

Тема 6

ЦИФРОВА ОБРОБКА ЗОБРАЖЕНЬ

ЗМІСТ

Вступ.....	4
1 ОСНОВНІ ПОНЯТТЯ ЦИФРОВОЇ ОБРОБКИ ЗОБРАЖЕНЬ.....	5
1.1 Основні поняття і визначення	5
1.2 Технічні засоби обробки зображень	22
2 МЕТОДИ КЛАСИФІКАЦІЇ ЦИФРОВИХ ЗОБРАЖЕНЬ	27
2.1 Параметричні та непараметричні методи класифікації цифрових зображень.....	27
3 МАТЕМАТИЧНІ ОСНОВИ ФІЛЬТРАЦІЇ ЗОБРАЖЕНЬ	40
3.1 Фільтрація та відновлення зображень	40
3.2 Лінійна просторово-інваріантна фільтрація та фільтрація у просторовій області	66
СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ.....	74

1 ОСНОВНІ ПОНЯТТЯ ЦИФРОВОЇ ОБРОБКИ ЗОБРАЖЕНЬ

1.1 Основні поняття і визначення

План

1. Области використання цифрової обробки зображень.
2. Реєстрація зображень.
3. Дискретизація та квантування зображення.
4. Лінійне контрастування зображення.
5. Соляризація зображення.
6. Препарування зображення.
7. Перетворення гистограм, еквалізація.
8. Застосування табличного методу до поелементного перетворення.
9. Види зображень.

Області використання цифрової обробки зображень

У даному підрозділі мова піде про найпопулярніші аспекти застосування технологій та методів цифрової обробки зображень, про їх актуальність та практичне значення. Аналіз областей використання методів цифрової обробки зображень показав, що вона проникла майже у всі види інформаційної діяльності людини [1].

Компресія цифрових зображень (стиснення, компактне подання) – одна з найактуальніших проблем цифрової обробки зображень, вона пов'язана з необхідністю економії місця на фізичних носіях інформації.

Прикладами практичного використання компресії зображень є:

- стиснення зображень на штучному супутнику Землі з метою збільшення об'єму інформації, що передається за сеанс зв'язку [2];
- зменшення об'ємів зображень для швидкого завантаження веб-сторінок;
- компактне подання зображень оптоелектронними пристроями (фотоапаратами, камерами) для економного використання дискового простору;
- застосування графічних архіваторів для збереження архівних, історичних, художніх та інших документів;
- стиснення зображень як окремих кадрів відеопотоку для зменшення об'ємів носіїв та скорочення вимог до якості каналів зв'язку.

Фільтрація шумів – одне із завдань ідентифікації об'єктів у військовій справі, важливе значення має при підвищенні чіткості зображень у цифровому телебаченні, фільтруванні сигналів у відповідних блоках сучасної апаратури.

Веб-дизайн та комп'ютерна графіка – актуальні для розробників штучної графіки (комп'ютерні ігри, анімація, зображення веб-сторінок) та, наприклад, розробників спеціалізованих програм для автоматизованих систем управління військами, об'єктів критичного застосування (управління атомними станціями, залізничним та авіаційним транспортом, центрами космічних польотів).

Розпізнавання образів – актуально у військовій справі, коли правильна ідентифікація військового об'єкта може кардинально змінити хід бою чи локальної операції, знаходить своє місце в археології та архівній справі, допомагає відтворити історично цінні документальні знахідки, а також при картографуванні місцевості і передачі інформації зі штучних супутників Землі.

Цифрова обробка зображень – це область обчислювальної техніки, що динамічно розвивається та охоплює як технічні, так і програмні засоби.

Цифрове зображення – модель реального або синтезованого зображення, що зберігається у файлі на машинному носії у вигляді сукупності кодів (цифр).

Область діяльності, що пов'язана з підготовкою і обробкою цифрових зображень, називається *комп'ютерною графікою*.

Спорідненими областями до цифрової обробки зображень є теорія інформації, теорія оптимального прийому сигналів і теорія розпізнавання образів. Термін «сигнал» ототожнюють із поняттями «дані» і «інформація».

Сигнал – це інформаційна функція, що несе повідомлення про фізичні властивості, стан або поведінку будь-якої фізичної системи, об'єкта або середовища.

Під сигналом ми розуміємо будь-яку змінну, яка передає або містить якийсь вид інформації, і яку можна, наприклад, переносити, виводити на екран або виконувати з нею якісь дії.

Обробку сигналів зображень здійснюють з метою зміни якості зображень, надання зображенню нових властивостей, аналізу інформації, що міститься у зображенні або для скорочення сигнального потоку, що передуює передаванню або запису сигналів зображень [3].

Незалежно від типу зображень все різноманіття принципів та способів обробки зображень можна розділити на такі напрямки:

- реставрація та поліпшення зображень;
- аналіз зображень (розпізнавання образів та аналіз сцен);
- синтез зображень;
- кодування сигналів зображень.

У рамках першого напрямку здійснюють зміну контрасту, придушення шумів, уточнення меж об'єктів, корекцію кольорів.

Коли мова йде про поліпшення зображення, то мають на увазі зміну його властивостей, яка призводить до більш комфортного суб'єктивного сприйняття цього зображення, а не про досягнення повної ідентичності з реальним зображенням.

У рамках другого напрямку здійснюють ідентифікацію об'єктів, що є на досліджуваній сцені, оцінюють взаємозв'язок фрагментів зображення, а також визначають характеристики зображених об'єктів.

Синтез зображень набув останнім часом надзвичайно значного поширення. Методи та способи синтезу зображень використовують в абсолютно різних галузях діяльності. Так, наприклад, синтез тривимірних зображень здійснюють за плоскими фотознімками земної поверхні або поверхні інших космічних об'єктів з метою дослідження властивостей цих об'єктів.

Синтез здійснюють під час двовимірного або тривимірного моделювання об'єктів під час автоматизованого проектування споруд, транспортних засобів.

Процес кодування сигналів зображень здійснюють з метою зменшення сигнального потоку необхідного для запису або передачі інформації про зображення.

Реєстрація зображень – це процес трансформування різних наборів даних в одну координатну систему. Даними можуть бути серія фотографій, дані з різних сенсорів, моментів часу, глибини або точок спостереження [4].

Алгоритми реєстрації зображень використовуються в комп'ютерному баченні, методах медичної візуалізації, у військовій справі для автоматичного розпізнавання цілей, для впорядковування та аналізу зображень із супутників. Реєстрація необхідна для того, щоб мати можливість порівнювати або інтегрувати дані, що отримані з різних пристроїв реєстрації.

У сучасних інформаційних системах зображення передають, зберігають та обробляють переважно у цифровій формі, але первинні зображення здебільшого існують у вигляді безперервних двовимірних полів розподілу яскравості та кольору.

Перетворення первинних зображень у цифрові сигнали є обов'язковою операцією, якщо передбачається використати цифрову обробку, передачу, зберігання. Таке перетворення складається із двох процедур, які здійснюють одночасно.

Перша полягає у заміні безперервного зображення набором дискретних елементів та називається *дискретизацією*, а друга виконує заміну безперервного розподілу яскравості та забарвлення множиною квантованих значень для кожного елемента зображення і називається *квантуванням* [5].

Двовимірний характер зображення у порівнянні зі звичайними одновимірними сигналами надає додаткові можливості щодо оптимізації цифрового сигнального потоку з метою скорочення обсягу цифрових даних.

Растром називають структуру поля зображення утворену в результаті поелементного розкладання або синтезу зображення. У сучасній термінології елемент зображення називають «піксел» або «піксель», а в англійській літературі можна зустріти кілька еквівалентних назв, які було утворено від сполучення слів picture element (елемент зображення) – pictel, pixel, pel.

На практиці застосовують дискретизацію, з використанням прямокутного растра, та рівномірне квантування яскравості. Такий підхід застосовують через простоту виконання відповідних операцій, а також із-за необхідності здійснення у подальшому операцій пов'язаних з перетвореннями зображень.

За умови використання прямокутного растра в остаточному вигляді сигнали оцифрованого зображення подають у вигляді матриці (монохромне зображення) або набору матриць (кольорове зображення), рядки та стовпці яких містять квантовані значення параметрів відповідних елементів дискретизованого зображення.

Цифрова обробка зображень шляхом поелементних перетворень

У різних інформаційних системах результати обробки даних подають у вигляді зображення, що формують на екрані пристрою відображення.

Процедуру, що забезпечує перетворення електричних сигналів у оптичне зображення називають *візуалізацією*.

Відтвореному зображенню бажано надати таких властивостей, за яких сприйняття окремих, найбільш суттєвих фрагментів зображення, або зображення в цілому буде зручним для спостерігача з огляду на художні або технологічні властивості цього зображення.

Якість зображення не завжди оцінюють за точністю відтворення первинного образу або лінійністю передачі зміни параметрів.

Наприклад, зображення, яке було отримано у сутінках або тумані буде характеризуватись малою контрастністю, нечіткими обрисами, блідим забарвленням.

Зображення, яке було отримано за допомогою інфрачервоного перетворювача, за умови безпосереднього відтворення сигналів, може мати малу кількість градацій яскравості, що не дозволить виділити малоконтрастні деталі. Часто буває корисним підкреслити або підсилити якісь особливості відтвореної сцени з метою поліпшення її суб'єктивного сприйняття.

Суб'єктивність сприйняття зображення у разі його візуального аналізу та оцінки значно ускладнює застосування формалізованих підходів щодо формування властивостей відтвореного зображення. Тому під час здійснення візуалізації зображень набули поширення методи, для яких відсутні строгі математичні критерії оптимальності.

Застосування тих чи інших методів, параметрів та режимів часто ґрунтується на досвіді, суб'єктивному сприйнятті та мистецьких здібностях.

Серед процедур, що використовують для обробки зображень можна виділити дві групи [6]:

- поелементна обробка;
- кореляційна обробка.

За умови використання методів обробки першої групи результат обробки будь-якої точки кадру зображення залежить тільки від значення відліку характерного параметра первинного зображення у цій же точці. Очевидною перевагою таких процедур є простота реалізації, вони призводять до значного суб'єктивного поліпшення візуальної якості. Поелементну обробку зображень застосовують як заключний етап процесу комплексної обробки зображень.

Друга група процедур базується на тому, що між елементами всього зображення або окремих його фрагментів є взаємні зв'язки – *кореляція*.

Для одержання результату перетворення для кожної точки зображення у цьому випадку залучають дані про параметри певної множини точок первинного зображення, що знаходяться навколо обробленої точки.

Сутність поелементної обробки зображень полягає у встановленні певної функціональної залежності між остаточним значенням відліку сигналу зображення та його первинним значенням або статистичною характеристикою.

Нехай $x(i, j) = x_{i,j}$ та $y(i, j) = y_{i,j}$ – значення яскравості вхідного і вихідного сигналів зображень у точці, що має координати відповідно до i -го відліку в напрямку вертикальної осі та j -го відліку в напрямку горизонтальної осі.

Поелементна обробка базується на тому, що між значеннями яскравості існує однозначна функціональна залежність [7]:

$$y_{i,j} = f_{i,j}(x_{i,j}), \quad (1.1)$$

що дозволяє на підставі значень первинного сигналу визначити значення кінцевого.

Параметри функції $f_{i,j}(\cdot)$, що описує процес обробки, можуть залежати від поточних координат. Якщо така залежність існує, то обробку, яку здійснюють, називають *неоднорідною*.

У більшості процедур, що знайшли практичне застосування, здійснюють *однорідну поелементну обробку*, у цьому випадку індекси i та j з виразу (1.1) можуть бути відсутні:

$$y = f(x). \quad (1.2)$$

Залежність між яскравістю елементів вхідного та вихідного зображень буде визначено функцією (1.2) однаковою для всіх точок зображення.

Процедури поелементної обробки можна розділити на:

- процедури пов'язані зі зміною контрасту зображення;
- процедури бінаризації зображень.

Контраст зображення – це параметр, що дорівнює відношенню максимальної яскравості до мінімальної в полі зображення.

Контрастність зображення – параметр, що визначає перепад яскравості відносно середнього рівня та чисельно дорівнює відношенню різниці максимальної та мінімальної яскравості до їх суми [8]. Часто остаточне значення подають у відсотках.

Лінійне контрастування зображень

Лінійне контрастування використовують для узгодження динамічного діапазону вхідного сигналу зображення та динамічного діапазону яскравості пристрою відображення. Якщо для цифрового значення кожного відліку зображення виділяється 1 байт (8 біт) пристрою пам'яті, то вхідний або вихідний сигнали можуть набувати одного з 256 значень. Діапазон можливих значень аналогового сигналу знаходиться у межах від 0 до 255. Зазвичай, значення 0 відповідає рівню чорного, а значення 255 – рівню білого.

Припустимо, що мінімальна та максимальна яскравості первинного зображення мають значення $x_{\min}$ та $x_{\max}$ відповідно.

Якщо ці параметри або один з них істотно відрізняються від граничних значень діапазону сигналів яскравості, то відтворене зображення виглядає як малоконтрастне, незручне для сприйняття та дослідження.

Під час лінійного контрастування здійснюють лінійне поелементне перетворення вхідної величини x відповідно до рівняння

$$y = a \cdot x + b, \quad (1.3)$$

де параметри a та b визначаються бажаними значеннями мінімальної $y_{\min}$ та максимальної $y_{\max}$ яскравості в кінцевому зображенні.

Для знаходження зазначених параметрів потрібно розв'язати систему рівнянь [8]:

$$\begin{cases} y_{\min} = a \cdot x_{\min} + b, \\ y_{\max} = a \cdot x_{\max} + b. \end{cases} \quad (1.4)$$

Після підстановки знайдених значень у (1.3) рівняння для лінійного контрастування набуває остаточного вигляду:

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} (y_{\max} - y_{\min}) + y_{\min}. \quad (1.5)$$

При порівнянні двох зображень (вхідного та вихідного) встановлено значно кращу візуальну якість обробленого зображення. Поліпшення обумовлено узгодженням динамічного діапазону вхідного зображення та динамічного діапазону екрана внаслідок здійснення лінійного контрастування.

Соляризація зображення

Даний різновид нелінійного контрастування отримав свою назву від слова «солярій» (лат. solarium, від sol – сонце).

Результат здійснення даного перетворення призводить до освітлення певної частини зображення, що має рівні яскравості, що відносяться до середини динамічного діапазону.

Перетворення здійснюють відповідно до співвідношення

$$y = k \cdot x \cdot (x_{\max} - x), \quad (1.6)$$

де $x_{\max}$ – максимальне значення вихідного сигналу;

k – константа, що управляє динамічним діапазоном перетвореного зображення.

Функція, що описує дане перетворення, є квадратичною.

Якщо $k = \frac{4}{x_{\max}}$, то при $y_{\max} = x_{\max}$ динамічні діапазони зображень

збігаються. Інший спосіб нормалізації динамічного діапазону може бути здійснено шляхом застосування одного із методів контрастування, наприклад, описаного вище лінійного контрастування.

Внаслідок соляризації ділянки вхідного зображення з рівнем білого (або близьким до нього рівнем яскравості) набувають чорного рівня. Рівень білого у вихідному зображенні набувають ділянки, що мали на вході середній рівень яскравості (рівень сірого).

Зональне контрастування зображення

Цей різновид контрастування називають препаруванням зображення, тобто здійснення такого перетворення яскравостей певної частини динамічного діапазону, що дозволяє дослідити або підкреслити певні властивості зображення.

Препарування – це клас поелементних перетворень зображень, більшість із яких має власну назву. Амплітудні характеристики процедур препарування, що найчастіше застосовують на практиці [9]:

- бінаризація;
- зріз за яскравістю;
- інверсне контрастування;
- зональне лінійне контрастування.

Бінаризація – це перетворення з пороговою характеристикою.

За допомогою бінаризації напівтонове зображення, що містить всі рівні яскравості, перетворюють у таке, що містить тільки два рівні яскравості $y = 0$ або $y = y_{\max}$.

Операцію застосовують, коли метою обробки є виділення характерних контурів об'єктів зображення. Основною проблемою під час здійснення такої обробки є визначення порогового значення x_0 , порівняння з яким яскравості вхідного зображення дозволяє визначити значення кінцевого зображення в кожній його точці.

Найбільш виправданим підходом для визначення порогового значення є застосування математичного апарату теорії ймовірностей, випадкових процесів і випадкових полів для опису зображень.

За цієї умови визначення оптимального порогу бінарного квантування є статистичне завдання.

Іноді під час обробки зображення доводиться мати справу із зображеннями, збереженими у напівтоновому вигляді, але за своєю сутністю вони є такими, що мало відрізняються від бінарних. До таких належать текст, штрихові малюнки, креслення, зображення відбитка пальця.

Щільність імовірності $w(x)$, що описує розподіл яскравості такого зображення, може містити два виділені максимуми. Поріг бінарного квантування варто вибирати посередині між цими максимумами.

Заміна вихідного напівтонового зображення *бінарним препаратом* вирішує два основні завдання.

По-перше, буде досягнуто більшої наочності під час візуального сприйняття у порівнянні з первинним зображенням.

По-друге, буде відчутно скорочено обсяг пам'яті для збереження зображення, оскільки бінарний препарат для запису кожної точки бінарного зображення потребує лише 1 біт пам'яті, у той час як для запам'ятовування напівтонового зображення частіше за все слід використовувати 8 біт.

Іноді наочність зображення підвищують шляхом здійснення контрастування із застосуванням пилкоподібної характеристики перетворення. В цьому випадку різні піддіпазони яскравості одночасно зазнають локального лінійного контрастування. Однак, треба взяти до уваги ту обставину, що дане перетворення, як і деякі інші, може супроводжуватись появою хибних контурів у кінцевому препараті.

Перетворення гістограм, еквалізація

У результаті всіх поелементних перетворень відбувається зміна закону розподілу ймовірностей яскравості пікселів, що описує зображення. Розглянемо механізм цієї зміни на прикладі довільного перетворення з монотонною характеристикою.

Нехай функція перетворення $y = f(x)$, а однозначна обернена до неї функція $x = \varphi(y)$. Припустимо, що випадкова величина x розподілена відповідно до щільності ймовірності $\omega_x(x)$.

Нехай Δx – довільний малий інтервал значень випадкової величини x , а Δy – відповідний інтервал перетвореної випадкової величини y .

Знаходження величини x в інтервалі Δx обумовлює знаходження величини y в інтервалі Δy , що означає імовірнісну еквівалентність цих двох подій. З огляду на малі розміри обох інтервалів, можна записати таку рівність:

$$w_x(x)|\Delta x| \approx w_y(y)|\Delta y|. \quad (1.7)$$

Модулі у (1.7) враховують залежність ймовірностей від абсолютних довжин інтервалів (незалежність від знаків Δx та Δy).

Для обчислення щільності ймовірності перетвореної величини, підставимо замість x її вираз через обернену функцію та виконаємо граничний перехід $\Delta x \rightarrow 0$ (що обумовлює $\Delta y \rightarrow 0$). У результаті можна записати:

$$w_y(y) = w_x(\varphi(y)) \cdot \left| \frac{d\varphi(y)}{dy} \right|. \quad (1.8)$$

Цей вираз дозволяє обчислити щільність імовірності результату перетворення, що *не збігається* із щільністю розподілу вихідної випадкової величини. Зрозуміло, що вигляд закону розподілу щільності ймовірності $w_y(y)$ залежить від характеристики перетворення, оскільки вираз (1.8) містить обернену функцію перетворення та її похідну.

Співвідношення набувають більш складного вигляду тоді, коли перетворення описується не взаємо-однозначною функцією.

У результаті кожного перетворення щільність імовірності кінцевого зображення не буде співпадати із щільністю ймовірності вихідного зображення.

Неважко переконатися, що за умови лінійного контрастування буде збережено вигляд закону розподілу щільності ймовірності, однак, параметри щільності ймовірності перетвореного зображення будуть іншими.

Перетворення щільності ймовірності передбачає знання інтегрального розподілу для первинного зображення. Як правило, достовірні відомості про нього відсутні.

Використання аналітичних апроксимацій для опису функцій розподілу також малопридатне, тому що їхні невеликі відхилення від дійсних розподілів можуть призводити до істотної відмінності кінцевих результатів.

Обробку зображень перетворення розподілів виконують у два етапи.

На першому етапі вимірюють гістограму вхідного зображення.

Для цифрового зображення, шкала яскравості якого знаходиться в межах від 0 до 255, гістограма є таблицею з 256 чисел, кожне з них показує кількість пікселів певної яскравості у даному кадрі.

Для знаходження оцінки розподілу ймовірностей яскравості пікселів зображення необхідно розділити всі числа цієї таблиці на загальний розмір вибірки, що дорівнює загальному числу пікселів у зображенні.

Позначимо цю оцінку як $w_x^*(j)$, $0 \leq j \leq 255$.

Тоді оцінку інтегрального розподілу буде визначено за формулою:

$$F^*(j) = \sum_{i=0}^j w_x^*(i). \quad (1.9)$$

На другому етапі виконують безпосередньо нелінійне перетворення (1.2), що забезпечує необхідні властивості вихідного зображення. Під час такого перетворення замість невідомого інтегрального розподілу використовують його оцінку, що базується на гістограмі.

Методи поелементного перетворення зображень, метою яких є зміна законів розподілу, одержали назву *гістограмних методів*. Зокрема, перетворення, у результаті якого кінцеве зображення має рівномірний розподіл, називається *еквалізацією (вирівнюванням) гістограм*.

Процедури перетворення гістограм можна застосувати як до зображення в цілому, так і до окремих його фрагментів.

Застосування даної процедури до окремих фрагментів зображення може бути корисним для обробки нестационарних зображень, зміст яких істотно розрізняється за своїми характеристиками на різних ділянках. У цьому випадку кращого ефекту можна досягти шляхом застосування гістограмної обробки до окремих ділянок зображення.

Використання співвідношень (1.5) – (1.9), що є справедливими для зображень із безперервним розподілом яскравості, є не цілком коректним для цифрових зображень.

Необхідно брати до уваги, що в результаті обробки не вдається одержати ідеальний розподіл ймовірностей вихідного зображення, тому корисно здійснювати контроль його гістограми.

Застосування табличного методу до поелементного перетворення зображень

Під час поелементних перетворень зображень обчислення згідно (1.2) необхідно виконувати для всіх пікселів вхідного зображення.

У тих завданнях, де функція $f(x)$ (1.2) передбачає трудомісткі обчислення (множення, ділення, розрахунки тригонометричних функцій, статистичних та інших функцій), застосування безпосередньо прямого методу перетворення яскравості може виявитись взагалі неприйнятним.

Час виконання обробки залежить від її обчислювальної складності. Для запобігання таких незручностей застосовують табличний метод, що отримав широке розповсюдження у практиці цифрової обробки зображень.

Сутність табличного методу полягає в тому, що шляхом попереднього розрахунку створюють таблицю функції $y = f(x)$.

Під час обробки зображення замість обчислень використовують готові результати шляхом звернення до цієї таблиці. Значення вхідної яскравості x використовують для визначення номера стовпчика, з якого треба прочитати значення перетвореного сигналу y .

Виконання цієї нескладної операції у порівнянні з безпосереднім обчисленням значень y дозволяє зробити обробку технологічною, а тривалість обчислень стає незалежною від складності перетворення.

Необхідно взяти до уваги той факт, що всі реальні таблиці, які можна записати в оперативній пам'яті комп'ютера, мають обмежену величину. Якщо множина значень вхідного сигналу перевищує розміри таблиці, то при розташуванні значення x між точками, для яких у таблиці зафіксовано певні значення, доводиться застосовувати інтерполяцію – наближене визначення відсутніх значень функції $y = f(x)$ за наявними сусідніми значеннями.

Часто для забезпечення цих обчислень використовують лінійну інтерполяцію – на проміжку між заданими вузлами невідому функцію замінюють відрізком прямої.

Слід зазначити, що у разі, якщо вхідне зображення подано у вигляді цілого числа з діапазону від 0 до 255, то розміри повної таблиці, що містить всі значення функції є незначними. Значення яскравості вхідного сигналу є адресою, що визначає номер стовпчика у таблиці функції. Обробка з використанням цього методу виявляється дуже зручною та швидкою.

Види зображень

Будемо називати *зображенням* функцію двох змінних $f(x, y)$, визначену у деякій області S площини (x, y) , що має певну множину своїх значень.

Наприклад, звичайну чорно-білу фотографію можна подати, як [8]:

$$f(x, y) \geq 0, \quad 0 \leq x \leq a, \quad 0 \leq y \leq b, \quad (1.10)$$

де $f(x, y)$ – яскравість (іноді називається оптичною щільністю або ступенем білизни) зображення в точці (x, y) ;

a – ширина;

b – висота кадру.

З урахуванням особливостей функції f виділяють такі класи зображень:

– *напівтонові (сірі) зображення* (рис. 1.1). Прикладами таких зображень можуть бути звичайні чорно-білі фотографії. Множина значень функції в області S може бути дискретною або неперервною. В один клас з напівтоновими відносять кольорові фото- і телевізійні зображення.

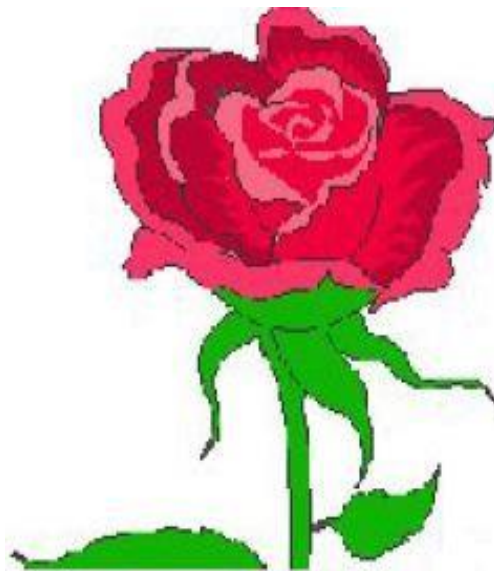


Рисунок 1.1 – Напівтонове зображення

Експериментально доведено, що практично будь-який колір, видимий людиною (але не всякий), може бути поданий як сума певних кількостей m_1, m_2, m_3 (лінійно-незалежних), так званих основних кольорів:

$$f = m_1 \cdot E_1 + m_2 \cdot E_2 + m_3 \cdot E_3. \quad (1.11)$$

Основними є кольори з довжиною хвилі:

- $\lambda_1 = 0,7$ мкм (червоний – R);
- $\lambda_2 = 0,5461$ мкм (зелений – G);
- $\lambda_3 = 0,4358$ мкм (блакитний – B).

Множина значень функції f є тримірним простором інтенсивностей основних кольорів.

Якщо $m_1 = m_2 = m_3 = m$, то, у залежності від величини m , виходять різні відтінки сірого кольору – від чорного до білого.

Якщо $m_1 \neq m_2 \neq m_3$, то зображення забарвлене у різні кольори.

При космічній та аерозйомці застосовують багатоспектральну фото-, теле- і скануючу апаратуру.

Множина значень функції f – це n -мірний простір інтенсивностей $m_1, m_2, m_3, \dots, m_n$, де n від 5 до 10.

При візуалізації таких зображень кожному вектору з цього простору можна присвоїти будь-який видимий людським оком колір, не обов'язково основний. У результаті отримаємо аеро- або космічне зображення в умовних кольорах.

За допомогою комп'ютера різним рівням яскравості напівтонового зображення можна зіставити деяку кольорову шкалу і забарвити зображення в так звані помилкові кольори;

– *бінарні (дворівневі) зображення* – множина значень функції $f \in (0,1)$. Зазвичай, 0 – це чорний колір, а 1 – білий (рис. 1.2);



Рисунок 1.2 – Бінарне (дворівневе) зображення

– *лінійні зображення* (рис. 1.3). Типовий приклад – карта ізоліній, зображення берегової лінії, різні контури. Зображення – одна крива або сімейство кривих;



Рисунок 1.3 – Лінійне зображення

– *точкові зображення* (рис. 1.4). Типовий приклад – кадр ділянки зоряного неба.

Зображення – k точок з координатами (x_i, y_i) і яскравістю $f_i, i = 1, \dots, k$.



Рисунок 1.4 – Точкове зображення

Все різноманіття задач обробки зображень можна подати як перетворення між класами зображень, а також як перетворення всередині класу.

Поняття «поліпшення якості» відносять до перетворення напівтонових або кольорових зображень, сегментація полягає у перетворенні напівтонового зображення в бінарне, лінійні зображення отримують з бінарних зображень, криві та лінії перетворюють у точкові зображення (рис. 1.5).

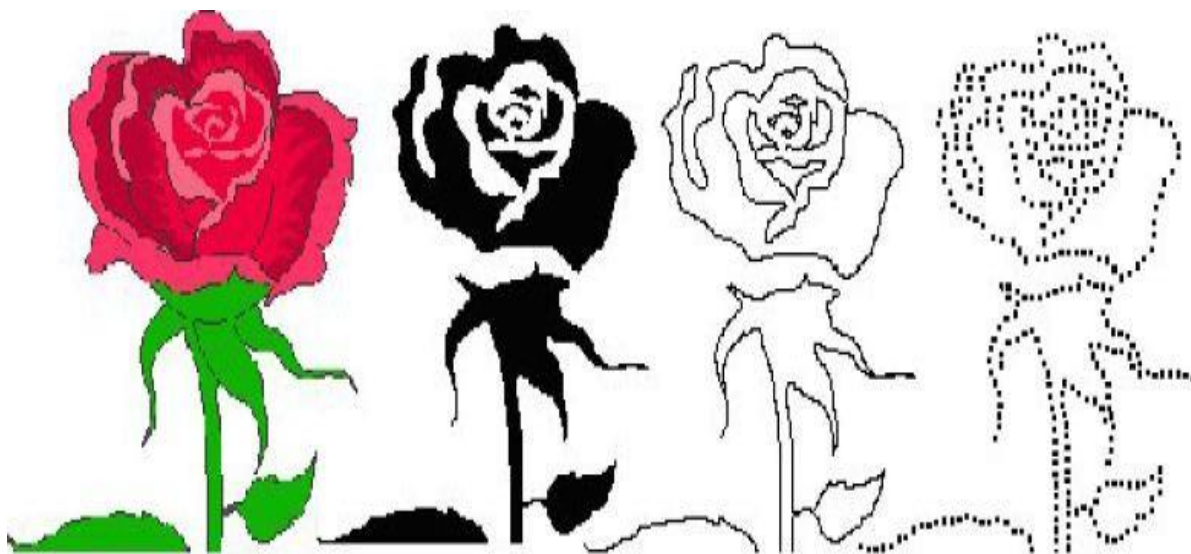


Рисунок 1.5 – Перетворення зображень

Під час обробки зображень, одержаних під час дистанційного зондування Землі з космосу, критичні чи опорні точки використовуються для поєднання зображення з картою, під час корекції геометричних спотворень зображення.

Перетворення зображення із класу з меншим номером у клас з більшим номером належить до задач *стиснення зображень та розпізнавання образів*.

Після проведення операції розпізнавання, коли об'єкт ідентифікований, часто буває необхідно провести вимірювання геометричних та інших параметрів об'єктів на зображенні.

Розглянемо різні види цифрових зображень, з якими доводиться мати справу на практиці.

За типом даних зображення поділяють на:

- бітові (булевські, логічні);
- байтові (зі знаком і без знаку);
- цілочисельні (зі знаком і без знака);
- дійсні (з фіксованою і плаваючою крапкою);
- кольорові (спеціальний тип даних);
- векторні (піксель є масивом або списком чисельних значень).

У семантичному плані про ці типи зображень можна сказати таке [8, 9]:

– *бінарні зображення* (бітові, булевські, логічні). 0 позначає піксель «фону», 1 – піксель «об'єкта», «символу» (або навпаки – залежить від знаку контрасту «об'єкт / фон»). Як вхідне бінарне зображення формує тільки один тип пристроїв введення – оптичний сканер у режимі сканування текстових документів;

– *напівтонові зображення стандартного дозволу* (байтові без знаку). Дозвіл 8 біт (діапазон значень [0 ... 255] – 256 градацій сірого) – фактичний стандарт для систем введення, здатних працювати у реальному часі. У такому форматі подаються і результати фільтрації напівтонових зображень. Системний тип даних TBitmap підтримується операційною системою Windows як частина графічного інтерфейсу на системному рівні;

– *результати обробки напівтонових зображень*, в яких можливі від'ємні значення пікселів (байтові зі знаком).

Коротке ціле зі знаком використовується для подання проміжних результатів обробки з метою економії використаного об'єму пам'яті.

Довге ціле зі знаком стандартної і подвійної довжини використовується при реалізації всіх можливих цілочислових операцій над зображеннями, якщо немає необхідності в спеціальному обмеженні потрібних об'ємів пам'яті.

Дійсні значення пікселів використовуються під час реалізації різних нецілочисельних операцій і перетворень над зображеннями. Якщо потрібно обмежити об'єм задіяної пам'яті або збільшити швидкодію алгоритмів, замість подавань з плаваючою точкою використовуються подання дійсних чисел з фіксованою точкою;

– *напівтонові зображення підвищеного дозволу* (цілочисельні без знаку). Використовуються у тих випадках, коли діапазон у 256 градацій сірого не дозволяє відобразити все багатство початкової інформації;

– *позначкові зображення* – використовуються під час автоматичного виділення зв'язкових областей і об'єктів. Кожен піксель такого зображення позначений номером області, якій він належить. Байтові зображення тут не підходять, так як у них можна закодувати тільки 255 різних областей, а на зображеннях високого дозволу їх може виявитися десятки і навіть сотні тисяч;

– *двовимірні частотні характеристики* – комплексні зображення, що складаються з дійсної та уявної частини. Формуються у результаті двовимірного косинусного перетворення, двовимірного перетворення Фур'є, швидкого перетворення Фур'є, перетворення зображення з просторової області у частотну. На програмному рівні реалізуються не як двовимірний масив комплексних чисел (двокомпонентних векторів), а як пара двовимірних масивів, один з яких представляє дійсну частину образу, а другий – уявну;

– *кольорові зображення* – спеціальний тип даних, запис формату TcolorRef = {Red, Green, Blue}. Дозвіл по кожному з каналів – 8 біт.

Кольорове зображення – системний тип даних, він підтримується усіма пристроями введення кольорових зображень. Крім того, стандартний тип даних TRGBBitmap підтримується операційною системою Windows як частина графічного інтерфейсу на системному рівні;

– *багатозональні і гіперспектральні зображення* – векторні, піксель є масивом цілочисельних значень. Формуються спеціальними пристроями введення. Використовуються для попінкельної класифікації та сегментації зображень. На програмному рівні, як правило, реалізуються не як двовимірний масив векторів, а як набір двовимірних зображень, кожне з яких відповідає одній зональній спектральній компоненті;

– *ознакові зображення* – скалярні та векторні, піксель є скаляр, масив або список дійсних значень, є результатом ознакового аналізу зображень. Використовуються для попінкельної класифікації та сегментації зображень. На програмному рівні, як правило, реалізуються не як двовимірний масив векторів, а як набір двовимірних зображень, кожне з яких відповідає одному типу ознак.

1.2 Технічні засоби обробки зображень

План

1. Введення зображень за допомогою настільного сканера.
2. Сучасні монітори.
3. Струменевий і лазерний принтери.
4. Залежність дозволу при друку від числа градацій яскравості.
5. Нейрокомп'ютери в обробці зображень.
6. Метод головних компонентів.

Введення зображень за допомогою настільного сканера

У датчиках зображень – сканерах авіаційного та космічного базування, телевізійних камерах, настільних сканерах, у цифрових фотоапаратах спосіб подання зображення називається *растровою графікою*. Дані датчики дозволяють отримувати напівтонові і кольорові зображення в аналоговій формі.

Для введення в електронно-обчислювальну машину зображення необхідно перетворити в деяку *числову матрицю*. Процес перетворення називається *дискретизацією* і складається з *вибірки* і *квантування*.

Вибірка полягає у виборі на полі спостереження деякої множини точок, у кожній з яких вимірюється величина яскравості $f(x, y)$.

Під час використання електронно-обчислювальної машини для обробки зображень нерідко застосовується формат 512×512 , згадана множина точок складається з $512 \cdot 512 = 262\,144$ пікселів.

Вибірка проводиться з деяким кроком, який і характеризує просторовий дозвіл. Під час використання *сканерів космічного базування* дозвіл, зазвичай, визначається швидкістю супутника і частотою сканування.

У даний час пристроєм введення зображень в електронно-обчислювальну машину з плоских оригіналів є планшетний сканер (Desk Scanner), він застосовується для введення топографічних карт [2].

Типовий планшетний сканер містить джерело світла, систему дзеркал, одне з них разом з лампою переміщається за допомогою крокового двигуна уздовж сканованого зображення (рис. 1.6).

Для *квантування* і *перетворення* $f(x, y)$ в числовий код у пристроях введення використовується аналого-цифровий перетворювач.

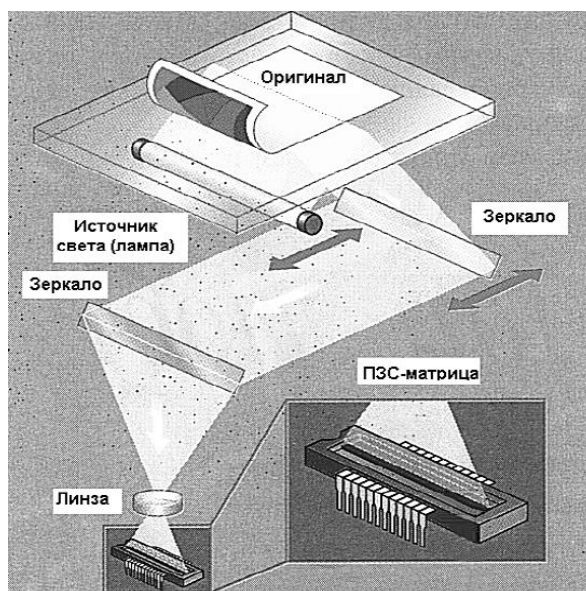


Рисунок 1.6 – Типовий планшетний сканер

Оскільки людське око не в змозі розрізняти рівні освітленості, що відрізняються один від одного менше, ніж на 1% чи 2%, для якісного подання півтонування цілком достатньо мати 256 рівнів квантування, тобто по 1 байту на піксель. Ця величина характеризує *тоновий дозвіл*.

Під час *дискретизації кольорового зображення* витрачають по 1 байту на кожен із трьох основних кольорів, тобто 3 байта на піксель (24-бітове кодування, 24-бітна глибина кольору).

Чорно-білі сканери реєструють не менше 256 градацій сірого, кольорові – не менше 256 градацій кожного з трьох основних кольорів, тобто не менше 16700000 відтінків. У деяких системах, наприклад в сканерах космічного базування, число рівнів квантування вибирається, виходячи з конкретного завдання. До числа пристроїв, спеціально призначених для введення зображень в електронно-обчислювальну машину, відносяться *цифрові фото-і відеокамери*.

Типова цифрова фотокамера містить аналого-цифровий перетворювач матрицю з роздільною здатністю не менше 5 мегапікселів та процесор, зображення у цифровій формі записується на напівпровідникову пам'ять.

До електронно-обчислювальної машини зображення вводиться або через кабель, або через безконтактний пристрій введення, або через флеш-карту.

За якістю зображення тверді копії невеликого формату, отримані за допомогою цифрової камери, не поступаються звичайній фотографії розміром 9×13см, однак, 35-міліметрові фотонегативи допускають 10-кратне збільшення без істотної втрати якості. На збільшеному в 3 або 4 рази зображенні з цифрової камери проглядається піксельна структура.

Сучасні монітори

Для візуалізації зображень використовуються кольорові монітори на електронно-променевих трубках, рідкокристалічних плазмових дисплеях.

Рідкокристалічні дисплеї менше стомлюють зір, ніж електронно-променеві. Рідкокристалічний дисплей має дві тонкі пластини з кварцового скла, на внутрішній стороні яких нанесені матриці з електродів, кожна чарунка відповідає одному пікселю.

Наприклад, монітори формату 1024×768 містять 1024 елемента по горизонталі і 768 по вертикалі. Якщо користувач побажає перевести цей монітор у режим 640×480, то зображення буде виводитися на середину екрану, а по краях залишатимуться темні поля.

Можливе застосування інтерполяції, тоді зображення займає весь екран, проте воно буде виглядати не дуже якісним, наприклад, літери можуть мати не плавні обриси, а ступінчасті. Рідкокристалічні монітори дають більш бляклі кольори, ніж електронно-променеві.

Плазмові екрани поєднують технології електронно-променевих трубок і рідкокристалічних екранів, вони складаються з окремих чарунків, у яких знаходиться інертний газ при низькому тиску, а на стінку, повернену до глядача, нанесено люмінофор.

Для професійної обробки зображень використання рідкокристалічних моніторів небажано.

Струменевий і лазерний принтери

Тверді копії отримують за допомогою принтерів і плотерів.

Матричні принтери мало придатні для виведення зображень.

Струменеві принтери забезпечують значно кращу якість при тій же ціні, що і матричні.

Сучасні струменеві принтери друкують від 2 до 10 сторінок формату А4 за хвилину з роздільною здатністю до 1400 точок на дюйм. Деякі струменеві принтери реалізують таку ж якість зображень, яка досягається на кольорових фотографіях невеликого розміру. Практично всі струменеві принтери дозволяють друкувати кольорові зображення.

Більш досконалі *лазерні принтери*.

Спеціальний світлочутливий барабан заряджають статичною електрикою, через барабан «прокочується» аркуш паперу, тонер переходить на папір і закріплюється нагріванням.

Лазерні принтери можуть забезпечувати більш високий дозвіл, ніж струменеві.

Залежність дозволу під час друку від числа градацій яскравості

Зазначений вище дозвіл реалізується під час друку не напівтонових, а бінарних зображень, до числа яких можна віднести літери тексту, чорні лінії на білому фоні. Для друку напівтонових зображень, як правило, доводиться імітувати півтони за допомогою створення растру. Растр півтонування може друкуватися у вигляді чарунків 8×8 пікселів. Залежно від заповнення чарунку точками чорного кольору можна імітувати 64 градації півтонів. Незаповнений чарунок має білий колір, заповнений 8 точками, що імітує 12%-й рівень почорніння, заповнений 32 точками – 50%-й рівень.

Отже, при друці зображень дозволу принтера по півтонам та за елементами залежать один від одного. Якщо дозвіл принтера складає 600 точок на дюйм, то імітують 64 півтони, маємо $600 / 8 = 75$ чарунків на дюйм.

Кращого результату можна досягти, якщо відмовитися від строгої відповідності між напівтоновими чарунками та друкованими точками, тобто використовувати так званий дифузійний метод друку.

Нейрокомп'ютери в обробці зображень

У 90-х роках було розроблено нейромережеві модулі обробки інформації, що дають можливість поліпшити характеристики систем розпізнавання образів і систем пошуку інформації у відповідності з інтересами користувача. Створено також апаратні засоби підтримки нейромережних модулів – нейрокомп'ютери.

Розроблені нейромережеві модулі використані для створення системи пошуку інформації, яка самостійно будує модель інтересів користувача і з поточної інформації вибирає таку, що найбільше його цікавить. Модулі, що містять нейромережеві класифікатори, були використані для створення систем розпізнавання рукописних слів.

Відбувається нове підвищення складності об'єктів обробки інформації, з якими людина може лише шляхом підвищення рівня інтелектуалізації технічних засобів. Це спричиняє необхідність розвитку таких кібернетичних понять, як самоорганізація, самокерування, розумна, адаптивна поведінка, самовдосконалювальні та генераційні системи. Дедалі більшого значення набувають методи та інструментарій системного аналізу.

Метод головних компонентів

До недавнього часу метод головних компонент вважали різновидом факторного аналізу. Сьогодні даний метод відносять до групи самостійних статистико-математичних методів багатомірного аналізу. Виникла необхідність у стисненні інформації, тобто описати досліджуване явище (об'єкт) більш укрупненими показниками, так званими «головними компонентами».

Вихідним ступенем тут є кореляційна матриця, на підставі якої з використанням методу головних компонент може бути продовжено аналіз значень спостережуваних ознак. Правильно відібрані у кореляційну модель ознаки пов'язані між собою. Наявність таких зв'язків між ними дозволяє на основі одного фактора мати інформацію про інший. Існування тісного зв'язку між ознаками дає підставу для виключення однієї з них. Ідея обліку однієї ознаки на підставі другої лежить в основі методу головних компонент.

З двома ознаками метод головних компонент малоефективний. Його використовують, як правило, при десятках взаємопов'язаних ознак, ставиться мета «набрати» певну частину загальної варіації результативної ознаки мінімальною кількістю змінних, які підбирають до тих пір, поки сума їх дисперсій не досягне заданої частки у дисперсії досліджуваного явища.

Метод головних компонентів розв'язує такі завдання:

- відшкодування скритих закономірностей у зміні явищ;
- характеристика явища, що вивчається. Число головних компонентів, виділених у процесі дослідження, буде вміщувати у компактній формі більше інформації, ніж початково-виміряні ознаки;
- виявлення ознак, найбільш тісно пов'язаних з головною компонентою. Інакше кажучи, вивчення стохастичного зв'язку між ними (зв'язок, при якому зі зміною однієї змінної змінюється закон розподілу другої);
- прогнозування рівнів досліджуваних явищ на підставі рівняння регресії, яке одержане по інформації головних компонентів.

Переваги такого методу прогнозування на відміну від класичного регресійного аналізу можна пояснити тим, що в модель намагаються включити максимально можливу кількість факторів, які в економічних явищах часто характеризуються істотною корельованістю (мультилінеарністю). Прогноз за такими змінними, як правило, буває не точним. Тому виникає завдання про заміну вихідних взаємопов'язаних змінних сукупністю некорельованих параметрів. Це завдання вирішується математичним апаратом – методом головних компонентів, який є характеристиками, що побудовані на підставі первинно виміряних ознак.

Недоліком методу головних компонентів слід назвати складність математичного апарату, зумовлену абсолютністю знань теорії ймовірностей, математичної статистики, лінійної алгебри, а також математичного забезпечення електронно-обчислювальної машини. Формальне використання стандартних програм без розуміння математичної суті обчислювальних процедур може призвести до необґрунтованих висновків.

2 МЕТОДИ КЛАСИФІКАЦІЇ ЦИФРОВИХ ЗОБРАЖЕНЬ

2.1 Параметричні та непараметричні методи класифікації цифрових зображень

План

1. Роль нормального розподілу в задачах класифікації.
2. Метод максимальної правдоподібності.
3. Метод мінімальних відстаней.
4. Метод паралелепіпедів.
5. Особливості непараметричних методів класифікації.
6. Робастні алгоритми.
7. Рангова статистика.
8. Двовибірний алгоритм Вілкоксона.
9. Декореляції фону.

Роль нормального розподілу в задачах класифікації

Під час дистанційного зондування яскравість кожного пікселя складається з яскравостей природних об'єктів, що потрапили під миттєве поле зору сканера. Об'єкти випадковим чином відображають і розсіюють сонячне випромінювання, так що яскравість пікселя є сумою випадкових величин.

З теорії ймовірностей відомо, що сума незалежних, однаково розподілених випадкових величин має нормальний (гаусівський) закон розподілу (центральна гранична теорема).

Таким чином, є підстави вважати, що для природних об'єктів яскравість пікселів може підпорядковуватися нормальному закону розподілу.

Названа вище причина не єдина, за якою у задачах параметричного розпізнавання зображень широко використовується гіпотеза про нормальний закон розподілу яскравості пікселів.

Під час розпізнавання монохроматичних зображень для розподілу яскравості майже завжди вдається підібрати один з відомих одновимірних законів (нормальний, біноміальний, експонентний, пуасонівський).

Інша справа – зображення, одержані за допомогою сканерів, що мають n -спектральних каналів, $n > 1$. У цьому випадку для опису статистичних характеристик потрібно n -мірний спільний розподіл величин яскравості в каналах.

n -мірна нормальна функція щільності ймовірностей яскравості має вигляд

$$w_n(f_{ij}) = (2\pi)^{-\frac{n}{2}} |R|^{-\frac{1}{2}} \cdot \exp \left[-\frac{1}{2} (f_{ij} - \mu)^T \cdot R^{-1} (f_{ij} - \mu) \right]. \quad (2.1)$$

У (2.1) вектор даних, що характеризує яскравість (i, j) -го пікселя у всіх n -спектральних каналах, позначений через f_{ij} , вектор математичного очікування – через μ , кореляційна матриця – через R , $|R|$ – визначник кореляційної матриці, R^{-1} – матриця, обернена до R , символ $(f_{ij} - \mu)^T$ означає транспонування вектора [8].

$$f_{ij} = \begin{pmatrix} f_{1ij} \\ f_{2ij} \\ \dots \\ f_{nij} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \dots \\ \mu_m \end{pmatrix}, \quad R = \begin{pmatrix} R_{11} & R_{12} & \dots & \dots & R_{1n} \\ R_{21} & R_{22} & \dots & \dots & R_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ R_{n1} & R_{n2} & \dots & \dots & R_{nn} \end{pmatrix}. \quad (2.2)$$

Компоненти вектора $f_{ij} : f_{1ij}, f_{2ij}, \dots, f_{nij}$ – яскравість пікселя у першому, другому та інших каналах. Компоненти вектора $\mu : \mu_1, \mu_2, \dots, \mu_n$ – математичне очікування яскравості пікселя для першого, другого та інших каналів.

На головній діагоналі кореляційної матриці (2.2) на перетині n -го рядка та n -го стовпця стоїть величина $R_{nn} = \sigma_p^2$, що показує значення дисперсії яскравості пікселів у n -му каналі (характеристика контрасту).

$R_{ns} = R_{sn}$ – це коефіцієнт кореляції значень яскравості пікселів у n -му та s -му каналах.

Методи класифікації, засновані на нормальному законі розподілу, задовільно працюють навіть при помітних відхиленнях від нормальності.

Нехай супутникове сканерне зображення земної поверхні, що отримане в n -спектральних каналах, є сукупністю пікселів, яскравість (i, j) -го пікселя описується вектором f_{ij} з n компонентів. Потрібно віднести цей піксель до одного з M класів. Попередньо створюється *навчальна вибірка* – супутникове зображення, на якому по наземним даним визначені ділянки, що відповідають різним класам. Користувач вказує межі цих класів, відбувається процес навчання, тобто комп'ютер оцінює вектор математичного очікування і кореляційну матрицю для кожного із класів.

Далі проводиться обробка даного супутникового зображення. Комп'ютер для кожного класифікованого пікселя визначає вектор вимірювань f_{ij} і в кожному k класі формує умовні щільності ймовірностей $w\left(\frac{f_{ij}}{C_k}\right)$.

Для цього у багатовимірний нормальний закон розподілу (2.1) підставляються μ_k і R_k – задані умови, тобто у даному випадку оцінки вектора математичного очікування і кореляційної матриці еталонного об'єкта « k ».

Таким чином, для кожного пікселя формується M умовних щільностей ймовірностей, які порівнюються з деяким порогом Q_{kl} . Порівняння дозволяє визначити, яка з гіпотез найбільш правдоподібна.

Відношення правдоподібності L_{kl} для класів k та l має вигляд [8]:

$$L_{kl} = \frac{w\left(\frac{f_{ij}}{C_k}\right)}{w\left(\frac{f_{ij}}{C_l}\right)} = \frac{|R_k|^{-1} \exp\left[-\frac{1}{2} (f_{ij} - \mu_k)^T R_k^{-1} (f_{ij} - \mu_k)\right]}{|R_l|^{-1} \exp\left[-\frac{1}{2} (f_{ij} - \mu_l)^T R_l^{-1} (f_{ij} - \mu_l)\right]}. \quad (2.3)$$

Замість відношення правдоподібності L_{kl} можна використати будь-яку монотонну функцію від нього, наприклад, натуральний логарифм $\ln L_{kl}$:

$$\ln L_{kl} = \frac{1}{2} \ln \frac{|R_l|}{|R_k|} + \frac{1}{2} \left[(f_{ij} - \mu_k)^T R_k^{-1} (f_{ij} - \mu_k) \right] - \frac{1}{2} \left[(f_{ij} - \mu_l)^T R_l^{-1} (f_{ij} - \mu_l) \right]. \quad (2.4)$$

У цьому випадку величина $\ln L_{kl}$ порівнюється з порогом $q_{kl} = \ln Q_{kl}$.

Перший член $\frac{1}{2} \ln \frac{|R_l|}{|R_k|}$ і множник $\frac{1}{2}$ у виразі для $\ln L_{kl}$ можуть бути включені до порогу q_{kl} . Таким чином, завдання з'ясування, до k -го або l -го класу відноситься (i, j) -й піксель, зводиться до порівняння квадратичних форм:

$$A = \left[(f_{ij} - \mu_k)^T R_k^{-1} (f_{ij} - \mu_k) \right], B = \left[(f_{ij} - \mu_l)^T R_l^{-1} (f_{ij} - \mu_l) \right]. \quad (2.5)$$

Нерідко вважають, що $q_{kl} = 0$. У цьому випадку вважається, що (i, j) -й піксель відноситься до k -го класу, якщо квадратична форма $A < B$, тобто коли компоненти вектора f_{ij} тісніше групуються навколо μ_k .

Квадратична форма $\left[(f_{ij} - \mu_k)^T R_k^{-1} (f_{ij} - \mu_k) \right]$ характеризує відстань класифікованого пікселя з вектором яскравості f_{ij} , вона часто застосовується у задачах розпізнавання зображень. У випадку, коли вектор яскравості пікселя f_{ij} складається з двох компонент, тобто для двох спектральних каналів, квадратична форма має вигляд (2.6).

$$P_k = \frac{1}{(1 - \rho_{12k}^2)} \left[\frac{(f_{1ij} - \mu_{1k})^2}{\sigma_{1k}^2} + \frac{(f_{2ij} - \mu_{2k})^2}{\sigma_{2k}^2} - 2\rho_{12k} \frac{(f_{1ij} - \mu_{1k})(f_{2ij} - \mu_{2k})}{\sigma_{1k}\sigma_{2k}} \right], \quad (2.6)$$

де $\rho_{12k} = \frac{R_{12k}}{\sigma_{1k}\sigma_{2k}}$ – нормований коефіцієнт кореляції двох спектральних каналів для об'єкта « k » еталонного зображення.

Метод максимальної правдоподібності

Розпізнавання за методом *максимальної правдоподібності* використовує значний обсяг статистичних даних та успішно справляється із складними завданнями. Метод максимальної правдоподібності забезпечує оптимальне розпізнавання. Однак, цей метод вимагає досить великої кількості операцій і є порівняно повільним. Бажано, щоб під час його використання враховувалися не тільки яскравості в спектральних каналах для кожного одиночного пікселя, але і його оточення. У деяких випадках розглянутий варіант методу максимальної правдоподібності за якістю програє більш простим методам, які беруть до уваги взаємні зв'язки сусідніх пікселів.

Квадратичну форму $\left[(f_{ij} - \mu_k)^T R_k^{-1} (f_{ij} - \mu_k) \right]$ можна спростити, якщо покласти $R_{mnk} = 0$, $m \neq n$, тобто не враховувати кореляцію яскравостей пікселя у спектральних каналах. У (2.6) залишаться тільки два члени, множник перед квадратними дужками стане рівним 1.

У разі двох класів k та s отримуємо такі вирази [8]:

$$r_k^2 = \sum_{m=1}^n \frac{(f_{mij} - \mu_{mk})^2}{\sigma_{mk}^2}, \quad r_s^2 = \sum_{m=1}^n \frac{(f_{mij} - \mu_{ms})^2}{\sigma_{ms}^2}, \quad (2.7)$$

де m – номер спектрального каналу;

n – загальне число каналів;

r_k^2 – квадрат відстані в просторі яскравості по евклідовій метриці між класифікованим пікселем з вектором f_{ij} і вектором середнього значення μ_k еталонного об'єкта « k ».

Під час використання (2.7) вважається, що піксель відноситься до того класу, де менший один із двох r_k^2 чи r_s^2 .

Точність r_k^2 істотним чином залежить від точності оцінювання σ_{mk}^2 , як і в методі максимальної правдоподібності.

Метод мінімальних відстаней

Подальше спрощення призводить до методу *мінімальних відстаней*, коли вирішення про те, до якого класу належить піксель, приймається на підставі порівняння відстаней між пікселями і середніми значеннями в просторі яскравості без урахування вектора дисперсії яскравості еталонного об'єкта:

$$r_k^2 = \sum_{m=1}^n (f_{mij} - \mu_{mk})^2. \quad (2.8)$$

Це самий простий і швидкий метод класифікації, він вимагає мінімальних відомостей про класи, але нерідко поступається методу максимальної правдоподібності у точності класифікації. Порівняння з результатами класифікації по методу максимальної правдоподібності показує, що у даному випадку краще всього виділилися згарища та інші елементи ландшафту, позбавлені рослинності, гірше виділилися листяні дерева.

Метод мінімальних відстаней не враховує оточення пікселя, тому іноді може поступатися методам кластерного аналізу, методам сегментації по нарощуванню областей та по виділенню кордонів.

Метод мінімальних відстаней заснований на евклідовій метриці та не пов'язаний з нормальним законом розподілу.

Метод паралелепіпедів

На практиці знаходить застосування ще один простий метод класифікації, пов'язаний з нормальним законом, – метод паралелепіпедів (метод прямокутників). Він використовує відомості про класи у вигляді векторів середніх значень яскравості μ_k і векторів дисперсій σ_{mk}^2 , що отримані у процесі навчання. Відомо, що для нормально-розподіленої випадкової величини 95,4% її значень лежать у межах відхилень від середнього значення, менших 2σ .

Розглянемо випадок двох спектральних компонент.

На рисунку 2.1 по вертикальній осі відкладені значення яскравості f_1 у першому каналі, по горизонтальній осі – значення яскравості f_2 у другому каналі.

Зазначені середні значення яскравості μ_{11} і μ_{12} для першого і другого класів у першому каналі, μ_{21} і μ_{22} для першого і другого класів у другому каналі. Від значення середньої яскравості в обидві сторони відкладені значення середньоквадратичного відхилення σ_{11} , σ_{12} , σ_{21} , σ_{22} для першого і другого класу в кожному із двох каналів.

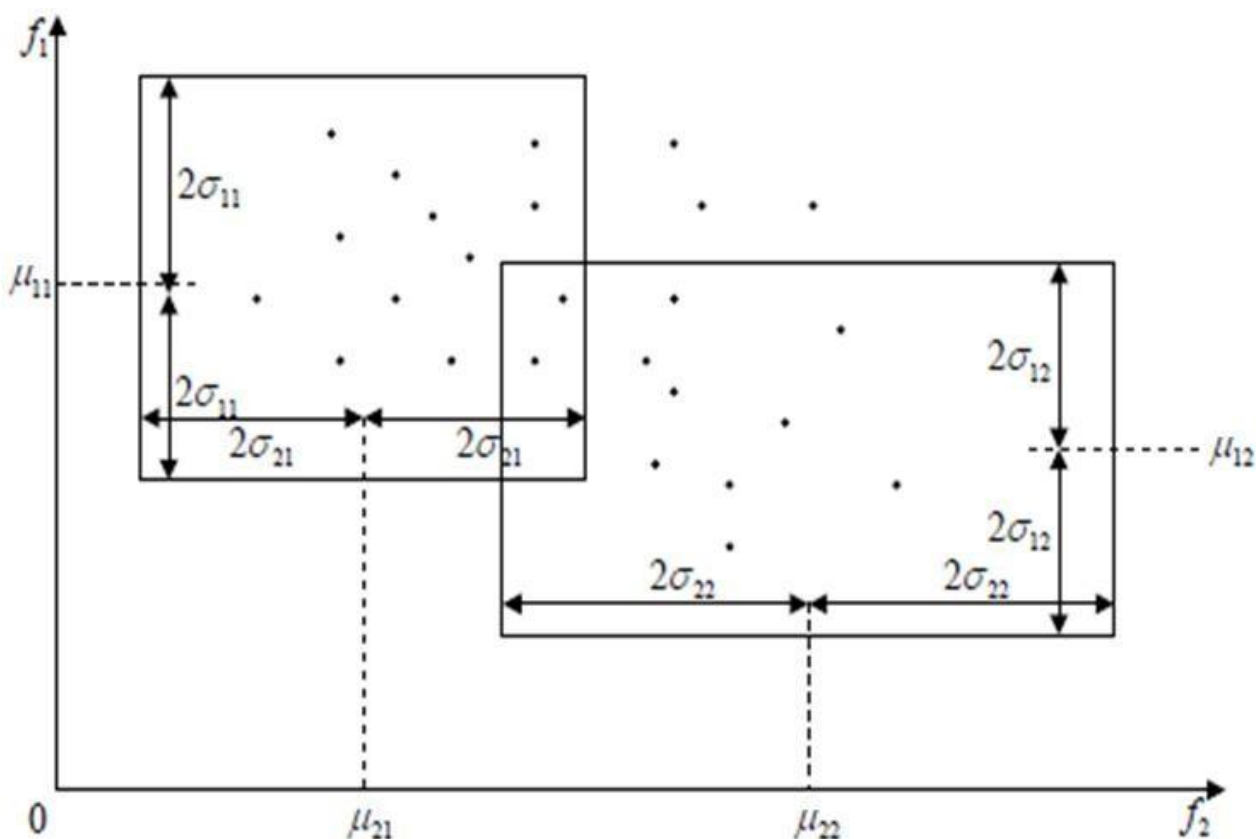


Рисунок 2.1 – Дві спектральні компоненти

У двовимірному випадку отримано прямокутники, у три- і більше мірному випадку маємо паралелепіпеди у відповідному просторі.

Якщо компоненти вектора яскравості f_{1ij} та f_{2ij} такі, що потрапляють в лівий прямокутник, приймається рішення про належність пікселя до першого класу, якщо вони потрапляють у правий прямокутник, то до другого.

Прямокутники можуть частково перекриватися, особливо це характерно для областей, що примикають до кутів. У цьому випадку точки виявляються нерозпізнаними.

Метод паралелепіпедів відноситься до числа швидких методів класифікації, однак, за якістю розпізнавання він поступається більшості з розглянутих методів, як видно із зображення.

Метод паралелепіпедів найдоцільніше використовувати для попередньої класифікації, він найбільш ефективний при нормальному законі, так як для нього виконується «правило двох сигма».

Особливості непараметричних методів класифікації

Параметричні методи статистичної класифікації зображень є потужним засобом тематичної обробки. Специфіка дистанційного зондування в оптичному діапазоні полягає у тому, що зображення подане у вигляді декількох шарів, кожен з яких отримано в одному із спектральних каналів [10].

Для використання параметричних методів необхідно знати багатовимірні функції розподілу ймовірностей, пов'язані з класами образів. Часто ці функції не відомі і повинні оцінюватися по множині навчальних образів. Використовується багатовимірний нормальний закон, хоча не відомі докази, що багатовимірні функції розподілу ймовірностей яскравостей багатоспектральних зображень мають саме такий вигляд.

В одновимірному випадку нормальний закон розподілу несправедливий хоча б через те, що яскравість – величина невід'ємна. У разі прийняття гіпотези про нормальність розподілу по навчальним образам необхідно оцінити тільки окремі параметри, пов'язані з цими функціями (такі як математичні сподівання, дисперсії, функції кореляції). Параметричні методи, зазвичай, легше реалізуються, але вимагають більшого обсягу апріорної інформації або фундаментальних припущень щодо природи образів. Якщо форма функцій розподілу ймовірностей не відома заздалегідь, метод є *непараметричним*.

Непараметричні методи мають великі потенційні можливості для точної оцінки функцій розподілу ймовірностей і для точного розпізнавання, але ця перевага, зазвичай, вимагає складних систем, що розпізнають, великого числа навчальних образів і, головне, великих часових витрат. Непараметричні алгоритми, зазвичай, синтезують евристично, часто використовуються локальні оцінки ймовірності за емпіричною частотою.

Робастні алгоритми

Серед непараметричних алгоритмів слід виділити робастні (стійкі) алгоритми, справедливі з тією чи іншою ефективністю при різних законах розподілу. Прикладом непараметричного алгоритму є метод мінімальних відстаней, метод використовує середні значення яскравості і дисперсії або тільки середні значення яскравості об'єктів, знайдені з навчальної вибірки.

При цьому припущення про функції розподілу не використовуються. Головне, щоб середні значення яскравості і дисперсії існували.

Прикладом непараметричного алгоритму, що використовується під час придушення шуму у вигляді крапок і ліній на зображенні, є алгоритм *медіанної фільтрації*.

Рангова статистика

Розглянемо типовий робастний алгоритм, так званий ранговий, який враховує відхилення елементів даної вибірки від елементів випадкової вибірки. Наприклад, є зображення лісу з дозволом 30 м, на якому видні крони дерев та просіка. Потрібно виявити і виділити зображення просіки (рис. 2.2) [10].



Рисунок 2.2 – Зображення лісу з просікою

Можна використати градієнтний фільтр, наприклад, фільтр Лапласа. Просіка стає більш помітною, але пікселі, що відносяться до неї, оточені іншими пікселями, які можна інтерпретувати як шум (рис. 2.3) [10].

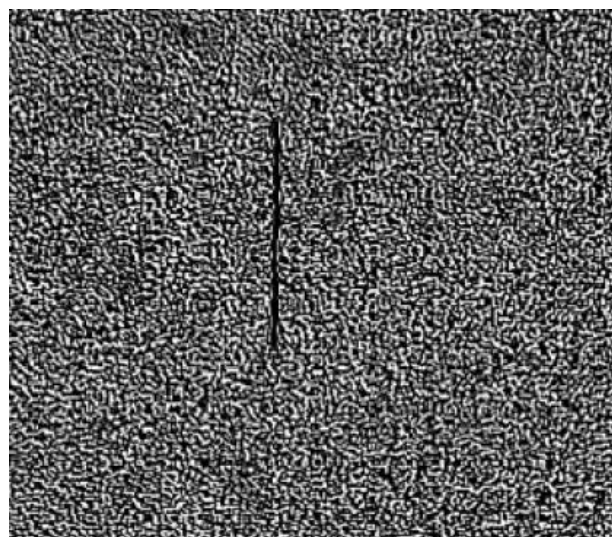


Рисунок 2.3 – Результат застосування фільтру Лапласа

Розглянемо один рядок зображення, який показано на рисунку 2.4.

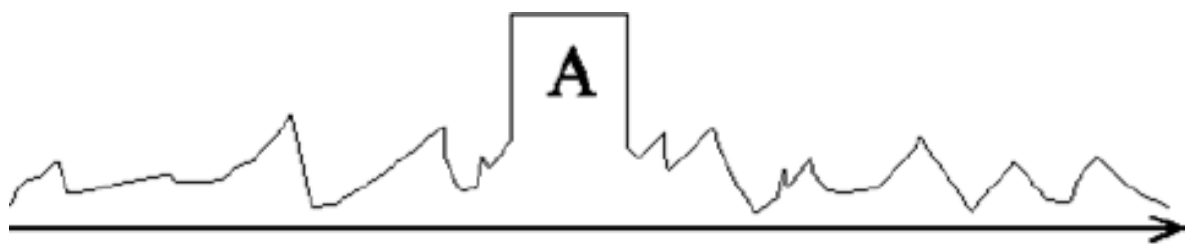


Рисунок 2.4 – Рядок зображення, що містить пікселі об'єкта та шуму

Прийmemo, що A – пікселі об'єкта (просіки), x_i – вибіркoві значення яскравості пікселів (табл. 2.1).

Таблиця 2.1 – Вибіркові значення яскравості пікселів

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
4	3	13	8	5	18	9	6	11	15

Рангом R_i елемента вибірки x_i називається число елементів n цієї вибірки, менших або рівних x_i . Якщо елементи вибірки розташувати у порядку зростання від меншого до більшого (сформувати варіаційний ряд за зростанням), то отримаємо таблицю 2.2.

Таблиця 2.2 – Вибіркові значення яскравості пікселів

R_i	1	2	3	4	5	6	7	8	9	10
	3	4	5	6	8	9	11	13	15	18
x_i	x_2	x_1	x_5	x_8	x_4	x_7	x_9	x_3	x_{10}	x_6

Ранг R_9 елемента x_9 дорівнює 7, ранг R_2 елемента x_2 дорівнює 1, ранг R_6 елемента x_6 дорівнює 10.

Коли спостерігається тільки шум, враховуючи незалежність та однорідність значення яскравості пікселів, то величини рангів рівноймовірні при будь-яких функціях розподілу.

Якщо ж у вибірці є суміш шуму і деякого не випадкового зображення, то величини рангу не будуть рівноймовірними.

Двовибірний алгоритм Вілкоксона

Для виявлення локальних контурних ознак можна використати двовибірний алгоритм Вілкоксона. Формуються дві вибірки, одна з яких $X = x_1, x_2, \dots, x_n$ є опорною і належить області фону, інша $Y = y_1, y_2, \dots, y_m$ – робоча. Вхідні в Y елементи можуть бути як пікселями фону, так і пікселями шуканого зображення, тобто до складу вектора Y включені ті пікселі, які відповідають очікуваному положенню об'єкта, що цікавить (табл. 2.3).

Таблиця 2.3 – Двовибірний алгоритм Вілкоксона

Без локальних контурних ознак										
R_i	1	2	3	4	5	6	7	8	9	10
x_i	3	4	5	6	8	9	11	13	15	18
З локальними контурними ознаками										
R_i	1	2	3	7	8	9	10	11	12	13
y_i	3	4	5	11	13	15	18	24	24	24

Складаються два варіаційних ряди – один для елементів опорної X вибірки, а другий для робочої Y вибірки. Ранги елементів робочої вибірки визначаються з урахуванням рангів елементів опорної вибірки, тобто для тих елементів робочої вибірки y_i , які збігаються (або близькі) до елементів опорної вибірки x_i , зберігаються ті ж самі ранги. Значення яскравості y_i пікселів, що відносяться до області A , перевищують x_i , тому вони потраплять у кінець варіаційного ряду з новими (більшими) номерами (рангами), а у загальній послідовності номерів відбудеться стрибок, зникнуть ранги 4, 5, 6.

Далі визначається рангова статистика $R = \sum R_i$, порівняння якої з порогом R_0 призводить до вирішення про наявність чи відсутність локальних контурних ознак згідно правила:

$$R \geq R_0 - \text{локальні контурні ознаки є,} \quad (2.9)$$

$$R < R_0 - \text{локальні контурні ознаки відсутні.} \quad (2.10)$$

У розглянутому прикладі $R = 55$, якщо локальні контурні ознаки відсутні і $R = 76$, якщо локальні контурні ознаки є. При відсутності локальних контурних ознак елементи обох векторів X та Y утворюють однорідну множину випадкових величин, що підкоряються одному і тому ж закону розподілу $w(x)$.

У цьому випадку елемент робочої вибірки з рівними шансами займає будь-яке положення у варіаційному ряді, а це означає, що випадкова величина R_i підпорядковується рівноймовірному закону розподілу:

$$w(R_i) = \frac{1}{(n+1)}, \quad R_i = 0, 1, 2, \dots, n. \quad (2.11)$$

На даний розподіл не робить впливу конкретний вид розподілу некорельованого фону $w(x)$.

Закон розподілу $w(R)$ вирішальної статистики R при фактичній відсутності в Y яскравісного перепаду не залежить від розподілу фону $w(x)$.

Отже, від $w(x)$ не залежить і ймовірність помилкового виявлення локальних контурних ознак, яким відповідає верхня умова у процедурі $R \geq R_0$.

Для практичного застосування особливо важливо, що і вибір порогу рангової процедури R_0 , який підпорядковують вимозі отримання бажаної величини ймовірності помилкового виявлення, також не пов'язаний із розподілом фону $w(x)$.

Таким чином, застосування рангового алгоритму автоматично стабілізує ймовірність помилкового виявлення, роблячи її незалежною від виду та параметрів закону розподілу фону. Ця властивість досягається тільки за умови, що всі випадкові величини, що утворюють вибірки X та Y , незалежні. В іншому випадку в закон розподілу $w(R)$ будуть входити параметри, що описують залежність елементів.

При наявності об'єкта A елементи вектора Y статистично перевищують елементи X , тому закон розподілу рангу $w(R_i)$ перестає бути рівномірним: імовірності великих значень R_i збільшуються за рахунок зменшення ймовірностей малих значень. Ця тенденція проявляється тим сильніше, чим більша величина яскравісного перепаду.

Декореляції фону

На зображенні ділянки лісу з просікою яскравості пікселів фону залежні статистично, це типово для багатьох зображень випадкового фону.

Тому для коректного використання рангового алгоритму необхідно перетворити зображення. Будемо вважати, що яскравості пікселів фону на рисунку 2.2 підкоряються нормальному закону розподілу, для якого незалежність і некорельованість елементів зображення еквівалентні один одному, тому перетворення зображення зводиться до декореляції фону.

Нехай фон є однорідним та ізотропним випадковим полем, у якого статистичні характеристики однакові по рядках і стовпцях.

Як модель фону (крона дерев) використовуємо чотириточкову модель:

$$f(i, j) = \sigma [f(i-1, j) + f(i, j-1) + f(i+1, j) + f(i, j+1)] + z(i, j), \quad (2.12)$$

де i та j – номери рядків і стовпців;

σ – параметр, що характеризує ступінь впливу сусідніх відліків на даний відлік поля $f(i, j)$;

$z(i, j)$ – випадкове поле з нормальним законом розподілу та некорельованими відрахунками.

Для декореляції фону використовуємо таку процедуру [11]:

$$g(i, j) = f(i, j) - \sigma [f(i-1, j) + f(i, j-1) + f(i+1, j) + f(i, j+1)] = z(i, j), \quad (2.13)$$

$$\begin{vmatrix} 0 & -\sigma & 0 \\ -\sigma & 1 & -\sigma \\ 0 & -\sigma & 0 \end{vmatrix}.$$

Декореляція здійснюється шляхом обробки зображення ковзаючим вікном та враховує конфігурацію сусідніх підрахунків.

Результат цієї обробки при відсутності локальних контурних ознак збігається з породжуючим процесом $z(i, j)$, який за властивостями своєї кореляційної функції близький до білого шуму. Для того, щоб при виявленні локальних контурних ознак використовувати незалежну вибірку, її відліки слід розташовувати так, щоб їх взаємна кореляція дорівнювала нулю.

Вид сигналу, що пройшов через процедуру декореляції, залежить від значень параметра σ , тому зупинимося на їх визначенні. Прийнемо умову однаковості характеристик зображення по рядках і стовпцях. Параметр σ бере участь у процедурі декореляції, тому його вибір необхідно підпорядкувати мінімізації помилок декореляції [12].

$$\langle [g(i, j) - z(i, j)]^2 \rangle, \quad (2.14)$$

враховуючи, що

$$\langle z(i, j) \rangle = 0, \quad \langle |z(i, j)|^2 \rangle = \sigma_z^2, \quad \langle [f(i, j) - \langle f(i, j) \rangle]^2 \rangle = \sigma_f^2,$$

причому коефіцієнт кореляції між сусідніми відліками дорівнює (2.15).

$$\langle (f(i,j) - \langle f(i,j) \rangle) \cdot (f(i,j+1) - \langle f(i,j) \rangle) \rangle = \sigma_f^2 \cdot \rho, \quad (2.15)$$

де ρ – нормований коефіцієнт кореляції між сусідніми пікселями фону.

Як модель коефіцієнта кореляції приймемо вираз

$$\langle (f(i,j) - \langle f(i,j) \rangle) \cdot (f(I+n, j+m) - \langle f(i,j) \rangle) \rangle = \sigma_f^2 \cdot \rho \sqrt{n^2 + m^2}, \quad (2.16)$$

при цьому

$$\langle [z(i,j) - g(i,j)]^2 \rangle = \sigma_z^2 + 4a^2 \sigma_f^2 - 2a 4\sigma_f^2 \rho + 2 \cdot 2a^2 \sigma_f^2 \rho^2 + 2 \cdot 4a^2 \sigma_f^2 \rho^{\frac{1}{2}}. \quad (2.17)$$

Прирівнюючи нулю похідну по a лівій частині цього виразу, знаходимо:

$$a = \frac{\rho}{\left(\frac{1}{1 + 2\rho^2 + \rho^2} \right)}. \quad (2.18)$$

Процедура декореляції впливає на сигнальну (A) складову зображення. Необхідно оцінити зміну яскравісного перепаду, що відбувається під час цього впливу. Уявімо повну модель вихідного зображення $h(i,j)$ у вигляді суми безперервної (фонові) складові $f(i,j)$ і ступінчастої функції $s(i,j)$, яка описує яскравісний перепад:

$$h(i,j) = s(i,j) + f(i,j). \quad (2.19)$$

Для зображень, поданих на рисунках 2.2 та 2.3, значення коефіцієнта кореляції лежать у межах $0,9 < \rho < 1$. Цьому відповідає $0,25 < a < 0,255$.

Зміна параметру a у таких невеликих межах незначно змінює відгук оператора декореляції на яскравісний перепад $s(i,j)$. Нижче наведено переріз вихідного сигналу, розрахований при $a = 0,25$ [12]:

$$\begin{vmatrix} 0 & -0,25 & 0 \\ -0,25 & 1 & -0,25 \\ 0 & -0,25 & 0 \end{vmatrix} = 0,25 \begin{vmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{vmatrix}.$$

Лінія локальних контурних ознак перетвориться фільтром Лапласа, причому яскравість всіх елементів ділиться на 4, множник 0,25 не є принциповим, його можна упустити.

3 МАТЕМАТИЧНІ ОСНОВИ ФІЛЬТРАЦІЇ ЗОБРАЖЕНЬ

3.1 Фільтрація та відновлення зображень

План

1. Формати графічних файлів. Растрова і векторна графіка.
2. Стиснення зображень без втрат і з втратами.
3. Структура графічного файлу.
4. Моделі зображень. Каузальні, півкаузальні і некаузальні моделі.
5. Авторегресійні моделі.
6. Перетворення яскравості та контрасту.
7. Розтяжка і еквалізація гістограми.
8. Сегментація зображень. Способи сегментації.
9. Розрахунок порогу при порогової сегментації.
10. Сегментація шляхом нарощування областей.
11. Сегментація шляхом виділення кордонів.
12. Розпізнавання зображень.
13. Кластерний аналіз.
14. Теорема Байєса.
15. Помилки класифікації.
16. Геометричні перетворення та прив'язка зображень.

Формати графічних файлів. Растрова і векторна графіка

Для подання графічної інформації на двовимірній площині (екрані монітору) застосовуються два підходи: растровий та векторний.

При *векторному підході* графічна інформація описується як сукупність абстрактних геометричних об'єктів – прямих, відрізків, кривих, прямокутників. Векторний опис припускає апріорні знання про структуру зображення.

Векторна модель – лінійний список параметрів (вектор), що математично визначають графічні об'єкти, це складові синтезованого зображення (об'єктно-орієнтована модель).

Растрова модель – прямокутна матриця однакових неподільних елементів-пікселів, кожен з яких у закодованому вигляді відображає певну ділянку реального або синтезованого зображення.

Елемент растру називають *пікселем*.

Стандартна ідентифікація пікселей [8, 9]:

$$f(i, j) = (A(i, j), C(i, j)), \quad (3.1)$$

де $A(i, j) \in R^2$ – область пікселя,

$C(i, j) \in C$ – атрибут пікселя (як правило, колір).

Частіше за все використовують два види атрибутів:

– $C(i, j) = I(i, j)$ – інтенсивність (яскравість) пікселя;

– $C(i, j) = \{R(i, j), G(i, j), B(i, j)\}$ – кольорові атрибути у моделі RGB.

Растрова графіка оперує з довільними зображеннями у вигляді растрів.

Растр – це опис зображення на площині шляхом розбиття (дискретизації) його на однакові елементи по регулярній сітці і присвоєння кожному елементу свого кольору та будь-яких інших атрибутів.

З математичних позицій растр – це кусково-постійна апроксимація на площині безперервної функції зображення.

Найпростіший растр – прямокутний, найекономічніший для передачі зображень – гексагональний.

Нехай зображення містить дві прямі з відомими координатами початку і кінця кожної, як показано на рисунку 3.1.

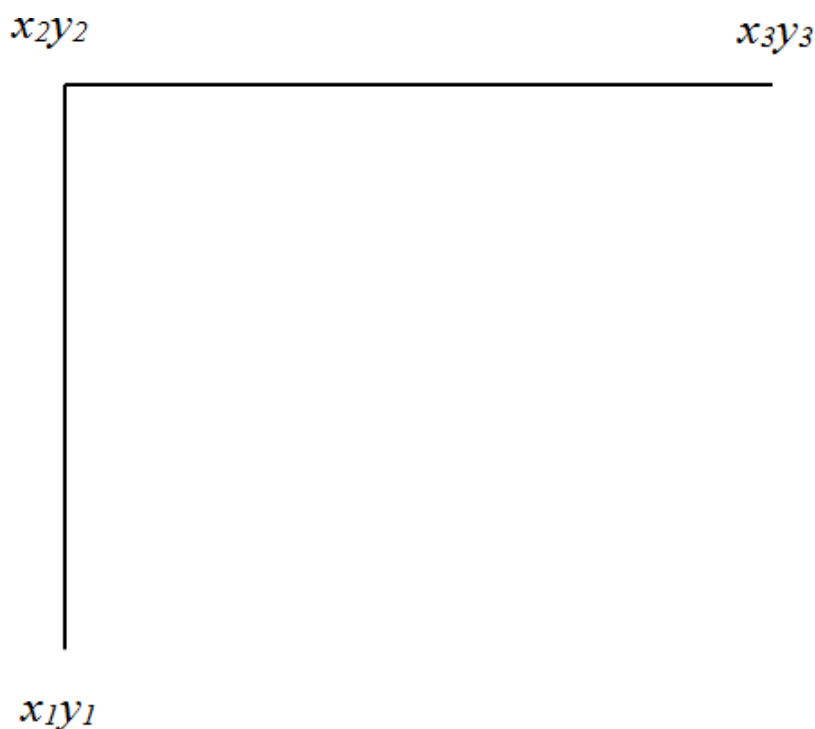


Рисунок 3.1 – Дві прямі з відомими координатами

Зображення (рис. 3.1) у векторній формі має вигляд $(x_1, y_1; x_2, y_2; x_3, y_3)$. Зображення у растровій формі, що відповідає цим прямим, представлено в умовному вигляді на рисунку 3.2.

<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>w</i>	<i>w</i>	<i>w</i>	<i>w</i>	<i>w</i>	<i>w</i>
<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>w</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>
<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>w</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>	<i>a</i>

Рисунок 3.2 – Зображення у растровій формі

Зображення (рис. 3.2), що містить дві прямі з яскравістю *w* на фоні *a* у вигляді 3 рядків по 10 пікселів, зазвичай, записується в пам'яті електронно-обчислювальної машини у вигляді: *aaaawwwwwwwaaaaawaaaaaaawaaaaa*.

Для виведення такого зображення на монітор необхідно навести дані про число рядків і пікселів у рядку.

Кожна з форм запису має свої переваги і недоліки.

Растрова форма запису більш універсальна, більш інформативна, ідеально відповідає архітектурі електронно-обчислювальної машини. Однак, якщо збільшити растрове зображення ліній, то вони стануть товщі.

Векторна графіка позбавлена цього недоліку. При будь-якому масштабі точка зображується пікселем, лінія має однакову товщину. Це надзвичайно важливо при роботі з геоінформаційними системами, коли необхідно послідовно переходити від дрібномасштабних до великомасштабних зображень, від карт місцевості до планів, що містять зображення міст, житлових кварталів та окремих будинків з різними комунікаціями, причому зображення супроводжуються текстовою та цифровою інформацією. Векторна графіка більш орієнтована на роботу з базами даних, ніж растрова.

Під час роботи з космічними зображеннями Землі доводиться одночасно використовувати растрову і векторну графіку, наприклад, накладати зображення земної поверхні на географічну карту, накладати координатну сітку на зображення. Тому в програмному забезпеченні станцій прийому та обробки космічної інформації передбачається перетворення вектор-растр.

Для збереження растрових зображень потрібний значний обсяг пам'яті електронно-обчислювальної машини.

Етапи отримання растрової моделі (рис. 3.3):

- а) дискретизація – розбиття на однакові елементи (дискрети);
- б) кодування – присвоєння дискрету параметрів, що характеризують відповідний фрагмент зображення.

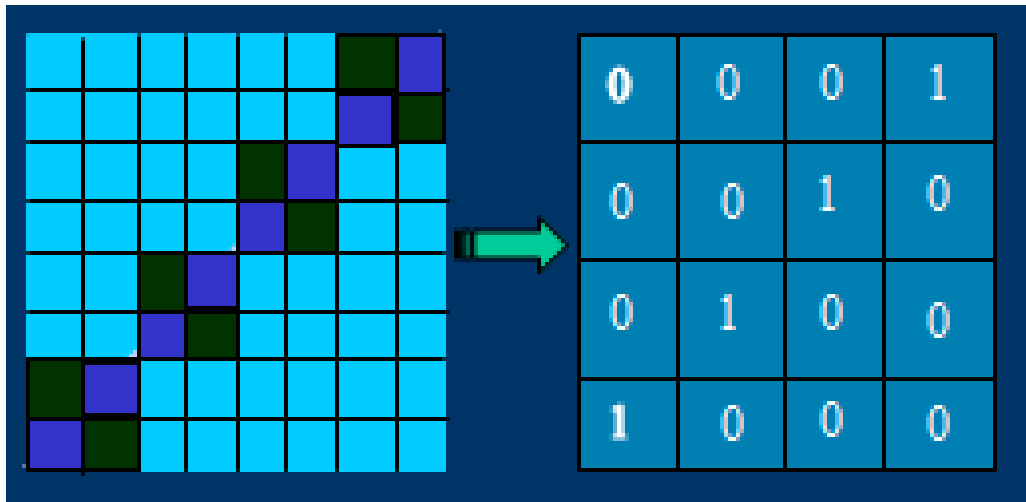


Рисунок 3.3 – Етапи отримання растрової моделі

Основні параметри растрових цифрових зображень:

- розмір – добуток ширини на висоту в пікселях (наприклад, 400×600).
- дозвіл – щільність інформації на одиницю довжини по горизонталі і вертикалі (ppi – pixels per inch – пікселів на дюйм, наприклад, 300 ppi).
- тип колірної моделі (наприклад, RGB).
- глибина кольору (біт на піксель, наприклад, 24 б/п).

Стиснення зображень без втрат і з втратами

У даний час розроблені різні методи економічного запису зображень у пам'ять електронно-обчислювальної машини та стиснення зображень, існує багато форматів запису.

Один з найпростіших методів стиснення – *групове кодування*. Відповідно до групового кодування повторювані величини замінюються однією величиною із зазначенням їх кількості. Наприклад, замість послідовності значень яскравості *abbbcccddeeeeeeeef* можна записати *1a3b3c2d9e1f*.

Цей метод легко реалізується, він добре працює з довгими серіями повторюваних величин, наприклад, при наявності великих областей з однаковою яскравістю або кольором, у цьому випадку він зручніший, ніж кодування кодами змінної довжини (статистичне кодування).

Групове кодування використовується у форматах TIF, PCX, BMP, воно не вимагає складання кодової таблиці.

Наприклад, космічне зображення природного об'єкту (місце впадання річки Селенга в Байкал) при 24-бітному поданні RGB, 780×780 пікселів має розмір 1,46 Мбайт. Групове кодування у форматі BMP стискає зображення до 479 Кбайт. Групове кодування добре справляється зі стисненням зображень об'єктів, що містять великі ділянки, які однорідні по тону (кольору).

Статистичне кодування більш ефективне, ніж групове [8, 13].

У теорії інформації доведено, що найбільшою середньою кількістю інформації на одне повідомлення володіє джерело, що видає повідомлення з випадковими, незалежними, рівномірно розподіленими значеннями.

Якщо інтервал зміни цих значень дорівнює $[a, b]$, то ймовірність появи будь-якого значення з цього інтервалу однакова, значення яскравості необхідно кодувати словами однакової довжини. З цієї точки зору, «найбільш інформативним зображенням» буде шум з рівномірним законом розподілу при статистичній незалежності яскравості всіх пікселів.

Однак, реальні зображення істотно відрізняються від випадкового двовимірного поля, вони містять фон і деякі зв'язкові об'єкти. Фон і об'єкти, як правило, мають повторювані значення яскравості, деякі значення яскравості зустрічаються часто, деякі – рідко. Реальне зображення володіє надмірністю, і якщо її усунути, для його збереження потрібний менший об'єм пам'яті. Надмірність усувається шляхом раціонального кодування, що враховує нерівну ймовірність (точніше, частоту) появи яскравості пікселів і їх взаємозалежність. Найпростіше врахувати нерівну ймовірність появи різних значень яскравості.

У даний час широко застосовується статистичне кодування *за методом Хаффмана*, для якого довжина кодового слова залежить від імовірності появи повідомлення. Кодування за Хаффманом вимагає під час першого етапу проведення аналізу зображень і визначення ймовірності (частоти) появи яскравості (або основних кольорів) пікселів зображення та складання кодової таблиці з урахуванням цих ймовірностей. Далі всім яскравостям присвоюються кодові слова з таблиці. При цьому відбувається стиснення зображення, так як для передачі цієї сукупності буде потрібно менше біт, ніж для запису рівномірним кодом [13].

Нехай растрове зображення має вигляд *abbbcccddeeeeeeeef*.

Частоти, з якими з'являються значення яскравості, рівні:

$a: 1, b: 3, c: 3, d: 2, e: 9, f: 1.$

Для кодування цих 6 значень можна використати рівномірний 3-бітний код та записати:

$a: 001; b: 010; c: 011; d: 100; e: 101; f: 110$.

Загальна довжина кодової комбінації в цьому випадку складе:

$$19 \cdot 3 = 57 \text{ (біт)}.$$

Для кодування по Хаффману використовуємо *двійкове дерево* (рис. 3.4).

Найбільш рідко зустрічаються значення у цьому прикладі – a та f , вони стають першою парою: a присвоюється 0-а гілка, f – 1-а. 0 і 1 стануть молодшими бітами кодів для a та f . Старші біти будуть отримані з дерева у міру його побудови [8].

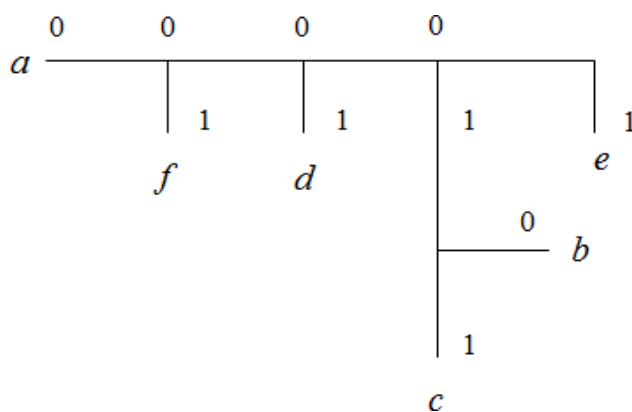


Рисунок 3.4 – Двійкове дерево

Підсумовуємо частоти a та f , у сумі отримуємо 2, ця пара далі об'єднується із символом d , який також має частоту 2.

Парі a та f присвоюється 0-а гілка цього дерева, а d присвоюється гілка 1. Код для a закінчується на 00, для f – на 01, для d – на 1.

Дерево продовжує будуватися так, що найменш поширені величини описуються більш довгими кодами, а найбільш поширені – одним (або більше) бітами. У нашому випадку маємо:

$a: 0000; b: 010; c: 011; d: 001; e: 1; f: 0001$.

Загальна довжина кодової комбінації тепер дорівнює:

$$1 \cdot 4 + 3 \cdot 3 + 3 \cdot 3 + 2 \cdot 3 + 9 \cdot 1 + 1 \cdot 4 = 4 + 9 + 9 + 6 + 9 + 4 = 41 \text{ (біт)}.$$

Коефіцієнт стиснення рівний 1,39:1.

Алгоритм Хаффмана дозволяє у певних випадках досягати стиснення більше, ніж 8:1. Однак, він вимагає два підходи: один для створення кодової таблиці, другий – для кодування. Кодування і декодування по Хаффману великих масивів інформації – порівняно повільні процеси, це характерно і для інших кодів змінної довжини.

Існують процедури, які не потребують попереднього створення кодової таблиці, у цих процедурах кодова таблиця створюється і уточнюється по мірі надходження даних – *адаптивні методи стиснення*. До числа таких процедур відноситься схема стиснення LZW (Лемпела-Зеева-Велч), яка використовується в форматах GIF та TIF. 24-бітове зображення гирла річки Селенга розміром 780×780 пікселів займає 1,46 Мбайт, а стиснене за методом LZW у форматі TIF займає 436 Кбайт проти 479 Кбайт при груповому кодуванні у форматі BMP.

Розглянуті методи не охоплюють всіх відомих схем стиснення зображень. Особливістю їх є відсутність втрати інформації при стискуванні. Це важливо при зберіганні вихідних даних.

Існують алгоритми *стиснення з втратами*, наприклад, *JPEG* (Joint Photographic Experts Group), розширення jpg. Цей алгоритм використовує ідею, реалізовану в сучасних системах кольорового телебачення (SECAM, PAL, NTSC) і застосовується у цифрових фотоапаратах та видавничих системах. Враховується, що людське око дуже чутливе до зміни яскравості зображення, але не помічає колір дрібних його деталей. У телебаченні дані про колір передаються у сильно стислому вигляді так, що кадр чорно-білого і кадр кольорового телевізійного зображення займає однакову смугу частот.

Алгоритм JPEG забезпечує дуже високий коефіцієнт стиснення, але мало придатний для збереження зображень, одержаних під час дослідження Землі з космосу і призначених для подальшої обробки на електронно-обчислювальній машині. Однак, він корисний при створенні рисунків до статей та презентацій.

Зображення гирла річки Селенга, стиснене за методом JPEG, займає всього 65 Кбайт замість 1,46 Мбайт вихідного зображення у форматі TIF.

Під час обробки зображень застосовуються дискретні ортогональні перетворення, зокрема, *вейвлет-перетворення*, на основі їх використання розроблені нові швидкі алгоритми стиснення зображень та нові формати, наприклад, *JPEG-2000*.

Структура графічного файлу

Кожний графічний файл складається з двох основних частин: заголовка та даних. Кольорові зображення містять також таблицю, відповідно до якої елементам зображення присвоюються значення основних кольорів.

Заголовок починається з ідентифікатора, який зазначає, в якому форматі записаний файл (TIF, BMP чи GIF), далі містяться загальні відомості про структуру файлу (ширина, висота зображення, чи кольорове, чи стиснене). Форма подання цих відомостей різна для кожного формату.

Найбільш розвиненим, але і найскладнішим є *TIF-формат* (Tag Image File Format). Кожна серйозна програма обробки зображень може читати і записувати TIF-файли. У цьому форматі можна зберігати всі види зображень. Поряд із основною інформацією про зображення (розміри зображення, дані про кольори), у заголовку TIF-файлу можна записати множину додаткових відомостей про зображення.

Формат файлу, придатний для обміну даними між комп'ютерами різних систем, – це *формат Targa* (TGA), він не створює ніяких проблем і практично виключає несумісність між програмами, але і цей формат має недолік, який полягає у тому, що дозвіл зображення у файлі, зазвичай, не запам'ятовується.

Формат GIF відомий всім користувачам Інтернету. Метою розробки формату було забезпечити максимальне стиснення відеоданих під час їх запису в пам'ять, щоб зменшити обсяг файлів, мінімізувати витрати на їх завантаження та передачу по каналах зв'язку. Стандартна версія формату GIF обмежується зображеннями з палітрою, що містить максимум 256 кольорів.

Моделі зображень. Каузальні, півкаузальні і некаузальні моделі

Під час моделювання випадкових полів на електронно-обчислювальній машині числа від датчика випадкових чисел перетворюються у числову матрицю, що володіє певними властивостями. Залежно від того, які відліки датчика беруть участь у формуванні поточного значення поля, моделі випадкових полів розділяються на *каузальні, півкаузальні і некаузальні*.

Основна відмінність між цими моделями обумовлена їх просторовими особливостями, які найчастіше призводять до принципово різних алгоритмів.

Якщо для формування відліку поля (біла точка на рисунку 3.5) з координатами (x, y) використовується поточний відлік датчика, попередні його відліки, то модель називають *каузальною* (рис. 3.5) [8], з латинської *causa* – причина [14].

Іноді під *каузальною моделлю* розуміють модель, що використовує дані області, геометрія якої визначається растровою розгорткою (рис. 3.6) [8].

Під час роботи в реальному часі, коли відомі лише ці дані, виконується умова причинності.

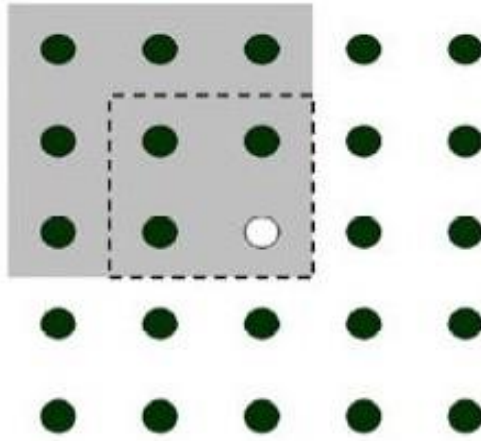


Рисунок 3.5 – Каузальна модель

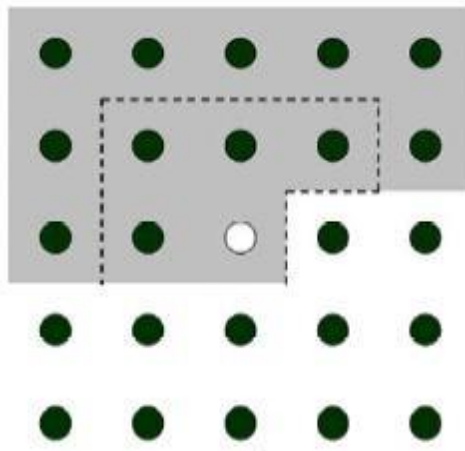


Рисунок 3.6 – Каузальна модель (растрова розгортка)

Півкаузальний фільтр (рис. 3.7) [8] використовує вхідні дані верхньої півплощини, а *некаузальний* – дані всієї площини (рис. 3.8) [8].

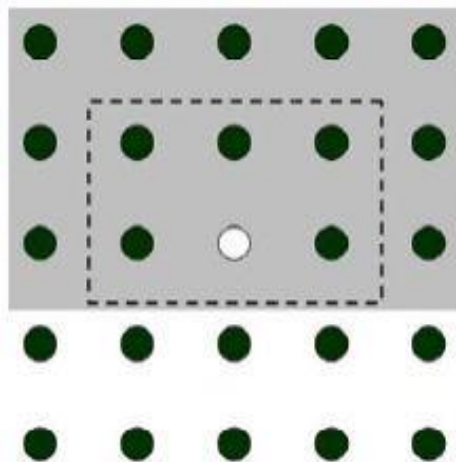


Рисунок 3.7 – Півкаузальна модель

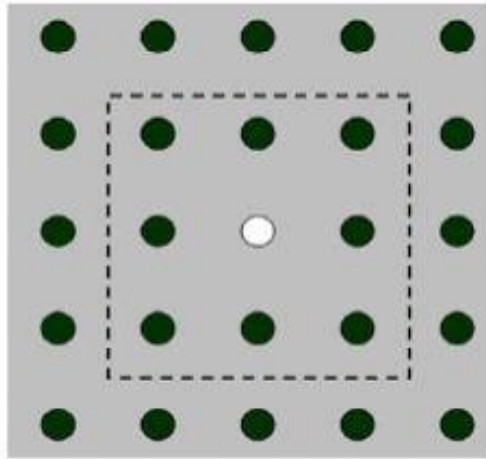


Рисунок 3.8 – Некаузальна модель

Можуть використовуватися не всі дані відповідної області, а лише деяка її частина (рис. 3.9) [8], наприклад, зазначена штриховий лінією.

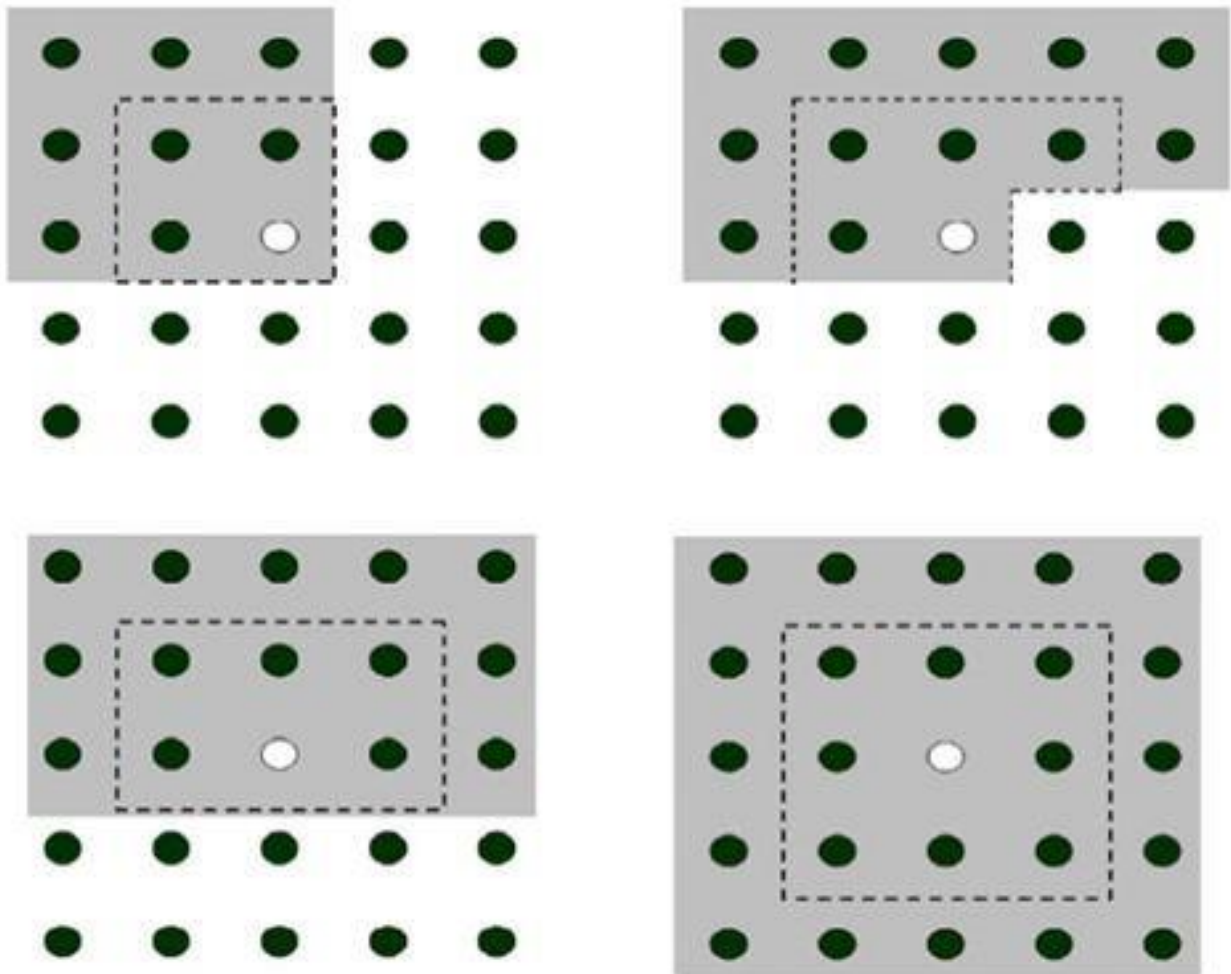


Рисунок 3.9 – Приклади використання моделей

Приклад зображення (рис. 3.10), побудованого за каузальною методикою, модель космічного знімку лісу у вигляді сукупності кругів різного діаметру, випадковим чином розміщених на площині, причому круги не перекриваються.

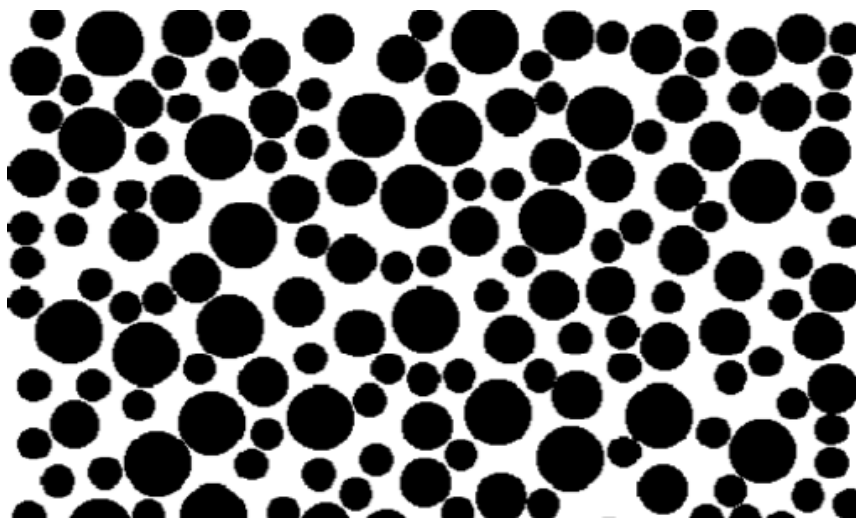


Рисунок 3.10 – Модель космічного знімку лісу

Під час побудови моделі спочатку, за випадковим законом, вибирається центр першого кола в околі лівого верхнього кута, зображується ця окружність. Далі в її околі випадковим чином ставляться наступні точки – центри інших кіл.

Окружність зображується, якщо виконується умова [8]:

$$(R_i + R_j)^2 > (x_i - x_j)^2 + (y_i - y_j)^2, \quad (3.2)$$

де $(x_i, y_i), (x_j, y_j)$ – координати центрів найближчих сусідів;

R_i, R_j – радіуси найближчих сусідів, $i \neq j$.

За цією схемою поступово, зверху вниз, заповнюється все поле.

У програмному забезпеченні сучасних електронно-обчислювальних машин є алгоритм моделювання випадкових чисел, розподілених рівномірно в інтервалі від 0 до 1, ці числа є статистично незалежними у досить довгій послідовності.

Математичне очікування таких чисел $\mu_1 = 0,5$, дисперсія $\sigma^2 = \frac{1}{12}$, якщо

відняти з цих чисел μ_1 , то отримаємо послідовність, розподілену рівномірно в інтервалі від мінус 0,5 до 0,5. З послідовності рівномірно розподілених випадкових чисел можна отримати послідовність нормально розподілених випадкових чисел.

З теорії ймовірностей відомо, що сума незалежних однаково розподілених випадкових величин має нормальний закон розподілу за умови, що доданки дають рівномірно малий вклад у суму (центральна гранична теорема). Щоб отримати один відлік випадкового числа з нормальним законом розподілу, досить підсумувати 8 чи 10 рівномірно розподілених випадкових чисел, отриманих за описаною методикою, зручніше використовувати 12 чисел, при цьому $\sigma^2 = \frac{1}{12}$.

Нехай вихідна послідовність чисел $\{x_i\}$ з рівномірним законом розподілу має $\mu_1 = 0$ та $\sigma^2 = \frac{1}{12}$.

Враховуючи, що значення цієї реалізації x_i, x_j при $i \neq j$ статистично незалежні, підсумуємо 12 таких чисел і отримаємо реалізацію $\{h_k\}$ з нормальним законом розподілу, з нульовим середнім і одиничною дисперсією. Помножимо всі h_k на постійне число σ і додамо до них постійну μ_1 . Нова реалізація f_j матиме задану дисперсію і середнє значення.

Використовуючи випадкові числа x_i або h_k , можна моделювати зображення, якщо послідовно, рядок за рядком, присвоювати пікселям ці випадкові значення яскравості. Необхідно враховувати, що яскравість – величина невід’ємна. Якщо її значення f_j лежать в інтервалі від 0 до 255, то при нормальному законі розподілу можна припустити, наприклад, $\mu_1 = 127$, $\sigma = 50$, при цьому значення $f_j < 0$ і $f_j > 255$ мало вірогідні. Можна задавати великі σ , у цих випадках від’ємним f_j слід присвоювати нульові значенням, а $f_j > 255$ присвоювати значення $f_j = 255$.

Отримане таким чином зображення – це сукупність непов’язаних одна з одною точок. Більш реалістичні зображення вдається отримати, якщо використовувати *авторегресійну модель*.

Авторегресійні моделі

Використання авторегресійної моделі дозволяє генерувати однорідне, однорідне і ізотропне або неоднорідне випадкове поле із заданими кореляційною функцією та математичним очікуванням.

Авторегресійні методики типово каузальні.

Розглянемо найпростішу лінійну авторегресійну модель зображення – авторегресії по рядках.

Це може бути, наприклад, модель рядка телевізійного сигналу, коли значення яскравості зчитуються уздовж рядка з деяким кроком [8]:

$$f_i = \rho \cdot f_{i-1} + (1-\rho) \cdot \mu_1 + h_i \sigma \sqrt{1-\rho^2}, \quad (3.3)$$

де f_i, f_{i-1} – значення яскравості на i -му та $(i-1)$ -му кроці;

h_i – однаково розподілені незалежні випадкові числа з нульовим середнім та одиничною дисперсією (що породжує випадковий процес);

ρ, μ_1, σ^2 – постійні параметри.

Таким чином, статистичні характеристики випадкової послідовності $\{f_i\}$ повинні бути інваріантні щодо зміни початку відліку пікселів у рядку, тобто рядок $\{f_i\}$ має властивості однорідного випадкового поля.

Наведене рекурентне співвідношення описує лінійне інерційне перетворення випадкової послідовності h_i .

Усереднимо праву і ліву частини виразу (3.3), дужки $\langle \rangle$ означають усереднення [8, 12]:

$$\langle f_i \rangle = \rho \cdot \langle f_{i-1} \rangle + (1-\rho) \cdot \mu_1,$$

звідси

$$\langle f_i \rangle = \langle f_{i-1} \rangle = \mu_1.$$

Перепишемо авторегресійне рівняння у вигляді:

$$f_i - \mu_1 = \rho(f_{i-1} - \mu_1) + (1-\rho) \cdot \mu_1 + h_i \sigma \sqrt{1-\rho^2}. \quad (3.4)$$

Помножимо праву і ліву частину на $(f_{i-1} - \mu_1)$ та усереднимо:

$$\langle (f_i - \mu_1)(f_{i-1} - \mu_1) \rangle = \rho \langle (f_{i-1} - \mu_1)^2 \rangle. \quad (3.5)$$

Середнє від добутку $f_{i-1}h_i$ дорівнює нулю, тому що f_{i-1} та h_i статистично незалежні (випадкове число h_i з'явилося пізніше, ніж f_{i-1}).

Вираз $\langle (f_i - \mu_1)(f_{i-1} - \mu_1) \rangle = R_1$ – це коефіцієнт кореляції між яскравостями двох сусідніх пікселів у рядку.

Зведемо ліву та праву частину авторегресійного рівняння у квадрат та усереднимо:

$$\langle (f_i - \mu_1)^2 \rangle = \rho^2 \langle (f_{i-1} - \mu_1)^2 \rangle + \sigma^2 \langle h_i^2 \rangle (1-\rho^2).$$

Так як $\langle h_i^2 \rangle = 1$, то $\langle (f_i - \mu_1)^2 \rangle = \langle (f_{i-1} - \mu_1)^2 \rangle = \sigma^2$ – дисперсія випадкової послідовності $\{f_i\}$. Параметр $\rho = \frac{R_1}{\sigma^2}$ – це нормований на σ^2 коефіцієнт кореляції між f_i та f_{i-1} , $0 \leq \rho \leq 1$.

Помножимо ліву і праву частину авторегресійного рівняння на $(f_{i-2} - \mu_1)$ та усереднимо. Отримаємо $\langle (f_i - \mu_1)(f_{i-2} - \mu_1) \rangle = R_2$ – коефіцієнт кореляції між значеннями яскравості двох пікселів, віддалених один від одного у рядку через один крок. Можна показати, що $R_2 = \rho^2 \sigma^2$. Продовжуючи цю процедуру, знайдемо, що коефіцієнт кореляції між значеннями яскравості пікселів, розділених вздовж рядка n кроками, $R_n = \rho^n \sigma^2 = \sigma^2 e^{-an}$, де $a = \ln\left(\frac{1}{\rho}\right)$.

Таким чином, випадкова послідовність $\{f_i\}$ має експоненціальну функцію автокореляції, f_i володіє нормальним законом розподілу, якщо h_i розподілено по нормальному закону або якщо h_i має відмінний від нормального закон розподілу, але близько до 1 (має місце нормалізація випадкового процесу f_i). Трансформація закону розподілу шляхом нелінійного перетворення $\{f_i\}$ призводить до «некерованої» функції кореляції, спроба зробити спочатку нелінійне перетворення послідовності незалежних відліків $\{h_i\}$, цього провести лінійне не завжди призводить до бажаного результату, так як лінійна система нормалізує випадковий процес при чималому ρ .

У такій моделі існує деяка залежність між яскравістю пікселів у рядку, але кожний сусідній рядок випадковим чином слідує за попереднім, зображення, скоріше, схоже на візерунок на тканині, ніж на знімок земної поверхні, отриманий із космосу.

Триточкова авторегресійна модель зображення записується так [8]:

$$f_{i,j} = \mu_1 + \rho_1 (f_{i-1,j} - \mu_1) + \rho_2 (f_{i,j-1} - \mu_2) - \rho_1 \rho_2 (f_{i-1,j-1} - \mu_1) + \sigma_1^2 \sigma_2^2 h_{i,j} \sqrt{1 - \rho_1^2 \rho_2^2},$$

де i – номер рядка;

j – номер пікселя у рядку;

ρ_1, ρ_2 – коефіцієнти кореляції значень яскравості поруч розташованих пікселів у сусідніх рядках і в одному рядку;

$h_{i,j}$ – однаково розподілені незалежні випадкові величини з нульовим середнім і одиничною дисперсією.

Триточкова модель при правильному підборі параметрів дозволяє створювати досить реалістичні зображення, дуже схожі на супутникові зображення земної поверхні (рис. 3.11). Ці зображення можна використовувати при моделюванні на різних алгоритмів обробки зображень.



Рисунок 3.11 – Триточкова модель зображення

Наблизитися до однорідного та ізотропного поля можна, використовуючи складніші авторегресійні моделі, наприклад, чотириточкові.

Схема авторегресії знаходить застосування і в інших задачах, наприклад, при відновлення пропущених пікселів.

Розтяжка і еквалізація гістограм

Гістограма яскравості зображення, побудованого за триточною моделлю, має вигляд кривої нормального закону розподілу. Гістограма показує, скільки пікселів n з близьким значенням яскравості f потрапляє в інтервал від f_i до $f + \Delta f_i$. Параметри, що входять до триточної моделі, мають певний фізичний зміст: μ_1 – це середня яскравість всіх пікселів, σ – величина стандартного відхилення, що характеризує контраст. Параметри $\frac{1}{\sigma}, \frac{1}{\beta}$ задають характерний розмір «плям» на зображенні (у пікселях). Збільшення яскравості призводить до зміщення гістограми вправо, зменшення – вліво. При збільшенні σ гістограма розширюється, контраст збільшується.

Під час цифрової обробки можлива зміна яскравості та контрасту. Слабкий контраст – досить поширений дефект сканерних, фотографічних і телевізійних зображень, обумовлений обмеженістю діапазону відтворених яскравостей. Під *контрастом* розуміють різницю між максимальним і мінімальним значенням яскравості.

Враховуючи специфіку цифрової обробки зображень, будемо називати середнє значення μ_1 яскравістю зображення, а стандартне відхилення σ – мірою контрасту.

Шляхом цифрової обробки контраст можна підвищити, змінюючи яскравість кожного елемента зображення та збільшуючи діапазон яскравостей.

Для цього розроблено декілька методів.

Нехай рівні деякого півтонування займають інтервал від 6 до 158 із середнім значенням яскравості 67 при можливому найбільшому інтервалі значень від 0 до 255. Зображення малокоонтрастне, переважає темний відтінок. Методом поліпшення контрасту є *лінійна розтяжка гістограми*, коли рівням вихідного зображення, що належать інтервалу $[f_{\min}, F_{\max}]$, присвоюються нові значення для того, щоб охопити весь можливий інтервал зміни яскравості, у даному випадку $[0, 255]$, контраст істотно збільшується (рис. 3.12) [8].

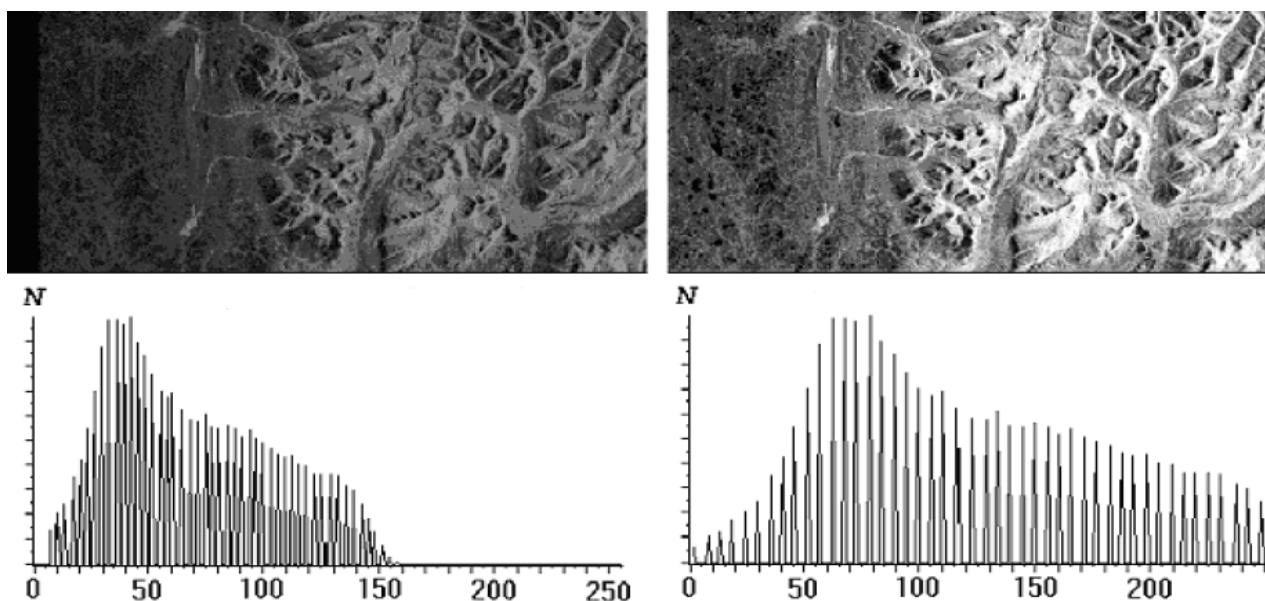


Рисунок 3.12 – Розтяжка гістограми

Перетворення рівнів яскравості здійснюється за формулою:

$$g_i = c + df_i, \quad (3.6)$$

де f_i – старе значення яскравості i -го пікселя;

g_i – нове значення яскравості i -го пікселя;

c, d – коефіцієнти.

Варіант, коли на весь максимальний інтервал рівнів яскравості $[0, 255]$ розтягується не вся гістограма, що лежить в межах $[f_{\min}, F_{\max}]$, а її найбільш інтенсивна ділянка, малоінформативному лівому «хвосту» присвоюється значення 0, правому «хвосту» 255, цю процедуру називають *нормалізацією гістограми*.

Метою *вирівнювання гістограми* (лінеаризація чи еквалізація) є таке перетворення, щоб всі рівні яскравості придбали б однакову частоту, а гістограма яскравостей відповідала б рівномірному закону розподілу. Початкове зображення та результат еквалізації (рис. 3.13) [8].

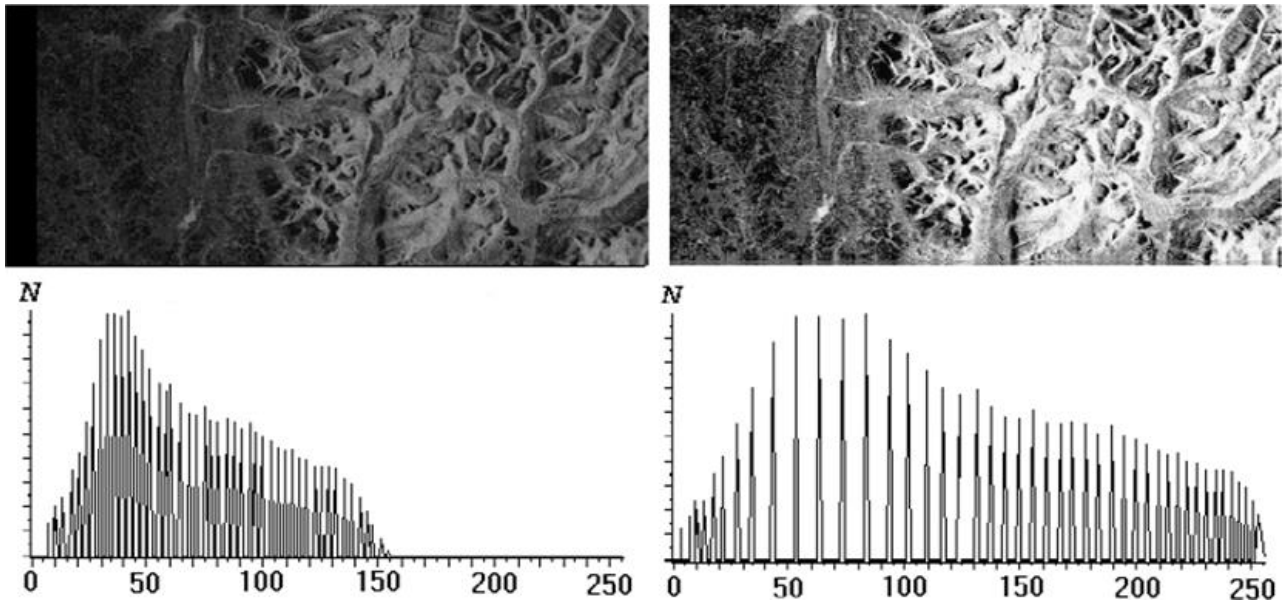


Рисунок 3.13 – Еквалізація гістограми

У кожному конкретному випадку вибирають той метод перетворення гістограм, який призведе до найкращого, з погляду користувача, результату.

Сегментація зображень. Способи сегментації

Одним з найпоширеніших методів виділення об'єктів на космічних зображеннях Землі є *сегментація*. Цей метод носить риси детермінованого і статистичного підходів.

Під *сегментацією*, у широкому сенсі, розуміють перетворення напівтонових або кольорових зображень в зображення, що мають менше число тонів або кольорів, ніж початкові. У вузькому сенсі *сегментацією* називають перетворення півтонування в дворівневе (бінарне), що містить всього два рівня яскравості – мінімальний (0) і максимальний (255).

При цьому об'єкт і фон розділені, легко визначити кількість об'єктів, характеристики їх місця розташування (координати, поворот виділеної осі об'єкта щодо координатних осей), геометричні характеристики (наприклад, площа кожного об'єкту, периметр, середній, мінімальний, максимальний розміри) та ідентифікувати об'єкт – вказати, що це таке.

Метою сегментації є виділення областей, однорідних в якомусь певному сенсі. Однорідність є ознакою належності області до певного класу.

Дуже часто сегментація використовується для виділення областей, близьких за тоном або кольором. Сегментація часто використовується для виділення областей, однорідних за деякою складною властивістю, що називаються *кластерами*.

Застосовують три основні способи сегментації зображень: порогова сегментація зображень, шляхом нарощування та шляхом виділення кордонів.

Розрахунок порогу під час порогової сегментації

Порогова сегментація полягає в об'єднанні близьких за характеристиками областей зображення у невелике число сегментів. Якщо яскравість перевищує поріг, то область відносять до одного сегменту, якщо вона нижче порогу – то до іншого. Найпростіший випадок – розбиття на два сегменти (бінаризація). Порогова сегментація може виконуватися на основі апріорно заданих порогів. Інший спосіб полягає в тому, що пороги вибираються як кордони мод гістограми зображення. Розглянемо цей спосіб на прикладі бінаризації півтонування, у якого гістограма містить дві моди. Цей випадок типовий для задання виділення площ покритих снігом і льодом на фоні лісу і відталої землі за результатами дистанційних досліджень (рис. 3.14) [8].

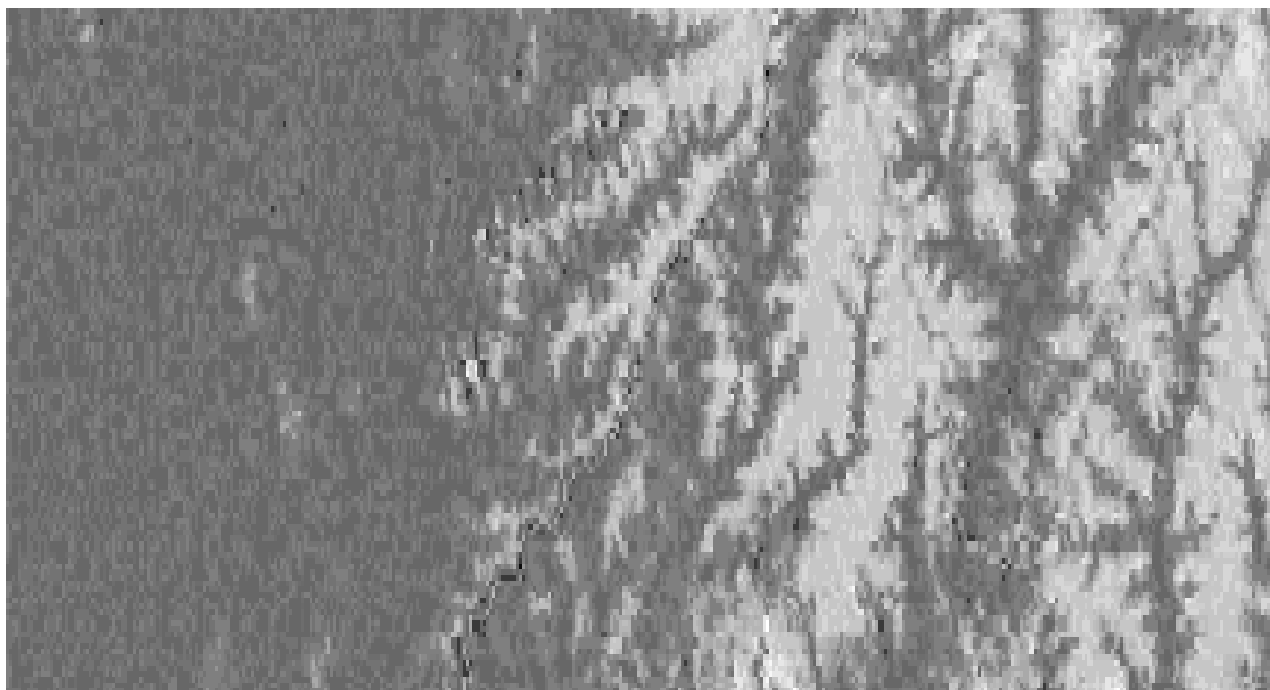


Рисунок 3.14 – Початкове зображення

Гістограма зображення має дві моди – одна відповідає більш темному фону, друга – об'єктам з більшою яскравістю, між модами видна зона мінімуму. Моди перекриваються слабо, поріг обраний посередині зони між максимумами мод. Бінаризоване зображення показано на рисунку 3.15 [8].

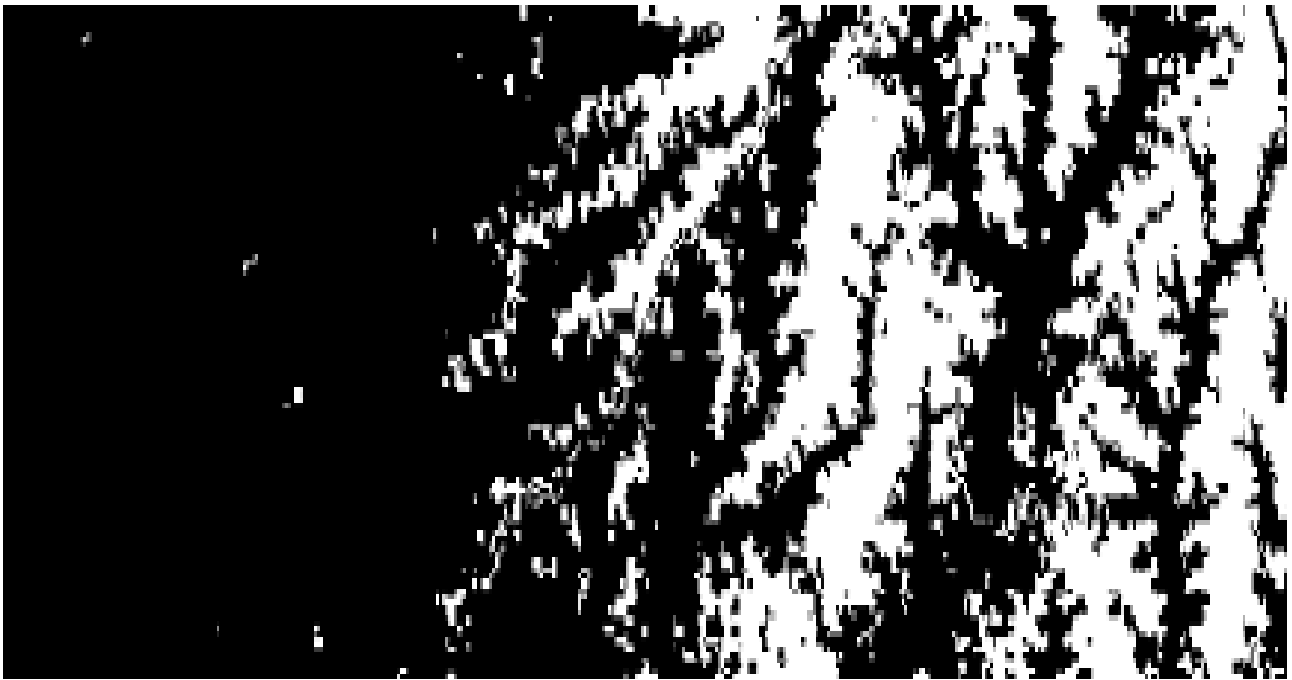


Рисунок 3.15 – Порогова сегментація

Складніше провести сегментацію, коли гістограми фону та об'єкта сильно перекриваються «хвостами» і при цьому частина пікселів об'єкта може бути віднесена до фону, а частина пікселів фону віднесена до об'єкта.

Процес прийняття рішення при цьому також зводиться до послідовного віднесення f до одного з двох класів поточного рівня, тобто реалізується *метод послідовних дихотомій*.

Сегментація шляхом нарощування областей

У способі *сегментація шляхом нарощування областей* виділяються однорідні області. Розглянемо сегментацію шляхом нарощування областей з використанням критерію однорідності за величиною вектора яскравості.

Схема алгоритму цього методу передбачає вибір стартового пікселя і розгляд суміжних з ним пікселів для перевірки близькості їх значень, наприклад, по евклідовій відстані. Якщо значення яскравості поточного і якого-небудь суміжного пікселів виявляються близькими, то ці пікселі зараховуються в одну область. Таким чином, формується область у результаті зрощування окремих пікселів. На наступному етапі область перевіряється на однорідність і, якщо результат перевірки виявляється негативним, то область розбивається на більш дрібні ділянки. Процес продовжується до тих пір, поки всі виділені області не пройдуть перевірку на однорідність.

Можлива реалізація алгоритму, що використовує формування областей як зрощенням окремих пікселів, так і зрощенням невеликих областей.

Загальна схема перевірки області на однорідність полягає у такому.

Нехай $F(R)$ – задана міра однорідності області R . Якщо $R_{12} = R_1 \cap R_2$, то критерій однорідності можна задати, вимагаючи, щоб виконувалася умова $F(R_{12}) \leq \varepsilon$, де ε – заданий поріг [14].

Різні алгоритми сегментації класифікуються за виглядом міри $F(R)$.

У деяких випадках F можуть використовувати як величину розмаху

$$F(R) = \max_{(i,j) \in R} f_{ij} - \min_{(k,m) \in R} f_{km}. \quad (3.7)$$

Таким чином, при сегментації шляхом нарощування областей враховується структура області та її однорідність. Це буває важливо при обробці даних дистанційного зондування, нерідко цей метод дає кращі результати, ніж інші методи, що не враховують зв'язності і розраховані на віднесення ізольованого пікселя до деякого класу.

Сегментація шляхом виділення кордонів

Сегментація шляхом виділення кордонів передбачає використання оператора градієнта. Для встановлення факту, що дійсно виявлена межа, застосовується процедура поділу по порогу. Пікселі, ідентифіковані як граничні, з'єднуються у замкнуті криві, що оточують відповідні області.

У даному методі, як і в інших методах сегментації, істотним є критерій однорідності області, за характеристикою якого і обчислюються значення градієнта. Прямі методи сегментації шляхом виділення кордонів застосування до початкового зображення градієнтних фільтрів (Робертса, Собела, Лапласа).

Задача побудови кордонів сегментів на зображенні градієнта виступає як самостійне завдання. Ця задача досить складна і може бути вирішена лише в найпростіших випадках, наприклад, можна виділяти локальні максимуми градієнта всіх рядків і стовпців зображення.

Для кордонів простої форми може бути використана процедура апроксимації початкового зображення градієнта якою-небудь параметричною функцією. Прикладом такого завдання є орієнтація штучного супутника Землі по кадру зображення частини краю земного диску.

Розпізнавання зображень. Кластерний аналіз

Під розпізнаванням образів (класифікацією) розуміють процедуру, що дозволяє винести рішення про належність даного зображення або його фрагмента до одного з n класів, $n > 1$. Це рішення виноситься на підставі наявності ознак того чи іншого класу. Результатом є виділення об'єктів на зображенні, розділених на класи.

Розпізнавання образів відноситься до тематичної обробки зображень. Перед цим, зазвичай, виконують попередню обробку – відновлюють і покращують вихідне зображення.

Завдання розпізнавання зображень вирішуються на основі загальної теорії розпізнавання образів. Розпізнавання образів – розділ кібернетики, який розробляє принципи і методи класифікації та ідентифікації предметів, явищ, процесів, сигналів, ситуацій. У загальній теорії розпізнавання образів розрізняють два типи завдань:

- завдання таксономії (навчання без учителя);
- завдання розпізнавання (навчання з учителем).

У задачах таксономії необхідно розділити запропоновані об'єкти за кількома групами тільки на основі їх описів. Завдання другого типу виникає при необхідності визначити клас, до якого належить опис деякого об'єкта.

Передбачається, що є набір з N об'єктів, для яких відомі образи (класи), до яких вони належать. Цю сукупність називають навчальною вибіркою, а складові її об'єкти – еталонами. Необхідно, виходячи з навчальної вибірки, визначити клас, до якого належить опис деякого об'єкта.

Формальне вирішення задачі розпізнавання прийнято записувати у вигляді вирішальної функції $i = d(x)$, де i – номер класу, до якого належить опис об'єкта – вектор ознак x .

Кластерний аналіз дозволяє вирішувати завдання обох класів. Кластер – це однорідна ділянка зображення, з точки зору деякого набору ознак. Пікселі, що належать до різних кластерів, відрізняються за цими ж ознаками.

Оцінка подібності у кластерному аналізі заснована на поняття відстані. Якщо відстань менше порогу, то елемент зображення відносять до відповідного кластеру. При використанні багатоспектральних даних дистанційного зондування у відстань входять значення яскравості (i, j) -го пікселя зображення у різних каналах. Сукупність цих значень можна записати у вигляді вектора f_{ij} .

Кластери формуються так, щоб відстань між окремими пікселями у кожному кластері була мінімальною, а відстані між пікселями, що відносяться до різних кластерів, були якомога більшими.

Найбільш розповсюдженою мірою подібності є евклідова відстань між векторами f_{ij} і f_{mn} . Якщо $\{f_{kij}\}$, $\{f_{kmn}\}$ – компоненти цих векторів, k – номер спектрального каналу, то евклідова відстань [8, 9]:

$$r_E = \sqrt{\sum_k (f_{kij} - f_{kmn})^2}. \quad (3.8)$$

Мірою подібності може бути також косинус кута між векторами, що визначається як відношення скалярного добутку векторів до добутку їх норм:

$$\cos\varphi = \frac{(f_{ij}, f_{mn})}{\|f_{ij}\| \cdot \|f_{mn}\|}. \quad (3.9)$$

Косинус максимальний при близькості напрямів векторів. Процедура кластеризації може ґрунтуватися на оптимізації якого-небудь показника якості, наприклад, критерію мінімуму суми квадратів помилки [8]:

$$\varepsilon = \sum_{k=1}^K \sum_{f \in S_k} \|f_{k-\mu_k}\|^2, \quad (3.10)$$

де K – число кластерів;

S_k – множина об'єктів (пікселів), що відносяться до k -го кластеру;

μ_k – вектор середніх значень для класу k .

У задачах таксономії методи визначення належності точок одного й того ж кластеру пов'язані з навчанням без вчителя. Деякі з цих методів використовують послідовне злиття наявних кластерів.

Спочатку кожна точка даних розглядається як окремий ембріональний кластер. На кожному кроці ітераційного процесу виявляються два кластери, що містять дві точки, розташовані одна біля одної ближче, ніж будь-які дві точки інших кластерів, ці два кластери зливаються. Ітераційний процес закінчується коли або знайдено очікуване число кластерів, або відстань до наступної точки, доданої до кластеру, перевищує заданий поріг. Для управління цими процесами розроблені численні евристики.

Протилежна стратегія роз'єднує наявні кластери уздовж ліній «розрідження». Спочатку весь набір точок розглядається як один великий кластер. На кожному етапі визначається кластер, який можна розбити на два. Ітераційний процес закінчується коли або досягнуте бажане число кластерів, або подальше розбиття неперспективне по деякому попередньо визначеному критерію. У більшості випадків ми знаємо, які точки належать одному класу.

Теорема Байєса

Статистичний підхід передбачає, що в одній точці простору ознак з відмінною від нуля апріорною ймовірністю можуть з'являтися реалізації, що відносяться до різних класів. Це може бути пов'язано з неминучими випадковими помилками, а також з тим, що ознаки є принципово випадковими величинами (характерно для задач дистанційного зондування).

Статистичні методи дозволяють виробляти класифікацію, яка в середньому є найбільш правильною. Методи статистичного розпізнавання образів використовують функції розподілу ймовірностей, пов'язані з класами образів. У деяких випадках форма функцій розподілу ймовірностей вважається відомою (наприклад, нормальної) і по навчальним вибіркам необхідно оцінити тільки окремі параметри, які пов'язані з цими функціями (математичне сподівання, дисперсія, функція кореляції). Метод називається *параметричним*.

Якщо форма функцій розподілу ймовірностей не відома, метод є *непараметричним*.

При статистичному підході до розпізнавання часто використовується теорема Байєса – одна з фундаментальних теорем теорії ймовірностей. Мова йде про випадкові події, тобто такі, для яких неможливо заздалегідь передбачити точний результат у кожному конкретному випадку. Однак, при великому числі реалізацій ці події можна характеризувати середніми результатами, стабільними і відтворюваними.

Якщо можуть відбутися дві події A і B , то слід говорити про три різні ймовірності:

- подія A відбувається з ймовірністю $P(A)$;
- подія B відбувається з ймовірністю $P(B)$;
- події A і B відбуваються одночасно з ймовірністю $P(A \cdot B)$, її називають ймовірністю спільної події (спільної ймовірності).

Нехай $A_i (i=1, 2, \dots, n)$ і B – випадкові події. Ймовірність того, що подія A_i відбудеться, за умови, що подія B вже відбулася (умовна ймовірність), прийнято записувати у вигляді [8]:

$$P\left(\frac{A_i}{B}\right). \quad (3.11)$$

Умовну ймовірність (3.11) називають *апостеріорною*, її можна обчислити за теоремою Байєса. Так як $P(B) = \sum P(A_i)P\left(\frac{B}{A_i}\right)$, то

$$P\left(\frac{A_i}{B}\right) = \frac{P(A_i B)}{P(B)} = \frac{P(A_i)P\left(\frac{B}{A_i}\right)}{\sum P(A_i)P\left(\frac{B}{A_i}\right)}. \quad (3.12)$$

У (3.12) апіорна ймовірність події A_i позначається як $P(A_i)$.

Використовуючи апостеріорні ймовірності, можна розробити різні методи автоматичної класифікації. Нехай X_k – вектор вимірювань, що представляє k -й клас. Апіорна ймовірність того, що ця реалізація відноситься до класу з номером k , є $P(X_k)$.

Зазвичай, апіорна ймовірність вважається заданою самою постановкою завдання. Наприклад, для лісових територій буває відомий відсоток площі, зайнятої листяними і хвойними породами. Необхідно віднести невідомий спостережуваний об'єкт Z (наприклад, деякий піксель зображення) до одного з відомих класів C_k з мінімальною помилкою. Яскравість пікселя задана в m спектральних каналах. Результатом є вектор вимірювань X_m (вектор яскравості пікселя), для якого можна знайти умовну ймовірність (або її щільність):

$$P\left(\frac{X_m}{C_k}\right). \quad (3.13)$$

Шукані ймовірності можуть бути обчислені згідно теореми Байєса з використанням умовних ймовірностей $P\left(\frac{X_m}{C_k}\right)$, отримаємо:

$$\frac{P\left(\frac{X_m}{C_1}\right)}{P\left(\frac{X_m}{C_2}\right)} > \frac{P(C_2)}{P(C_1)}. \quad (3.14)$$

У лівій частині нерівності дріб називають *відношенням правдоподібності*. Рішення про те, що спостерігається об'єкт першого класу, приймається тоді, коли відношення правдоподібності перевищує вираз у правій частині. Права частина містить тільки відомі члени і є величиною порогу, не залежного від спостережуваної величини.

Помилки класифікації

Процес класифікації пов'язаний з деяким ризиком, оскільки можливі помилки. Підрахуємо ймовірність помилки, яка може виникнути під час використання цієї процедури. Вирішуючи (2.14), знайдемо поріг X^* , такий, що при $X_m \leq X^*$ вважається, що об'єкт належить до першого класу, інакше – до другого.

Нехай через випадковий характер величини X_m прийнято помилкове рішення, що об'єкт належить до першого класу, насправді він належить до другого класу.

Виникає *помилкова тривога*, ймовірність якої виражається через умовну щільність ймовірності $w\left(\frac{X_m}{C_2}\right)$ [8]:

$$P_{n.m.} = \int_{-\infty}^{X^*} w\left(\frac{X_m}{C_2}\right) dX_m. \quad (3.15)$$

Можливо, що $X_m < X^*$, але об'єкт належить до першого класу, виникає *пропуск виявлення*, ймовірність якого виражається через умовну щільність ймовірності $w\left(\frac{X_m}{C_1}\right)$:

$$P_{n.v.} = \int_{-\infty}^{X^*} w\left(\frac{X_m}{C_1}\right) dX_m. \quad (3.16)$$

Сумарна *помилка класифікації* дорівнює сумі двох ймовірностей:

$$P_{ном} = P_{n.m.} + P_{n.v.}. \quad (3.17)$$

Геометричні перетворення та прив'язка зображень

Багато задач тематичного дешифрування зводяться до взаємного зіставленню між собою зображень, сформованих за допомогою датчиків різних фізичних полів. Яскравим прикладом може служити розвиток дистанційних методів контролю природних ресурсів та динаміки екосистем, що зводиться до співставлення знімків однієї і тієї ж території, отриманих у різний час і за допомогою різних датчиків. Найчастіше використовуються оптичне, радіолокаційне, радіотеплове, магнітне та інші поля. Спільне використання різних фізичних полів вимагає попередньої обробки відповідних їм зображень, наприклад, з метою переведення зображень в одну спектральну область.

На практиці зображення одного і того ж об'єкта або ділянки місцевості, отримані в різний час або за допомогою різних датчиків, можуть значно відрізнятися один від одного. Звідси впливає ряд важливих задач прив'язки, а також точної взаємної геометричної та амплітудної корекції для подальшого спільного аналізу.

У кожному разі це вимагає встановлення відповідності між елементами вихідних зображень, що зводиться до виділення так званих опорних (реперних або сполучених) точок на зображеннях, за якими можна здійснити координатну прив'язку знімків з одночасною геометричною корекцією.

Наприклад, аерокосмічний комп'ютерний моніторинг передбачає наявність дискретного за часом спостереження з невеликим тимчасовим інтервалом, тому, коли рухома камера фіксує яскравіший образ спостережуваного об'єкта (оптичну поверхню) у вигляді послідовності зображень, то цей образ від знімку до знімку деформується внаслідок перспективних спотворень і зміни положення камери. Геометрія відповідних деформацій моделюється проєктивними перетвореннями, які складають більш великий клас, ніж відомі перетворення евклідової геометрії (довжини і кути в проєктивній геометрії не зберігаються, а паралельні лінії можуть перетинатися).

Відновлення просторового рельєфу по стереознімку призводить до проблеми ідентифікації: встановлення точної координатної (поточечної) відповідності елементів стереозображень.

Вирішення цього завдання полягає у виділенні пар реперних фрагментів і оцінюванні параметрів «розбіжності» відповідних точок, за якими можна відновити функцію геометричного перетворення і оцінити поверхню тривимірної сцени (рельєф).

Після оцінювання параметрів геометричного перетворення виникає задача геометричної корекції або відновлення зображення у перетворених координатах.

У практиці обробки зображень завдання пошуку відповідності набуло великого поширення та відоме як проблема «пошуку за зразком». Формально його можна розглядати як процес ототожнення еталонного зображення (образу фрагмента) на першому знімку з одним із множини образів фрагментів, що лежить у деякій (що задається) області (зоні пошуку) другого знімка.

Алгоритми встановлення подібності у своїх основоположних варіантах у тій чи іншій мірі пов'язані з отриманням характеристик стохастичної взаємозв'язки порівнюваних фрагментів зображень. Всі вони ґрунтуються на ідеях кореляційної та спектральної теорії сигналів, для відповідних критеріїв отримані експериментальні характеристики основних процедур пошуку за зразком.

3.2 Лінійна просторово-інваріантна фільтрація та фільтрація у просторовій області

План

1. Модель спотвореного зображення.
2. Лінійні просторово-інваріантні фільтри.
3. Двовимірне перетворення Фур'є.
4. Глобальна фільтрація.
5. Інверсна фільтрація.
6. Лінійні згладжують фільтри.
7. Лінійні фільтри для виділення контурів.
8. Нелінійні фільтри.

Модель спотвореного зображення

Реальні зображення разом з корисною інформацією містять різні перешкоди. Джерелами перешкод є власні шуми фотоприймальних пристроїв, зернистість фотоматеріалів, шуми каналів зв'язку. Можливі геометричні та радіометричні спотворення, зображення може бути розфокусованим (не типово для супутникових зображень з роздільною здатністю 10 м та більше).

Для зображень з роздільною здатністю 1 м і менше турбулентність атмосфери призводить до розмивання дрібних деталей при коротких експозиціях. Модель спотвореного завадами безперервного зображення має вигляд [8, 9, 14]:

$$f(x, y) = m(x, y) \cdot Qs(x, y) + n(x, y), \quad (3.18)$$

де $f(x, y)$ – спотворене зображення;

$m(x, y)$ – мультиплікативна перешкода, що моделює зображення за яскравістю;

$s(x, y)$ – вихідне зображення;

Q – функціонал, що описує геометричні і радіометричні спотворення, а також розфокусування;

$n(x, y)$ – адитивна перешкода, що накладається на зображення.

Модуляція супутникового зображення за яскравістю може відбуватися через те, що атмосфера над різними точками Землі має різну прозорість, випромінювання від цих точок проходить різний шлях в атмосфері.

Під час реставрації зображень необхідно відновити початкове зображення. Якщо спотворення відсутні, $m(x, y) = 1$, то

$$f(x, y) = Qs(x, y) + n(x, y), \quad (3.19)$$

Результат реставрації $\hat{s}(x, y) + g(x, y)$ запишемо як наслідок впливу на $f(x, y)$ деякого оператора $g(x, y) = Tf(x, y)$.

Оператор T (системний оператор) вказує на правило, за яким «вхідному сигналу» $f(x, y)$ ставиться у відповідність «вихідний сигнал» $g(x, y)$.

Для того, щоб модель була повною, необхідно вказати області допустимих значень $f(x, y)$ та $g(x, y)$. Під час реставрації застосовують оператор T , що мінімізує відстань між $g(x, y)$ і $s(x, y)$ при заданих статистичних характеристиках випадкових полів $s(x, y)$, $n(x, y)$ і відомому F .

Як критерій близькості $g(x, y)$ і $s(x, y)$ часто використовують критерій мінімуму середньоквадратичної помилки:

$$\min \langle [g(x, y) - s(x, y)]^2 \rangle. \quad (3.20)$$

Лінійні просторово-інваріантні фільтри

У задачах поліпшення зображень, зазвичай, вважається, що $n(x, y) = 0$, функцією оператора T є згладжування різких перепадів яскравості, підкреслення або виділення контурів.

Розглянемо просторово-інваріантні оператори, вихідна реакція яких не залежить від зміни початку відліку по x і по y , від орієнтації об'єктів на зображенні. Перша умова означає, що оператор переводить однорідне випадкове поле в однорідне. Друга умова означає, що оператор переводить в ізотропне поле. Властивості просторової інваріантності виконуються строго, якщо області допустимих значень координат x, y потрапляють в інтервал від $-\infty$ до ∞ . Реальні зображення мають кінцеві розміри $A \leq x \leq B$, $C \leq y \leq D$, умова просторової інваріантності виконується наближено.

Оператор називається лінійним, якщо для нього справедливий принцип суперпозиції – реакція на суму сигналів $f_1(x, y)$ і $f_2(x, y)$ дорівнює сумі реакцій на кожне з впливів окремо

$$T[f_1(x, y) + f_2(x, y)] = Tf_1(x, y) + Tf_2(x, y). \quad (3.21)$$

Для будь-якого довільного числа a справедливо

$$Ta f(x, y) = a T f(x, y). \quad (3.22)$$

Властивості лінійності виконуються строго, якщо область допустимих значень яскравості f, g потрапляють в інтервал $(-\infty; \infty)$. Під час цифрової обробки яскравість – величина речова, невід’ємна і обмежена, зазвичай, $0 \leq f, g \leq 255$. Якщо кожному $g(x, y)$ відповідає єдине $f(x, y)$, то оператор T може бути подано у вигляді:

$$g(x, y) = \iint f(x', y') \cdot h(x, y, x', y') dx' dy'. \quad (3.23)$$

В (3.23) інтегрування ведеться по всій області, де визначені x та y , характеризує перетворення всього зображення цілком – глобальну фільтрацію. Можна обробляти зображення по частинам, у цьому випадку здійснюється локальна фільтрація.

Ядро перетворення (3.23) $h(x, y, x', y')$ в оптиці іменують *функцією розсіювання точки*, це зображення точкового джерела на виході оптичної системи, що вже є не точкою, а деякою прямою. Всі точки зображення $f(x', y')$ перетворюються у плями, відбувається підсумовування (інтегрування) усіх плям. Не слід думати, що ця процедура обов’язково приводить до розфокусування зображення, навпаки, можна підібрати таку функцією розсіювання точки, яка дозволить сфокусувати розфокусовані зображення.

Для того, щоб для функції розсіювання точки виконувалася умова просторової інваріантності, тобто, щоб вона не змінювалася при зміні початку відліку по x і по y , функція розсіювання точки повинна мати вигляд:

$$h(x, y, x', y') = h(x - x', y - y'). \quad (3.24)$$

Функція розсіювання точки повинна володіти осьовою симетрією.

При обробці растрових зображень на прямокутній сітці простіше всього реалізувати функцію розсіювання точки кінцевих розмірів у вигляді прямокутної матриці форматом $N \times N$, наприклад, 3×3 .

Нехай $f(x, y)$ – функція двох змінних, визначена на інтервалах $(-\infty < x < \infty)$, $(-\infty < y < \infty)$, яка задовольняє умові абсолютної інтегрованості:

$$\int \int_{-\infty}^{\infty} |f(x, y)| dx dy < \infty. \quad (3.25)$$

Тоді існує інтеграл Фур'є, це означає таке [8]:

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp[-2\pi i(ux + vy)] dx dy, \quad (3.26)$$

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) \exp[2\pi i(ux + vy)] du dv, \quad (3.27)$$

де x і y – це координати;

u і v – це просторові частоти;

$F(u, v)$ – спектр просторових частот або спектр.

Комплексна експонента є лінійною комбінацією синуса і косинуса.

Двовимірне перетворення Фур'є

Вираз (3.26) носить назву прямого, а (3.27) – зворотного перетворення Фур'є. Використовуючи перетворення Фур'є, переходимо від координатної площини (x, y) до частотної площини (u, v) . Перетворення Фур'є лінійне, так як інтеграл – лінійна функція. Перехід в частотну площину має сенс, оскільки деякі властивості у спектра простіше, ніж у функції $f(x, y)$, яка описує розподіл яскравості в координатній площині. Нехай, наприклад, потрібно знайти результат глобального лінійного перетворення деякого зображення:

$$g(x, y) = \iint f(x', y') \cdot h(x - x', y - y') dx' dy'. \quad (3.28)$$

У координатній площині для цього потрібно обчислити інтеграл типу згортки (3.28), що часто буває справою досить складною. Якщо ввести частотний коефіцієнт передачі $K(u, v)$, який на практиці для функції розсіювання точки завжди існує і який пов'язаний парою перетворень Фур'є з функцією розсіювання точки [8]:

$$K(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) \exp[-2\pi i(ux + vy)] dx dy, \quad (3.29)$$

$$h(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(u, v) \exp[2\pi i(ux + vy)] du dv, \quad (3.30)$$

то в площині просторових частот (3.28) зводиться до перемноження функцій $F(u, v)$ і $K(u, v)$:

$$G(u, v) = F(u, v) \cdot K(u, v), \quad (3.31)$$

де $G(u, v)$ – спектр після лінійного перетворення.

Спектр $F(u, v)$ від функції $f(x, y)$ є комплексною функцією. Для просторово-інваріантної функції розсіювання точки частотний коефіцієнт передачі $K(u, v)$ завжди дійсний та інваріантний щодо поворотів навколо початку координат.

Реальне растрове зображення f_{nm} має кінцеві розміри $A \leq x \leq B, C \leq y \leq D$ і складається з окремих пікселів, розташованих з деяким кроком у вузлах прямокутної сітки. У цьому випадку для переходу в частотну площину застосовується *подвійне дискретне перетворення Фур'є*.

Глобальна фільтрація

Під час лінійної фільтрації зображень в частотній площині потрібно помножити спектр просторових частот на частотний коефіцієнт передачі (3.31) і виконати два двовимірні перетворення Фур'є з використанням алгоритму *швидкого перетворення Фур'є* – пряме і зворотне. Перетворення Фур'є здійснюється від всього зображення цілком, спектр $F(u, v)$ зберігає інформацію про все зображенні цілком – це *глобальна фільтрація*. Залежно від вибору коефіцієнта передачі $K(u, v)$, можна виділяти зображення на фоні перешкод, покращувати і погіршувати різкість зображення, виділяти контури об'єктів на зображенні. Якість обробки при цьому трохи краща, ніж при локальній лінійній фільтрації.

Типовою задачею глобальної фільтрації, не характерної для цифрової обробки космічних зображень земної поверхні, є відновлення розфокусованих зображень. Запишемо спектр такого зображення у вигляді [8, 9, 11]:

$$F_1(u, v) = F(u, v) \cdot K(u, v), \quad (3.32)$$

де $F(u, v)$ – спектр початкового зображення;

$K(u, v)$ – коефіцієнт передачі оптичної системи, що відповідає функції розсіювання точки.

Інверсна фільтрація

Очевидно, що для того, щоб визначити $F(u, v)$ по відомим $F_1(u, v)$ та $K(u, v)$, необхідно помножити $F_1(u, v)$ на $K_1(u, v) = \frac{1}{K(u, v)}$. Ця процедура носить назву *інверсної фільтрації*.

Задача інверсної фільтрації відноситься до числа зворотних некоректних задач математичної фізики. По-перше, рішення може не існувати. По-друге, якщо рішення існує, то може бути не єдиним.

Функція $K(u, v)$ у деяких точках може дорівнювати нулю, тоді рішення перетворюється у нескінченність.

По-третє, рішення може бути нестійким, тобто невеликі варіації початкових даних можуть призвести до істотних змін рішення. Виходом з положення є *регуляризація рішення*. Коефіцієнт передачі записується у формі:

$$K_1(u, v) = \left[\frac{1}{K(u, v) + \sigma \gamma(u, v)} \right], \quad (3.33)$$

де $\gamma(u, v)$ – стабілізуюча функція;

σ – параметр регуляризації.

Підбором σ вдається досить якісно відновити розфокусовані зображення.

Лінійні згладжують фільтри

На практиці глобальна фільтрація застосовується рідко. Частіше використовують локальну фільтрацію, коли інтегрування і усереднення проводиться не по всій області визначення x і y , а по порівняно невеликому околі кожної точки зображення. Функція розсіювання точки при цьому має обмежені розміри. Перевагою такого підходу є гарна швидкодія.

Під час обробки растрових зображень, які складаються з окремих пікселів, інтегрування замінюють підсумовуванням. Лінійне перетворення у разі локальної фільтрації приймає вигляд [11, 14]:

$$g_{ij} = \sum_D a_{kl} f_{i+kj+l}, \quad (3.34)$$

підсумовування ведеться за деяким околom D точки (i, j) , а a_{kl} – значення функції розсіювання точки у цьому околі.

Яскравості пікселів f в точці (i, j) і в її околі множаться на коефіцієнти a_{kl} , перетворена яскравість (i, j) -го пікселя є сумою цих добутків. Зазвичай, набір коефіцієнтів a_{kl} подають у вигляді прямокутної матриці (маски), наприклад розмірності 3×3 .

Елементи матриці задовольняють умові просторової інваріантності. Фільтрація здійснюється переміщенням зліва направо (або зверху вниз) маски на один піксель. При кожному положенні апертури виробляються згадані вище операції, а саме перемноження вагових множників a_{kl} з відповідними значеннями яскравостей початкового зображення і підсумовуванням добутків. Отримане значення присвоюється центральному (i, j) -му пікселю.

Маска містить непарне число рядків і стовпців, щоб центральний елемент визначався однозначно.

Розглянута вище процедура фільтрації характеризувалася тим, що вихідні значення фільтра g визначалися тільки через вхідні значення фільтра f . Такі фільтри називаються *нерекурсивними*. Фільтри, в яких вихідні значення g визначаються не тільки через вхідні значення f , але і через відповідні вихідні значення, називаються *рекурсивними*.

Лінійні фільтри для виділення контурів

Лінійні фільтри можуть бути призначені не для придушення шуму, а для підкреслення перепадів яскравості і контурів.

Виділення вертикальних перепадів здійснюється диференціюванням по рядках, горизонтальних – по стовпцях.

Горизонтальний перепад можна виділити шляхом обчислення приросту різниці яскравості пікселів уздовж рядка, що рівноцінно обчисленню другої похідної по напрямку (оператор Лапласа).

Оператори Лапласа реагують на перепади яскравості у вигляді ступеневого перепаду і на «дахоподібний» перепад. Вони також виділяють ізольовані точки, тонкі лінії, їхні кінці та гострі кути об'єктів. Лінія підкреслюється у 2 рази яскравіше, ніж ступінчастий перепад, кінець лінії у 3 рази яскравіше, а точка – у 4 рази. Оператор Лапласа виділяє в основному неструктуровані елементи, тому він чутливий до шуму. Підкреслення перепадів за допомогою оператора Лапласа відбувається без урахування їх орієнтації, оператор Лапласа інваріантний до орієнтації перепадів, наприклад, відгук оператора на похилий перепад у діагональному напрямку майже вдвічі більший, ніж у горизонтальному і вертикальному.

Нелінійні фільтри

Для підвищення стійкості при виділенні та накладення контурів, для усунення ефекту розмивання контурів під час придушення шуму слід переходити до нелінійної обробки. Прикладом нелінійного фільтру для придушення шуму, побудованого евристично, служить *медіанний фільтр*. Метод медіанної фільтрації, а також алгоритми *екстремальної фільтрації*, що використовують значення мінімуму і максимуму поточного околу, відносяться до так званих *рангових методів фільтрації*.

Зрозуміло, що при медіанній фільтрації може відбуватися спотворення об'єкта на зображенні, але тільки на межі або поблизу неї, якщо розміри об'єкта більші розмірів маски. Фільтр має високу ефективність при придушенні імпульсних перешкод, однак, ця якість досягається підбором розмірів маски, коли відомі мінімальні розміри об'єктів і максимальні розміри перекручених перешкодою локальних областей.

На випадковий шум з нормальним законом розподілу такий фільтр впливає слабше, ніж лінійний усереднюючий фільтр (приблизно на 60% менша ефективність).

Крім рангових алгоритмів фільтрації, під час нелінійної обробки зображень знаходять застосування *порогові алгоритми*, найбільш відомий з яких σ -фільтр. Ідея побудови σ -фільтру заснована на тому, що закон розподілу шуму можна наближено вважати гаусівським. Часто шум зумовлений багатьма незалежними або слабо залежними факторами.

Відомо, наприклад, що сума невеликого числа (5 або 6) незалежних рівномірно розподілених доданків досить добре підпорядковується гаусівському закону розподілу. Властивість нормалізації закону розподілу суми незалежних випадкових величин (за умови рівномірно малого вкладу доданків у суму) впливає з центральної граничної теореми теорії ймовірностей. Припущення про гаусівський характер закону розподілу суми виконується тим краще, чим більше число факторів її обумовлюють.

Для гаусівського розподілу 95,5% його значень лежить у межах відхилень від середнього значення, менших 2σ . Відповідно до алгоритму σ -фільтру оцінюється середнє значення і середньоквадратичне відхилення яскравостей у межах досить великої маски. Пікселі, що мають яскравість вище цього допуску, виключаються – це нелінійна операція. Замість виключених пікселів підставляється середнє по околу, можлива інтерполяція з використанням схеми авторегресії. Обраний допуск не так великий, щоб спотворити пікселі, що належать крутим елементам, дрібним деталям і слабо спотворені шумом.

Найбільш успішно σ -фільтр працює з імпульсними перешкодами у вигляді точок невеликої площі. Ймовірність того, що така перешкода потрапить на контур невелика, при використанні σ -фільтру спотворення контурів практично не відбувається. У цьому його перевага у порівнянні з лінійними і медіанними фільтрами. Однак, переваги σ -фільтру найпомітніше виявляються, коли розподіл шуму на зображенні близький до нормального. Сигма-фільтр може бути реалізований у рекурсивній формі.

Нелінійні фільтри, як і лінійні, можуть застосовуватися для виділення контурів і перепадів яскравості.

Відмінність нелінійних алгоритмів у тому, що вони використовують нелінійні оператори дискретного диференціювання (фільтр Робертса та фільтр Собела). На відміну від фільтра Лапласа, при використанні фільтрів Робертса і Собела контури об'єктів і перешкоди у вигляді точок і ліній мають однакову яскравість.

СПИСОК РЕКОМЕНДОВАНИХ ДЖЕРЕЛ

1. Методичні вказівки для виконання практичних та самостійної робіт з навчальної дисципліни «Цифрова обробка зображень» (для студентів 4 курсу денної форми навчання напряму 6.080101 – Геодезія, картографія та землеустрій) / Харків. нац. ун-т міськ. госп-ва ім. О. М. Бекетова ; уклад. І. С. Творошенко. – Харків : ХНУМГ ім. О. М. Бекетова, 2016. – 55 с.
2. Берлянт А. М. Картография : учебник для вузов / А. М. Берлянт. – М. : Аспект Пресс, 2001. – 336 с.
3. Шапиро Л. Компьютерное зрение / Л. Шапиро, Дж. Стокман. – М. : Бинум. Лаборатория знаний, 2006. – 716 с.
4. Форсайт Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс. – М. : Вильямс, 2004. – 928 с.
5. Савиных В. П. Аэрокосмическая фотосъемка / В. П. Савиных, А. С. Кучко, А. Ф. Стеценко. – М. : КартоГеоЦентр Геоиздат, 1997. – 378 с.
6. Янтуш Д. А. Дешифрирование аэрокосмических снимков / Д. А. Янтуш. – М. : Недра, 1991. – 240 с.
7. Лисицин В. З. Практикум по фотограмметрии и дистанционному зондированию / В. З. Лисицин. – Харьков : ХНАГХ, 2006. – 200 с.
8. Кашкин В. Б. Дистанционное зондирование Земли из космоса. Цифровая обработка изображений / В. Б. Кашкин, А. И. Сухинин. – М. : Логос, 2001. – 264 с.
9. Цифровая обработка изображений в информационных системах / И. С. Грузман, В. С. Киричук и др. – Новосибирск : НГТУ, 2002. – 352 с.
10. Трифонов Т. А. Геоинформационные системы и дистанционное зондирование в экологических исследованиях / Т. А. Трифонов. – М. : Академический проект, 2005. – 252 с.
11. Бондарев В. Н. Цифровая обработка сигналов: методы и средства : учеб. пособие / В. Н. Бондарев, Г. Трестер, В. С. Чернега. – Севастополь : СевГТУ, 1999. – 398 с.
12. Творошенко І. С. Конспект лекцій з дисципліни «Цифрова обробка зображень» (для студентів 5 курсу денної та заочної форм навчання спеціальності 7.08010105 – Геоінформаційні системи та технології) / І. С. Творошенко ; Харків. нац. ун-т міськ. госп-ва ім. О. М. Бекетова. – Харків : ХНУМГ ім. О. М. Бекетова, 2015. – 75 с.
13. Ватолин Д. Методы сжатия данных / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М. : Диалог-Мифи, 2002. – 384 с.
14. Прэтт У. Цифровая обработка изображений / У. Прэтт. – М. : Мир, 1982. – 480 с.

