

УДК 004.93'1

Інформаційні технології

Скірко Ілля Олегович

студент

Національний технічний університет України

“Київський політехнічний інститут”

Скирко Илья Олегович

студент

Национальный технический университет Украины

“Киевский политехнический институт”

Skirko I.O.

student

National Technical University of Ukraine “Kyiv Polytechnic Institute”

ПОРІВНЯЛЬНИЙ АНАЛІЗ МЕТОДІВ РОЗПІЗНАВАННЯ ТЕКСТУР

СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ РАСПОЗНАВАНИЯ

ТЕКСТУР

COMPARATIVE ANALYSIS OF TEXTURE RECOGNITION

METHODS

Анотація: Проведено дослідження методів розпізнавання текстур в контексті зображень, що отримані з безпілотних літальних апаратів.

Ключові слова: мультифрактальний спектр, фільтр, нейронна мережа.

Аннотация: Проведено исследование методов распознавания текстур в контексте изображений, полученных с помощью беспилотных летательных аппаратов.

Ключевые слова: мультифрактальный спектр, фильтр, нейронная сеть.

Summary: Conducted research of texture recognition methods in context of images from drones.

Key words: multyfractal spectrum, filter, neural network.

Вступ

Сьогодні проблема реалізації комп'ютерного зору є дуже популярною і такою, що стрімко розвивається багатьма науковими інститутами, окремими спеціалістами та великими компаніями. За роки розвитку та досліджень в області комп'ютерного зору було виявлено велику кількість окремих проблем та задач. Однією з таких задач є розпізнавання текстур.

За своєю природою кожна текстура є набором відносно простих структур (текстонів), що повторюються за деяким визначеним специфічним законом. Наприклад поверхня водойм є сукупністю хвиль, що йдуть одна за одною та в орієнтовані в одному напрямку. Таким чином поверхню води можна розглядати як текстуру, де елементарними частинами є окремі хвилі, а законом є правила їх розповсюдження. Іншими прикладами є: текстури лісу, дерева, трави, шкіри, тканини, поверхні асфальту, бетону, цегляної стіни та багато інших. Якщо звернути увагу предмети, що оточують нас у повсякденному житті, то практично на всіх можна помітити текстурні структури, які може бути корисно розпізнавати за допомогою комп'ютерних засобів для вирішення специфічних задач.

Оскільки задача розпізнавання текстур є перспективною та потрібною, існує велика кількість підходів та методів для її розв'язання, що відрізняються за представленням елементів текстур, законів їх розміщення відносно сусідніх.

Ціллю дослідження є пошук таких методів та підходів до розпізнавання текстур, які можуть стати складовою частиною системи розпізнавання текстурних регіонів на зображеннях, що отримані з безпілотних літальних апаратів. Текстурними регіонами можуть бути ліси, водойми тощо. Оскільки такі області не мають постійної форми, кольору чи кута освітлення, проводиться дослідження методів саме розпізнавання текстур.

1 Алгоритми розпізнавання текстур

1.1 Розпізнавання за допомогою мультифрактального спектру

Мультифрактальний спектр (MFS) — дескриптор текстур, що має за основу фрактальний аналіз та надає ефективний фреймворк для роботи із сутністю структури текстур. Даний дескриптор має властивість залишатись інваріантним до будь-якого гладкого перетворення, включно до змін точки зору, нежорстких деформацій поверхні текстури та змін у освітленні [1]. MFS можна розглядати як варіант гістограми яка охоплює деяку геометричну інформацію, що закодована у текстурі, і є дуже стійкою до змін навколишнього середовища.

Основою даного підходу є поняття фрактальної розмірності, яка є статистичною величиною, що дає опис того, наскільки складним або наскільки “нерівним” є геометричний об’єкт.

Фрактальна розмірність набору точок E на площині визначається як:

$$D = \dim(E) = \lim_{\delta \rightarrow 0} \frac{\log N(\delta, E)}{-\log \delta},$$

де $N(\delta, E)$ - найменша кількість множин діаметру δ якими можна покрити множину точок E .

Фрактальна розмірність D будь-якого об’єкту на площині знаходиться в інтервалі $[0; 2]$. Наприклад фрактальна розмірність точки дорівнює 0, гладкої кривої — 1, повністю заповненого прямокутника — 2.

Мультифрактальний спектр — розширення фрактальної розмірності. Для його знаходження спочатку визначається категоризація точок об’єкту відповідно до деякого критерію, і потім обчислюють фрактальну розмірність для кожної множини точок з цієї категоризації. Отриманий вектор фрактальних розмірностей дає глибокий опис структури, притаманної даній текстурі. MFS може бути визначений на різних функціях інтенсивності зображення — на результатах різних фільтрів.

У порівнянні з гістограмою, найбільш популярним засобом глобального статистичного оцінювання, мультифрактальний спектр відрізняється додатковим розрахунком фрактальних розмірностей замість звичайного підрахування пікселів у деякій множині, як це роблять для побудови гістограм. MFS може бути розглянутий, як гістограма, яка кодує інформацію про просторове розподілення множини точок, та інваріантна до геометричних перетворень.

Алгоритм розпізнавання текстур полягає у наступних етапах:

1. Визначити критерій категоризації. Їх може бути декілька.
2. Розділити вихідне зображення за критерієм категоризації на декілька множин точок.
3. Для кожної множини розрахувати фрактальну розмірність
4. Побудувати гістограму фрактальних розмірностей.
5. Порівняти отриману гістограму з відомими. Визначити ступінь відповідності. Якщо критеріїв категоризації декілька, тоді точність розпізнавання може бути покращена через побудову більшої кількості гістограм, та їх порівнювання з відомими для відповідних критеріїв категоризації.

1.2 Розпізнавання за допомогою нейронних мереж

Існує велика кількість видів нейронних мереж та їх застосування для розв'язання проблем комп'ютерного зору, і розпізнавання різноманітних об'єктів — одна із них. У випадках, коли поставлена задача навчити нейронну мережу розпізнати конкретний об'єкт, для обчислення параметрів мережі часто в якості вектору ознак використовують зображення з об'єктом (або без нього) з мінімальними перетвореннями. Такий підхід виправданий для “звичайних” предметів — предметів, для розпізнавання яких не важлива наявність великої кількості повторень у їх

структурі. Проте для розпізнавань саме текстур, необхідно мати вектор ознак, який буде містити у собі вичерпну інформацію про “закони” утворення конкретної текстури — про елементарний візерунок (текстон, частина текстури, що повторюється), частоту його повторення, розподіл у просторі, тощо. Вектор звичайних інтенсивностей пікселів для цієї задачі не підходить.

Одним із варіантів визначення вектору ознак є використання матриці сумісної появи (GLCM, gray-level co-occurrence matrix) для математичного представлення просторової залежності появи рівнів градацій сірого текстури на зображенні. Матриця сумісної появи C для зображення I розміром $n \times m$ розраховується наступним чином:

$$C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1, \text{ якщо } I(p, q) = i \text{ та } I(p + \Delta x, q + \Delta y) = j \\ 0, \text{ інакше} \end{cases},$$

де i, j — інтенсивності зображення, p, q — просторові позиції на зображенні, зміщення $\Delta x, \Delta y$ задаються як параметри. [2]

Використовуючи матрицю сумісної появи запропоновано [3] наступні метрики ознак текстур:

$$\text{Ентропія} = \sum_i \sum_j C(i, j) \log(C(i, j))$$

$$\text{Енергія} = \sum_i \sum_j C^2(i, j)$$

$$\text{Контраст} = \sum_i \sum_j (i - j)^2 C(i, j)$$

$$\text{Однорідність} = \sum_i \sum_j \frac{C(i, j)}{1 + |i - j|}$$

Дані ознаки об’єднуються у вектор разом з іншими ознаками, наприклад ознаками розподілу кольору, та подаються на вхід нейронної мережі для розрахування її параметрів та виконання процедури класифікації.

Для обрання структури нейронної мережі потрібно прийняти до уваги той факт, що текстури за своєю сутністю доволі прості тому немає

сенсу будували складні та глибокі нейронні мережі. Для деяких випадків достатньо нейронної мережі із прямим поширенням сигналу та зі зворотнім поширенням помилки, та невеликою кількістю шарів.

1.3 Розпізнавання динамічних текстур за рухом

Такі приклади текстур як вода, дерева, трава, вогонь та багато інших за своєю природою динамічні. Тобто структура даних текстур змінюється з часом за різними законами, притаманними різним конкретним явищам. Для розпізнавання таких текстур за рухом необхідно працювати з послідовностями зображень конкретних текстур, замість одиничних зображень для випадку статичних текстур.

Одним із підходів до вирішення даної проблеми є застосування моделей простору станів, які інтерпретують динамічні текстури як реалізацію моделі часових рядів, такої як процес авторегресії. Визначивши параметри моделі для таких послідовностей зображень можна сподіватись розпізнати однакові рухи порівнюючи дані моделі.[4]

Багато типових процесів розпізнавання ґрунтуються на сигналах зовнішнього вигляду. Проте в динамічних текстурах поєднано зовнішній вигляд із рухом, і це поєднання досить запутане. Лише зовнішній вигляд не надає повної інформації про конкретну текстуру. Таким чином модель розпізнавання повинна комбінувати розпізнавання за зовнішнім виглядом, та за рухом.

Розглянемо підхід з акцентом на власне рух.

Модель динамічної текстури представляє зображення використовуючи модель простору станів. Зображення представлені вектором-стовпцем y . Послідовність T зображень формують матрицю Y . За моделлю простору станів припускається, що кожне зображення є лінійним відображенням вектору станів малої розмірності x . Кожне зображення

додатково зашумлюється нормальним випадковим процесом з нульовим середнім та коваріаційною матрицею R .

$$y_t = Cx_t + w_t, w_t \sim N(0, R)$$

Матриця C іноді називається вихідною матрицею. Зміна стану моделюється за допомогою моделі авторегресії:

$$x_{t+1} = Ax_t + v_t, v_t \sim N(0, Q),$$

де A — матриця станів, Q — матриця коваріації шуму. Таким чином послідовність зображень представляється наступними параметрами $\theta = (C, A, Q)$, де C моделює зовнішній вигляд послідовності, а A та Q — її рух. Для навчання моделі необхідно провести оцінювання зазначених параметрів. Оскільки акцентується розпізнавання руху тому R не необхідно оцінювати.

Оцінювання параметрів виглядає наступним чином:

$$x_t = C^T y_t, t = 1 \dots T$$

$$A = \operatorname{argmin} \sum_{t=2}^T \|x_t - Ax_{t-1}\|$$

$$Q = \frac{1}{T-1} \sum_{t=2}^T (x_t - Ax_{t-1})(x_t - Ax_{t-1})^T$$

На рисунку 3 зображено приклади траєкторій x_t з двовимірного простору станів.

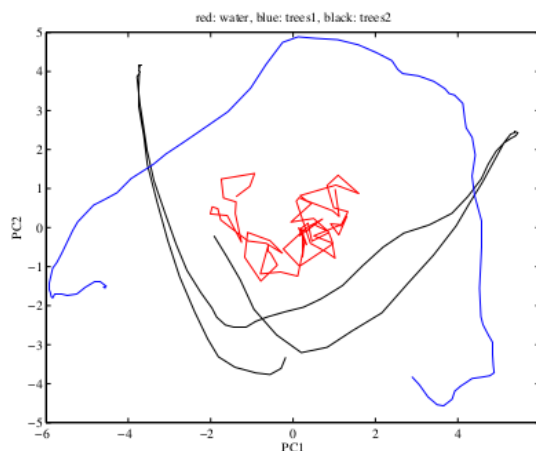


Рисунок 3. Траєкторія червоного кольору — рух води, синього та чорного — коливання дерев.

Розпізнавання конкретних послідовностей зображень полягає у порівнянні моделей простору станів. Звичайне поелементне порівнювання параметрів не є правильним, оскільки для будь-якої матриці M , що має обернену, моделі $(CM^{-1}, MAM^{-1}, MQM^T)$ та (C, A, Q) можуть згенерувати однакові послідовності зображень.

Пропонуються наступні способи розрахування відстані між моделями динамічних текстур:

Порівняння щільностей імовірностей серед всіх можливих послідовностей зображень, які можуть бути згенеровані моделями, що порівнюються.

Та розрахування відстані між двома моделями динамічних текстур є відстані основою яких є виділення особливостей. Розходження між двома моделями може бути виміряне як евклідова відстань між векторами особливостей або характеристик. Вектор особливостей це деяка векторно-значна функція від параметрів (C, A, Q) .

1.4 Розпізнавання текстур шляхом аналізу результатів роботи множини фільтрів

В загальному випадку обробки зображень, фільтр це деяка визначена функція, що застосовується до кожного пікселя (i, j) зображення та множини сусідніх пікселів. Після застосування фільтру на виході отримується матриця результатів фільтру, що відповідає за розмірами вихідному зображенню. Функції фільтрів можуть бути як лінійні — тоді вони є просто зваженими сумами значень пікселів; так і нелінійні.

Існує фундаментальна, добре вивчена проблема, що пов'язана з лінійними фільтрами. В загальному випадку вони реагують на будь-який стимул. Наприклад наявність реакції орієнтованого непарно-симетричного фільтра ще не означає наявності границі в цьому місці. Може бути, що є

смуга з вищою контрастністю в іншому місці та з іншою орієнтацією, яка спричинила цю реакцію. Такі ознаки як границі або смуги або кути не можуть бути пов'язані з результатом роботи одного фільтра [5].

В роботах [5-6] запропоновано використовувати результати роботи великої кількості фільтрів та розглядати ці множини результатів для кожного пікселя зображення як точки у просторі високої розмірності. Оскільки текстури за своєю природою складаються з частин, що повторюються, то будуть з'являтися однакові вектори результатів фільтрів так само як з'являються однакові ознаки текстури. Таким чином застосувавши множину фільтрів до навчального набору зображень та виконавши кластеризацію отриманих векторів результатів, можна отримати так звані "прототипні" вектори — вектори, що відповідають загальним особливостям серед усіх навчальних зображень.

Ядра фільтрів можна інтерпретувати як моделі чутливих областей у клітинах зорової кори головного мозку. Тоді їх можна класифікувати наступним чином:

Фільтри з радіально-симетричними чутливими областями. Зазвичай використовують різницю двох функцій Гауса з різними значеннями σ .

Орієнтовані, непарно-симетричні фільтри, чутливі області яких можуть бути представлені як повернуті копії горизонтальної непарно-симетричної області. Придатною для цього функція розподілу точок $\epsilon f(x, y) = G'_{\sigma_1}(y)G_{\sigma_2}(x)$, де $G_{\sigma}(x)$ представляє нормальний розподіл зі стандартним відхиленням σ . Відношення $\sigma_2:\sigma_1$ показує довжину фільтру

Орієнтовані, парно-симетричні фільтри, чутливі області яких можуть бути представлені як повернуті копії горизонтальної парно-симетричної області. Придатною для цього функція розподілу точок $\epsilon f(x, y) = G''_{\sigma_1}(y)G_{\sigma_2}(x)$ [6]

Фільтри цих типів зображені на рисунку 4: парні та непарні з шістьма варіантами повороту та у трьох масштабах, а також чотири

масштаби радіального фільтру.

Коли визначений словник “прототипних” векторів (текстонів), з’являється можливість аналізувати зображення у термінах текстонів. Тобто кожен піксель зображення відноситься до того чи іншого текстону, в залежності від отриманого вектору результатів роботи фільтрів з центром у даному пікселі. Таким чином можна порахувати кількість пікселів, що відповідають кожному текстону, та побудувати відповідні гістограми.

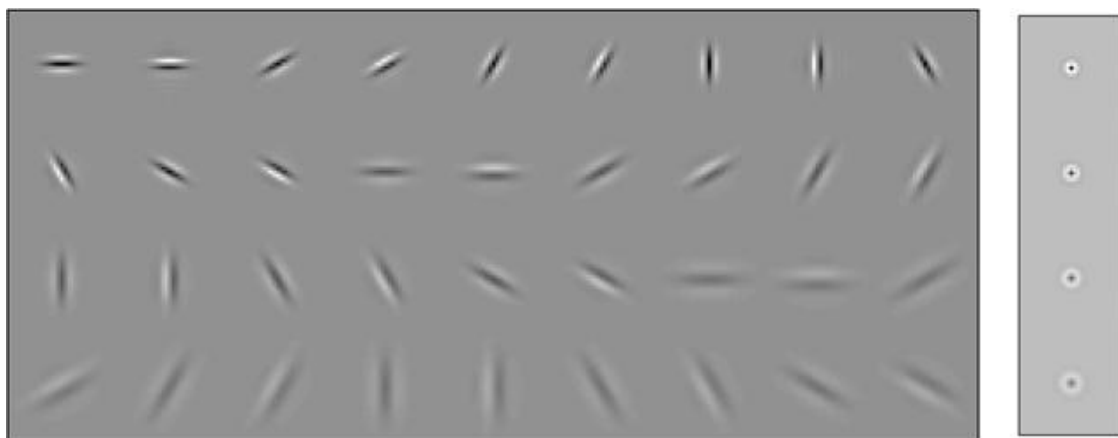


Рисунок 4. Приклади фільтрів.

Для розпізнавання зображення потрібно отримати результати роботи фільтрів у кожному пікселі та побудувати гістограму в термінах текстонів. Відповідність зображення тому чи іншому класові визначається за допомогою розрахування значення критерію схожості з універсальною гістограмою для кожного класу. [5]

1.5 Порівняння розглянутих методів розпізнавання текстур

Як і будь-який об’єкт розпізнавання, текстури потребують деякого вичерпного опису за допомогою деяких дескрипторів для побудови системи класифікації та розпізнавання. В даній статті розглянуті такі дескриптори, як: Мультифрактальний спектр, метрики ознак на основі матриці сумісної появи, моделі динамічних текстур, результат роботи множини фільтрів. В ролі класифікатора можуть виступати, наприклад,

нейронні мережі.

Побудова мультифрактального спектру надає можливість отримати вектор фрактальних розмірностей з деякого зображення, тим самим закодувати геометричне розміщення одних пікселів відносно інших, що знаходяться в одній “категорії”. Категоризація може відбуватись за інтенсивністю, енергією границь (energy of edges), енергією лапласіану (energy of the Laplacian), тощо.

Іншим варіантом опису текстур є використання таких ознак, як ентропія, енергія, контраст, однорідність, що розраховуються за допомогою матриці сумісної появи.

Використання банку фільтрів для виявлення текстонів також є потужним способом опису текстури.

У випадку динамічних текстур, єдиним логічним способом опису є побудова деяких статистичних моделей руху тої чи іншої текстури.

В таблиці 1 наведено порівняння основних особливостей розглянутих дескрипторів.

Застосування мультифрактального спектру та ознак на основі матриці сумісної появи для зображень, що містять тільки одну текстуру може бути реалізовано шляхом розрахування цих ознак для всього зображення. У інших випадках, коли на зображенні наявно більше текстур, є сенс розраховувати значення цих ознак для деякого вікна невеликого розміру, та рухати його по всьому зображенню.

Фільтри застосовуються також до невеликого вікна, що рухається в межах всього зображення, незалежно від вмісту зображення. Це потрібно, щоб виявити ключові текстони текстури, які можуть бути дуже малі за розміром.

Як можна помітити, дані дескриптори мають схожі сценарії застосування. Тому є сенс у комбінуванні цих підходів до опису текстури та виконання класифікації на основі одразу декількох дескрипторів.

Порівняння дескрипторів текстур

Дескриптор текстур	Основна особливість	Спосіб виявлення якісних особливостей текстур	Стійкість
Мульти-фрактальний спектр	Розраховуються фрактальні розмірності для підмножин пікселів зображення.	Фрактальна розмірність кодує “складність” розміщення пікселів. В загальному випадку може бути не унікальною.	Інваріантний до зсуву, повороту. Значна зміна освітлення може призвести до великих втрат якості.
Метрики на основі матриці сумісної появи	Використовуються GLCM для представлення просторової залежності появи пікселів на зображенні.	GLCM — оцінка щільності розподілу імовірностей сумісної появи градацій сірого. На основі цього розраховуються такі показники, як контраст, ентропія, тощо.	Матриця сумісної появи буде змінюватись разом із зображенням. Тому і показники не володітимуть властивістю інваріантності.
Застосування банку фільтрів	Накладаються фільтри, та розраховуються їх значення для кожного пікселя.	Кожен фільтр реагує на прості патерни. Застосувавши велику кількість фільтрів до однієї області, можна отримати вичерпний опис складної структури.	Інваріантні до зсуву, зміни яскравості. Поворот може викликати проблеми з орієнтованими фільтрами.
Статистична модель динамічної текстури	Послідовність зображень представляється у вигляді моделі простору станів.	Модель простору станів описує зміну зображення з часом.	Самі моделі простору станів не є інваріантними. Проте, можна застосувати стійкі способи їх порівняння.

Висновок

Проведене дослідження методів розпізнавання текстур показало, що головною проблемою є знаходження такого способу опису текстури, який би зміг описати структури текстур та виявити зв'язки всередині них. Одними з таких методів опису є: побудова мультифрактального спектру, розрахування метрик ознак на основі матриці сумісної появи, побудова моделей динамічних текстур, розрахування результат роботи множини фільтрів.

Для розпізнавання текстур на зображеннях з безпілотних літальних апаратів, найбільший інтерес представляють такі способи опису текстур: розрахування мультифрактального спектру та застосування банку фільтрів. По-перше, мультифрактальний спектр кодує "складність" форми фігури, що утворена пікселями зображення. І, хоча фрактальні розмірності в загальному випадку не є унікальними значеннями, на практиці розраховується декілька таких розмірностей для різних категорій пікселів одного зображення. По-друге, фільтри є досить простими структурами, але їх комбінації можуть дати представлення і про складні структури. Також існує гнучкість у виборі окремих фільтрів та їх параметрів, що дає можливість налаштувати банк фільтрів на реагування на специфічні текстури.

Література

1. Y Xu, H Ji, C. Fermüller, "Viewpoint invariant texture description using fractal analysis". *International Journal of Computer Vision*, 83 (1), 85 - 100 (2009).

2. Robert M Haralick, K Shanmugam, Its'hak Dinstein (1973). "Textural Features for Image Classification". *IEEE Transactions on Systems, Man, and*

Cybernetics. SMC-3 (6): 610–621.

3. Nidhal K. Al abbadi , Nizar Saadi Dahir, Zaid Abd Alkareem. "Skin Texture Recognition Using Neural Networks"

4. Franco Woolfe, Andrew W. Fitzgibbon. "Shift-Invariant Dynamic Texture Recognition". ECCV (2) 2006: 549-562

5. Laura Walker Renninger, Jitendra Malik. "When is scene recognition just texture recognition? ". *Vision Research*, 44, 2004, pp. 2301-2311.

6. Jitendra Malik, Serge Belongie, Thomas Leung, Jianbo Shi. "Contour and Texture Analysis for Image Segmentation". *International Journal of Computer Vision*, 43(1), 7-27, June 2001.