і квантувати кожний відлік сигналу окремо або ралом з іншими відліками, об’єднуючи їх у вектор Відповідно розрізняють скалярне та векторне квантування.

Скалярний квантувач, як уже описувалось раніше, характеризується двома параметрами: кількістю рівнів квантування та кроком квантування. Найпростішим видом скалярного квантувача є рівномірний квантувач.

У результаті квантування, яке здійснюється аналого-цифровим перетворювачем, у канал зв’язку передається послідовність двійкових чисел. Кількість таких чисел визначається частотою дискретизації fд, а швидкість С їх передавання в каналі залежить ще і від розрядності Ь представлення чисел:

$$C = bf_д,$$

і під час передавання по каналу, наприклад, мовного сигналу за стандартних b = 12, fд =8000 Гц отримуємо С = 96000 біт/с, що перешкоджає використанню стандартних телефонних каналів.

У таких випадках використовують *векторне квантування* (рис. 1.5): у послідовності, яка кодується, визначають L блоків по k відліків у кожному за умови, що ці відліки не сильно відрізняються один від одного. Згадані k відліки можуть бути відображені із заданою похибкою своїм представником y. Такі блоки називають *кластерами*, а представника i-го кластера у i - центроїдом. Тоді кожний відлік $x[n]$, який належить 2-му кластеру, заміняється на відповідний центроїд, і в канал зв'язку передається тільки номер кластера (центроїда). Під час передавання та прийому необхідно мати множину із L центрощів, яка називається кодовою книгою, а параметр L - розміром кодової книги. Типові розміри кодових книг умовних технологіях - 256, 512, 1024, 2048 центроїдів (кластерів). Нехай, наприклад, L = 1024 = 2^{10}. Це означає, що кодова книга містить 1024 центроїди розмірністю k = 10 кожний. Тобто замість набору із десяти параметрів (на передавання яких необхідно було б відвести близько 40 бітів) передається лише номер кластера (а тому і номер центроїда), якому належить цей набір, на що необхідно всього 10 бітів, тобто забезпечується стиснення сигналу в 4 рази.

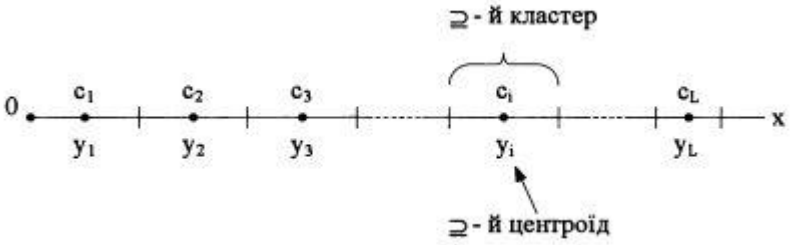


Рисунок 2.11 – Кластери та центроїди одномірної послідовності

На практиці коефіцієнт стиснення може бути значно більшим, що визначається співвідношенням:

$$K_{ст} = b/R, \quad R=\log 2L/K.$$

де b - розрядність одного відліку $к$ вимірного вектора; R - кількість бітів, які затрачаються на передавання одного відліку.

ЛЕКЦІЯ 3

Тема: Z-перетворення та його властивості.

План

1. Лінійна система.
2. Імпульсна характеристика лінійних систем.
3. Z-перетворення та його властивості.

Перетворення і обробка сигналів здійснюється в системах. Поняття сигналу і системи нерозривні, так як любий сигнал існує в будь-якій системі його звернення. Система обробки сигналів може бути реалізована як у матеріальній формі (спеціальний пристрій, вимірювальний прилад і т.п.), так і програмно на ЕОМ або на будь-якому іншому обчислювальному пристрої. Існують і комплексні вимірювально-обчислювальні системи (ІОС), які виконують як реєстрацію та первинну обробку сигналів безпосередньо в матеріальній формі їх подання, так і перетворення сигналів в цифрову форму, і подальшу програмну обробку. Форма реалізації систем істотного значення не має і визначає тільки їх можливості при аналізі і обробці сигналів. Основна увага при розгляді даної теми будемо приділятися цифровим системам і дискретної математики їх відображення та аналізу, застосовуючи аналітичну математику при розгляді загальних питань, якщо останнє спрощує виклад і розуміння теоретичного матеріалу.

Загальні поняття систем. Незалежно від призначення і виконання система завжди має вхід, на який подається вхідний сигнал або вхідний вплив, в загальному випадку багатовимірний, і вихід, з якого знімається оброблений вихідний сигнал. Якщо пристрій системи і внутрішні операції перетворень принципового значення не мають, то система в цілому може сприйматися як

"чорний ящик", в формалізованому вигляді. Формалізована система являє собою певний системний оператор (алгоритм) перетворення вхідного сигналу - впливу s (t), в сигнал на виході системи y (t) - відгук або вихідну реакцію системи. Символічне позначення операції перетворення (трансформації):

$$y(t) = T[s(t)].$$

Системний оператор T - це правило (набір правил, алгоритм) перетворення сигналу s(t) в сигнал y(t). Для загальновідомих операцій перетворення сигналів застосовуються також розширені символи операторів трансформації, де другим символом і спеціальними індексами позначається конкретний вид операції (як, наприклад, TF - перетворення Фур'є, TF-1 - зворотне перетворення Фур'є).

Вхідний сигнал системи може являти собою m - мірний вектор (m вхідних сигналів), а вихідний сигнал n - мірний вектор, при цьому система буде мати m входів і n виходів. Приклад такої системи з геофізики: трьохканальний гамма-спектрометр, на три входи вирішального блоку якого надходять сигнали від калієвого, радієвого і торієвого каналів спектрометра, а на три виходи виводяться сигнали складів калію, урану і торію, при цьому системний оператор реалізує алгоритм розв'язання системи трьох лінійних рівнянь з трьома невідомими.

Для детермінованих вхідних сигналів співвідношення між вихідними і вхідними сигналами однозначно задається системним оператором. У разі реалізації на вході системи випадкового вхідного процесу також існує однозначна відповідність процесів на виході і вході системи, однак при цьому одночасно відбувається зміна статистичних характеристик вихідного сигналу (математичного очікування, дисперсії, кореляційної функції та ін.), Яке також визначається системним оператором.

Для визначення системи необхідно задати характер, тип і області допустимих величин вхідних і вихідних сигналів. Як правило, системи виконуються на сигнали одного типу по входу / виходу та поділяються на системи безперервного часу (аналогові або дискретні сигнали на вході і виході)

і цифрові системи. Сукупність системного оператора T і простору сигналів утворює математичну модель системи.

Будь-які перетворення сигналів супроводжуються зміною їх спектру і за характером цих змін поділяються на два види: лінійні і нелінійні. До нелінійних відносять зміни, при яких в складі спектра сигналів з'являються нові гармонійні складові. При лінійних змінах сигналів змінюються амплітуди і / або початкові фази гармонічних складових спектру. Обидва види змін можуть відбуватися як зі збереженням корисної інформації, так і з її спотворенням. Це залежить не тільки від характеру зміни спектра сигналів, але і від спектрального складу найкориснішої інформації.

Лінійні системи складають основний клас систем обробки сигналів. Термін лінійності означає, що система перетворення сигналів повинна мати довільну, але в обов'язковому порядку лінійний зв'язок між вхідним сигналом (збудженням) і вихідним сигналом (відгуком). У нелінійних системах зв'язок між вхідним і вихідним сигналом визначається довільним нелінійним законом.

Система вважається лінійною, якщо в межах встановленої області вхідних і вихідних сигналів її реакція на вхідні сигнали адитивна (виконується принцип суперпозиції сигналів) і однорідна (виконується принцип пропорційного подібності).

Принцип адитивності вимагає, щоб реакція на суму двох вхідних сигналів дорівнювала сумі реакцій на кожен сигнал окремо:

$$T[a(t)+b(t)] = T[a(t)]+T[b(t)].$$

Принцип однорідності або пропорційної подібності вимагає збереження однозначності масштабу перетворення при будь амплітуді вхідного сигналу:

$$T[c * a(t)]= c* T[a(t)].$$

Іншими словами, відгук лінійної системи на зважену суму вхідних сигналів повинен бути рівний зваженій сумі відгуків на окремі вхідні сигнали незалежно від їх кількості і для будь-яких вагових коефіцієнтів, в тому числі комплексних.

При програмній реалізації лінійних систем на ЕОМ особливих труднощів

із забезпеченням лінійності в розумних межах значень вхідних і вихідних сигналів, як правило, не виникає. При фізичнії (апаратнії) реалізації систем обробки даних діапазон вхідних і/або вихідних сигналів, в якому забезпечується лінійність перетворення сигналів, завжди обмежений і повинен бути спеціально обумовлений в технічній документації або методичної інструкції.

Основні системні операції. До базових лінійних операцій, з яких можуть бути сформовані будь-які лінійні оператори перетворення, відносяться операції скалярного множення, зсуву і складання сигналів:

$$y(t) = b \times x(t),$$

$$y(t) = x(t-\Delta t),$$

$$y(t) = a(t)+b(t).$$

Графічне відображення операцій (цифрова форма) наведено на рис. 3.1.

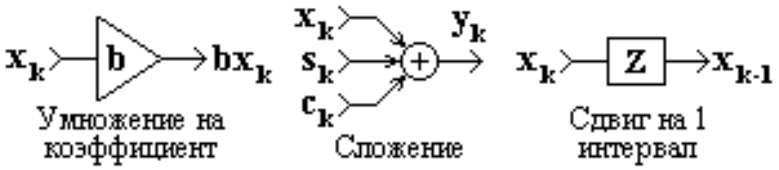


Рисунок 3.1 – Графіка системних операцій

Відзначимо, що операції додавання і множення є лінійними тільки для аналогових і дискретних сигналів. У разі цифрових сигналів вони лінійні щодо самих цифрових сигналів, але якщо останні - результат операції амплітудно-цифрового перетворення, то додавання і множення не може вважатися лінійним абсолютно точно по відношенню до вихідних сигналів.

Для систем, з розмірністю 2 і більше існує також ще одна базова операція, яка називається операцією просторового маскування, яка може розглядатися як узагальнення скалярного множення. Так, для двовимірних систем:

$$z(x,y) = c(x,y)\cdot u(x,y),$$

де u(x, y) - двовимірний вхідний сигнал, с (x, y) - просторова маска постійних (вагових) коефіцієнтів. Просторове маскування являє собою поелементне відтворення значень сигналу з коефіцієнтами маски.

Інваріантність систем до зрушення. Система називається інваріантною до зрушення (інваріантною в часі, а так само по будь-яких інших аргументах), якщо зсув вхідного сигналу по аргументах викликає відповідний зсув вихідного сигналу:

$$s(x,t) = T[a(x,t)], \quad T[a(x-\Delta x,t-\Delta t)] = s(x-\Delta x,t-\Delta t).$$

Лінійність і інваріантність до зрушення є незалежними властивостями систем і не визначають один одного. Так, наприклад, операція квадратування сигналу (зведення в квадрат всіх значень сигналу) інваріантна до зрушення, але нелінійна.

У теорії аналізу і обробки даних основне місце займають системи, лінійні і інваріантні до зсуву (ЛІС - системи). Вони мають досить широкі практичні можливості при відносній простоті математичного апарату. Надалі, якщо це спеціально не обмовляється, будемо мати на увазі саме такі системи.

Перевага, яка віддається ЛІС - системам в методах обробки інформації, базується на можливості розкладання вхідного сигналу будь-який, як завгодно складної форми, на складові найпростіших форм, відгук системи на які відомий і добре вивчений, з подальшим обчисленням вихідного сигналу у вигляді суми відгуків на всі складові вхідного сигналу. Як найпростіші форми розкладання сигналів використовуються, як правило, поодинокі імпульси і гармонійні складові. Перша застосовується при поданні сигналу в динамічній формі і використовує перетворення згортки, друга - частотне подання сигналу і перетворення Фур'є.

Іншою важливою особливістю ЛІС - систем є те, що будь-які їх комбінації також є ЛІС - системами, а будь-яку складну ЛІС - систему можна розкласти на комбінації простих систем. Так, наприклад, при послідовному (каскадному) з'єднанні систем, коли вихідний сигнал однієї системи служить

вхідним сигналом для другої і т.д., утворена система в цілому також є ЛІС - системою, якщо лінійні і інваріантні до зсуву всі системи, які в неї входять , при цьому по відношенню до загальної системної операції перетворення порядок з'єднання входять до неї систем значення не має.

Стаціонарні та нестаціонарні системи. Система вважається стаціонарною і має постійні параметри, якщо її властивості (математичний алгоритм оператора перетворення) в межах заданої точно не залежать від вхідного і вихідного сигналів і не змінюються ні в часі, ні від будь-яких інших зовнішніх факторів. Математично це означає завдання системи рівняннями типу з постійними значеннями коефіцієнтів a_j і b_i і реакція системи на будь-який вплив не залежить від часу (координат) його застосування. В іншому випадку система є нестаціонарною або параметричною (системою зі змінними параметрами). Серед останніх велике значення мають так звані адаптивні системи обробки даних. У цих системах виробляється, наприклад, оцінювання певних параметрів вхідних і вихідних сигналів, за результатами порівняння яких здійснюється підстроювання параметрів перетворення (перехідної характеристики системи) таким чином, щоб забезпечити оптимальні по продуктивності умови обробки сигналів або мінімізувати похибку обробки.

Імпульсний відгук системи. За визначенням, імпульсними характеристиками систем (другий широко використовуваний термін - імпульсний відгук систем) називаються функції h (t) для аналогових і h (kΔt) для цифрових систем, які є реакцією (відгуком) систем на одиничні вхідні сигнали: дельта-функцію Δ (t) для аналогових і імпульс кронекера δ (kΔt) для цифрових систем, що надходять на вхід систем відповідно при t = 0 і k = 0. Ця реакція однозначно визначається оператором перетворення:

$$y(t) = T[\delta(t)] \equiv h(t).$$
$$y(k\Delta t) = T[\delta(k\Delta t)] \equiv h(k\Delta t).$$

Імпульсний відгук аналогової системи, як результат операції над дельта-функцією, в певній мірі є математичною абстракцією ідеального перетворення.

З практичної точки зору під імпульсним відгуком аналогової системи можна розуміти математичне відображення реакції системи на вхідний сигнал довільної форми з площею, яка дорівнює 1, якщо тривалість сигналу дуже мала в порівнянні з часом реакції системи або з періодом її власних коливань. Під часом (довжиною) реакції системи зазвичай розуміють інтервал, на якому значення функції h (t) істотно відрізняються від нуля після припинення дії одиничного сигналу на її вході.

Для цифрових систем імпульсний відгук однозначно визначається реакцією системи на імпульс кронекера δ (kΔt) = 1 при k = 0. Функцію імпульсного відгуку називають також ваговій функцією системи.

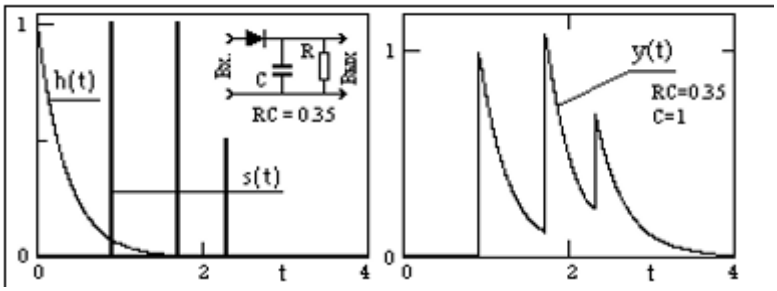


Рисунок 3.2 – Імпульсний відгук системи, вхідний сигнал та реакція системи.

Z-перетворенням (перетворенням Лорана) називають згортання вихідного сигналу, заданого послідовністю дійсних чисел у часовій області, в аналітичну функцію комплексної частоти. Якщо сигнал являє імпульсну характеристику лінійної системи, то коефіцієнти Z-перетворення показують відгук системи на комплексні експоненти $E(n) = z^{-n} = r^{-n}e^{-i\omega n}$, тобто на гармонійні осциляції з різними частотами і швидкостями наростання / загасання.

Z-перетворення, як і багато інтегральних перетворень, може бути як одностороннє, так і двостороннє.

Двостороннє Z-перетворення X (z) дискретного часового сигналу x [n]

задається як:

$$X(z) = Z\{x[n]\} = \sum_{n=-\infty}^{\infty} x[n]z^{-n}.$$

де n - ціле, z - комплексне число.

$z = Ae^{j\varphi}$, де A - амплітуда, а φ - кутова частота (у радіанах на відлік).

У випадках, коли x [n] визначена тільки для $n \geqslant 0$, одностороннє Z-перетворення задається як:

$$X(z) = Z\{x[n]\} = \sum_{n=0}^{\infty} x[n]z^{-n}.$$

ЛЕКЦІЯ 4

Тема: Фур'є аналіз сигналів.

План

1. Амплітудний та фазовий спектри сигналу.
2. Алгоритм виконання перетворення Фур'є.
3. Дійсне дискретне перетворення Фур'є.

Складна періодична функція f (t) визначається сукупністю величин A_n та φ_n Сукупність величин A_n називається амплітудним спектром, а сукупність величин φ_n спектром фаз. Для багатьох застосувань достатньо знати спектр амплітудю Він використовується настільки часто, що коли говорять 'спектр', то розуміють саме амплітудний спектр. В інших випадках роблять відповідні зауваження.

Амплітудний та фазовий спектри зручно зображати у вигляді дискретного графіку. Приклад зображення амплітудного спектру наведеного рисунку 4.1.

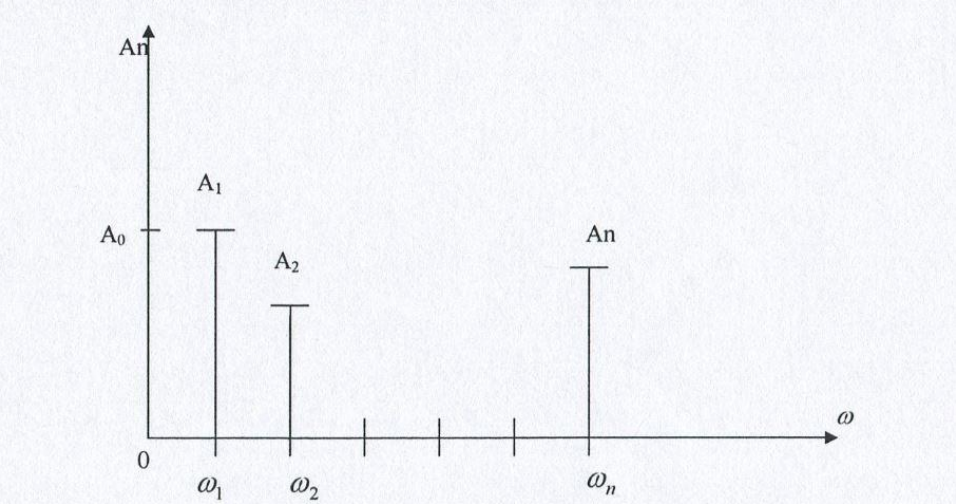


Рисунок 4.1 – Амплітудний спектр сигналу

В ідеалі спектр повинен би був зображатися рядом точок на площині. Проте, так як таке зображення є незручним, то з відповідних точок опускають перпендикуляри на горизонтальну вісь, отримуючи на графіку таким чином множину вертикальних ліній. З огляду на таке зображення спектру, його часто називають лінійчастим. Для періодичних функцій лінійчастий спектр обов'язково є гармонічним.

Амплітудний і фазовий спектри використовуються при розв'язуванні багатьох задач систем управління, зокрема, при визначенні амплітудного і фазового спектрів вхідного та вихідного сигналів системи; при визначенні комплексної передаточної функції системи та її ланок.

Розглянемо спектри неперіодичних функцій. Внаслідок граничного переходу від ряду до інтегралу Фур'є відстані між окремими спектральними лініями необмежено зменшуються, лінії зливаються і замість дискретних точок спектр повинен зображатися неперервною послідовністю точок, тобто неперервною кривою. Такий спектр називається суцільним.

Принципово важливим є факт, що спектральна щільність S (ω одночасно несе в собі інформацію про амплітуди, так і про фази елементарних синусоїд. Спектральну щільність S(ω) функції іноді називають ще комплексним

спектром неперіодичної функції, а модуль цієї величини M (ω) =|S(ω)| спектром.

Рівність Парсеваля. Нервність Бесселя. Для Фур'є має місце рівність Парсеваля:

$$||f(t)||^2 = \int_T^2 f\ (t)dt = \frac{T}{2}\left(\frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2\right)$$

Остання формула має чітку фізичну інтерпритацію: для функції f (t інтеграл $\int_T^2 f$ (t)dt який в прикладних дослідженнях називається енергією функції (сигналу), може бути виражений через коефіцієнти Фур'є цієї функції.

Рівність іноді записують у вигляді

$$\frac{a_0^2}{2} + \sum_n^{\infty} a_n^2 + b_n^2 = \frac{2}{T}\int_T f^2\ (t)dt$$

Якщо в останній формулі взяти скінченну кількість доданків, то отримаємо нерівність:

$$\frac{a_0^2}{2} + \sum_{n=1}^{\infty} a_n^2 + b_n^2 \leq \frac{2}{T}\int_T f^2\ (t)dt$$

яку називають нерівністю Бесселя.

Дискретне перетворення Фур'є. Припустимо тепер, що у нас замість функції неперервного аргументу f (t) задано функцію дискретного аргументу f_k, k= $\overline{0,N-1}$, тобто задані значення функції f (t) для скінченної послідовності значень її аргументу t_k=kh, k=$\overline{0,N-1}$ h=$\frac{T}{N}$ - крок дискретизації. Очевидно, що якщо вихідна функція f (t) – періодична, то, T – період функції, в протилежному випадку T – інтервал, на якому задана функція. Проте при роботі з функціями дискретного аргументу зазвичай оперують номерами відліків і спектральних гармонік без прив'язки до реального масштабу часу і частоти. В такому випадку частоту дискретизації вважають рівною одиниці.

Можна показати, що для дискретної послідовності f_k, k= $\overline{0,N-1}$ справедливе наступне перетворення:

$$f_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{n-1} y_n \exp\left(-i\frac{2\pi nk}{N}\right), k = \overline{0,N-1}$$

де y_n, n= $\overline{0,N-1}$ – коефіцієнти перетворення обчислюються за формулою:

$$y_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{n-1} f_k \exp\left(i - \frac{2\pi nk}{N}\right), k = \overline{0,N-1}$$

Формули описані вище задають пару дискретних перетворень Фур'є: обернене та пряме. В більш загальному випадку функцію дискретного аргументу f_k, k= $\overline{0,N-1}$ можна розглядати як елемент евклідового простору R^N а формулу (3.22) трактувати як розклад в цьому ж просторі за ортогональним базисом дискретних експоненціальних функцій (ДЕФ):

$$def(n,l) = \exp(i\frac{2\pi}{N})nl = \cos\left(\frac{2\pi}{N}nl\right) + i\sin\left(\frac{2\pi}{N}nl\right), n,l = \overline{0,N-1}$$

де n – номер функції в базисі, I – номер відліку самої функції. Отже, множина ДЕФ є ортогональним базисом в R^N причому норма кожної функції з цього базису рівна $\|\mathrm{def}(n,l)\| = \sqrt{N}\forall_n$

При парному N система ДЕФ складається з двох дійсних функцій $def(0,l)$ та $def\left(\frac{N}{2},l\right)$та$\left(\frac{N}{2}-1\right)$ пар комплексно-спряжених функцій. При непарному N система ДЕФ складається тільки з однієї дійсної функції $def(0,l)$ а решта утворюють $\left(\frac{N}{2}-1\right)$ комплексно-спряжених пар. У всіх випадках комплексно-спряжені функції розміщені в системі симетрично, тобто $def(n,l) = \overline{def(N-n,l)}$

Дійсне дискретне перетворення Фур'є. При використанні програмних засобів, що не підтримують комплексних чисел, може бути доцільним використання дійсної форми дискретного перетворення Фур'є. Формулу прямого перетворення можна записати у вигляді:

$$f_k = \frac{2}{\sqrt{N}} \sum_{n=0}^{N-1} y_n^c \cos(\frac{2\pi\, nk}{N} + y_n^s \sin(\frac{2\pi\, nk}{N}))k = \overline{0, N-1}$$

Коефіцієнти цього перетворення визначаються рівнями:

$$y_n^c = \frac{2}{\sqrt{N}} \sum_{k=0}^{N-1} f_k \ \cos\frac{2\pi\, nk}{N}$$

$$y_n^s = \frac{2}{\sqrt{N}} \sum_{k=0}^{N-1} f_k \ \sin\frac{2\pi\, nk}{N}$$

Очевидно, що дійсне перетворення Фур'є вимагає обчислення вдвічі більшої кількості коефіцієнтів. Зменшення кількості коефіцієнтів можна сподіватись за рахунок симетрії тригонометричних функцій. Врахувавши періодичність, парність та непарність тригонометричних функцій можна показати, що $y_n^c = y_{N-n}^c$ та $y_n^s = y_{N-n}^s$ Таким чином, якщо кількість коефіцієнтів непарна, то незалежним є лише $\frac{N-1}{2} + 1$ коефіцієнтів, а у випадку парної кількості незалежним будуть $\frac{N}{2} + 1$ коефіцієнтів.

ЛЕКЦІЯ 6

Тема: Згортка та кореляція сигналів, обчислення та їх застосування.

План

1. Поняття згортки сигналів.
2. Алгоритм виконання згортки та застосування.
3. Кореляція сигналів та її застосування.

Довільний сигнал на вході системи з використанням виразів розкладання сигналу за одиничними імпульсами Дірака може бути представлений у вигляді послідовно-котельної лінійної комбінації зважених одиничних імпульсів:

$$y(t) = T[s(t)] = T[\int_{-\infty}^{\infty} s(\tau)\ \delta(t-\tau)\ d\tau]. \qquad (5.1)$$

На підставі принципу суперпозиції лінійний оператор T може бути внесений під знак інтеграла, тому що останній являє собою граничне значення суми. При цьому операція перетворення діє тільки по змінної t. Звідси випливає:

$$y(t) = \int_{-\infty}^{\infty} s(\tau)\, T[\delta(t-\tau)]\, d\tau = \int_{-\infty}^{\infty} s(\tau)\, h(t-\tau)\, d\tau. \qquad (5.2)$$

Цей варіант динамічного подання сигналу і є інтегралом згортки (конволюції) вхідного сигналу з імпульсною характеристикою системи. Заміною змінних $t-\tau = \tau$ можна переконатися в тому, що згортка комутативна:

$$\int_{-\infty}^{\infty} s(\tau)\, h(t-\tau)\, d\tau \equiv \int_{-\infty}^{\infty} h(\tau)\, s(t-\tau)\, d\tau. \qquad (5.2')$$

Функція $h(\tau)$ називається *ядром згортки (kernel)* або імпульсною характеристикою линів-ної системи.

Сенс інтеграла згортки полягає в тому, що вхідний сигнал подається зімкнутої по-послідовністю наступних один за одним коротких імпульсів, площі яких дорівнюють значення сигналу в моменти їхнього проходження при тривалості імпульсів, що прагне до нульової. Така послідовність імпульсів умовно може розглядатися у вигляді послідовності дельта-функцій з площами, рівними площам відповідних імпульсів. Реакція системи (5.2) знаходиться як сума реакцій на кожен імпульс, що становить вхідний вплив.

Аналогічно, для дискретних сигналів, де значення t, як правило, приймається рівним 1, а індекси k і n виконують роль номерів відліків числових рядів:

$$y(k) = \sum_{n=-\infty}^{\infty} h(n)\, s(k-n). \qquad (5.2'')$$

У цифрових методах обробки сигналів функцію h (n) зазвичай називають оператором згортки, а його розмір за кількістю відліків - вікном оператора згортки.

Вирази (5.2) мають спеціальну форму спрощеної математичної записи в символічному вигляді:

$$y(t) = s(t-\tau) * h(\tau) \equiv s(t) * h(t).$$

Порівнянням виразів (5.2 'і 5.2' ') з виразами неважко переконатися в їх повній ідентичності, за винятком нижньої межі інтегрування (підсумовування). Це і зрозуміло, так як вираження були отримані при розгляді реальної фізичної системи, що працює в реальному масштабі часу, імпульсний відгук яких є одностороннім (дорівнює нулю при $\tau<0$). Для таких систем інтегрування (і підсумовування) від $-\infty$ до 0 не має сенсу. Крім того, в реальних фізичних системах імпульсний відгук, як правило, відрізняється від нуля тільки на певному інтервалі, і, відповідно, межі інтегрування (підсумовування) у виразах (5.2) обмежуються значеннями, на яких функції $h(\tau)$ и $h(n)$ існує або має значущі значення.

Сигнали, оброблювані на комп'ютері, мають кінцеву тривалість. Допустимо, сигнал s (k) відмінний від нуля тільки на відрізку від 0 до K включно ("має довжину K + 1"). Нехай вікно оператора згортки h (n) відмінне від нуля на відрізку від - N до N (2N + 1 відлік). При підстановці цих сигналів в рівняння згортки, ми отримаємо сигнал y (k), який відрізняється від нуля на відрізку від - N до K + N включно. Таким чином, довжина вихідного сигналу дорівнює 2N + K-1, тобто сумі довжин вихідного сигналу і ядра згортки мінус один.

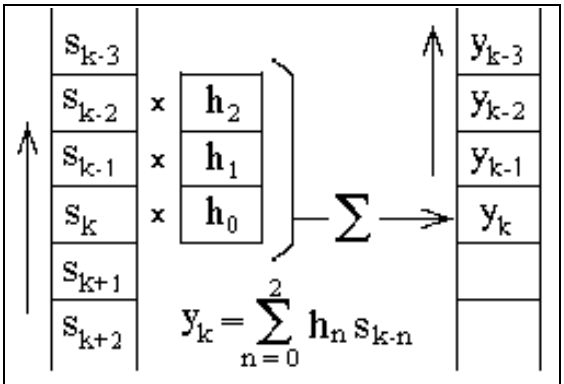


Рис. 5.1 – Графічне зображення техніки виконання операції згортки

Техніка виконання згортки для цифрової системи приведена на рис. 3.3.1. Для обчислення згортки масив однією з функцій (s_k - вхідного або згортається сигналу) розташовується по ходу зростання номерів. Масив другий функції (більш короткого оператора h_n), будується паралельно першому масиву в зворотному порядку (в режимі зворотного часу, по ходу зменшення номерів першого масиву). Для обчислення y_k значення h_0 розташовується проти s_k, все значення s_{k-n}

перемножуються з розташованими проти них значеннями h_n і підсумовуються. Результати підсумовування є вихідним значенням функції y_k, після чого оператор h_n зсувається на один номер k вперед (або функція s_k зсувається йому назустріч) і обчислення повторюється для номера k + 1 і т.д.

Для обчислення аналогової згортки за виразом (5.2') функція імпульсного відгуку реверсують по своїй координаті, тобто будується в режимі зворотного часу, і рухається відносно функції вхідного сигналу в бік зростання значень t. У кожен поточний момент часу значення обох функцій перемножуються, а твір інтегрується в межах вікна імпульсного відгуку. Отриманий результат відноситься до тієї координатної точці, проти якої знаходиться значення імпульсного відгуку h (0).

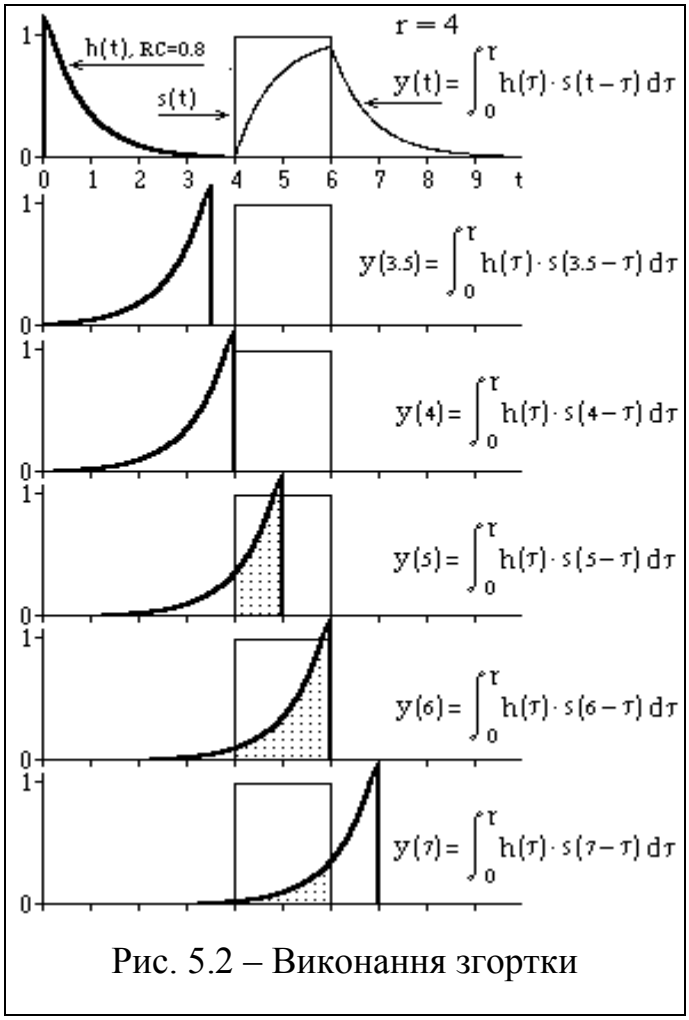


Рис. 5.2 – Виконання згортки

На рис. 3.3.2. наведено приклад виконання згортки прямокутного імпульсу з імпульсним відгуком RC-ланцюга, площа якого нормована до 1. Якщо площа імпульсного відгуку h (t) дорівнює 1, то площа вихідного сигналу згортки завжди повинна дорівнювати площі вхідного сигналу, що можна бачити на верхньому графіку малюнка , при цьому одномасштабне порівняння вхідного і вихідного сигналів наочно демонструє характер перетворення сигналу в даній системі. На наступних графіках малюнка демонструється обчислення результатів згортки в ряді послідовних точок t_i = {3.5, 4, 5, 6, 7} тимчасової осі. В силу негативного знака в аргументах функції $s(t-\tau)$ інтегрування твори $h(\tau)s(t-\tau)$ виконується назад по

часу і може обмежуватися тільки певною довжиною значущих значень імпульсного відгуку (яка в даному випадку встановлена рівній r = 4), а результат відноситься до початкової точки h (0) імпульсного відгуку. Так як вхідний сигнал, розглянутий на малюнку, є прямокутним імпульс з амплітудою 1, то інтеграл згортки в кожній поточній точці розрахунку дорівнює площі імпульсного відгуку в межах кордонів вхідного прямокутного імпульсу (заповнене точками).

Властивості згортки. Для згортки характерні наступні властивості:

1. Дистрибутивність: $h(t) * [a(t)+b(t)] = h(t) * a(t)+h(t) * b(t)$.

2. Комутативність: $h(t) * a(t) * b(t) = a(t) * b(t) * h(t)$.

3. Асоціативність: $[a(t) * b(t)] * h(t) = h(t) * a(t) * b(t)$.

Перетворення згортки однозначно визначає вихідний сигнал y (t) для встановленого значення вхідного сигналу s (t) при відомому значенні функції імпульсного відгуку системи h (t). Зворотній завдання деконволюції - визначення функції s (t) за функціями y (t) і h (t), відноситься до розряду некоректних, і має рішення тільки при цілком певних умовах. Це пояснюється тим, що згортка може істотно змінити частотний спектр сигналу y (t) щодо s (t) і відновлення функції s (t) стає неможливим, якщо певні частоти її спектру в сигналі y (t) повністю втрачені.

Будь-яка практична система повинна бути *стійкою*, тобто для сигналів, кінцевих по енергії або середньої потужності, вихідні сигнали також повинні бути кінцевими за цими параметрами. Стійкість забезпечується при виконанні умови абсолютної інтегрованості імпульсного відгуку системи:

$$\int_{-\infty}^{\infty} |h(t)|\, dt < \infty.$$

Для систем з m входами і n виходами аналогічно визначаються парціальні імпульсні відгуки hij (t), i = {1,2, ..., n}, j = {1,2, ..., m}, кожним з яких відображається сигнал на i-му виході при надходженні сигналу δ(t) на j-й вхід. Повна сукупність імпульсних відгуків утворює матрицю:

$$\overline{\mathbf{h}}(\tau) = \begin{vmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,m} \\ h_{2,1} & h_{2,2} & \dots & h_{2,m} \\ . & . & . & . \\ h_{n,1} & h_{n,2} & \dots & h_{n,m} \end{vmatrix},$$

а вираз згортки набуває вигляду:

$$\overline{\mathbf{Y}}(t) = \int_{-\infty}^{\infty} \overline{\mathbf{h}}(\tau)\, \overline{\mathbf{S}}(t-\tau)\, d\tau.$$

Тут (і в подальшому тексті) жирним шрифтом з "кришкою" виділяються векторні величини.

Системи згортки. Згортка виконується системою (фізичним або програмним пристроєм). Фізичні системи, що працюють в реальному часі, обчислюють поточне значення вихідного сигналу по всьому минулим значенням вхідного сигналу, і не можуть мати в своєму розпорядженні майбутніх значень вхідного сигналу. Оператори таких систем є одне-сторонніми (каузальних). Вищенаведена, нормована до 1 по площі, функція RC-ланцюга $h(t) = (1/RC)\cdot\exp(-t/RC)$, прийнята в якості системного оператора на рис. 5.2, є саме таким одностороннім каузальних оператором. При порівнянні вихідного сигналу такої системи з вхідним неважко помітити, що вихідний сигнал зсувається щодо вхідного сигналу. Для каузальних систем такий "зсув по фазі" існує завжди і не може бути вико-укладений (сигнал на виході системи не може бути раніше сигналу на її вході).

Вхідним сигналом програмних систем є сигнал в цілому, записаний в пам'ять обчислювального пристрою. При обробці таких даних у розпорядженні системи при обчисленні будь поточної точки вихідного сигналу є як "минулі" для даної точки, так і "майбутні" значення вхідного сигналу. Це дозволяє створювати системи без зсуву фази вихідного сигналу щодо вхідного. Для створення таких систем без зсуву фази може використовуватися два способи:

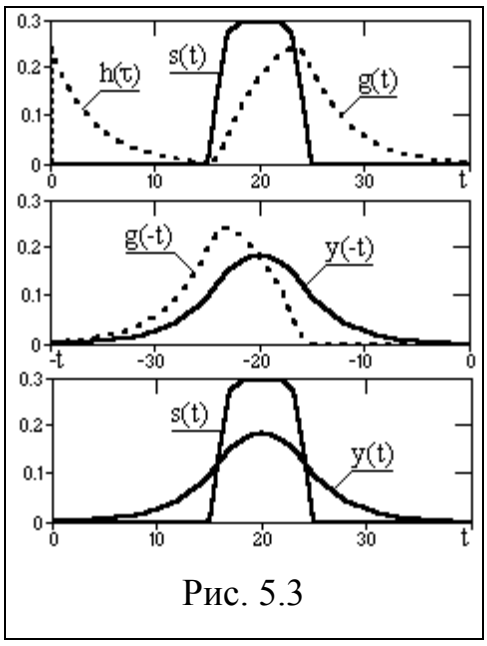


Рис. 5.3

1. Перший спосіб ілюструє рис. 5.3. Здається система з одностороннім каузальних оператором $h(\tau)$. Вхідний сигнал s (t) пропускається через систему в звичайному порядку, і виконується згортка g (t) = $h(\tau)$* s (t). Потім вихідний сигнал g (t) реверсують (g (t) => g (-t), кінець сигналу стає його початком в порядку зростання t) і повторно пропускається через систему, тобто виконується згортка y (-t) = $h(\tau)$ * g (-t). Отриманий сигнал знову реверсують y (-t) => y (t), і результат є остаточним вихідним сигналом y (t) системи.

Три останні операції (реверс g (t) $\Rightarrow$ згортка з $h(\tau)$ $\Rightarrow$ реверс вихідного сигналу) еквівалентні-валентні згортку сигналу g (t) з реверсувати відгуком системи $h(-\tau)$, і зсув по фазі при пакунку реверсувати сигналу компенсує зсув по фазі сигналу, отриманий при першій пакунку. Загальний результат операції $y(t) = h(\tau)*h(-\tau)*s(t)$ не має зсуву по фазі вихідного сигналу щодо вхідного. Таку операцію доводиться виконувати для виключення зсуву фази при застосуванні рекурсивних систем, які завжди є односторонніми.

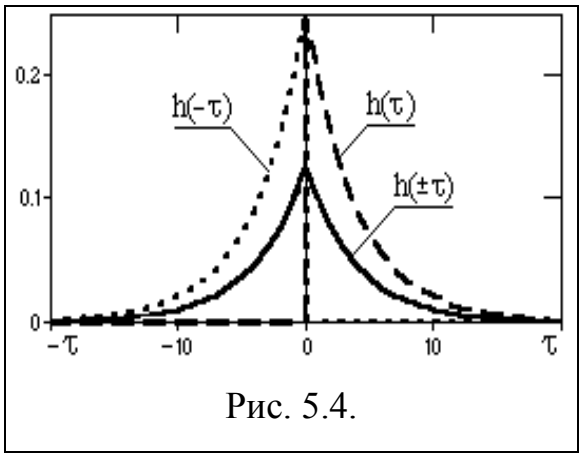


Рис. 5.4.

2. Другий спосіб. Вихідний результат y (t) = $h(\tau)*h(-\tau)*s(t)$ попередньої операції дозволяє, використовуючи властивість комутативності згортки, спочатку виконати згортку $h(\tau)*h(-\tau) = h(\pm\tau)$ і отримати один системний оператор $h(\pm\tau)$ (див. рис. 5.4), що забезпечує згортку без зсуву фази. Цей системний оператор є двостороннім і симетричним щодо $\tau = 0$. Але використання його можливо тільки для попередньо записаних сигналів, тому що при виконанні згортки у $y(t)= h(\pm\tau)*s(t-\tau)$ для від'ємних значень τ

потрібні "майбутні" значення вхідного сигналу $s(t+\tau)$. Результат згортки з симетричним оператором повністю аналогічний першої операції (сигнал y (t) на рис. 5.3).

На практиці цілком природним є розрахунок безпосередньо симетричних двосторонніх операторів під необхідні завдання обробки числових даних (сигналів, зареєстрованих в дискретної числової формі).

Початкові умови згортки. У початковий момент згортки, при обчисленні значень $y(t_i)$ для значень $t_i < \tau_{max}$ оператора $h(\tau)$, функція оператора, побудована в режимі зворотного часу, при $\tau > t_i$ "зависає" для значень $t_i-\tau$ проти відсутніх значень вхідної функції.

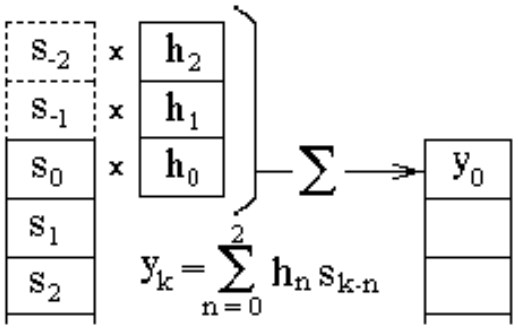


Рис. 5.5 – Техніка виконання згортки

Приклад такого зависання оператора дискретної згортки проти неіснуючих відліків s_{-1} и s_{-2} вхідного масиву даних при обчисленні відліку y0 наведено на рис. 5.5. Звісно виключають або завданням початкових умов - додаткових відліків, найчастіше нульових або рівних першому відліку вхідний функції, або початком згортки з відліку вхідний функції $k_i = n_{max}$ з відповідним скороченням інтервалу вихідний функції на інтервал завдання системного оператора. Для симетричних операторів зі значеннями -n (вперед за часом) такий же момент настає і в кінці вхідного масиву, і вимагає завдання кінцевих умов або скорочення розміру вихідного сигналу.

ЛЕКЦІЯ 6

Тема: Цифрова фільтрація сигналів**.**

План

1. Поняття фільтрації сигналу.
2. Побудова фільтрів.
3. Основні типи фільтрів.

Фільтрація є однією із найпоширеніших операцій обробки сигналів. Мета фільтрації полягає в усуненні завад, що містяться у сигналі, або у виділенні окремих складових сигналу, які відповідають тим чи іншим властивостям досліджуваного процесу.

У техніці цифрового оброблення сигналів все частіше застосовують цифрові фільтри (ЦФ). За допомогою останніх у сигналі виділяють, заглушуються (перериваються) або послаблюються певні частоти з метою покращання відношення сигнал/шум. Крім цього, ЦФ широко використовують як структурні елементи пристроїв та систем оброблення сигналів. Зважаючи на це, ЦФ мають велике практичне значення.

Сучасні активні фільтри та цифрові фільтри - це пристрої, які використовують у різних сферах техніки. В останні десятиліття інтенсивно розвиваються також методи фільтрації, специфічні саме для вимірювальної та обчислювальної техніки. Ці методи, основані на реалізації спеціальних вагових функцій. Отримані завдяки їм фільтри дуже близькі за властивостями до ЦФ, але можуть бути використані як в цифровій, так і в аналоговій частині пристрою вимірювання. В останньому випадку можуть бути отримані деякі вигоди порівняно з цифровою фільтрацією: відсутність похибки квантування, зменшення частини вимірювального каналу, для якого необхідно передбачувати розширений діапазон змін сигналу, що враховує можливу заваду.

Оскільки цифрова фільтрація сигналу полягає у цілеспрямованій зміні співвідношення між різними компонентами спектра сигналу, то фільтрувальні пристрої необхідно розглядати з позиції теорії спектрів.

Фільтрацію сигналу, тобто зміну його спектра, звичайно здійснюють з метою

збільшення відношення сигнал - шум, зменшення впливу завад або виділення (підсилення) якої-небудь корисної якості сигналу. Наприклад, для вимірювання сигналів, отриманих від термопар, найчастіше доводиться застосовувати фільтри, що послаблюють завади мережі. Вихідний корисний сигнал термопар становить, як правило, декілька мілівольт і завада від силової мережі може бути порівнянна з корисним сигналом або навіть перевищувати його. Другим прикладом фільтрації є фільтрування сигналу, одержаного від давана моменту, який розвиває двигун деякого транспортного засобу. Виділяючи за допомогою фільтра постійну складову цього сигналу, отримують інформацію про середню потужність двигуна. Якщо виділити і проаналізувати високочастотні складові сигналу, то можна зробити висновок про якість системи регулювання, а також про вібрацію, зумовлену роботою двигуна.

Фільтри можуть класифікуватися за різними особливостями. Здебільшого для поділу фільтрів використовують такі ознаки.

Перша ознака - вид вхідного і вихідного сигналу фільтра. Якщо ці сигнали аналогові, то фільтри називаються аналоговими, якщо ж сигнали представлені цифровим кодом, то фільтри належать до цифрових. Можливі і проміжні випадки: аналогово-цифровий фільтр (вхід аналоговий, вихід цифровий) та цифроаналоговий (вхід цифровий, вихід аналоговий).

Друга ознака - вид частотної характеристики. За цією ознакою фільтри поділяються на такі групи: фільтри низьких частот (ФНЧ) - пропускають низькочастотні складові спектра і затримують високочастотні; фільтри високих частот (ФВЧ) - пропускають тільки високочастотні складові; смугові фільтри (СФ) - пропускають складові сигналу лише в певній смузі частот; смугозагороджувальні або режекторні фільтри (РФ) - пропускають усі складові сигналу, окрім тих, частоти яких входять в певну смугу; всепропускні фільтри (ВФ) - пропускають усі без винятку складові сигналу, але змінюють фазові співвідношення між ними. Графіки амплітудно-частотних характеристик (АЧХ) згаданих фільтрів наведено на рис. 6.1 (а-д).

Крім перерахованих, основних за цією ознакою груп фільтрів, існують ще

і інші їх різновидності. Наприклад, резонансний фільтр є окремим випадком СФ фільтра, але з дуже вузькою смугою пропускання (штрихована АЧХ на рис. 6.1, ***в).*** Фільтр-корок на визначену частоту - це РФ фільтр с вузькою смугою загороджування (штрихова АЧХ на рис. 6.1, г). Гребінчастий фільтр - фільтр, який має декілька смуг пропускання (рис. 6.1, ***е).*** У назву фільтра входить звичайно та частотна смуга, яку фільтр пропускає. Так, фільтр нижніх частот - це фільтр, що пропускає нижні частоти сигналу.

Необхідно зазначити, що під час аналізу і синтезу фільтрів фільтр ФНЧ звичайно використовують як базовий. Саме ФНЧ, як правило, розглядають в різноманітних публікаціях, для нього розробляють методики синтезу. Інші види фільтрів можуть бути побудовані на основі ФНЧ. Так, якщо з повного сигналу вирахувати вихідний сигнал ФНЧ, то в результаті отримаємо ФВЧ (рис. 6.2, ***а),*** СФ можна побудувати, якщо ввімкнути паралельно ФНЧ1 і ФНЧ2 з різними частотами зрізу (рис. 6.2, ***б).*** Для побудови РФ достатньо з'єднати послідовно відповідно розраховані ФНЧ та ФВЧ фільтри.

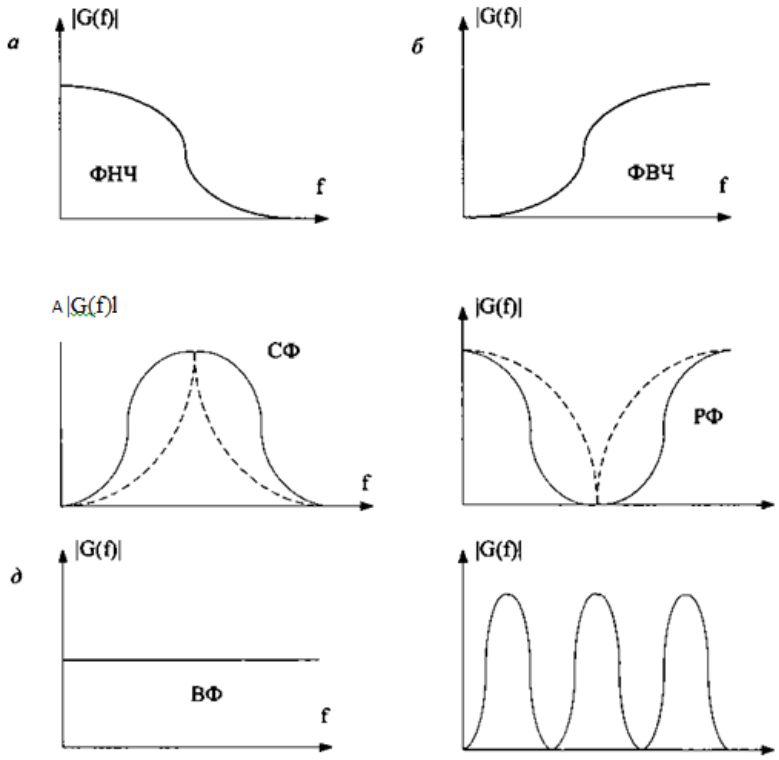


Рисунок 6.1 – Амплітудно-частотні характеристики фільтрів

За третьою ознакою розрізняють різні типи фільтрів за їх імпульсною

характеристикою. Неперервний фільтр - це фільтр з неперервною імпульсною характеристикою (ІХ), дискретний фільтр - це фільтр, ІХ якого подана набором 5-імпульсів. Нарешті, імпульсний фільтр має ІХ, що складається з послідовності однакових за формою імпульсів скінченної довжини різної амплітуди. Існують фільтри, у класифікації яких за цією ознакою виникають деякі ускладнення, але такі фільтри на практиці трапляються рідко.

Четверта ознака, за якою класифікують фільтри, - це протяжність імпульсної характеристики. Якщо ІХ фінітна, тобто має межу в часі, то такі фільтри називають фільтрами зі скінченним імпульсним відгуком, або, коротко, СІВ-фільтрами *(рос.* КИХ-фильтри - фильтри с конечной импульсной характеристикой; *англ.* FIR-filter - finite impulse response filter). Якщо ІХ, хоча і загасає в часі, але має теоретично необмежену в часі тривалість, то такі фільтри називають НІВ-фільтрами, тобто фільтрами з нескінченним імпульсним відгуком *(рос*. БИХ-фильтрьі - фильтрьі с бесконечной импульсной характеристикой; англ. IIR-filter - interminable impulse response filter).

Для прикладу на рис. 6.2 наведена ІХ двох видів фільтрів: імпульсного СІВ-фільтра (рис. 6.2, *а)* та дискретного НІВ-фільтра (рис. 6.2, *б).* Теорія фільтрації сигналів та методи побудови фільтрів сьогодні добре розвинені.

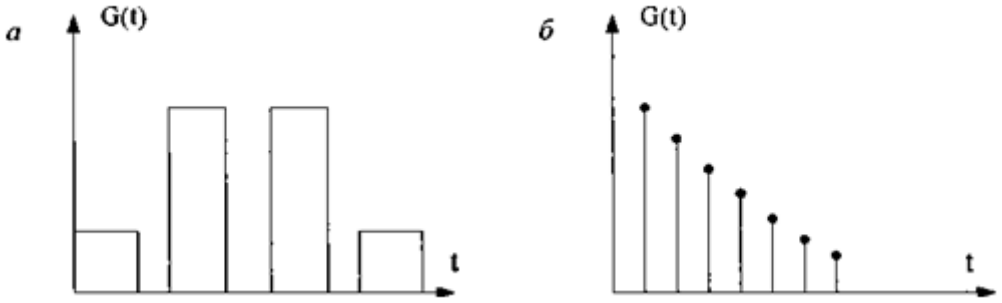


Рисунок 6.2 – Приклади імпульсних характеристик імпульсного (а) та дискретного (б) фільтрів

Розроблена велика кількість різноманітних фільтрів та методів їхнього проектування. Типовою завадою, що впливає на пристрої перетворення і оброблення, є періодична завада, викликана впливом мережевої напруги.

Спектр такої завади має лінійний характер: потужність завади сконцентрована на основній частоті *fn1* (частоті мережі) і кратних їй частотах вищих гармонік *fПк =kfП1>* де *k≥1* - ціле число. Для усунення такої завади необхідно, щоби значення амплітудно-частотної характеристики вимірювального пристрою було, близьким до нуля, у вузьких областях, що розміщені поблизу частот основної та вищої гармонік мережевої напруги.

У діапазоні частот корисного сигналу АЧХ фільтра повинна мати постійне, бажано велике значення, а фазочастотна характеристика (ФЧХ) фільтра в цьому діапазоні повинна мати, за змогою, лінійний характер.

Якщо до швидкодії пристрою вимірювання ставлять порівняно високі вимоги, то перевагу віддають СІВ-фільтрам, оскільки НІВ-фільтри за стрибка вхідного сигналу мають теоретично нескінченний час встановлення вихідного сигналу.

Алгоритм обчислення. У виразах ЦФ використовують вагові коефіцієнти $a_к$. Вхідний дискретизований сигнал фільтра формується з неперервного вхідного сигналу JC(/) за допомогою оцінки його вибірок й квантування [1]

$$x[(n-k)T_a] = x_{n-k} \tag{6.1}$$

де n - номер вибірки; к- індекс вагового коефіцієнта, який вказує порядок фільтра; T_a - період дискретизації сигналу.

На виході фільтра формується вихідний сигнал

$$y(nT_a) = y_n \tag{6.2}$$

який пов’язаний з вхідними даними та ваговими коефіцієнтами фільтра такою залежністю.

$$y_n = \sum_{k=0}^{N} a_k * x_{n-k} \tag{6.3}$$

Вираз (6.3) є рівнянням нерекурсивного несиметричного ЦФ N-то порядку.

Вагові коефіцієнти фільтра можуть індексуватися симетрично, тобто змінюватися від к = -N до к = N . У цьому випадку рівняння фільтра визначається так

$$y_n = \sum_{k=-N}^{N} a_k * x_{n-k} \quad (6.4)$$

Згідно з (6.4) нерекурсивний ЦФ є симетричним і має порядок 2N . Отже, кількість зміщених у часі вхідних сигналів характеризує ступінь або порядок фільтра. Для симетричного нерекурсивного фільтра шостого порядку рівняння матиме такий вигляд

$$y_n = a_{-3}x_{n+3} + a_{-2}x_{n+2} + a_{-1}x_{n+1} + a_0x_n + a_1x_{n-1} + a_2x_{n-2} + a_3x_{n-3}$$

Особливістю ЦФ є те, що вихідний сигнал у таких фільтрах можна отримати лише після обчислення усіх вибірок вхідного сигналу x_{n-k}. Обчислення y_n пов'язане з деякою затримкою, яка є особливо критичною для швидкодіючих систем з великою кількістю вибірок та використаними в них ЦФ високих порядків.

Зазначимо, що вагові коефіцієнти фільтра $a_к$ індексують в інший спосіб, оскільки вхідні вибірки x_n-$_k$ не перемножуються з ваговими коефіцієнтами фільтра, а над їхньою послідовністю виконують операцію згортки ($a_к$ • x_n-$_k$) *.

З наведеного можна зробити висновок, що до нерекурсивних фільтрів належать фільтри, вихідний сигнал яких залежить лише від значень вхідного сигналу, на відміну від рекурсивних фільтрів, в яких вихідний сигнал залежить від попередніх значень вихідного сигналу. Нерекурсивний фільтр за відсутності зворотного зв'язку не збуджується і завжди є стабільним.

Як випливає з означення згортки, якщо нерекурсивний фільтр стимулювати одиничним імпульсним сигналом $x_n = 1$ при $n = 0$ і $x_n = 0$ при $n \neq 0$, то вагові коефіцієнти $a_к$ фільтра з'являться як вихідний сигнал y_n . Після (2N +1)

тактів вплив вхідного сигналу на вихідний зникає. Саме тому нерекурсивний фільтр називається ще фільтром зі скінченним імпульсним відгуком. Отже, нерекурсивний фільтр має скінченну пам'ять і у нього можуть бути дуже добрі фазові характеристики. Однак в нерекурсивному фільтрі переважно необхідна велика кількість членів, щоб отримати порівняно гострий зріз частотної характеристики. Натомість рекурсивний фільтр має безмежну пам'ять і здебільшого потребує меншої кількості членів. За такої економії членів фазові характеристики, як правило, є порівняно гіршими. Фільтри з

гострим зрізом частотної характеристики значно легше проектувати, використовуючи рекурсивну структуру. Щоб отримати передаточну дискретну функцію ЦФ, на його вхід подають дискретний періодичний сигнал такого вигляду $x[\{n-k)T_a] = e^{jw(n-k)}$

ЛЕКЦІЯ 7

Тема: Основи цифрового представлення зображень.

План

1. Зображення як багатовимірні сигнали.
2. Застосування цифрової обробки зображень.
3. Цифрове представлення зображення.

В попередніх лекціях була представлена інформація про сигнали, які були функціями одної змінної, як правило, часу *s(t)* . Такі сигнали називається одновимірними. До цього класу сигналів відносяться всі сигнали, які вимірюються з використанням одного вимірювального каналу. Багатовимірними називаються сигнали, значення яких залежить не від одної змінної, а від кількох. Найбільш розповсюдженим прикладом багатовимірного сигналу є зображення. В такому сигналі значення сигналу є функцією двох координат в просторі I_1 *(x, y)*. Наприклад, якщо мова іде про сірошкальне зображення, то значення сигналу є яскравістю точки в просторі. Як правило, яскравість – число від 0 до 1, де 0 відповідає точці чорного кольору (відсутність яскравості), а 1 – відповідає точці білого кольору (максимальна яскравість). Проміжні значення між 0 та 1 говорять про насиченість кольору: 0.5 відповідає сірому кольору, 0.2 – темно-сірому, 0.8 – світло-сірому і т.д.

Всі методи аналізу та обробки одновимірних сигналів, яки вивчалися до цього, можуть бути застосовані до багатовимірних сигналів, але з урахуванням залежності значення сигналів від більше ніж одної змінної.

Отримання, обробка та аналіз зображень розповсюджені в багатьох

галузях науки і техніки, і набувають все більш широкого розвитку.

Застосування аналізу зображень:

– аналіз та обробка двовимірних зображень (фотографій) при стисненні, кодуванні, розпізнаванні об’єктів, покращенні якості зображень;

– аналіз та обробка рухомих зображень (відеопослідовностей) при стисненні відео, автоматичному контролі якості продукції, стеженні за транспортними засобами;

– біомедичні застосування (аналіз та розпізнавання зображень, які отримані в результаті УЗД, МРТ, КТ та ін. технік медичної інтроскопії).

Основні напрями досліджень з обробки та аналізу зображень:

– вивчення сприйняття зображень людиною та технічними системами;

– розробка та удосконалення техніки реєстрації зображень;

– розробка методів покращення якості зображень (підвищення контрастності, чіткості, позбавлення від шумів);

– розробка ефективних та швидких методів стиснення зображень;

– сегментація зображень (розділення зображення на складові частини);

– морфологічний аналіз та розпізнавання об’єктів на зображеннях.

Основними методами отримання зображень є такі:

–отримання зображень у видимому діапазоні довжин хвиль (фотографія, відеозйомка);

– термографія – в результаті вимірювання температурного поля отримують інфрачервоні зображення, які показують розподіл тепла по поверхні об’єкта. Зокрема, в медицині такі зображення використовуються при діагностуванні запалень, захворювань судин та онкологічних захворювань шкіри. В електроніці методом термографії можна визначити області локального розігріву плати, мікросхеми або електронного чіпа;

– ультразвукова інтроскопія – при цьому система п’єзовипромінювачів генерує і передає в об’єкт ультразвукові імпульси. Ці звукові хвилі проходять скрізь об’єкт і на границях між областями із різним акустичним опором

відбувається часткове віддзеркалення (частина хвилі повертається в бік зонду), частина затухає, а частина проходе далі, поки не досягнуть іншої границі і т.д. Віддзеркалені хвилі реєструються приймачем та передаються в комп'ютер, де на основі часу приходу кожного віддзеркаленого імпульсу та швидкості розповсюдження звуку в об'єкті розраховують відстань від зонду до границі розподілу. Обчислені відстані та інтенсивності будуються у вигляді зображення на дисплеї. Часто такі зображення будуються в реальному часі та, наприклад, дозволяють спостерігати за роботою серця чи активністю плоду;

– рентгенівська комп'ютерна томографія – при цьому дослідженні набір детекторів випромінювання та джерело рентгенівських променів розташовуються по кільцю, всередину якого поміщають об'єкт (наприклад, пацієнта). Кільце обертається навколо пацієнта, і випромінювання, яке пройшло через пацієнта, реєструється детекторами, що розташовуються на протилежному боці кільця. За допомогою математичного перетворення Радонаможна побудувати зображення перерізу тіла пацієнта в площині кільця використовуючи інтенсивності випромінювання, яке пройшло через пацієнта. Цей процес відбувається неперервно при повороті кільця та при зміщенні кільця вздовж пацієнта (по спіралі). При зміщенні пацієнта вздовж вісі кільця отримують багато перерізів, які у сукупності утворюють тривимірне зображення;

– ядерно-магнітний резонанс або магнітно-резонансна томографія (ЯМР або МРТ) – при цьому дослідженні пацієнта поміщають в сильне постійне магнітне поле. Через його тіло пропускають радіохвилі у вигляді коротких імпульсів, на які реагують спіни електронів в атомах водню. У відповідь на радіоімпульси атоми водню в тканинах пацієнта випромінюють радіосигнали, сила та місце походження яких визначаються комп'ютерною системою обробки. В результаті будується тривимірне зображення тіла пацієнта;

– гамма-ізотопна візуалізація (однофотонна емісійна комп'ютерна томографія, ОФЕКТ) – при цьому пацієнту вводять радіоактивний ізотоп, розпад якого супроводжується гамма-випроміненням. Це випромінювання

реєструється детекторами, підсилюється та використовується для побудови тривимірних карт та будується зображення розподілу ізотопу в тілі. В основному цей метод використовують для знаходження ділянок патологій кісток, при інфекційних та онкологічних захворюваннях.

– позитронно-емісійна томографія – при ній пацієнту вводять радіоактивний ізотоп, розпад якого супроводжується позитронним випромінюванням. При зустрічі позитрона з електроном вони анігілюють з утворенням двох гамма-квантів. Результуюче гамма-випромінення реєструється детекторами та формується томографічне зображення.

З точки зору теорії сигналів, *зображення* – це функція двох змінних $I_1(x, y)$, де x та y – координати на площині.

Значення $I_1(x, y)$ в довільній точці площини називається *яскравістю* (інтенсивністю) або *рівнем сірого* для зображення. Як і у випадку одновимірних сигналів, $I_1(x, y)$ називається неперервним зображенням в разі, коли координати x та y можуть приймати довільні значення, а дискретним – в разі, коли x та y визначені лише для деякої множини значень. Можна провести дискретизацію зображень, якщо ввести поняття частоти дискретизації F_s або відповідного кроку дискретизації T_s. В загальному випадку, частота дискретизації може бути різною по різним координатам (x, y): F_{sx} , F_{sy} (з відповідними кроками дискретизації T_{sx} , T_{sy}). Як правило, дискретизація зображень проводиться не в часовій області, а в просторовій, і частота дискретизації показує, скільки відліків зображення буде виміряне на одиницю довжини по кожній координаті. Розмірність частоти дискретизації в цьому випадку буде $\frac{1}{\text{м}}$, а крок дискретизації буде виражений в одиницях довжини.

Дискретне зображення задане своїми значеннями $I_1[n, m]$ в деяких точках площини. В реальних застосуваннях дискретне зображення містить скінченну кількість елементів, які називаються пікселями (pixel, від англійського picture

element – елемент зображення). В цьому випадку дискретна функція I_1 [n, m] має лише скінченну кількість значень:

$$I_1[n,m], n = \overline{0, N-1},\ m = \overline{0, M-1}.$$

Таке зображення можна представити у вигляді матриці, кожний елемент якої дорівнює значенню яскравості відповідного пікселя:

$$I1[n,m] = \begin{bmatrix} I1[0,0] & I1[0,1] & \cdots & I1[0,M-1] \\ I1[1,0] & I1[1,1] & \cdots & I1[1,M-1] \\ \vdots & & \ddots & \vdots \\ I1[N-1,0] & I1[N-1,1] & \cdots & I1[N-1,M-1] \end{bmatrix}.$$

Нумерація пікселів подвійна, перший індекс відповідає номеру рядка, другий – номеру стовпця. Лівий верхній піксель зображення має номер $[0,0]$. Всі методи аналізу та обробки зображень зводяться до математичних операцій над неперервною функцією *f (x, y)*, яка описує неперервне зображення, або над пікселями $I_1[n,m]$ (у випадку дискретного зображення).

Якщо значення яскравості кожного пікселя $I_1[n,m]$ може приймати лише значення 0 або 1, то таке зображення називається бінарним, і пікселі можуть бути тільки двох кольорів (як правило, чорного та білого). У випадку, коли значення $I_1[n,m]$ може бути дійсним числом від 0 до 1, то говорять, що таке зображення є сірошкальним (grayscale image), або зображенням в градаціях сірого.

Зображення, які одержуються в інформаційних системах, дуже рідко мають цифрову форму. Тому їх перетворення у цей вигляд є обов'язковою операцією, якщо передбачається використовувати цифрову обробку, передавання, зберігання. Як і у випадку одновимірних сигналів, це перетворення поєднує дві процедури: дискретизацію, яка полягає у заміні

неперервного кадру дискретним, і квантування, яка виконує заміну неперервної множини значень яскравості множиною квантованих значень. За цифрового представлення кожному із квантованих значень яскравості ставиться у відповідність двійкове число.

Заміну неперервного зображення дискретним можна виконати різними способами. Можна, наприклад, вибрати яку-небудь систему ортогональних функцій і, вирахувавши коефіцієнти представлення зображення за цією схемою (за цим базисом), замінити ним зображення. Різноманітність базисів уможливлює утворення різних дискретних представлень неперервного зображення. Однак найвживанішою є періодична дискретизація, а саме дискретизація з прямокутним растром. Такий спосіб дискретизації може розглядатись як один із варіантів використання ортогонального базису, у якому використовуються як елементи зсунуті δ –функції.

Основне питання, яке виникає під час заміни неперервного зображення дискретним, полягає у визначенні умов, за яких така заміна є повноцінною, тобто не супроводжується втратою інформації, яка міститься в неперервному сигналі. Втрати відсутні, якщо, маючи в розпорядженні дискретний сигнал, можна відновити неперервний. З математичного погляду питання полягає у відновленні неперервного сигналу в двовимірних проміжках між вузлами, у яких його значення відомі, тобто в здійсненні двовимірної інтерполяції.

Під час цифрової обробки зображень динамічний діапазон значень яскравості [dl, dL + 1] (припускається, що він обмежений) ділять на інтервали. Якщо яскравість x відліку належить інтервалу(d_j,d_j+1), тобто

(d_j <x≤ d_j+1) вихідний відлік заміняється на рівень квантування r_j, де d_j, j=1 L+1 - пороги квантування. Задача побудови кваигувача полягає у визначенні значень порогів d_j та рівнів r_j. Найпростіший спосіб розв’язання цієї задачі полягає у розбиті динамічного діапазону на однакові інтервали. Однак таке рішення не буде найкращим, оскільки, у разі наявності групування значень яскравості відліків зображення в певних областях доцільно квантувати нерівномірно. Це дасть змогу зменшити помилку квантування. Отже, задачу

побудови квантувача можна сформулювати як задачу знаходження оптимальних значень d_j та r_j, що задовольняли б певний критерій оптимізації. Як правило, за фіксованої кількості рівнів квантувач оптимізуються за критерієм мінімальної середньоквадратичної похибки.

У системах цифрової обробки зображень прагнуть зменшити кількість рівнів та порогів квантування, оскільки від їхньої кількості залежить довжина двійкового кодового слова, яким представляються проквантовані відліки в ЕОМ. Однак за порівняно невеликої кількості рівнів L на проквантованому зображенні з'являються хибні контури, які виникають внаслідок стрибкоподібної зміни яскравості проквантованого зображення і особливо помітні на пологих ділянках її зміни. Вони значно погіршують візуальну якість зображення, оскільки зір людини особливо чутливий якраз до контурів.

Щоб уникнути нерівномірного квантування, яке не можна реалізуваїи за допомогою стандартного АЦП. використовують нелінійне перетворення. Здійснюють нелінійне перетворення відліку вихідного зображення, щоб густина розподілу ймовірностей перетворених відліків була рівномірною, тобто виконується процедура еквалізації. Потім відліки квантують з рівномірним кроком та виконують зворотне нелінійне перетворення. Для руйнування хибних контурів перед рівномірним квантуванням до відліків яскравості додають шум з рівномірною густиною розподілу ймовірностей, який переводить одні відліки зображення на рівень вище, а інші - на рівень нижче. Дисперсія шуму, який додається, повинна бути невеликою, щоб не призвести до спотворень, які сприймаються як “сніг" на зображенні, і водночас достатньою для руйнування хибних контурів.

Визначення реального часу (рис. 7.1) залежить від конкретної задачі і пов'язане з обсягом обчислень алгоритму, точністю обчислень та частотою (періодом) дискретизації. Нехай T - період дискретизації, τ_a - час виконання алгоритму. Цифрова система працюватиме в режимі реального часу, якщо час виконання алгоритму τ_a не перевищуватиме періоду дискретизації. Це означає, що залишається іще деякий запас часу, який називають часом очікування.

Знайти час виконання алгоритму можна, знаючи час виконання елементарної одноциклової команди τ_k, який називається командним циклом, та кількість командних циклів N_a які необхідні для виконання алгоритму. Тоді згідно з [64]:

$$\tau_a = \tau_k \cdot N_a, t_{оч} = T - \tau_a$$

Однак виконання основної умови роботи в режимі реального часу (час виконання алгоритму, яким би складним він не був, не повинен перевищувати періоду дискретизації) може бути забезпечено лише у випадку, коли тактова частота обчислювача істотно перевищуватиме частоту дискретизації.

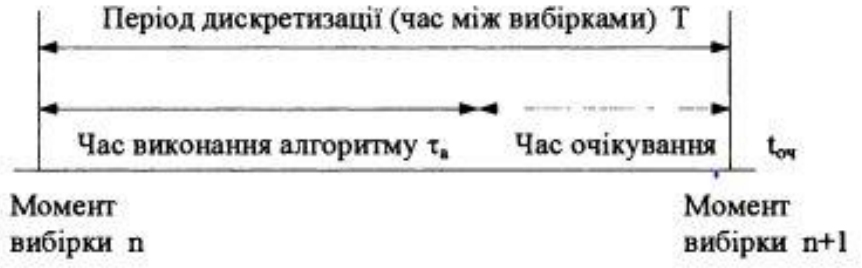


Рисунок 7.1 – Визначення реального часу

Тактова частота показує, як швидко процесор виконує найпростішу операцію, наприклад пересилання в регістрі з розряду в розряд. Відношення тактової частоти до частоти дискретизації належить до найважливіших характеристик, які визначають, як буде реалізована система. Це відношення частково визначає кількість апаратних засобів, необхідних для реалізації алгоритму заданої складності в режимі реального часу. Якщо відношення указаних частот зменшується, то кількість і складність апаратних засобів, необхідних для реалізації алгоритму, збільшується.

Час виконання алгоритму τ_a залежить не лише від процедури, яка представляє цей алгоритм, а й від способу реалізації алгоритму. Можливі три способи реалізації алгоритмів цифрової обробки сигналів: апаратний, програмний і апаратно-програмний.

Апаратна реалізація передбачає використання різноманітних функціональних блоків: регістрів, суматорів, шифраторів і дешифраторів, лічильників, ліній затримки, пристроїв пам’яті, помножувачів, логічних елементів, інтегральних та великих інтегральних схем, програмованих логічних

матриць тощо. Сукупність функціональних блоків та зв’язків між ними визначає алгоритм, який реалізується.

Перевага апаратної реалізації полягає у дуже високій швидкодії, що дає змогу обробляти сигнали за частоти дискретизації у десятки мегагерц. Це досягається використанням функціональних блоків на базі ТТ-логіки, розпаралелюванням операцій та вузькою спрямованістю (спеціалізацією) створюваних пристроїв.

З іншого боку, апаратна реалізація орієнтована на розв'язання вузькоспеціаліюваних задач, передбачає створення систем із жорсткою логікою, коли будь-яка зміна алгоритму вимагає зміни структури пристрою, тобто введення додаткових функціональних блоків. Крім цього, апаратна реалізація приводить до значного споживання енергії та необхідності організації тепловідведення. Усе це разом визначає високу вартість апаратної реалізації, причому проектування, виготовлення та налаштування виявляються дуже трудомісткими за великих часових затрат.

Програмна реалізація передбачає представлення алгоритму у вигляді програми, яку послідовно від команди до команди виконує один або одночасно декілька незалежних блоків. Програма повинна бути написана мовою програмування. яка відповідає конкретному операційному блоку. Так, для персонального комп’ютера це буде будь-яка з мов високого рівня (Pascal, C++, Java тощо), а для мікропроцесорного комплекту або цифрового процесора - відповідна мова асемблера.

Усі команди, які утворюють програму, повинні бути подані у вигляді, зрозумілому процесору та безпосередньо ним сприйманому. Машинна мова це команди у вигляді комбінацій нулів та одиниць. Тому для полегшення підготовки та налагодження програм для процесора створюються спеціальні засоби, коли кожній машинній команді або групі команд ставиться у відповідність скорочений символьний запис так, щоб він відображав їх смисловий зміст. Така мова програмування називається мовою асемблера.

Переклад програми на машинну мову називається трансляцією і виконується автоматично за допомогою спеціальної програми - асемблера (від англ, assembler - складальник). Розроблені та набувають усе більшого значення спеціальні дуже ефективні програми-трансляторн з мов високого рівня на мови асемблерів. Ці програми називаються компіляторами.

Перевагами програмної реалізації є:

- незмінна структура системи при різних алгоритмах та сферах використаній;
- хороша гнучкість, яка дає змогу достатньо легко змінювати алгоритми роботи системи за рахунок корекції або зміни програми;
- істотне прискорення, полегшення та здешевлення проектування, виготовлення та налаштування системи, оскільки замість приладу розробляють програму.

Недоліком програмної реалізації с порівняно низька швидкодія через послідовне виконання операцій програми в одному процесорі. І. як би не збільшували швидкодію виконання команд, вона залишатиметься нижчою від продуктивності відповідного пристрою, реалізованого апаратно. Це породжує проблему забезпечення роботи в режимі реального часу, яка передбачає дві обставини:

час обробки одного відліку або групи відліків сигналу повинен бути меншим від допустимого часу затримки. Контроль за виконанням цієї умови повинен здійснюватись як під час написання програми, так і під час налаштування;

цикл роботи програми та момент надходження відліків вхідного сигналу повинні бути чітко погоджені за часом, тобто початок обробки чергового відліку повинен збігатись або слідувати за надходженням цього відліку.

Апаратно-програмна реалізація передбачає апаратне виконання однієї частини функцій системи цифрової обробки сигналів (аналогово-цифрове та цифроаналогове перетворення, множення, прийом-передача даних тощо), та

програмне виконання другої частини функцій. Апаратно-програмна реалізація поєднує переваги апаратної та програмної реалізації. Розумне поєднання програмних та апаратних засобів дає змогу знизити вимоги до обчислювальних можливостей елементної бази та спростити реалізацію системи цифрової обробки сигналів загалом, для налаштування якої необхідні спеціальні засоби.

ЛЕКЦІЯ 8

Тема: Просторові методи покращення зображень.

План

1. Просторова область обробки зображень.
2. Градаційні перетворення зображень.
3. Еквалізація гістограми.

Термін «просторова область» відноситься до площини, в якій задане зображення. При просторовій обробці всі методи маніпулюють безпосередньо величинами яскравостей пікселів зображення. Їх можна описати загальним виразом

$$I2[n,m] = T\,(I1[n, m]),$$

де T – деяке правило, яке ставить у відповідність кожному пікселю початкового зображення $I_1[n,m]$ відповідний піксель результуючого зображення $I_2[n,m]$, який отриманий в результаті обробки. Для отримання $I_2[n,m]$ може використовуватися не тільки значення пікселя $I_1[n,m]$ але також і інші пікселі в деякому околі $I_1[n \pm \Delta_n, m \pm \Delta_m]$. Як правило, це прямокутний або квадратний окіл.

При обробці зображення центр цього околу зміщують від пікселя до пікселя і проводять розрахунки для кожного центрального пікселя.

Якщо для того, щоб отримати яскравість пікселя обробленого зображення, використовується яскравість лише одного пікселя початкового

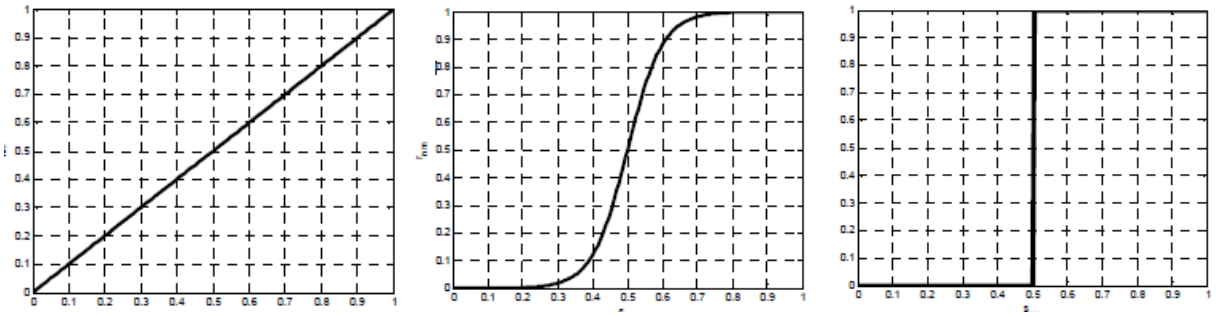
зображення, то говорять про градаційні перетворення. Іншими словами, обробка проводиться в околі розміром 1x1 піксель. Якщо яскравості пікселів початкового та результуючого зображення рівні відповідно $s_{nm} = I_1[n, m]$, та $r_{nm} = I_2[n, m]$,, то можна записати, що

$$r_{mn} = T(s_{nm})$$

Тут T описує залежність, яка пов'язує яскравості кожного пікселя початкового та результуючого зображення.

Нехай ми маємо сірошкальне зображення, яскравості пікселів якого лежать в межах від 0 до 1. На рисунку 8.1 наведені приклади деякихзалежностей T. Перетворення яскравості згідно з кривою на рис. 8.1а не приведе до зміни яскравості пікселя, оскільки крива T переводить значення яскравості s_{nm} в таке саме значення r_{nm} . Якщо змінювати яскравість за кривою на рис. 8.1б, то зміни будуть помітні. Видно, що яскравості s_{nm} в околі нуля (дуже темні пікселі) будуть перетворені в пікселі r_{nm} нульової яскравості. Так само, дуже світлі пікселі (в з яскравостями близькими до 1) стануть білими (будуть мати яскравість 1). Достатньо вузький діапазон яскравостей початкового зображення s_{nm} від 0.4 до 0.6 (сірі пікселі) перетвориться на дуже широкий діапазон яскравостей в результуючому зображенні від 0.1 до 0.9, тобто розтягнеться. Це буде приводити до того, що деталі, які були на початковому зображенні сирими та слабко розрізнялися по кольору, будуть більш чітко видимі на зображенні після обробки.

При перетворенні за залежністю з рисунку 8.1в всі пікселі з яскравістю від 0 до 0.5 будуть мати яскравість 0, а пікселі з яскравістю більше 0.5 матимуть яскравість 1. Отже, зображення з сірошкального стане чорно-білим: пікселі темніші за сірий колір стануть чорними, а світліше за сірий – білими.



а) б) в)

Рисунок 8.1 – Приклади перетворень *T* яскравості пікселів при градаційній обробці

Широковживаним видом градаційних перетворень є степеневі перетворення. При цьому яскравості пікселів результуючого та початкового зображення пов’язані за таким законом:

$$r_{nm} = c\,(s_{nm} + \varepsilon)^{\gamma},$$

де c, ε , γ – деякі додатні константи.

Часто використовується більш проста залежність степеневого перетворення:

$$r_{nm} = c s_{nm}^{\gamma} .$$

Графіки для залежностей при різних показниках γ подані на рисунку 8.2.

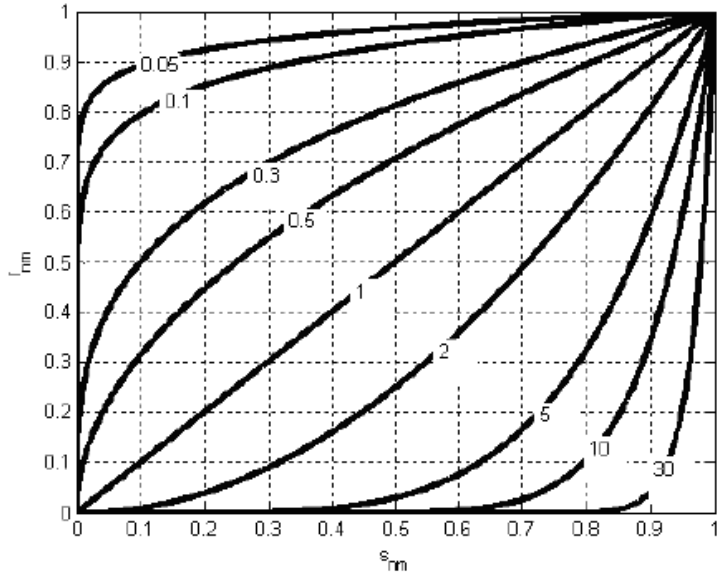


Рисунок 8.2 – Залежності між яскравостями пікселів вхідного та вихідного зображення при гамма-корекції для різних значень γ

Видно, що значеннях γ менше за одиницю, криві таких степеневих залежностей відображають вузький діапазон малих значень яскравостей пікселів вхідного зображення у широкий діапазон яскравостей пікселів результуючого зображення. Якщо значення показника степеню γ більше за одиницю, то відбувається протилежний ефект: вузький діапазон великих яскравостей відображається у широкий діапазон яскравостей пікселів результуючого зображення. Ця процедура називається *гамма-корекцією*.

Еквалізація гістограми. Гістограмою дискретного зображення називається дискретна функція $H(b_k) = \frac{N_k}{N}$, де b є k - им рівнем яскравості пікселя, N_k – кількість пікселів, які мають яскравість b_k, а N - кількість пікселів у всьому зображенні. Значення $H(b_k)$ є оцінкою імовірності появи пікселя яскравості b_k в зображенні.

Еквалізацію (лінеаризацію) гістограми проводять в тому випадку, коли в зображенні є багато пікселів зі схожими яскравостями, і мало пікселів з іншими яскравостями. На гістограмі ми будемо бачити, що на деяких проміжках яскравостей згруповано багато пікселів, в той час як деякі проміжки яскравостей майже не зайняті. При цьому деталі зображення, які зображені цими кольорами, складно розрізнити. Натомість існують такі проміжки яскравості, пікселів з якими взагалі немає на зображенні. Ці вільні проміжки яскравості можна «зайняти» для покращення якості зображення. Для цього роблять еквалізацію гістограм.

Якщо маємо піксель початкового зображення з яскравістю b_k, яка є k - им рівнем яскравості на гістограмі ($k = 0 \ldots N - 1$) то яскравість відповідного пікселя результуючого зображення буде розраховуватися

$$r_k = \sum_{p=0}^{k} H(b_p) = \sum_{p=0}^{k} \frac{N_p}{N}.$$

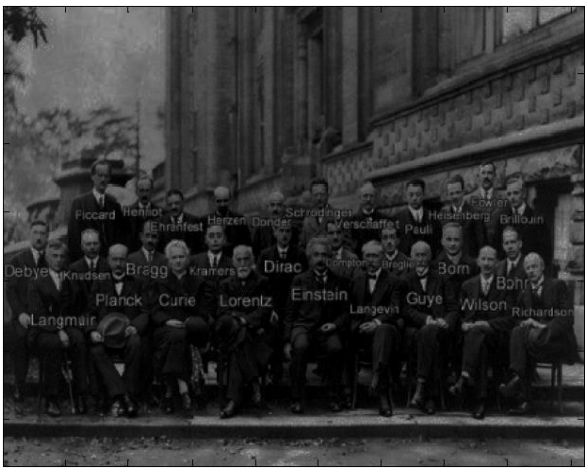
В результаті еквалізації гістограми яскравості пікселів на ній будуть розподілені рівномірно по всій шкалі яскравостей.

Наприклад, на рис. 8.3а наведене зображення, яке виглядає дуже темним.

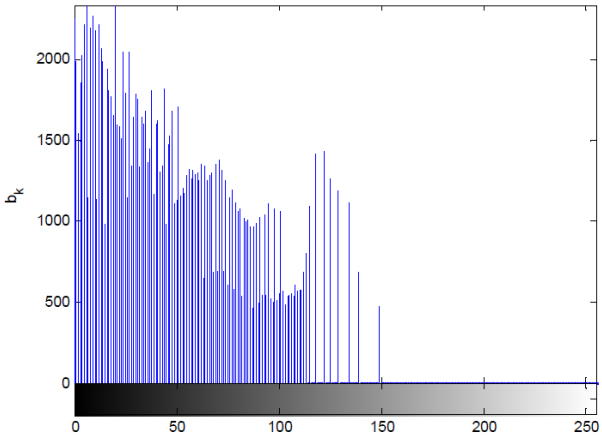
Дрібні деталі предметів та людей на ньому розрізнити складно, оскільки вони зображені схожими темними кольорами, які мало відрізняються один від одного. Гістограма цього зображення наведена на рис. 8.3б. На ній видно, що багато пікселів знаходяться в лівій частині шкали кольорів, що відповідає темним кольорам. Водночас, права частина шкали майже не зайнята, тобто світлих пікселів на зображенні немає. Цей вільний проміжок гістограми можна використати, щоб перенести туди яскравості деяких пікселів. Якщо гістограму цього зображення «розтягнути» на весь доступний діапазон яскравостей, то пікселі, які раніше мали дуже схожі кольори (їх яскравості знаходились близько на шкалі яскравостей), будуть віддалені один від одного на більшу величину яскравості.

Якщо подивитись на зображення, видно, що діапазон яскравостей пікселів, які присутні на зображенні, розширився: на зображенні тепер є і темні, і світлі пікселі (рис. 8.4а). Тепер стало легше розрізнити деталі зображення, оскільки вони зображені більш контрастно. На гістограмі видно, що з зображенні присутні пікселі всіх яскравостей, і весь діапазон яскравостей тепер зайнятий.

Перевагою еквалізації гістограм є те, що цей метод легко автоматизується і не вимагає задавання ніяких додаткових параметрів для отримання покращеного зображення. Розрахунки для еквалізації гістограм такою достатньо нескладні.

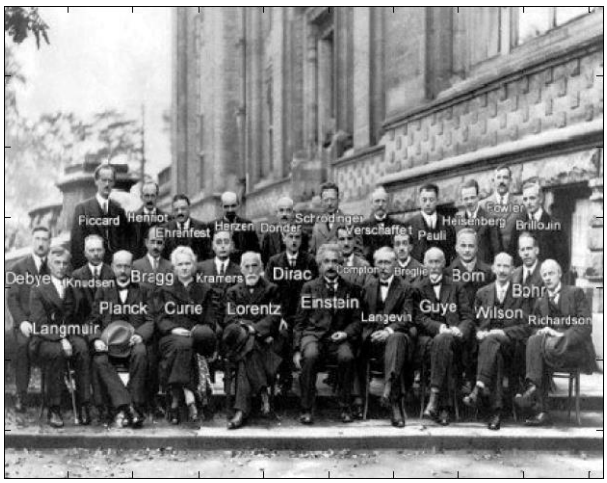


а)

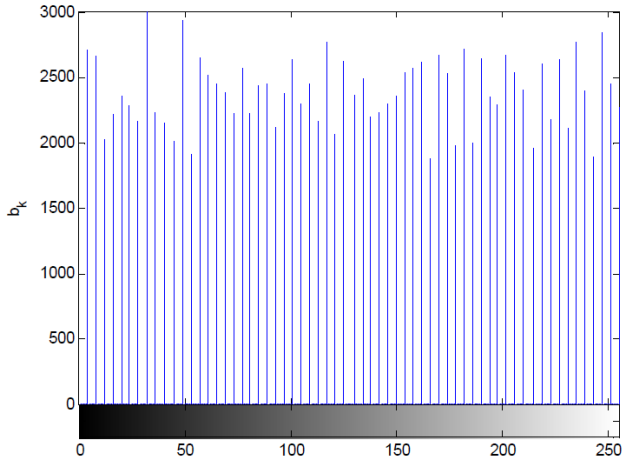


б)

Рисунок 8.3 – Початкове сірошкальне зображення (а) та його гістограма (б). градації сірого кольору кодуються числами від 0 до 255, оскільки для запису кольору пікселя використовується 8 біт



а)



б)

Рисунок 8.4 – Сірошкальне зображення (а) та його гістограма (б) після еквалізації

Просторова фільтрація зображень – це різновид обробки зображень в просторовій області, на відміну від класичного поняття «фільтрації», яка виконується із залученням спектральних характеристик фільтрів та спектрів зображень. При просторовій фільтрації для розрахунку яскравості пікселя результуючого зображення використовуються яскравості пікселів в деякому околі початкового зображення.

Схематично процес просторової фільтрації проілюстрований на рис. 7.5. Обробка виконується послідовно для кожного пікселя зображення. В початковому зображені обирається окіл розміром 3×3 пікселя з центром в деякому пікселі $I_1[n,m]$. На основі значень яскравості пікселя $I_1[n,m]$та його околу буде розрахована яскравість одного пікселя результуючого зображення $I_2[n,m]$. Для цього обирають «маску» коефіцієнтів, яка містить числа, на які мають смисл коефіцієнтів, з якими яскравість кожного пікселя з околу на початковому зображенні буде використана для отримання яскравості пікселя в результуючому зображенні. Розрахунок яскравості ведеться за формулою:

$$I_2[n,m] = c_{-1,-1}\, I_1[n-1, m-1] + c_{-1,0}\, I_1[n-1, m] + c_{-1,1}\, I_1[n-1, m+1] +$$

$$+ c_{0,-1}\, I_1[\text{n, m - 1}] + c_{0,0}\, I_1[\text{n, m}] + c_{0,1}\, I_1[\text{n, m + 1}] +$$
$$+ c_{1,-1}\, I_1[\text{n +1, m - 1}] + c_{1,0}\, I_1[\text{n + 1, m}] + c_{1,1}\, I_1[\text{n + 1, m + 1}] .$$

Видно, що маска центрується в пікселі з номером [*n*,*m*], і яскравість цього пікселя множиться на коефіцієнт $c_{0,0}$, що знаходиться в центрі маски.

При використанні маски розміром *A* x *B* (*A* і *B* повинні бути непарними) загальний вираз для отримання яскравості кожного пікселя результуючого зображення має вигляд:

$$I_2[n,m] = \frac{\sum_{p=-A}^{A} \sum_{s=-B}^{B} c_{p,s}\, I_1[\text{n + p,m + s}]}{\sum_{p=-A}^{A} \sum_{s=-B}^{B} c_{p,s}} . \qquad (8.1)$$

Константа *C* в попередньому вирази є нормувальним множником. Вона дорівнює сумі всіх коефіцієнтів маски, тому при діленні на неї сума коефіцієнтів при всіх яскравостях пікселів дорівнює одиниці. Це приводить до того, що яскравість пікселя відфільтрованого зображення буде не більше, ніж максимальна припустима для даного зображення яскравість. Нормувальний множник застосовується в тому випадку, якщо сума коефіцієнтів маски не дорівнює нулю.

При просторовій фільтрації з допомогою масок, вираз (8.1) повинен бути застосований по черзі для кожного пікселя початкового зображення. Треба мати на увазі, що при проходженні маски по краю зображення деякі пікселі можуть бути відсутні. Тому розрахунок результуючих пікселів повинен виконуватися за особливою процедурою для уникнення крайових ефектів. При цьому зображення або дзеркально продовжується, або добудовується деякий окіл по краю зображення (наприклад, він заповнюється пікселями з нульовими або середніми яскравостями). Також можна не проводити розрахунок яскравостей для таких пікселів, але в цьому випадку розмір отриманого після фільтрації зображення буде меншим на два стовпця та два рядка.

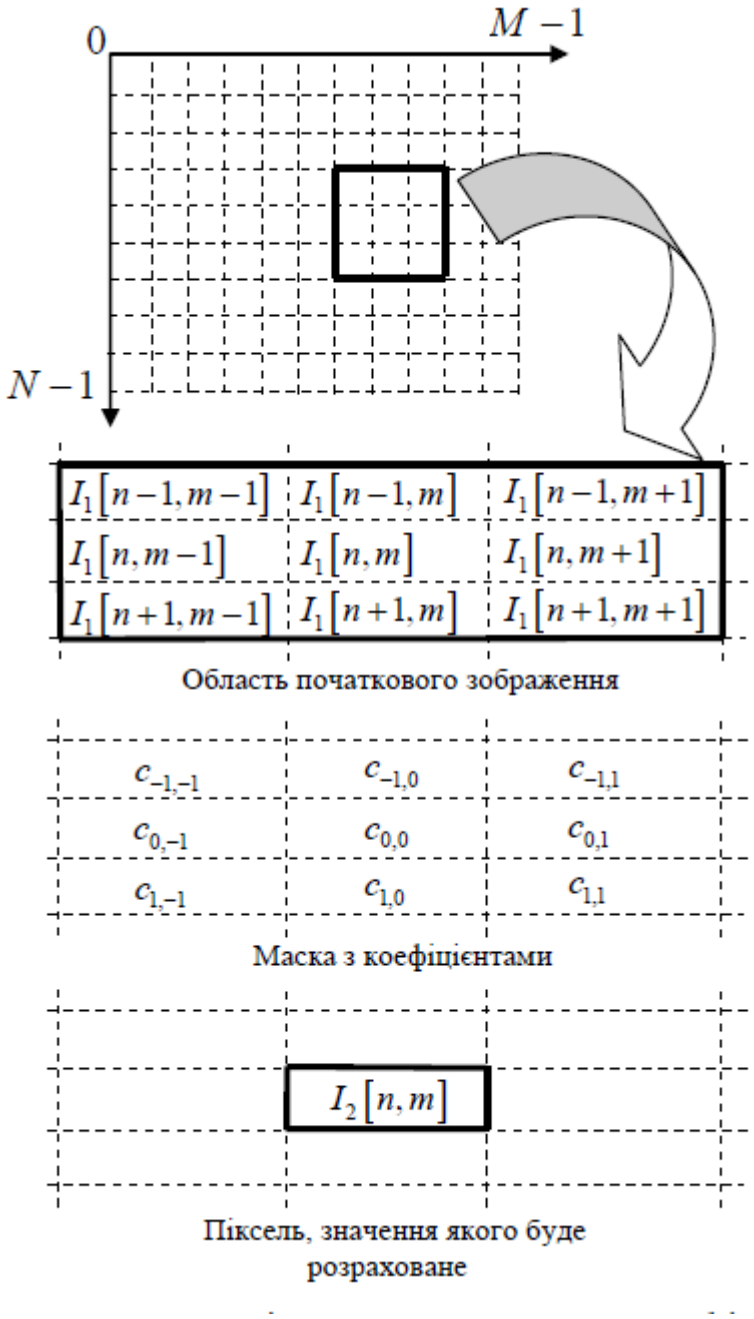


Рисунок 8.5 – Ілюстрація процесу просторової фільтрації

На рис. 8.6 наведені приклади масок, які часто використовуються на практиці. При використанні маски з рис. 8.5а виконується згладжування зображення, оскільки центральний піксель результуючого зображення буде мати яскравість, яка є середньою для околу відповідного пікселя на початковому зображенні. Маска на рис. 8.5б реалізує оператор Лапласа для зображення (взяття другої часткової похідної по двом напрямкам). Її застосування підкреслює розриви рівнів яскравостей на зображенні, і в

результаті підвищується різкість наявних на зображенні границь. Послідовне використання двох масок з рис. 8.5в та 8.5г еквівалентне розрахунку градієнта зображення; застосування цих масок ще називають оператором Собеля. Розрахунок градієнтів використовують для покращення видимості контурів на зображенні.

$$\frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$$

а)

$$\begin{bmatrix}1 & 1 & 1\\ 1 & -8 & 1\\ 1 & 1 & 1\end{bmatrix}$$

б)

$$\begin{bmatrix}-1 & -2 & -1\\ 0 & 0 & 0\\ 1 & 2 & 1\end{bmatrix}$$

в)

$$\begin{bmatrix}-1 & 0 & 1\\ -2 & 0 & 2\\ -1 & 0 & 1\end{bmatrix}$$

г)

Рисунок 8.6 – Приклади масок: а) згладжувальна; б) підвищення різкості; в) і г) оператор Собеля

При медіанній фільтрації для отримання яскравості пікселя результуючого зображення використовується окіл відповідного пікселя початкового зображення. Із значень яскравостей пікселів околу отримують послідовність, в якій вони розташовані від найменшої до найбільшої. Для цієї послідовності знаходиться медіана, тобто визначається, який піксель знаходиться тому місці в послідовності, яке відповідає половині її довжини. Наприклад, якщо в околі 9 пікселів, то медіаною буде той піксель, який іде п’ятим у ранжованому ряді. Яскравість цього пікселя буде значенням яскравості пікселя відфільтрованого зображення.

ЛЕКЦІЯ 9

Тема: Частотні методи покращення зображень.

План

1. Двовимірне перетворення Фурє.

2. Фільтрація зображення.

Аналогічно до того, як для одновимірного сигналу в часовій області можна знайти його спектральне представлення за Фурьє, можна знайти спектр зображення.

Якщо маємо неперервне зображення $I_1(x, y)$, то його спектр за Фурьє буде визначатися:

$$F(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} I_1(x, y)\, e^{-j(ux+vy)}\, dxdy, \qquad (9.1)$$

де u,v – просторові частоти, $\frac{1}{\text{м}}$.

Обернене перетворення Фур'є для зображення:

$$I_1(x, y) = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v)\, e^{j(ux+vy)}\, dudv, \qquad (9.2)$$

Якщо зображення дискретне розмірністю N x M , то аналоги формул будуть мати вигляд:

$$F[k, p] = \frac{1}{NM} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} I_1[n, m] e^{-j2\pi\left(\frac{kn}{N}+\frac{pm}{M}\right)}, \qquad (9.3)$$

$$I_1[n, m] = \sum_{k=0}^{N-1} \sum_{p=0}^{M-1} I_1[n, m] e^{-j2\pi\left(\frac{kn}{N}+\frac{pm}{M}\right)}. \qquad (9.4)$$

Тут k, p – номери гармонічних спектральних складових зображення, $k = \overline{0, N-1}$, $p = \overline{0, M-1}$.

Як і для одновимірного випадку, спектр дискретного зображення є періодичною функцією частот k, p з періодами, які дорівнюють розмірностям зображення: $F[k \pm qN, p + cM] = F[k, p]$ для цілих q,c .

Аналогічно до одновимірного випадку, для комплексного спектру зображення можна отримати амплітудний та фазовй спектри:

$$|F[u, v]| = \sqrt{Re(F(u, v))^2 + Im(F(u, v))^2},$$

$$\varphi(u, v) = \arg\big(F(u, v)\big) = arctg\frac{Im(F(u,v))}{Re(F(u,v))}.$$

Фільтрація зображень в частотній області полягає у наступному. В двовимірній частотній області задається комплексна частотна характеристика фільтра $H(u,v)$. Як і для одновимірного випадку, можна створити фільтри

нижніх (ФНЧ), верхніх (ФВЧ) частот, смугові (СФ) та загороджувальні (ЗФ) фільтри. Фільтрація зображення еквівалентна проходженню його через відповідний лінійний фільтр. В частотній області для того, щоб отримати спектр результуючого зображення $H(u,v)$, необхідно перемножити спектр початкового зображення $F_1(u,v)$ на КЧХ фільтра:

$$F_2(u,v) = H(u,v) \cdot F_1(u,v).$$

Для отримання відфільтрованого зображення необхідно виконати обернене перетворення Фур'є.

ЛЕКЦІЯ 10

Тема: Цифрова обробка кольорових зображень.

План

1. Колірна модель RGB.
2. Колірна модель CMYK.
3. Колірна модель HSB.

RGB (скорочено від англ. *Red, Green, Blue* — червоний, зелений, синій) — адитивна колірна модель, що описує спосіб синтезукольору, за якою червоне, зелене та синє світло накладаються разом, змішуючись у різноманітні кольори. Широко застосовується в техніці, що відтворює зображення за допомогою випромінення світла.

У даній моделі колір кодується градаціями складових каналів (Red, Green, Blue). Тому за збільшення величини градації котрогось каналу — зростає його інтенсивність під час синтезу.

Кількість градацій кожного каналу залежить від розрядності бітового значення RGB. Зазвичай використовують 24-бітну модель,

у котрій визначається по 8 біт на кожен канал, і тому кількість градацій дорівнює 256, що дозволяє закодувати 256^3 = 16 777 216 кольорів.

Колірна модель RGB призначена сприймати, представляти та відображати зображення в електронних системах, таких як телебачення та комп'ютери, хоча її також застосовували у традиційній фотографії. Вже до електронного віку, модель RGB мала за собою серйозну теорію, засновану на сприйнятті кольорів людиною.

RGB — апаратно-залежний простір кольорів. Типовими приладами із RGB-входом є кольоровий телевізор і відеокамера, сканер і цифровий фотоапарат. Переваги моделі

- Апаратна близкість із монітором, сканером, проектором, іншими пристроями;
- Велика кольорова гама, близька до можливостей людського зору;
- Доступність багатьох функцій обробки зображення (фільтрів) у програмах растрової графіки;
- Невеликий (порівняно до моделі CMYK) обсяг, проте ширший спектр кольорів.

Недоліки: Збереження імовірності помилки відображення кольорів на екрані монітора — невідповідно до кольорів, отриманих у результатікольоропроби.

CMYK (скорочено від англ. Cyan, Magenta, Yellow, BlacK color) — субтрактивна колірна модель, використовується у поліграфії, перш за все при багатофарбовому (повноколірному) друці. Вона застосовується у друкарських машинах і кольорових принтерах. Українською перші три кольори називають так: блакитний, пурпуровий, жовтий; але професіонали мають на увазі ціан, маджента тажовтий (про значення *K* див. далі). Ці кольори візуально не ідентичні із загальноприйнятими назвами кольорів. Так, маджента — це лише один з пурпурових відтінків; жовтий і блакитний — абсолютно певні відтінки, а не цілі діапазони, як у веселці.

Оскільки модель CMYK застосовують в основному в поліграфії при кольоровому друці, а папір і інші друкарські матеріали є поверхнями, що

відображають світло, зручніше рахувати яку кількість світла (і кольори) відбилося від тієї або іншої поверхні, ніж скільки поглинається. Таким чином, якщо відняти з білого три первинні кольори,RGB, ми отримаємо трійку доповнюючих кольорів CMY. «Субтрактивний» означає той, що «віднімається» — ми віднімаємо первинні кольори з білого. Модель CMYK забезпечує менше колірне охоплення, ніж адитивна модель RGB.

HSB — колірна модель, що використовується тільки для оформлення векторних і текстових об'єктів документа. Описує колірний простір, заснований на трьох характеристиках кольору: колірному тоні (Hue), насиченості (Saturation) і яскравості (Brightness).

- Hue — колірний тон, (наприклад, червоний, зелений або синьо-блакитний). Варіюється в межах 0-360°, але іноді приводиться до діапазону 0-100 або 0-1. У Windows весь колірний спектр ділиться на 240 відтінків (що можна спостерігати в редакторі палітри MS Paint), тобто тут «Hue» зводиться до діапазону 0-240 (відтінок 240 відсутній, оскільки він дублював би 0).
- Saturation — насиченість. Варіюється в межах 0-100 або 0-1. Чим більший цей параметр, тим «чистіший» колір, тому цей параметр іноді називають чистотою кольору. А чим ближчий цей параметр до нуля, тим ближчий колір до нейтрального сірого.
- Value (значення кольору) або Brightness — яскравість. Також задається в межах 0-100 або 0-1.

Модель була створена Елві Реєм Смітом, одним із засновників Pixar, в 1978 році. Вона є нелінійним перетворенням моделі RGB. Колір, представлений в HSV, залежить від пристрою, на який він буде виведений, так як HSV — перетворення моделі RGB, яка теж залежить від пристрою. Для отримання коду кольору, не залежного від пристрою, використовується модель Lab. Слід зазначити, що HSV (HSB) і HSL — дві різні колірні моделі.

Найпростіший спосіб відобразити HSV в тривимірний простір — скористатися циліндричною системою координат. Тут координата H визначається полярним кутом, S — радіус-вектором, а V — Z-координатою.

Тобто, відтінок змінюється при русі вздовж кола циліндра, насиченість — вздовж радіуса, а яскравість — вздовж висоти. Незважаючи на математичну точність, у такої моделі є істотний недолік: на практиці кількість помітних оком рівнів насиченості і відтінків зменшується при наближенні яскравості (V) до нуля (тобто, на відтінках, близьких до чорного). Також на малих S і V з'являються суттєві помилки округлення при перекладі RGB в HSV і навпаки. Тому частіше застосовується конічна модель.

Інший спосіб візуалізації колірного простору — конус. Як і в циліндрі, відтінок змінюється по колу конуса. Насиченість кольору зростає з віддаленням від осі конуса, а яскравість — з наближенням до його основи. Іноді замість конуса використовують правильну шестикутну піраміду.

Обидва ці способи є зручною тривимірною ілюстрацією простору HSV. Але через тривимірність вони в прикладному ПЗ не застосовуються.

ЛЕКЦІЯ 11

Тема: Стиснення зображення.

План

1. Методи стиснення зображень.
2. Алгоритм RLE.
3. Алгоритм JPEG.
4. Фрактальний стиск.
5. Алгоритм JBIG.

Растрові файли мають дуже великі розміри . Якщо знехтувати заголовками файла й іншими неграфічними даними, те його розмір пропорційний кількості пікселів у зображенні і кількості бітів, необхідних для представлення кожного піксела. Повнокольорове зображення розміром 1024x768 пікселів займає більш

двох мегабайт пам'яті, а одна секунда відеофільма телевізійної якості в растровому виді вимагає біля тридцятьох мегабайт. Тому жорсткий диск можна заповнити миттєво. Навіть компакт-диск, що вміщає біля 700 мегабайт даних, не настільки великий, щоб помістити такий об'єм інформації.

Використовуючи метод, який називається стисненням зображень, можна різко зменшити в розмірі графічні файли. При стисненні графічної інформації використовуються спеціальні прийоми, що зменшують кількість байтів, необхідних для представлення зображення. Степінь стиснення залежить від методу стиснення і вмісту графічного файла . Як правило графічний файл стискується в п'ять і більш разів. Існують методи, що стискують ще сильніше, але з втратами якості . При відновленні зображення втрачається деяка частина колірної інформації. У підсумку, розпаковане зображення може стати злегка розмитим або знебарвленим.

Методи стиснення зображень діляться на дві великі групи: стиснення з втратами і стиснення без втрат. Методи стиснення без втрат дають більш низький коефіцієнт стиснення, але зате зберігають точне значення пікселів вихідного зображення. Методи з втратами дають більш високі коефіцієнти стиснення, але не дозволяють відтворити початкове зображення з точністю до піксела. Для файлів, які формуються програмами автоматизованого проектування, дуже важливо зберегти всю інформацію, тому що втрата хоча б одного біта може змінити зміст усього файла. Зовсім інша справа з растровими даними. Людське око не сприймає всі відтінки кольору в звичайному растровому зображенні. Таким чином, деякі деталі можуть бути опущені без видимого порушення інформаційного змісту зображення.

Розглянемо два найбільш розповсюджені методи стиснення зображень. Спочатку познайомимося з одним із варіантів групового кодування (run-lenght encoding - RLE). Ідея методу полягає в тому, що послідовність значень, що повторюються, заміняється парою чисел: одне з них вказує на довжину групи (число повторень даного значення), а інше - на власне це значення. Це дуже

загальний і дуже простий метод без втрат. Він використовується в багатьох популярних сьогодні форматах графічних файлів і, зокрема, у PCX і BMP. У його основі лежить той факт, що багато зображень надлишкові, оскільки містять велику кількість суміжних пікселів одного кольору. Розглянемо, наприклад, як за допомогою групового кодування стискається зображення, у якому зустрічається підряд 100 пікселів із нульовим значенням. Ця послідовність із 100 нулів кодується парою чисел (100,0). Отже такий фрагмент картинки скоротиться в п'ятдесят разів.

Іншій методом, яким користуються досить часто, - JPEG (метод, що стискує з утратами) одержав свою назву від абревіатури об'єднаної групи експертів в області фотографії (Joint Photographic Expert Group - JPEG), що його і розробила. JPEG широко використовується при стиснення статичних зображень . Цей метод істотно складніший, чим RLE. Основна ідея методу перебуває в поділі інформації в зображенні за рівнем важливості, і потім відкиданні менше важливої її частини, зменшуючи тим самим загальний об'єм збережених даних. Це досягається перетворенням матриці колірних значень у матрицю амплітуд, що відповідають визначеним частотам розкладання зображення. (Звукові коливання, наприклад, можна розкласти математичними методами на прості синусоїдальні гармоніки різних амплітуд і частот, що при додаванні відтворюють вихідний сигнал). Рядок або стовпець пікселів зображення теж можна представити амплітудами і частотами. Мова в даному випадку йде не про спектральний склад світла, а про форму представлення кривих, що утворять графіки, якщо значення пікселів служать ординатами. Відзначимо, що формула перетворення матриці пікселів у матрицю амплітуд не проста. JPEG-стиснення відкидає частину високочастотних компонент зображення, залишаючи компоненти з низькими частотами. Людське око менше критичне до високочастотних варіацій кольору, оскільки загальний вид зображення визначається низькими частотами. Значення піксела, отримане при відновленні зображення, дещо відрізняється від вихідного значення, тому що частина інформації була загублена, хоча звичайно вони дуже близькі.

У методу JPEG є дуже цікава особливість: користувач може задавати коефіцієнт якості. Високий коефіцієнт якості дозволяє зберегти більше деталей, але при цьому зменшується ступінь стиснення. При низькому коефіцієнті якості степінь стиснення збільшується, але зображення стає менше чітким. Чим нижче коефіцієнт якості, тим більша кількість інформації відкидається. Коли один із методів (RLE або JPEG) застосовується до повнокольорового зображення, то червона, зелена і синя компоненти стискуються незалежно. Якщо в растровому зображені використовується палітра або просто відтінки сірого, то значення пікселів можливо закодувати в один прохід.

Розглянемо більш детально методи RLE і JPEG.

Алгоритм групового кодування

Починаючи з першого рядка, програма групового кодування переглядає значення пікселів зліва праворуч і шукає відрізки пікселів, що повторюються. Всякий раз, коли зустрічаються три або більше пікселів, що йдуть підряд, з однаковим значенням, програма заміняє їх парою чисел: перше число вказує на довжину відрізка, друге - на значення пікселів. Число, що визначає довжину відрізка, називають міткою відрізка. Щоб ідентифікувати серії значень пікселів, що не повторюються, програма також уставляє мітки , що вказують на кількість таких значень у серії. Зарезервований біт необхідний для того, щоб можна було відрізнити мітку відрізка від мітки серії значень, що не повторюються. Наприклад, у 8-ми бітах можна задати послідовності довжиною до 127 пікселів; восьмий біт у кожній мітці може відрізняти відрізок від серії пікселів, що не повторюються . Точно так само обробляється кожний рядок пікселів і відрізки однакових значень пікселів стискуються у всьому зображенні.

Графічна програма декодує зображення, зчитуючи стиснутий файл і відновлює відрізки повторюваних значень пікселів. Зауважимо, що відновлене зображення цілком збігається з оригіналом.

Алгоритм JPEG

Насамперед програма поділяє зображення на блоки - матриці розміром 8x8 пікселів. При використанні методу JPEG час, що затрачається на стиснення зображення, пропорційний квадрату числа пікселів у блоці. Обробка декількох блоків меншого розміру робиться значно швидше, чим обробка всього зображення цілком.

До значень пікселів застосовується формула, названа дискретним косинусоїдальним перетворенням (Discrete Cosine Transform - DCT). DCT переводить матрицю значень пікселів 8x8 у матрицю значень амплітуд тієї ж розмірності, що відповідає визначеним частотам синусоїдальних коливань. Лівий верхній кут матриці відповідає низьким частотам, а правий нижній - високим.

Коефіцієнт якості, введений користувачем, використовується в простій формулі, що генерує значення елементів іншої матриці 8x8, яка називається матрицею квантування. Чим нижче коефіцієнт якості, тим більші значення будуть мати елементи матриці.

Кожне значення в матриці, яка була сформована після DCT-перетворення, ділиться на відповідне значення з матриці квантування, потім округляється до найближчого цілого числа. Оскільки великі числа знаходяться в правій нижній половині матриці квантування, то основна частина високочастотної інформації зображення буде відкинута. Тому нижня права частина матриці пікселів буде перебувати в основному з нулів.

Далі програма зчитує елементи матриці і кодує їх послідовно методами без втрат. Зауважимо, що стиснення істотно залежить від нулів у правій нижній половині матриці. Чим нижче коефіцієнт якості, тим більше нулів у матриці і, відповідно, тим вище ступінь стиснення.

Декодування JPEG-зображення починається з кроку зворотного кодуванню без втрат, у результаті чого відновлюється матриця квантування пікселів.

Значення з матриці пікселів перемножується на значення з матриці квантування, щоб відновити, наскільки це можливо, матрицю, що була

обчислена на кроку застосування DCT. На етапі квантування була загублена деяка частина інформації, тому числа в матриці будуть близькі до початкових, але не буде абсолютного збігу.

Зворотна до DCT формула (IDCT) застосовується до матриці для відновлення значень пікселів вихідного зображення. Ще разом відзначимо, що отримані кольори не будуть цілком відповідати початковим через втрату інформації на кроку квантування. Відновлене зображення, при порівнянні з оригіналом, буде виглядати декілька розмитим і знебарвленим.

Алгоритм Хаффмана

Один з класичних алгоритмів. Використовує тільки частоту появи однакових байт в зображенні. Порівнює символів вхідного потоку, які зустрічаються більше число разів, ланцюжок біт меншої довжини. І навпаки - зустрічаються рідко - ланцюжок більшої довжини. Для збору статистики вимагає двох проходів по зображенню. Коефіцієнти стиснення: 1 / 8, 2 / 3, 1. Вимагає запису у файл таблиці відповідності кодуються символів і кодують ланцюжків. На практиці використовуються його різновиди. Так, в деяких випадках резонно або використовувати постійну таблицю, або будувати її "адаптивно", тобто в процесі архівації / розархівації. Ці прийоми позбавляють від двох проходів по зображенню і необхідності зберігання таблиці разом з файлом. Кодування з фіксованою таблицею застосовується як до останнього етапу архівації в JPEG.

Близька модифікація алгоритму використовується при стисненні чорно-білих зображень. Послідовності поспіль йдуть чорних і білих крапок замінюються числом, рівним їх кількості з ознакою кольору. А цей ряд вже, у свою чергу, стискується по Хаффману з фіксованою таблицею. Алгоритм реалізований у форматі TIFF.

JBIG

Алгоритм розроблений групою експертів ISO (-Joint Bi level Experts Group) спеціально для стиснення однобітних чорно-білих зображень. Наприклад, факсів або відсканованих документів. У принципі може

застосовуватися і до 2-х, і до 4-х двійкового картинок. При цьому алгоритм розбиває їх на окремі бітові площині. JBIG дозволяє управляти такими параметрами, як порядок розбиття зображення на бітові площині, ширина смуг в зображенні, рівні масштабування. Остання можливість дозволяє легко орієнтуватися в базі великих за розмірами зображень, переглядаючи спочатку їх зменшені копії. Настроюючи ці параметри, можна використовувати цікавий ефект при отриманні зображення по мережі або з будь-якого іншому каналу, пропускна здатність якого мала в порівнянні з можливостями процесора. Розпаковуватися зображення на екрані буде поступово, як би повільно "проявляючись". При цьому людина починає аналізувати зображення задовго до кінця процесу розархівації.

Алгоритм побудований на базі Q-кодувальника, патент на який володіє IBM. Q-кодер також, як і алгоритм Хаффмана, використовує для частіше з'являються символів короткі ланцюжки, а для рідше з'являються довгі. Однак, на відміну від нього, в алгоритмі використовуються і послідовності символів. Характерною особливістю JBIG є різке зниження ступеня стиснення при підвищенні рівня шумів вхідний картинки.

Фрактальний стиск

Ця група алгоритмів, мабуть, є найбільш перспективною і розвивається зараз найбільш бурхливо. Перші практичні результати були отримані зовсім недавно - у 1992 році - і виробили приголомшливе враження. Коефіцієнт стискування у фрактальних алгоритмів варіюється в межах 2-2000. Причому великі коефіцієнти досягаються на реальних зображеннях, що, взагалі кажучи, нетипово для попередніх алгоритмів. Крім того, при розархівації зображення можна масштабувати. Унікальна особливість цього алгоритму полягає в тому, що збільшене зображення не дробиться на квадрати. Під фрактального стиснення використовується принципово нова ідея - не близькість квітів у локальній області, а подібність різних за розміром областей зображення. Це, безумовно, найбільш прогресивний підхід на сьогоднішній день. Алгоритм орієнтований на повнокольорові зображення і зображення в градаціях сірого

кольору. Його особливістю є потреба в колосальних обчислювальних потужностях за архівації. При цьому розпаковування вимагає менше обчислень, ніж у JPEG.

ЛЕКЦІЯ 12

Тема: Основи сегментації зображення.

План

1. Сегментація зображення.
2. Сфера застосування сегментації зображення.
3. Методи сегментації зображень.

З комп'ютерної точки зору, сегментація — це процес розділення цифрового зображення на декілька сегментів (множина пікселів, які часто називають суперпікселями). Ціль сегментації полягає у спрощенні і/або зміні представлення зображення, щоб його було простіше і легше аналізувати.[1] Сегментація зображень зазвичай використовується для того, щоб виділити об'єкти і межі (лінії, криві, і т. д.) на зображеннях. Більш точно, сегментація зображень — це процес присвоєння таких міток кожному пікселю зображення, що пікселі з однаковими мітками мають спільні візуальні характеристики.

Результатом сегментації зображення є множина сегментів, які разом покривають все зображення, або множина контурів, виділених з зображення. Всі пікселі в сегменті схожі за деякою характеристикою або за визначеною властивістю, наприклад колір, яскравість або текстура. Сусідні сегменти значно відрізняються за цими характеристиками

Деякими практичними застосуваннями сегментації зображень є:
Медичні зображення.

- Виявлення пухлин та інших патологій;

- Визначення обсягів тканин;
- Хірургія за допомогою комп'ютера;
- Діагностика;
- Планування лікування;
- Вивчення анатомічної структури;

Виділення об'єктів на супутникових світлинах.

Розпізнавання обличчя.

Розпізнавання відбитків пальців.

Системи управління дорожнім рухом.

Виявлення стоп-сигналів.

Комп'ютерний зір.

Для сегментації зображень було розроблено декілька універсальних алгоритмів і методів. Так як загального рішення для задачі сегментації зображень не існує, часто ці методи доводиться поєднувати зі знаннями з предметної області, щоб ефективно вирішувати цю задачу в її предметній області.

Методи, засновані на кластеризації. Метод k-середніх — це ітераційний метод, який використовуєтсья для того, щоб розділити зображення на K кластерів. Базовий алгоритм наведений нижче:

1. Вибрати K центрів кластерів, випадково або на основі деякої евристики
2. Помістити кожен піксель зображення в кластер, центр якого найближче до цього пікселя
3. Знову визначити центри кластерів, усереднюючи всі пікселі в кластері
4. Повторювати кроки 2 і 3 до збіжності (наприклад, коли пікселі будуть залишатися в тому ж кластері)

Тут в якості відстані зазвичай береться сума квадратів або абсолютних значень різниць між пікселем і центром кластера. Різниця зазвичай базується на кольорі, яскравості, текстурі і місцезнаходження пікселя, або на зваженій сумі цих чинників. K-середніх може бути вибране вручну, випадково чи евристично.

Цей алгоритм гарантовано сходиться, але він може не привести до оптимального рішення. Якість рішення залежить від початкового множини кластерів і значення K.

Методи з використанням гістограми дуже ефективні порівняно з іншими методами сегментації оскільки вони вимагають тільки один прохід по пікселях. У цьому методі гістограма обчислюється за всіма пікселям зображення і її мінімуми і максимуми використовуються, щоб знайти кластери на зображенні.[1] Колір або яскравість можуть бути використані при порівнянні.
Покращення цього методу — рекурсивно застосовувати його до кластерів на зображенні для того, щоб поділити їх на дрібніші кластери. Процес повторюється з усе меншими і меншими кластерами до тих пір, коли перестануть з'являтися нові кластери.[1][3]

Один недолік цього методу — те, що йому може бути важко знайти значні мінімуми і максимуми на зображенні. У цьому методі класифікації зображень схожі метрика відстаней і зіставлення інтегрованих регіонів.

Підходи, засновані на використанні гістограм можна також швидко адаптувати для кількох кадрів, зберігаючи їх перевага в швидкості за рахунок одного проходу. Гістограма може бути побудована кількома способами, коли розглядаються декілька кадрів. Той самий підхід, який використовується для одного кадру, може бути застосований для кількох, і після того, як результати об'єднані, мінімуми і максимуми, які було складно виділити, стають помітніші. Гістограма також може бути застосована для кожного пікселя, де інформація використовується для визначення найчастішого кольору для даного положення пікселя. Цей підхід використовує сегментацію, засновану на рухомих об'єктах і нерухомому оточенні, що дає інший вид сегментації, корисний в відео трекінгу.

Виділення країв — це добре вивчена область в обробці зображень. Межі та краї областей сильно пов'язані, оскільки часто існує сильний перепад яскравості на кордонах областей. Тому методи виділення країв використовуються як основа для іншого методу сегментації.

Виявлені краю часто бувають розірваними. Але щоб виділити об'єкт на зображенні, потрібні замкнуті межі області.

Першим був метод розростання областей з насіння. В якості вхідних даних цей метод приймає зображень і набір насіння. Насіння відзначають об'єкти, які потрібно виділити. Області поступово розростаються, порівнюючи всі незайняті сусідні пікселі з областю. Різниця δ між яскравістю пікселя і середньою яскравістю області використовується як міра схожості. Піксель з найменшою такою різницею додається у відповідну область. Процес триває доки всі пікселі не будуть додані в один з регіонів.

Метод розростання областей з насіння вимагає додаткового вводу. Результат залежить від вибору насіння. Шум на зображенні може призвести до погано розміщення насіння. Метод розростання областей без використання насіння — це змінений алгоритм, який не вимагає явного насіння. Він починає з однієї області A_1 — піксель, обраний тут незначно впливає на кінцеву сегментацію. На кожній ітерації він розглядає сусідні пікселі так само, як метод розростання областей з використанням насіння. Але він відрізняється тим, що якщо мінімальна δ не менша, чим заданний поріг T, то він додається у відповідну область A_j. В іншому випадку піксель вважається таким, що сильно відрізняється від всіх поточних областей A_i і створюється нова область A_{n+1}, яка містить цей піксель.

Один з варіантів цього методу, запропонований Хараліком і Шапіро заснований навикористанні яскравості пікселів. Середнє и дисперсія області та яскравість пікселя-кандидата використовується для побудови тестової статистики. Якщо тестова статистика достатньо мала, то піксель додається до області, і середнє і дисперсія області перераховується. Інакше, піксель ігноруються і використовується для створення нової області.

Методи розрізу графа можуть бути ефективно застосовані для сегментації зображень. У цих методах зображення представляється як зважений неорієнтований граф. Зазвичай, піксель або група пікселів асоціюється

вершиною, а ваги ребер визначають (не) схожість сусідніх пікселів. Потім граф (зображення) розрізається відповідно до критерію, створеному для отримання «хороших» кластерів. Кожна частина вершин (пікселів), одержувана цими алгоритмами, вважається об'єктом на зображенні. деякі популярні алгоритми цієї категорії — це нормалізовані розрізи графів, випадкове блукання, мінімальний розріз, ізоперіметричний поділ і сегментація з допомогоюмінімального зваженого дерева.

Метод водоподілу — це заснований на областях метод математичної морфології. У географії, вододіл — це хребет, який ділить області різних річкових систем.

Розглядаючи зображення, як геологічний ландшафт, можна сказати, що лінії вододілу — це кордони, що розділяють ділянки зображень. У топографічному поданні зображення, чисельні значення (наприклад, рівні сірого) кожного пікселя виступають як висоти цієї точки. Перетворення водоподілу обчислює водозбірні басейни та лінії хребтів, при тому що водозбірні басейни — відповідні області зображення, а лінії хребтів — це межі цих областей. Основною проблемою даного алгоритму є надмірна сегментація, оскільки всі межі і шуми подаються в градієнті, що робить необхідним процес видалення.

Перший етап видалення шуму в початковому зображенні полягає у застосуванні морфологічних операцій закриття / розкриття, потім обчислюється морфологічний градієнт зображення без шуму і виконується нелінійне перетворення для рівнів сірого на градієнті зображення за допомогою принципу Вебера, останній етап — обчислення вододілу по нелінійному, розбитому на області, градієнтному зображенню

Сегментація за допомогою моделі. Основне припущення цього підходу — те, що структури які нас цікавлять або органи мають повторювані геометричні форми. Отже, можна знайти імовірнісну модель для пояснення змін форми органу і потім, сегментуючи зображення, накладати обмеження, використовуючи цю модель як апріорну. Таке завдання включає в себе (1)

приведення тренувальних прикладів до загальної пози, (2) ймовірнісне представлення змін наведених зразків і (3) статистичний висновок для моделі і зображення. Сучасні методи сегментації літературі, заснованої на знанні, містять активні моделі форми і зовнішності, активні контури, деформаційні шаблони і методи встановлення рівня.

Багатомасштабна сегментація. Сегментація зображень виконується в різних масштабах у масштабному просторі й іноді поширюється від дрібних масштабів до великих. Критерій сегментації може бути безпідставно складним і може приймати до уваги як локальні, так і глобальні критерії. Загальна вимога — те, що кожна область повинна бути пов'язана в деякому сенсі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

Базовий

1. Гольденберг Л.М. и др. Цифровая обработка сигналов: Учебное пособие для вузов. - М.: Радио и связь, 1990.- 256 с.

2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. Пер. с англ. – М.: Техносфера, 2006. – 1072 с.

3. Дьяконов В., Абраменкова И. MATLAB. Обработка сигналов и изображений. Специальный справочник. – СПб.: Питер, 2002, 608 с.

4. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. – М.: Связь, 1979. – 416 с.

5. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. – М.: Мир, 1978. – 848 с.

6. Сергиенко А.Б. Цифровая обработка сигналов. - СПб.: Питер, 2003. – 604 с.: ил.

7. Сиберт У.М. Цепи, сигналы, системы. – М.: Мир, 1988. – 336 с.

8. Солонина А.И., Улахович Д.А., Арбузов С.М., Соловьева Е.Б. Основы цифровой обработки сигналов: Курс лекций. Изд. 2-е испр. И перераб.- СПб.: БХВ-Петербург, 2005.

Допоміжний

1. Блейхут Р. Быстрые алгоритмы цифровой обработки сигналов. – М.: Мир, 1989. – 448 с.

2. Даджион Д., Мерсеро Р. Цифровая обработка многомерных сигналов. – М.: Мир, 1988. – 488 с.

3. Купер Дж., Макгиллем А. Вероятностные методы анали за сигналов и систем. – М.: Мир, 1989. – 376 с.

4. Марпл. С.Л. Цифровой спектральный анализ и его приложения. – М.: Мир, 1990. – 584 с.

5. Претт Е. Цифровая обработка изображений: Пер. с англ. под ред. Д. С. Лебедева, – М.: Мир, 1982. – в 2-х книгах.

6. Хемминг Р.В. Цифровые фильтры. – М.: Недра, 1987. – 221 с.

7. Шапиро Л., Стокман Дж. Компьютерное зрение - Москва: Бином. Лабо- ратория знаний, 2006 - 716 с.

8. Pattern recognition, fourth edition / Sergios Theodoridis, Konstantinos Koutroumbas. – Elsevier Inc., 2009. – 961 p.

Інформаційні ресурси

1. http://www.mathworks.com/products/matlab/ - MATLAB. The Language of Technical Computing.

2. http://www.chemometrics.ru/materials/textbooks/matlab.htm - MatLab. Руководство для начинающих.

3. http://algolist.manual.ru/compress/standard/ - Общие алгоритмы сжатия и кодирования.

4. http://matlab.exponenta.ru/ml/book1/index.php - Введение в Matlab.

5. http://ocw.mit.edu/resources/res-6-008-digital-signal-processing-spring-2011/ - Digital Signal Processing.

КОНСПЕКТ ЛЕКЦІЙ
з дисципліни «Обробка сигналів та зображень»
для студентів денної форми навчання напрямку підготовки 6.170101 «Безпека інформаційних і комунікаційних систем»

Укладачі: Фриз Михайло Євгенович, Стадник Марія Андріївна

Підписано до друку _____ Формат 60x84 1/16. Ум. др. арк. 1. Друк лазерний. Замовлення № _____. Наклад 100 пр.
Віддруковано у видавництві ТНТУ.