МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ В. Н. КАРАЗІНА
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ КОМП'ЮТЕРНОЇ ФІЗИКИ ТА ЕНЕРГЕТИКИ
Кафедра інформаційних технологій в фізико-енергетичних системах

ПОЯСНЮВАЛЬНА ЗАПИСКА

до дипломної роботи

за освітньо-кваліфікаційним рівнем магістр

«Використання прихованих марковських моделей для класифікації зображень»

Виконав: студент 2 курсу, групи НФ-63
Рак К. Д.
спеціальність:
105 Прикладна фізика та наноматеріали
освітня програма: Прикладна фізика
енергетичних систем
Керівник: к. ф.-м. н. Лісіна О. Ю.

Рецензент:

Харків – 2020

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Харківський національний університет імені В.Н.Каразіна

Навчально-науковий інститут комп'ютерної фізики та енергетики
Кафедра інформаційних технологій у фізико-енергетичних системах
Рівень вищої освіти (освітньо-кваліфікаційний рівень) магістр
Спеціальність 105 Прикладна фізика та наноматеріали

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ доц. Сухов Р.В.
підпис

"____"_____2020 року

З А В Д А Н Н Я
НА ДИПЛОМНУ РОБОТУ (ПРОЕКТ)

Рак Катерина Дмитріївна

1.Тема роботи «Використання прихованих марковських моделей для класифікації зображень»; керівник роботи: Лісіна Ольга Юліївна, кандидат фізико-математичних наук.

затверджені наказом по університету від "23" жовтня 2020 року №0210-05/1907

2. Строк подання студентом роботи 27.11.2020

3. Перелік питань, які потрібно розробити:

1) Провести аналіз методів обробки зображень;

2) Розробити математичне формулювання основних припущень двовимірної прихованої марковської моделі;

3) Обгрунтувати обчислювальну складність;

4) Розробити алгоритм класифікації зображення за допомогою двовимірної прихованої марковської моделі;

5) Провести тестування розробленого алгоритму класифікації.

4. План роботи

№ з/п	Назви етапів роботи
1	Вибір теми, узгодження, а також затвердження її на кафедрі
2	Робота над літературою та іншими джерелами
3	Затвердження теми на кафедрі
4	Написання вступу до роботи та першого розділу
5	Написання другого розділу
6	Написання третього розділу
7	Завершення написання дипломної роботи (висновки та література)
8	Подача науковому керівнику
9	Відгук, рецензія, підготовка до захисту
10	Попередній захист
11	Захист

5.Дата видачі завдання 26.10.2020

Студент _____ Рак Катерина Дмитріївна
підпис

Керівник роботи _____ Лісіна Ольга Юліївна
підпис

РЕФЕРАТ

Робота виконана Рак Катериною Дмитріївною на тему: Використання прихованих марковських моделей для класифікації зображень.

Дипломна робота представлена на 46 сторінках; складається зі вступу, 3 розділів та висновків; містить 13 рисунків, список використаної літератури з 35 найменувань.

Мета даної роботи – розробити алгоритм класифікації зображення за допомогою двовимірної прихованої марковської моделі та провести тестування розробленого алгоритму класифікації.

Об’єктом дослідження дипломної роботи є двовимірна прихована марковська модель.

Предметом дослідження є класифікація зображень за допомогою двовимірної прихованої марковської моделі.

Класифікація зображень, мабуть, найважливіша частина аналізу цифрових зображень. Застосування даного алгоритму можливо в різних областях обробки зображень, в тому числі і їх класифікацій. Наприклад, обробка і класифікація супутникових знімків, застосування в області медицини, а також для використання в будь-яких базах даних.

Алгоритм на базі двовимірної прихованої марковської моделі є ефективним і швидким, в порівнянні з іншими алгоритмами для класифікації. Основні труднощі застосування двовимірної прихованої марковської моделі полягають в обчислювальній складності.

КЛЮЧОВІ СЛОВА: ОБРОБКА ЗОБРАЖЕНЬ, ДВОВИМІРНА ПРИХОВАНА МАРКОВСЬКА МОДЕЛЬ, КЛАСИФІКАЦІЯ ЗОБРАЖЕНЬ, АЛГОРИТМ ВІТЕРБІ.

РЕФЕРАТ

Работа выполнена Рак Екатериной Дмитриевной на тему: Использование скрытых марковских моделей для классификации изображений.

Дипломная работа представлена на 416 страницах; состоит из введения, 3 глав и выводов; содержит 13 рисунков, список использованной литературы из 35 наименований.

Цель данной работы – разработать алгоритм классификации изображений с помощью двумерной скрытой марковской модели и провести тестирование разработанного алгоритма классификации.

Объектом исследования дипломной работы является двумерная Скрытая Марковская Модель.

Предметом исследования является классификация изображений с помощью двумерной скрытой марковской модели.

Классификация изображений является, пожалуй, самой важной частью анализа цифровых изображений. Применение данного алгоритма возможно в различных областях обработки изображений, в том числе и их классификации. Например, для обработки и классификации спутниковых снимков, применения в области медицины, а также для использования в любых базах данных.

Алгоритм на базе двумерной скрытой марковской модели является эффективным и быстрым, по сравнению с другими алгоритмами для классификации. Основные трудности применения двумерной скрытой марковской модели заключаются в вычислительной сложности.

КЛЮЧЕВЫЕ СЛОВА: ОБРАБОТКА ИЗОБРАЖЕНИЙ, ДВУМЕРНАЯ СКРІТАЯ МАРКОВСКАЯ МОДЕЛЬ, КЛАССИФИКАЦИЯ ИЗОБРАЖЕНИЙ, АЛГОРИТМ ВИТЕРБИ.

ABSTRACT

The work was done by Ekaterina Dmitrievna Rak on the topic: Using hidden Markov models for image classification.

Thesis is presented on page 46; consists of an introduction, 3 chapters and conclusions; contains 13 figures, a list of used literature of 35 titles.

The purpose of this work is to develop an image classification algorithm using a two-dimensional hidden Markov model and to test the developed classification algorithm.

The research object of the thesis is a two-dimensional Hidden Markov Model.

The subject of this research is the classification of images using a two-dimensional hidden Markov model.

Image classification is perhaps the most important part of digital image analysis. The application of this algorithm is possible in various areas of image processing, including their classifications. For example, processing and classification of satellite images, applications in the field of medicine, as well as for use in any databases.

The algorithm based on the two-dimensional hidden Markov model is efficient and fast compared to other algorithms for classification. The main difficulties in applying the two-dimensional hidden Markov model are computational complexity.

KEYWORDS: IMAGE PROCESSING, TWO-DIMENSIONAL HIDDEN MARKOV MODEL, IMAGE CLASSIFICATION, VITERBY ALGORITHM.

ЗМІСТ

ВСТУП.....9
1 АНАЛІЗ МЕТОДІВ ОБРОБКИ ЗОБРАЖЕНЬ.....11
1.1. Методи обробки зображень.....12
1.2. Техніка обробки зображень.....14
2 КЛАСИФІКАЦІЯ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ ДВОВИМІРНОЇ ПРИХОВАНОЇ МАРКОВСЬКОЇ МОДЕЛІ.....20
2.1. Двовимірна прихована марковська модель.....20
2.2. Математичне формулювання основних припущень двовимірної ПММ 21
2.3. Обчислювальна складність.....26
2.4. Алгоритм Вітербі зі змінним станом.....30
3 ЕКСПЕРИМЕНТАЛЬНІ РЕЗУЛЬТАТИ.....34
3.1. Сегментація аерофотознімків.....35
3.2. Сегментація зображення документа.....40
ВИСНОВКИ.....42
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....43

ПЕРЕЛІК СКОРОЧЕНЬ ТА АБРЕВІАТУР

ПММ – Прихована Марковська Модель

ДКП – Дискретно-косинусне перетворення

ВСТУП

Для блокової класифікації зображення ділиться на блоки, і для кожного блоку формується вектор ознак шляхом групування статистики, витягнутої з блоку. Традиційні алгоритми класифікації на основі блоків визначають клас блоку, досліджуючи тільки вектор ознак цього блоку і ігноруючи контекстну інформацію.

Для поліпшення класифікації по контексту пропонується алгоритм, що моделює зображення двовимірними прихованими марковськими моделями (ПММ). ПММ розглядає вектори ознак, статистично залежні від базового процесу стану, який, як передбачається, є марковською сіткою, ймовірності переходу якої залежать від станів сусідніх блоків як в горизонтальному, так і у вертикальному напрямках. Таким чином, залежність відбивається одночасно в двох вимірах. Параметри ПММ оцінюються за допомогою алгоритму. Щоб класифікувати зображення, класи з максимальною апостеріорною ймовірністю розшукуються спільно для всіх блоків.

Для більшості алгоритмів класифікації зображень на основі блоків, зображення діляться на блоки, і рішення приймаються незалежно для класу кожного блоку. Такий підхід призводить до проблеми вибору розмірів блоку. Вибір занадто великого розмір блоку, тягне за собою грубу класифікацію. З іншого боку, при виборі невеликого розміру блоку, при класифікації будуть вивчені тільки дуже локальні властивості, що належать невеликому блоку.

Мінусом є втрата інформації про оточуючі регіони. Добре відомий метод обробки сигналів для вирішення цієї проблеми – використання контекстної інформації. Розміри блоків, і правила класифікації можуть відрізнятися в залежності від контексту. Досягнуте поліпшення демонструє потенціал контексту для допомоги в класифікації. Двовимірна прихована марковська модель вводиться в якості загальної основи для контекстно-залежних класифікаторів.

Мета даної роботи – розробити алгоритм класифікації зображення за допомогою двовимірної прихованої марковської моделі та провести тестування розробленого алгоритму класифікації.

Об’єктом дослідження дипломної роботи є двовимірна прихована марковська модель.

Предметом дослідження є класифікація зображень за допомогою двовимірної прихованої марковської моделі.

Класифікація зображень, мабуть, найважливіша частина аналізу цифрових зображень. Застосування даного алгоритму можливо в різних областях обробки зображень, в тому числі і їх класифікацій. Наприклад, обробка і класифікація супутникових знімків, застосування в області медицини, а також для використання в будь-яких базах даних.

Алгоритм на базі двовимірної прихованої марковської моделі є ефективним і швидким, в порівнянні з іншими алгоритмами для класифікації. Основні труднощі застосування двовимірної прихованої марковської моделі полягають в обчислювальній складності.

Для досягнення поставленої мети були розроблені такі завдання:

- Провести аналіз методів обробки зображень;
- Розробити математичне формулювання основних припущень двовимірної прихованої марковської моделі;
- Обгрунтувати обчислювальну складність;
- Розробити алгоритм класифікації зображення за допомогою двовимірної прихованої марковської моделі;
- Провести тестування розробленого алгоритму класифікації.

1 АНАЛІЗ МЕТОДІВ ОБРОБКИ ЗОБРАЖЕНЬ

Обробка зображень – це метод поліпшення необроблених зображень, отриманих з камер, встановлених на супутниках, космічних апаратах і літаках, або зображень, зроблених в звичайному житті, для різних додатків.

По суті, всі операції обробки зображень можна згрупувати в декілька методів: представлення зображення, попередня обробка зображення, поліпшення зображення, відновлення зображення, аналіз зображення, реконструкція зображення, стиснення даних зображення.

У минулому йшлося про початкову обробку необроблених даних зображення для корекції геометричної деформації, радіометричної стандартизації даних та усунення шуму, присутнього в даних. Метою операцій вилучення інформації є заміна візуального аналізу даних зображення кількісними методами для автоматизації ідентифікації тону. Це включає в себе аналіз даних мультиспектрального зображення і застосування статистично обґрунтованих правил прийняття рішень для визначення ідентичності земного покриву кожного пікселя зображення.

Мета процесу категоризації – віднести всі пікселі цифрового зображення до одного з декількох класів або тем земного покриву. Проводиться аналіз різніх типів існуючих методів і алгоритмів обробки зображень.

За останні кілька десятиліть в обробці зображень були розроблені різні методи. Більшість методів розроблені для поліпшення зображень, отриманих з безпілотних космічних апаратів, космічних зондів і військових інспекційних польотів. Системи обробки зображень стають популярними завдяки доступності потужних персональних комп'ютерів, пристроїв пам'яті великого розміру і графічного програмного забезпечення.

Обробка зображень використовується в різних сферах, таких як дистанційне зондування, медична візуалізація, текстильна промисловість, матеріалознавство, військова промисловість, кіноіндустрія, обробка документів, графіка. Звичайними етапами обробки зображень є сканування, збереження, поліпшення та інтерпретації зображень.

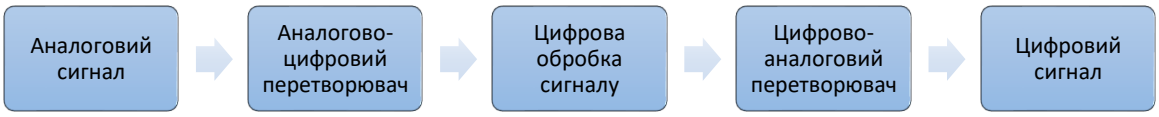


Рис. 1.1. Блок-схема цифрової обробки сигналів

1.1. Методи обробки зображень

Зазвичай обробка зображення (рис. 1.1.) складається з декількох етапів: імпорт зображення, аналіз, обробка та виведення зображення. Є два методи обробки зображень: цифровий та аналоговий.

Обробка аналогового зображення відноситься до зміни зображення за допомогою електричних засобів, наприклад, телевізійного зображення.

Телевізійний сигнал – це рівень напруги, який змінюється по амплітуді, щоб представити яскравість зображення. Змінюючи сигнал електрично, вид відображуваного зображення змінюється. Яскравість і контрастність на екрані телевізора служать для регулювання амплітуди і посилання відеосигналу, в результаті чого змінюється висвітленність, потемніння і діапазон яскравості зображення, що відображається.

В цьому випадку для обробки зображення використовуються цифрові комп'ютери. Зображення буде перетворено в цифрову форму за допомогою методу цифрового перетворення [1-4] і обробки. Він визначається як розрахунок числових уявлень об'єктів у серії операцій для отримання бажаного результату. Він починається з одного зображення і виробляє його модифіковану версію.

Отже, це процес, який перетворює зображення в інше. Цифрова обробка зображень відноситься до обробки двовимірного зображення на цифровому комп'ютері [5]. У більш широкому контексті мається на увазі цифрова обробка будь-яких двовимірних даних. Цифрове зображення – це масив дійсних чисел, представлених кінцевим числом бітів. Важлива перевага методів цифрової

обробки зображень є їх універсальність, повторюваність і збереження вихідної точності даних.

Комп'ютерні алгоритми грають вирішальну роль в обробці цифрових зображень. Розробники використовують різні алгоритми для вирішення різних завдань, включаючи виявлення цифрового зображення, аналіз, реконструкцію, відновлення, стиснення даних зображення, поліпшення зображення, оцінку зображення і спектральну оцінку зображення[6-7].

Основні способи обробки цифрових зображень:

• Редагування зображень, що в основному означає зміну цифрових зображень за допомогою графічних програмних засобів.

• Відновлення зображення, яке відноситься до оцінки чистого вихідного зображення з отриманого пошкодженого зображення, щоб повернути втрачену інформацію.

• Незалежний компонентний аналіз, який розділяє багатовимірний сигнал на додаткові подкомпоненти.

• Анізотропна дифузія, часто відома як дифузія Перона-Маліка, дозволяє зменшити шум зображення без видалення важливих частин зображення.

• Лінійна фільтрація. Це ще один метод обробки цифрових зображень, який відноситься до обробки вхідних сигналів, що змінюються в часі і створення вихідних сигналів, які підкоряються обмеженню лінійності.

• Нейронні мережі, які представляють собою обчислювальні моделі, широко використовувані в машинному навчанні для вирішення різних завдань.

• Пікселізація, яка часто відноситься до перетворення друкованих зображень в оцифровані (наприклад, GIF).

• Аналіз основних компонентів – метод обробки цифрових зображень, який можна використовувати для отримання ознак.

• Рівняння з частковими похідними, які також дозволяють ефективно усувати шум в зображеннях.

• Приховані марковські моделі, метод, який використовується для аналізу зображень в двох вимірах.

• Вейвлети, що означає математичну функцію, яка використовується при стисненні зображення.

• Самоорганізовуючі карти – метод обробки цифрових зображень для класифікації зображень по декількох класах.

1.2. Техніка обробки зображень

Класифікація способів обробки зображень наведена нижче.

1. Представлення зображення.

2. Попередня обробка зображень.

3. Поліпшення зображення.

4. Аналіз зображення.

5. Стиснення зображення.

Зображення, зазначене як «реальний світ», вважається функцією двох реальних змінних, наприклад, $f\ (x, y)$, де f – амплітуда зображення в дійсній координатної позиції (x, y). Операція обробки зображення зазвичай визначає нове зображення g в термінах існуючого зображення f. Ефект оцифровки наведено на рис. 1.2.

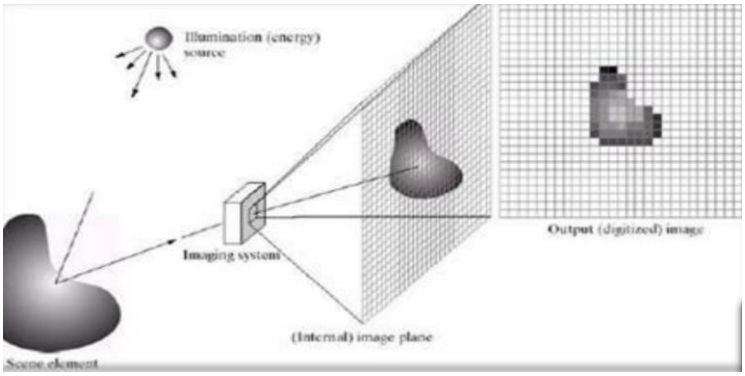


Рис. 1.2 Оцифрування зображення

Попередня обробка показує, що один і той же тип може мати різну шкалу інтенсивності сигналу для різних зображень. Функції попередньої обробки включають в себе ті операції, які звичайно потрібні перед аналізом основних даних і витяганням інформації і зазвичай групуються як радіометричні або геометричні поправки. Радіометричні корекції включають в себе виправлення даних на наявність несправностей датчика і небажаного датчика або

атмосферного шуму і перетворення даних, щоб вони точно відображали відбите випромінювання або те, що випускається, аби точно визначити перетворення між двома зображеннями.

Попередньо оброблені зображення будуть мати певний шум, який необхідно видалити для подальшої обробки зображення. Шум зображення найбільш помітний в областях зображення з низьким рівнем сигналу, таких як області тіні або недоекспоновані зображення. Існує так багато типів шуму, як шум «солі» і «перцю», «зернистість» плівки. Всі ці шуми усуваються за допомогою алгоритмів. Серед кількох фільтрів використовується медіанний фільтр.

Методи поліпшення зображення можна розділити на дві широкі категорії:

1. Методи просторової області, які працюють безпосередньо з пікселями.

2. Методи частотної області, які працюють з перетворенням Фур'є у зображенні.

Значення пікселя з координатами (x, y) в покращеному зображенні F є результатом виконання деякої операції над пікселями в околиці (x, y) у вхідному зображенні F. Околиці можуть мати будь-яку форму, але зазвичай вони прямокутні [8]. Є два методи для категорії просторової області:

1) Маніпуляція сірою шкалою

Найпростіша форма операції – це коли оператор T діє тільки на піксельну околицю 1x1 у вхідному зображенні $F(x, y)$, яка залежить тільки від значення F в точці (x, y). Це перетворення або відтворення сірої шкали. Спрощена форма випадку – це порогова обробка, при якій профіль інтенсивності замінюється ступінчастою функцією, активної при обраному пороговому значенні. У цьому випадку будь-який піксель з рівнем сірого нижче порогового значення у вхідному зображенні порівнюється з 0 в вихідному зображенні «1.3.».

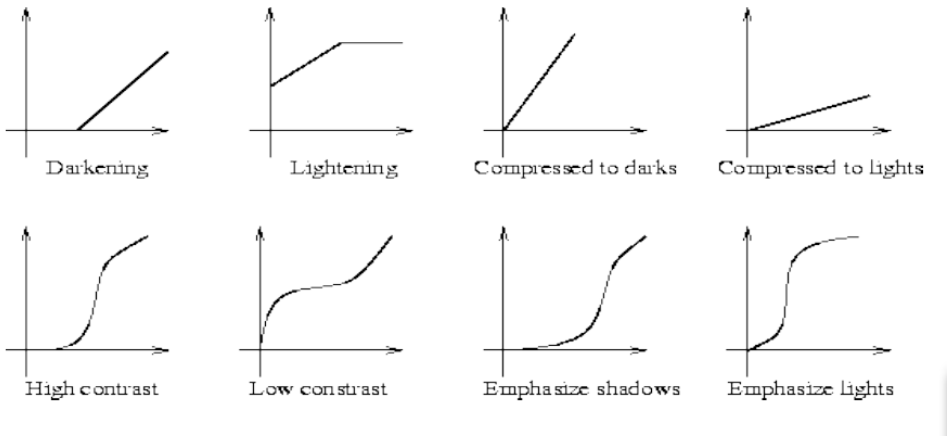


Рис. 1.3 Перетворення сірої шкали

2) Вирівнювання гістограми

Вирівнювання гістограми – поширений метод поліпшення зовнішнього вигляду зображень. Припустимо, у нас є зображення, яке переважно темне. Тоді його гістограма буде зміщена до нижнього краю сірої шкали, і всі деталі зображення будуть стиснуті в темному кінці гістограми. Якби ми могли «розтягнути» рівні сірого на темному кінці, щоб отримати більш рівномірно розподілену гістограмму, тоді зображення стало б набагато чіткіше «1.4.».

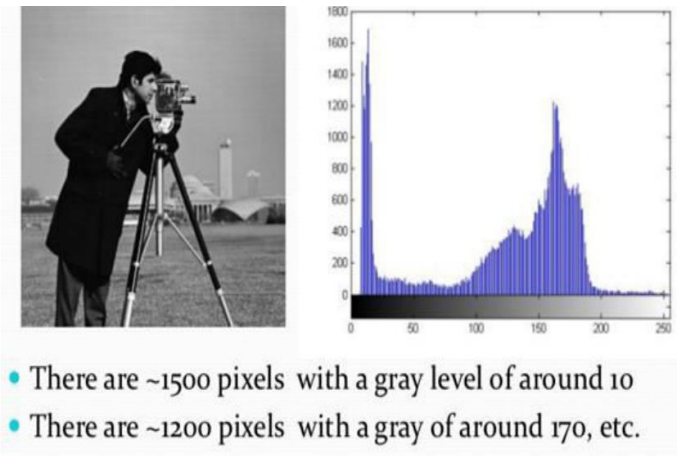


Рис. 1.4 Оригінальне зображення оператора і його гістограми та вирівнювання версії. Обидва зображення квантовано до 64 рівня сірого

Поліпшення зображення в частотній області просте. Ми обчислюємо перетворення Фур'є у зображенні, яке потрібно поліпшити, множимо результат на фільтр і виконуємо зворотне перетворення, щоб отримати покращене

зображення. Ідея розмиття зображення шляхом зменшення його високочастотних компонентів або підвищення різкості зображення шляхом збільшення величини його високочастотних компонентів інтуїтивно легко зрозуміти.

Однак з обчислювальної точки зору часто більш ефективно реалізувати ці операції як згортки за допомогою невеликих просторових фільтрів в просторовій області. Розуміння концепцій частотної області важливе і веде до методів поліпшення, про які, можливо, не доводилося думати, обмежуючи увагу просторовою областю.

Фільтрація нижніх частот включає у себе усунення високочастотних компонентів зображення. Це призводить до розмиття зображення. Ідеальний фільтр нижніх частот зберігав би всі низькочастотні компоненти і усував би все високочастотні компоненти. Однак ідеальні фільтри страждають двома проблемами: розмиванням і дзвоном. Ці проблеми викликані формою відповідного фільтра просторової області, який має велику кількість коливань. Більш плавні переходи в фільтрі частотної області, такому як фільтр Баттерворта, дають набагато кращі результати «рис. 1.5».



Рис. 1.5 Функція перетворення оригінального зображення на розмите зображення

Метод аналізу зображень виводить інформацію із зображення за допомогою автоматичних або напівавтоматичних методів, що називаються

аналізом сцен, описом зображень, розумінням зображень, розподіленням зображень, комп'ютерним/машинним зором. Аналіз зображень відрізняється від інших типових методів обробки зображень, таких як поліпшення або відновлення, тим, що в результаті аналізу результатів аналізу зображень представлені числові вихідні дані, а не зображення.

Метою стиснення зображень є зменшення розміру цифрових зображень для економії місця для зберігання і часу передачі. Стиснення без втрат переважно використовують для штучних зображень, таких як технічні креслення, значки, а також для цінних матеріалів, таких як медичні зображення або скановані зображення, зроблені для архівних цілей. Методи з втратами особливо підходять для природних зображень, таких як фотографії, в додатках, де допустима незначна втрата точності для досягнення значного зниження швидкості передачі даних.

Стиснення з втратами, яке створює непомітні відмінності, можна назвати візуально без втрат. Кодування довжин серій, кодування Хаффмана і кодування Лемпеля Зива – це методи стиснення зображень без втрат. Кодування з перетворенням: застосовується вейвлет-перетворення, за яким слідує квантування, кодування символів можна назвати методом стиснення зображень з втратами.

Обробку цифрових зображень супутникових даних в першу чергу можна розділити на кілька категорій, таких як виправлення зображень, поліпшення і зняття інформації. Виправлення зображень – це попередня обробка супутникових даних для геометричних і радіометричних зв'язків. До даних зображень застосовується поліпшення, щоб ефективно відображати дані для подальшої візуальної інтерпретації. Вивід інформації заснований на цифровій класифікації і використовується для створення цифрової тематичної карти.

Порівнювати продуктивність методів обробки зображень складно, якщо не використовуються ідентичні набори даних і показники продуктивності. Вивчені в різних статтях, деякі методи обробки зображень використовуються для поліпшення техніки для різних типів введення даних. Існують різні типи

методів кодування символів, які можна використовувати для обробки зображень.

Після вивчення всіх методів було виявлено, що методи стиснення зображень без втрат є найбільш ефективними в порівнянні з методами стиснення з втратами.

2 КЛАСИФІКАЦІЯ ЗОБРАЖЕНЬ ЗА ДОПОМОГОЮ ДВОВИМІРНОЇ ПРИХОВАНОЇ МАРКОВСЬКОЇ МОДЕЛІ

2.1. Двовимірна прихована марковська модель

Дана модель є «псевдо-двовимірною» в тому сенсі, що це не повністю двомірна ПММ. Основне припущення полягає в тому, що існує набір марковських «супер станів». Усередині кожного супер стану є набір простих марковських станів.

Для двовимірних зображень спочатку вибирається супер стан з використанням ймовірності марковського переходу першого порядку, заснованої на попередньому супер стані. Це супер стан визначає простий ланцюг Маркова, що буде використовуватися повністю усім рядком. Потім для створення спостережень в рядку використовується простий ланцюг Маркова. Таким чином, супер стани відносяться до рядків, а прості стани – до стовпців. Очікується, що псевдо-2D ПММ буде набагато ефективнішою зі звичайними зображеннями, такими як документи. Оскільки вплив стану пікселя на стан під ним розподіляється по всьому рядку, псевдо-двовимірна модель занадто обмежена для нормальної класифікації зображень.

Зусилля, присвячені застосуванню дійсно двовимірної ПММ для сегментації зображень, були вперше зроблені Девіджвером [9]. Девіджвер запропонував представляти зображення в вигляді прихованих марковських моделей з процесами стану, які є марковськими сітками, зокрема, марковськими сітками другого і третього порядків, перша з яких буде в центрі уваги наступних розділів. Були вивчені додатки для сегментації, відновлення і стиснення зображень [10]. Було відзначено, що складність оцінки моделей або їх використання для виконання максимальної апостеріорної класифікації експоненціально залежить від розміру зображення. Були розроблені обчислювально-здійсненні алгоритми [11] шляхом додаткових припущень щодо моделей або використання локально оптимальних рішень. Звертає на себе увагу детермінований релаксаційний алгоритм [12] для пошуку максимальних

апостеріорних станів. Алгоритм ітеративно оптимізує стан, вносячи локальні зміни в поточні стани таким чином, щоб збільшити ймовірність всього зображення. Результат критично залежить від початкових станів.

Інша робота, заснована на двовимірних ПММ, включає алгоритм розпізнавання символів, розроблений Левіним і Піерачіні [13-16], і систему декодування зображень по каналах з шумом, створену Парком і Міллером [17-19], де двовимірні ПММ з марковскими сітками використовуються для моделювання каналів з шумом , і в цьому випадку базові стани, відповідні істинним індексам, переданим кодером, спостерігаються з навчальних даних. Отже, оцінити моделі нескладно, тоді як оцінка є основною складністю для ситуацій, коли стани є неспостережними.

2.2. Математичне формулювання основних припущень двовимірної ПММ

Як і у всіх системах класифікації на основі блоків, зображення, яке підлягає класифікації, поділяється на блоки, а вектори ознак оцінюються як статистика блоків. Потім зображення класифікується відповідно до векторів ознак.

Двовимірна ПММ передбачає, що вектори ознак генеруються марковською моделлю, яка може змінювати стан один раз в кожному блоці. Припустимо, що є M станів, {1,…..,M}, стан блоку (i, j) позначається $s_{i,j}$.

Вектор ознак блоку (i, j) – це $u_{i,j}$, а клас – $c_{i,j}$. Позначимо $(i`, j`) < (i, j)$ або $(i, j) > (i`, j`)$, якщо $i` < i$ або $i`= i$ та $j` < j$, в цьому випадку ми говоримо, що блок $(i`, j`)$ стоїть перед блоком (i, j). Наприклад, на лівій панелі «рис. 2.1» блоки перед (i, j) є заштрихованими. Цей порядок такий, як і порядок растрів по рядках.

Відзначимо, однак, що цей порядок введений тільки для викладу припущень. При класифікації, блоки не класифікуються по одному в такому порядку. Алгоритм класифікації намагається знайти оптимальну комбінацію класів спільно для багатьох блоків одночасно.

Одновимірний підхід спільної класифікації, що передбачає порядок сканування при класифікації, зазвичай неоптимальний.

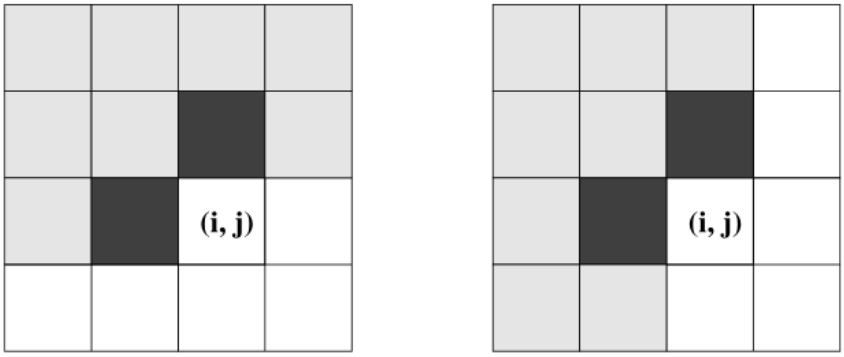


Рис. 2.1 Марківська властивість переходів між станами

Першим припущенням є таке:

$$P\left(s_{i,j}\middle|s_{i`,j`},u_{i`,j`}:(i`,j`)\in\Psi\right)=a_{m,n,l}, \qquad (2.1)$$

де $\Psi=\{(i`j`):(i`,j`)(i,j)\}$ та $m=s_{i-1,j}$, $n=s_{i,j-1}$, $l=s_{i,j}$.

Вищевикладене припущення може бути узагальнено двома моментами. По-перше, стан $s_{i`,j`}$ є достатньою статистикою для $(s_{i`,j`},u_{i`,j`})$ оцінки ймовірності переходу, тобто U умовно не має пам'яті.

По-друге, перехід стану є марковським в двовимірному сенсі першого порядку. Імовірність переходу системи в певний стан залежить від стану системи при сусідніх спостереженнях як в горизонтальному, так і у вертикальному напрямках. Дозволено перехід з будь-якого стану в будь-який стан. Як показано на лівій панелі «рис. 2.1», знаючи стан всіх заштрихованих блоків, нам потрібні тільки стан двох сусідніх блоків більш темного відтінку, щоб обчислити вірогідність переходу в наступний стан. Також передбачається, що існує унікальне відображення станів в класи. Таким чином, класи блоків визначаються після того, як стан відомий.

Друге припущення полягає в тому, що для кожного стану вектори ознак слідують гауссовому розподілу суміші [20]. Як тільки стан блоку стане відомим, вектор ознак умовно не залежить від інших блоків. Оскільки будь-який стан з М-компонентною гауссовою сумішшю може бути розділений на М підстанцій з одинарними гауссовими розподілами, модель обмежує нас

єдиними гауссовими розподілами. Для блоку зі станами та вектором ознак u розподіл має щільність:

$$b_s(u) = \frac{1}{\sqrt{(2\pi)^k |\Sigma_s|}} e^{-\frac{1}{2}(u-\mu_s)^t \Sigma_s^{-1}(u-\mu_s)}, \quad (2.2)$$

де Σ_s – матриця коваріації, а μ_s – середній вектор.

Марківське припущення про перехід станів може значно спростити оцінку ймовірності станів, тобто $P\{s_{i`j} : (i, j) \in \mathrm{N}\}$, де $N = \{(i,j) : 0 \le i < w,\ 0 \le j < z\}$, відноситься до всіх блоків на зображенні.

Щоб ефективно розширити цю ймовірність з допомогою формули умовної ймовірності [21-22], ми спочатку доведемо, що повернена форма двовимірної марківської властивості виконується при двох припущеннях.

Згадаймо визначення: (i`, j`) < (i, j), якщо i` < i або i` = i та j` < j. Потім ми визначаємо повернене відношення «<», що позначається «$\widetilde{<}$», яке вказує (i`, j`) $\widetilde{<}$ (i, j) або (i, j) $\widetilde{>}$ (i`, j`), якщо j`< j або j`= j та i`< i. Приклад показаний на правій панелі рис. 2.1. Щоб довести, що:

$$P\left(s_{i,j} \middle| s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \widetilde{\Psi}\right) = a_{m,n,l},$$

де $\widetilde{\Psi} = \{(i`j`) : (i`, j`) \widetilde{<} (i, j)\}$, та $m = s_{i-1,j}$, $n = s_{i,j-1}$, $l = s_{i,j}$,

використаємо попереднє визначення $\Psi = \{(i`j`) : (i`, j`) < (i, j)\}$ і введемо такі позначення:

$$\widetilde{\Psi} \cup \Psi = \{(i`, j`) : (i`, j`) \lhd (i, j) або (i`, j`) \widetilde{\lhd} (i, j)\},$$

$$\widetilde{\Psi} \cap \Psi = \{(i`, j`) : (i`, j`) \lhd (i, j) та (i`, j`) \widetilde{\lhd} (i, j)\},$$

$$\widetilde{\Psi} - \Psi = \{(i`, j`) : (i`, j`) \widetilde{\lhd} (i, j) та (i`, j`) \rhd (i, j)\}.$$

Зауважимо, що $\widetilde{\Psi} = (\widetilde{\Psi} \cap \Psi) \cup (\widetilde{\Psi} - \Psi)$.

Позначаємо $\gamma_0 = P(s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \widetilde{\Psi})$ та

$\gamma_1 = P(s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \widetilde{\Psi} \cap \Psi)$.

Потім ми можемо вивести:

$$P\left(s_{i,j} \middle| s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \widetilde{\Psi}\right) =$$

$$= \frac{1}{\gamma_0} \cdot P\left(s_{i,j}, s_{i`,j`}, u_{i`,j`}, s_{i``,j``}, u_{i``,j``} : (i`, j`) \in \tilde{\Psi} \cap \Psi, (i``, j``) \in \tilde{\Psi} - \Psi\right) =$$

$$= \frac{\gamma_1}{\gamma 0} \cdot P\left(s_{i,j} \middle| s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \tilde{\Psi} \cap \Psi\right).$$

$$P\left(s_{i``,j``}, u_{i``,j``} : (i``, j``) \in \tilde{\Psi} - \Psi \middle| s_{i,j}, s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \tilde{\Psi} \cap \Psi\right) =$$

$$= \frac{\gamma_1}{\gamma_0} \cdot P\left(s_{i,j} \middle| s_{i-1,j}, s_{i,j-1}\right). \tag{2.3}$$

$$P\left(s_{i``,j``}, u_{i``,j``} : (i``, j``) \in \tilde{\Psi} - \Psi \middle| s_{i,j}, s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \tilde{\Psi} \cap \Psi\right) =$$

$$= P\left(s_{i,j}, \middle| s_{i-1,j}, s_{i,j-1}\right). \tag{2.4}$$

$$\frac{\gamma_1}{\gamma_0} \cdot P\left(s_{i``,j``}, u_{i``,j``} : (i``, j``) \in \tilde{\Psi} - \Psi \middle| s_{i,j}, s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \tilde{\Psi} \cap \Psi\right) =$$

$$= P\left(s_{i,j}, \middle| s_{i-1,j}, s_{i,j-1}\right) = a_{m,n,l}. \tag{2.5}$$

де $m = s_{i-1,j}$, $n = s_{i,j-1}$, $l = s_{i,j}$.

Рівність (2.3) випливає з розкладу умовної ймовірності. Рівність (2.4) випливає з марковського припущення. Рівність (2.5) виконується як через марковське припущення, так і з-за припущення, що вектор ознак блоку умовно не залежить від інших блоків при його стані.

З цього випливає наступне твердження, а саме:

$$P\left(s_{i,j} \middle| s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \tilde{\Psi} \cup \Psi\right) = P\left(s_{i,j}, \middle| s_{i-1,j}, s_{i,j-1}\right). \tag{2.6}$$

Причина в тому, що при виведенні, якщо ми замінимо $\tilde{\Psi} \cap \Psi$ на Ψ і $\tilde{\Psi}$ на $\tilde{\Psi} \cup \Psi$, всі рівності збережуться.

Оскільки з рівняння (2.6), очевидно, слідує вихідне Марковське припущення і його обернена версія, ми показали еквівалентність двох припущень:

$P\left(s_{i,j} \middle| s_{i`,j`}, u_{i`,j`} : (i`, j`) \in \tilde{\Psi}\right) = P\left(s_{i,j}, \middle| s_{i-1,j}, s_{i,j-1}\right)$ та

$$P\left(s_{i,j}\middle|s_{i`,j`},u_{i`,j`}:(i`,j`)\in\tilde{\Psi}\right)=P\left(s_{i,j},\middle|s_{i-1,j},s_{i,j-1}\right).$$

Ми відзначаємо, що цей процес станів є окремим випадком марковского випадкового поля (МВП) [26-27], яке було названо марковською сіткою і запропоновано Абендом, Харлі і Каналом [23] для класифікації бінарних випадкових образів. Марковська сітка називається «причинною» [24], тому що стани в умові – це стани «минулих» блоків вгорі і зліва від поточного блоку. Причинність дозволяє отримати аналітичний ітераційний алгоритм для оцінки ПММ і для оцінки станів з максимальною апостеріорної ймовірністю.

Тепер ми готові спростити розширення $P\left\{s_{i`j}:(i,j)\in \mathrm{N}\right\}$:

$$P\left\{s_{i`j}:(i,j)\in \mathrm{N}\right\}=P(T_0)\cdot P\left(T_1\middle|T_0\right)\cdots P\left(T_{w+z-2}\middle|T_{w+z-3},T_{w+z-4},...,T_0\right). \qquad (2.7)$$

де T*i* позначає послідовність станів блоків на діагоналі *i*, $\{s_{i,0},s_{i-1,1},\cdots,s_{0,i}\}$ та *w* і *z* кількість рядків і стовпців відповідно « рис. 2.2.».

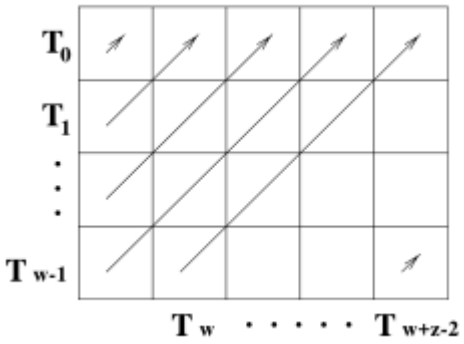


Рис. 2.2 Блоки по діагоналях зображення

Далі ми показуємо, що $P\left(T_i\middle|T_{i-1},...,T_0\right)=P\left(T_i\middle|T_{i-1}\right)$. Без обмеження спільності припустимо $T_i=\{s_{i,0},s_{i-1,1},\cdots,s_{0,i}\}$; тоді $T_{i-1}=\{s_{i-1,0},s_{i-2,1},\cdots,s_{0,i-1}\}$ та

$$P\left(T_i\middle|T_{i-1},...,T_0\right)=P\left(s_{i,0},s_{i-1,1},...,s_{0,i}\middle|T_{i-1},T_{i-2},...,T_0\right)=$$

$$=P\left(s_{i,0}\middle|T_{i-1},...,T_0\right)\cdot P\left(s_{i-1,1}\middle|s_{i,0},T_{i-1},...,T_0\right)$$

$$\cdots P\left(s_{0,i}\middle|s_{1,i-1},...,s_{i,0},T_{i-1},...,T_0\right)=$$

$$= P\left(s_{i,0}\middle|s_{i-1,0}\right)\cdot P\left(s_{i-1,1}\middle|s_{i-2,1},s_{i-1,0}\right)\cdots P\left(s_{0,i}\middle|s_{0,i-1}\right).$$

Остання рівність виходить з рівняння (2.6). Оскільки усі стани $s_{i,j}$, які з'являються в умовах, знаходяться в T_{i-1}, робиться висновок, що

$$P\left(T_i\middle|T_{i-1},...,T_0\right) = P\left(T_i\middle|T_{i-1}\right).$$

Рівняння (2.7) спрощується до

$$P\left\{s_{i`j} : (i,j)\in \mathrm{N}\right\} = P\left(T_0\right)\cdot P\left(T_1\middle|T_0\right)\cdots P\left(T_{w+z-2}\middle|T_{w+z-3}\right). \qquad (2.8)$$

Послідовність станів T_i, таким чином, служить «ізолюючим» елементом в розкладанні $P\left\{s_{i`j} : (i,j)\in \mathrm{N}\right\}$, який грає роль стану в одиницю часу в разі одновимірної марковської моделі. Як ми побачимо, ця властивість необхідна для розробки алгоритму. Ми можемо помітити, що крім діагоналей існують і інші геометричні форми, які можуть служити «ізолюючими» елементами, наприклад, послідовність станів в рядках або стовпцях. Послідовності станів T_i на діагоналях кращі з обчислювальних причин.

Завдання класифікатора полягає в тому, щоб оцінити двовимірну СММ на основі навчальних даних і класифікувати зображення шляхом знаходження комбінації станів з максимальною апостеріорної ймовірністю, що задана спостерігаючими векторами ознак.

2.3 Обчислювальна складність

Розрахунок ймовірностей $H_{m,n,l}^{(p)}$ і $L_m^{(p)}$ є ключем до ітераційної оцінки параметрів моделі. Якщо ми обчислюємо $L_m^{(p)}$ і $H_{m,n,l}^{(p)}$ безпосередньо відповідно до рівнянь (14) і (15), нам необхідно враховувати всі комбінації станів, які дають ті ж класи, що і в навчальній вибірці.

$$L_m^{(p)}(i,j) = \sum_s I\left(m = s_{i,j}\right)\cdot\frac{1}{\alpha}I\left(C(s)=c\right). \qquad (2.9)$$

$$H_{m,n,l}^{(p)}(i,j) = \sum_s I\left(m = s_{i-1,j}, n = s_{i,j-1}, l = s_{i,j}\right)\cdot\frac{1}{\alpha}I\left(C(s)=c\right). \qquad (2.10)$$

Велика кількість таких комбінацій станів призводить до неможливості обчислення.

Візьмемо, наприклад, $L_m^{(p)}$. Припустимо, що є M_0 станів для кожного класу і кількість блоків у зображенні $w \times z$, як передбачалося раніше, тоді кількість допустимих комбінацій станів, які задовольняють $C(s) = c$ і $s_{i,j} = m$, так само $M_0^{(w \times z - 1)}$. При застосуванні алгоритму ПММ, хоча одне зображення часто ділиться на безліч частин зображення, так що w або z – це кількість блоків в одному стовпці або одному рядку в допоміжному зображенні, нам потрібно зберегти значення w і z досить великим, щоб гарантувати включення адекватної кількості контекстної інформації в класифікацію. У межі, якщо $w = z = 1$, алгоритм являє собою просто алгоритм параметричної класифікації, що виконується незалежно для кожного блоку. Зазвичай $w = z = 8$. У цьому випадку, якщо є 4 стану для кожного класу, кількість комбінацій станів буде $M_0^{(w \times z - 1)} = 4^{63}$, що є неприпустимим для прямого обчислення $L_m^{(p)}$. Аналогічні труднощі виникають при оцінці одновимірної ПММ. Проблема вирішується рекурсивним обчисленням прямої і зворотньої ймовірностей [25]. Ідея використання прямих і зворотних ймовірностей може бути поширена на двовимірний ПММ для спрощення обчислень.

Нагадаємо рівняння (2.8).

$$P\{s_{i`j} : (i, j) \in \mathrm{N}\} = P(T_0) \cdot P(T_1 | T_0) \cdots P(T_{w+z-2} | T_{w+z-3})$$

Той факт, що послідовність станів T_i на діагоналі є «ізолюючим» елементом в розкладанні $P\{s_{i`j} : (i, j) \in \mathrm{N}\}$, дозволяє нам визначати прямі і зворотні ймовірності і оцінити їх за допомогою рекурсивних формул.

Пояснимо спочатку позначення:

1. Діагональ, на якій лежить блок (i, j), позначається $\Delta(i, j)$.

2. Вектори ознак на діагоналі d, $\{u_{i,j} : \Delta(i, j) = d\}$ позначається $u(d)$.

3. Послідовність станів на діагоналі d, $\{s_{i,j} : \Delta(i, j) = d\}$ позначається $s(d)$.

4. Для послідовності станів T на діагоналі d її значення в блоці (i, j) дорівнює $T(i, j)$.

Пряма ймовірність $\theta_T(d)$ для деякої моделі M визначається як:

$$\theta_T(d) = P\{s(d) = T, u(\tau): \tau \le d | M\}$$

Пряма ймовірність $\theta_T(d)$ – це ймовірність спостереження векторів, що лежать на діагоналі d або вище і мають послідовність станів T для блоків на діагоналі d.

Зворотня ймовірність $\beta_T(d)$ визначається як:

$$\beta_T(d) = P\{u(\tau): \tau \rangle d | s(d) = T, M\},$$

тобто $\beta_T(d)$ – це умовна ймовірність спостереження векторів, що лежать нижче діагоналі d, за умови, що послідовність станів на діагоналі d дорівнює T.

Як і в разі одновімірною ПММ, ми можемо отримати рекурсивні формули для обчислення $\theta_T(d)$ і $\beta_T(d)$.

$$\theta_T(d) = \sum_{T_{d-1}} \theta_{T_{d-1}}(d-1) \cdot P(T_d | T_{d-1}, M) \cdot P(u(d) | T_d, M). \quad (2.11)$$

$$\beta_T(d) = \sum_{T_{d+1}} P(T_{d+1} | T_d, M) \cdot P(u(d+1) | T_{d+1}, M) \cdot \beta_{T_{d+1}}(d+1). \quad (2.12)$$

Потім ми можемо обчислити $L_m(i, j)$ для даної моделі M за формулою:

$$L_m(i, j) = P(s_{i,j} = m | u, c, M) =$$

$$= \begin{cases} \Sigma_{T_d : T_d(i,j) = m} P(T_d | u, c, M) & C(m) = c_{i,j} \\ 0 & \textit{інакше}. \end{cases}$$

Розглянемо випадок $C(m) = c_{i,j}$. При виведенні нижче передбачається, що підсумовування по T_d охоплює тільки T_d, що дає узгоджені класи з навчальними даними.

$$L_m(i, j) = \sum_{T_d : T_d(i,j) = m} \frac{P(T_d, u | M)}{P(u, c | M)} =$$

$$= \sum_{T_d : T_d(i,j)=m} \frac{\theta_{T_d}(\Delta(i,j)) \cdot \beta_{T_d}(\Delta(i,j))}{P(u,c|M)}. \quad (2.13)$$

Індекс "*d*" в T_d позначає діагональ *d* блоку (*i*, *j*). У наступному обчисленні $H_{m,m,l}(i,j)$ підсумовування завжди виконується за послідовністю станів з тими ж класами, що і в навчальних даних.

$$H_{m,m,l}(i,j) = P(s_{i-1,j} m, s_{i,j-1} n, s_{i,j} = l | u, c, M) =$$

$$= \begin{cases} \sum_{T_d} \sum_{T_{d-1}} P(T_d, T_{d-1} | u, c, M) & C(m) = c_{i-1,j}, C(n) = c_{i,j-1}, C(l) = c_{i,j} \\ 0 & \textit{інакше}. \end{cases}$$

Розглянемо випадок $C(m) = c_{i-1,j}, C(n) = c_{i,j-1}, C(l) = c_{i,j}$. У наведеному нижче рівнянні підсумовування по T_d і T_{d-1} додатково обмежується T_d, що задовольняє $T_d(i,j) = l$, і T_{d-1}, що задовольняє $T_{d-1}(i-1, j) = m$, $T_{d-1}(i, j-1) = n$.

$$H_{m,m,l}(i,j) = \sum_{T_d} \sum_{T_{d-1}} \frac{\theta_{T_{d-1}}(\Delta(i,j)-1)}{P(u,c|M)} \cdot$$

$$\left[P(T_d | T_{d-1}, M) P(u(d) | T_d, M) \cdot \beta_{T_d}(\Delta(i,j)) \right]. \quad (2.14)$$

Хоча використання прямої і зворотньої ймовірностей значно скорочує обчислення для $L_m(i,j)$ і $H_{m,m,l}(i,j)$, обчислювальна складність все ще залишається високою завдяки двовимірним аспектам. Рівняння (2.11) і (2.12) для оцінки прямої і зворотньої ймовірностей – це підсумовування за всіма послідовностями станів на діагоналі *d – 1*, або *d + 1*, при цьому заняття відповідають навчальним даним. При збільшенні кількості блоків по діагоналі кількість послідовностей станів збільшується експоненціально. Та ж проблема виникає при обчисленні $L_m(i,j)$ і $H_{m,m,l}(i,j)$. Отже, при обчисленні $L_m(i,j)$ і $H_{m,m,l}(i,j)$ робиться наближення, щоб уникнути обчислення зворотної і прямої ймовірностей.

$$H^{(p)}_{m,n,l}(i,j) = \sum_s I(m = s_{i-1,j}, n = s_{i,j-1}, l = s_{i,j}) P(s | y, \phi^{(p)}),$$

$$L_m^{(p)}(i,j) = \sum_s I\left(m = s_{i,j}\right) P\left(s \middle| y, \phi^{(p)}\right).$$

Щоб спростити обчислення $L_m^{(p)}(i,j)$ і $H_{m,n,l}^{(p)}(i,j)$, передбачається, що єдина найбільш ймовірна послідовність станів враховує практично всю ймовірність спостережень. Таким чином, ми прагнемо знайти оптимальну послідовність станів, яка максимізує $P\left(s \middle| y, \phi^{(p)}\right)$, що досягається за допомогою алгоритму навчання Вітербі.

2.4 Алгоритм Вітербі зі змінним станом

Використання алгоритму Вітербі для максимізації $P\left(s \middle| y\right)$ еквівалентно максимізації $P\left\{s_{i,j}, u_{i,j} : (i,j) \in N\right\}$, обмеженого $C\left(s_{i,j}\right) = c_{i,j}$ під час навчання. Коли ми застосовуємо навчену модель для класифікації зображень (процес тестування), ми також прагнемо знайти стан $\left\{s_{i,j} : (i,j) \in N\right\}$ максимізації $P\left\{s_{i,j}, u_{i,j} : (i,j) \in N\right\}$.

Потім стан відображаються в класи. При тестуванні, оскільки $c_{i,j}$ повинен бути визначений, попереднє обмеження $C\left(s_{i,j}\right) = c_{i,j}$ знімається.

Розглядається необмежений (тестовий) випадок, оскільки в обмеженому випадку єдина відмінність полягає в скороченні діапазону пошуку $s_{i,j}$ до станів, що відповідають класу $c_{i,j}$. Розгорніть $P\left\{s_{i,j}, u_{i,j} : (i,j) \in N\right\}$ як:

$$P\left\{s_{i,j}, u_{i,j} : (i,j) \in N\right\} = P\left\{s_{i,j} : (i,j) \in N\right\} \cdot P\left\{u_{i,j} : (i,j) \in N \middle| s_{i,j} : (i,j) \in N\right\} =$$

$$= P\left\{s_{i,j} : (i,j) \in N\right\} \cdot \prod_{(i,j) \in N} P\left(u_{i,j} \middle| s_{i,j}\right) =$$

$$= P(T_0) \cdot P(T_1 | T_0) \cdot P(T_2 | T_1) \cdots P(T_{w+z-2} | T_{w+z-3}) \prod_{(i,j) \in N} P\left(u_{i,j} \middle| s_{i,j}\right). \quad (2.15)$$

де T_d позначає послідовність станів для блоків, що лежать на діагоналі *d*. Остання рівність випливає з рівняння (2.7).

Оскільки T_d служить «ізолюючим» елементом у розширенні $P\{s_{i,j}, u_{i,j} : (i, j) \in N\}$, алгоритм Вітербі може бути застосований прямо для пошуку комбінації станів, що максимізують ймовірність $P\{s_{i,j}, u_{i,j} : (i, j) \in N\}$. Відмінність від звичайного алгоритму Вітербі полягає в тому, що кількість можливих послідовностей станів у кожному положенні діаграми переходів Вітербі зростає експоненціально із збільшенням блоків у T_d. Якщо є M станів, обсяг обчислення і пам'ять знаходяться в порядку M^{ν}, де ν – кількість станів у T_d. На «рис. 2.3.» показаний приклад.

Отже, цю версію алгоритму Вітербі називають алгоритмом Вітербі зі змінним станом.

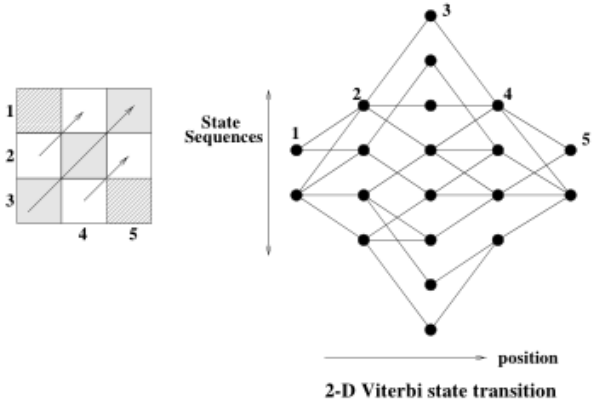


Рис. 2.3 Алгоритм Вітербі зі змінним станом

Той факт, що у двовимірному випадку лише послідовність станів на діагоналі, а не окремий блок, може служити ізолюючим елементом у розширенні $P\{s_{i,j}, u_{i,j} : (i, j) \in N\}$, що веде за собою обчислювальну неможливість для алгоритму Вітербі зі змінним станом.

Щоб зменшити кількість обчислень, в кожній позиції діаграми переходів Вітербі алгоритм використовує тільки N з усіх M^{ν} послідовностей станів, показаних на «рис. 2.4.».

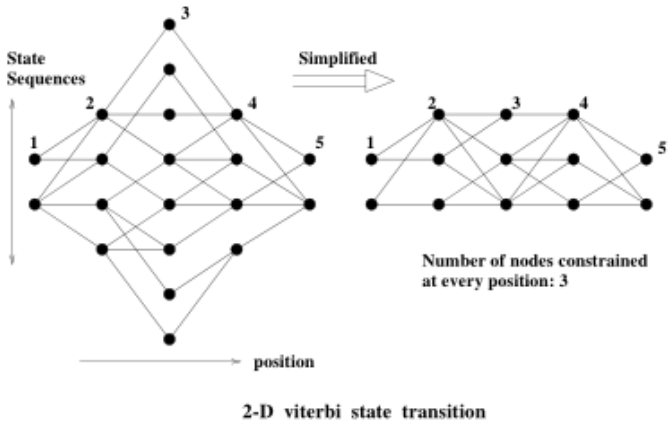


Рис. 2.4 Алгоритм Вітербі з обмеженнями по шляху

Шляхи обмежені для проходження одного з цих N пропустити необхідних. Шляхи обмежені проходженням одного з цих N вузлів. Щоб вибрати N послідовностей станів, алгоритм відокремлює блоки по діагоналі від інших блоків, ігноруючи їх статистичну залежність. Отже, апостеріорна ймовірність послідовності станів на діагоналі оцінюється як добуток апостеріорної ймовірності кожного блоку. Потім N послідовностей з найбільшими апостеріорними можливостями вибираються в якості N вузлів, дозволених в діаграмі переходів Вітербі [26-27]. При цьому мається на увазі, що оптимальна послідовність станів (вузол на оптимальному шляху діаграми переходів Вітербі) дає високу ймовірність, коли блоки обробляються незалежно. Також очікується, що, коли оптимальна послідовність станів відсутня серед N вузлів, обрана суб-оптимальна послідовність станів збігається з оптимальною послідовністю в більшості блоків. Суб-оптимальна версія алгоритму називається алгоритмом Вітербі з обмеженнями по шляху і змінним станом. Цей алгоритм відрізняється від М-алгоритму, введеного для кодування джерела Андерсоном [28], оскільки N вузли заздалегідь обрані, щоб уникнути обчислення апостеріорного ймовірностей всіх послідовностей M^{ν}.

Послідовності станів у рядках або стовпцях можуть також служити "ізолюючими" елементами в розгортанні $P\{s_{i,j}, u_{i,j} : (i,j) \in N\}$. Діагоналі

обираються для розширення, оскільки передбачається, що попередній вибір N вузлів, ігноруючи залежність між станами від діагоналі, погіршує продуктивність менше, ніж те ж саме для рядка або стовпця. Блоки на діагоналі геометрично не так близькі, як блоки на рядку чи стовпці.

Розроблено швидкий алгоритм вибору таких N послідовностей станів. Немає необхідності обчислювати апостеріорні ймовірності всіх M^{ν} послідовностей, щоб вибрати з них найбільше N.

Щоб гарантувати, що алгоритм Вітербі зі змінним станом та обмеженнями по шляху дає результати, досить близькі до алгоритму Вітербі зі змінним станом, параметр N повинен бути більше, ніж двомірний ланцюг Маркова. В результаті зображення зазвичай ділиться на фрагменти, щоб уникнути об'єднання багатьох блоків в один ланцюжок. Передбачається, що кожен фрагмент являє собою двомірний ланцюг Маркова, але залежність між фрагментами ігноруються. З іншого боку, щоб включити будь-який наперед заданий обсяг контекстної інформації для класифікації, фрагменти зображення повинні містити досить багато блоків.

3 ЕКСПЕРИМЕНТАЛЬНІ РЕЗУЛЬТАТИ

Вибір ознак є важливим питанням у класифікації, оскільки ознаки часто встановлюють межі ефективності класифікації. Для класифікатора, заснованого на двовимірній ПММ, використовуються як внутрішньоблокові, так і міжблокові ознаки. Внутрішньоблокові особливості визначаються відповідно до інтенсивності пікселів у блоці. Вони спрямовані на опис статистичних властивостей блоку. Вибрані функції сильно різняться для різних програм. Широко використовувані приклади включають моменти в просторовій області або частотній області та коефіцієнти перетворень, наприклад, дискретно-косинусне перетворення (ДКП).

Міжблочні ознаки визначені для подання відносин між двома блоками, наприклад, різниці між середньою інтенсивністю двох блоків. Використання міжблокових функцій аналогічно використанню дельта-коефіцієнтів і коефіцієнтів прискорення при розпізнаванні мови, в якому є достатнє емпіричне обгрунтування для включення цих функцій [29-30]. Мотивація до використання міжблочних функцій – це компенсувати строгість двомірної ПММ.

Двомірнна ПММ передбачає постійні ймовірності переходу між станами. Однак на практиці, перехід в стан може залежати від деяких взаємних властивостей двох блоків. Наприклад, якщо два блоки мають близьку інтенсивність, то з більшою ймовірністю вони знаходяться в одному і тому ж стані. Оскільки занадто складно оцінювати моделі, в яких ймовірності переходів є функціями, ми зберігаємо постійні ймовірності переходів і компенсуємо це припущення, додаючи взаємні властивості в вектори ознак таким чином, щоб вони могли впливати на визначення станів через апостеріорні ймовірності [31-32]. У двовимірній ПММ, оскільки стан сусідніх блоків справа над або зліва від блоку визначають ймовірність переходу в новий стан, взаємні властивості між поточним блоком і цими двома сусідніми блоками використовуються як міжблочні характеристики.

3.1 Сегментація аерофотознімків

Першим застосуванням двовимірного алгоритму ПММ є сегментація аерофотознімків на штучні і природні області. Зображення представляють собою напівтонові зображення 512x512 з 8 бітами на піксель. Це аерофотознімки району затоки Сан-Франциско [33]. Використовуваний набір даних містить шість зображень, чиї помічені вручну сегментовані зображення використовуються в якості істинного набору класів. Шість зображень і їх класи, помічені вручну, показані на рис.6.

Зображення були розділені на блоки 4x4, і коефіцієнти дискретно-косинусного перетворення або середні по деяким з них були використані в якості ознак. Всього таких функцій 6. Причина використання ДКП-коефіцієнтів полягає в тому, що різні розподілу енергії в частотної області краще розрізняють два класи. Позначимо ДКП-коефіцієнти для блоку 4x4 через $\{D_{i,j} : i, j \in (0,1,2,3)\}$, «див. Рис 3.1.»

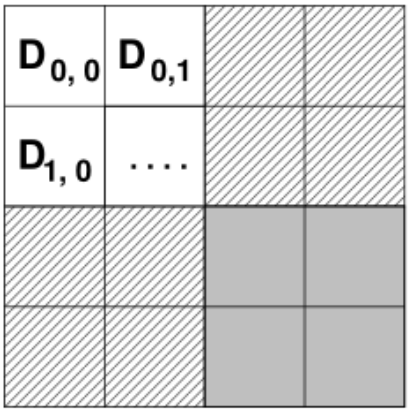


Рис. 3.1 Коефіцієнти ДКП для блоку зображення 4 x 4

Визначення 6 функцій:

1. $f_1 = D_{0,0}$; $f_2 = |D_{1,0}|$; $f_3 = |D_{0,1}|$;
2. $f_4 = \sum_{i=2}^{3} \sum_{j=0}^{1} |D_{i,j}| / 4$;
3. $f_5 = \sum_{i=0}^{1} \sum_{j=2}^{3} |D_{i,j}| / 4$;

4. $f_6 = \sum_{i=2}^{3} \sum_{j=2}^{3} \left| D_{i,j} \right| / 4$.

Крім того, в якості міжблокових характеристик використовувалися просторові похідні середніх значень інтенсивності блоків. Зокрема, просторова похідна відноситься до різниці між середньою інтенсивністю блоку і інтенсивністю верхнього або лівого сусіднього блоку.

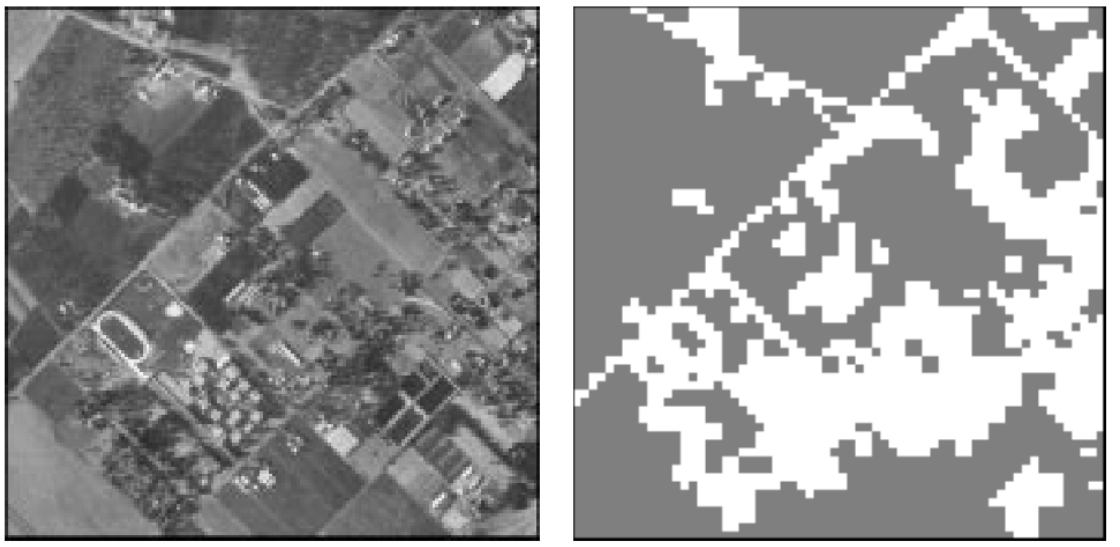
Отримані результати

Шестикратна перехресна перевірка [34] використовувалася для оцінки алгоритмів. Для кожної ітерації одне зображення використовувалося в якості тестових даних, а інше – в якості навчальних даних. Продуктивність оцінюється розподілених на всьому ітераціям. Були навчені і протестовані приховані марковські моделі з різним числом станів. Експерименти показують, що моделі з 4-6 станами для природного класу і від 7 до 10 станів для штучного класу дають дуже схожі результати. Для результату використовувалася модель з 5 станами для природного класу і 9 станів для штучного класу. Установка занадто великої кількості станів для кожного класу призводить до гіршої класифікації з двох причин: модель, найбільш близька до істини, може бути не такою складною; і більш складні моделі вимагають більшого навчального набору. При фіксованому навчальному наборі точність оцінки знижується зі збільшенням параметрів.

При навчанні і застосуванні ПММ з використанням двовимірного алгоритму Вітербі з обмеженнями по траєкторії зображення було розділене на квадратні фрагменти зображення, кожне з яких містить 16 блоків. Частини зображень вважалися окремими ланцюгами Маркова. Кількість вузлів, обмежених в кожній позиції на діаграмі переходів Вітербі, N, було вибрано рівним 32 для результату. Ми експериментували з кількома N_S. Для N від 2 до 16 продуктивність поступово збільшується. Якщо N більше 16, результати з невеликими відмінностями починають показувати тенденцію до збіжності. Частота помилок класифікації з $N = 16$ приблизно на 0,26% вище, ніж з $N = 32$. Оскільки час класифікації витрачається в основному на процес пошуку по

Вітербі, а час пошуку по Вітербі збільшується на порядок другого ступеня числа вузлів на кожному кроці переходу; час класифікації приблизно пропорційно N^2.

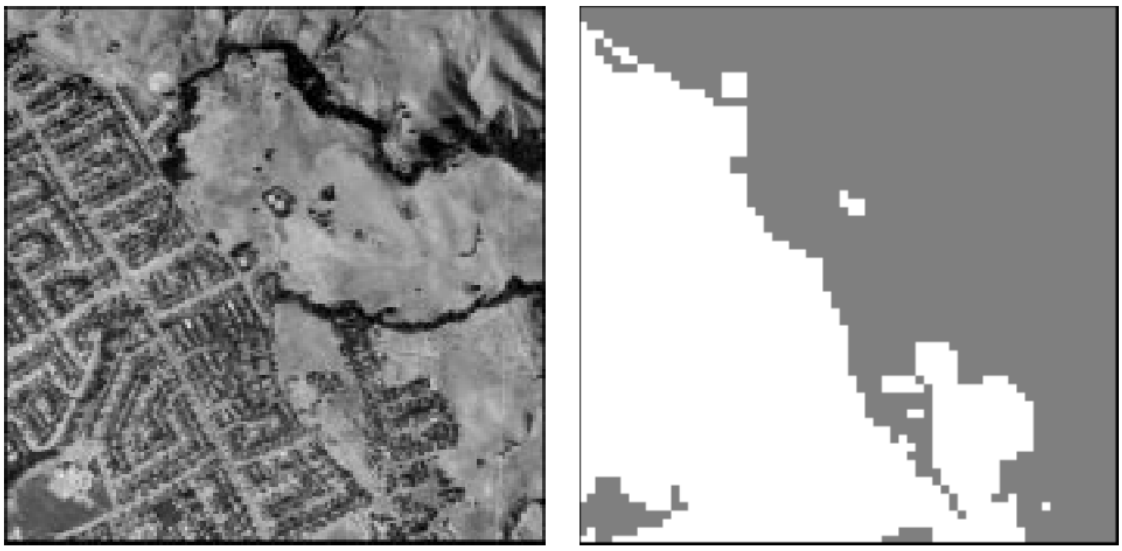
Результати класифікації для зображення на «рис. 3.2 (е)», наведені на «рис. 3.3». Результат двовимірного алгоритму ПММ забезпечує кращу відправну точку і з меншою ймовірністю буде мати поширення помилки.



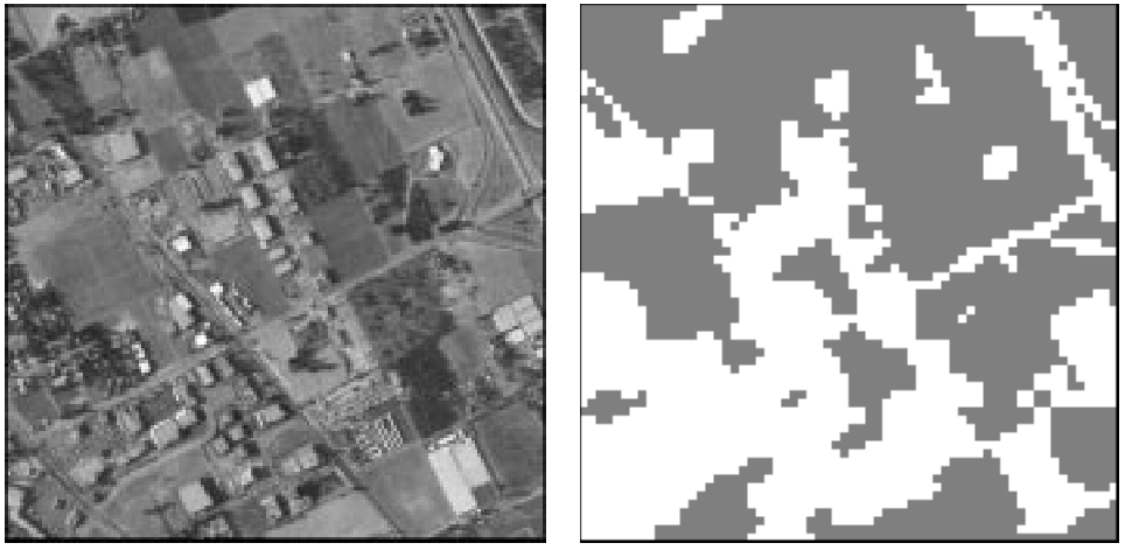
(а)



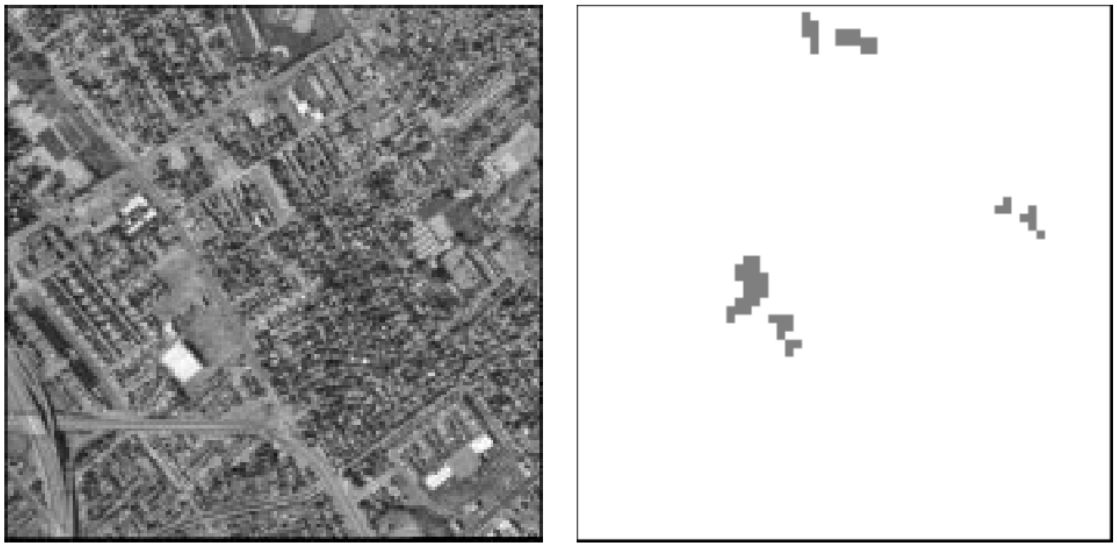
(б)



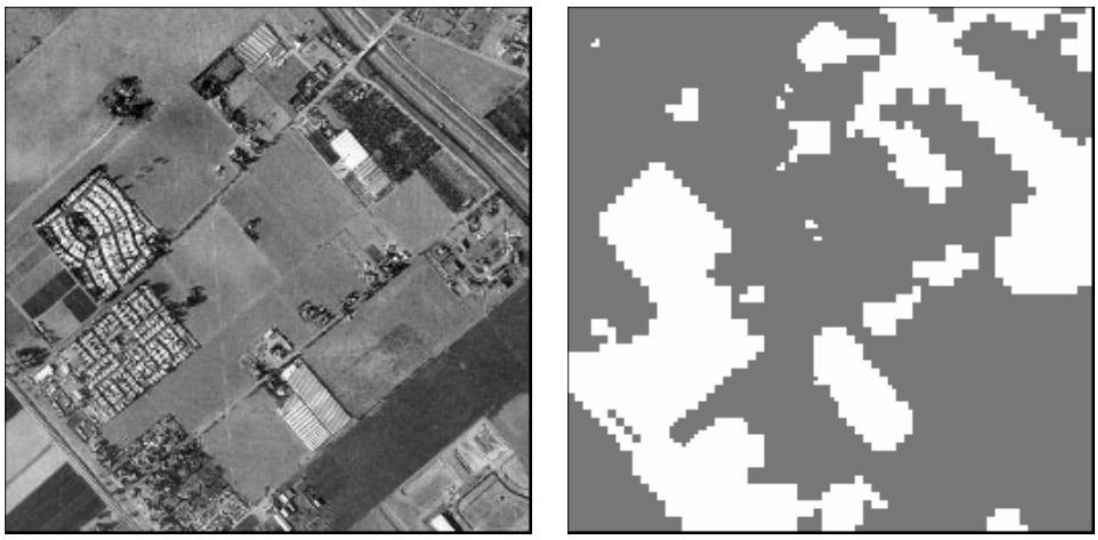
(в)



(г)



(д)



(e)

Рис. 3.2 Коефіцієнти ДКП для блоку зображення 4 x 4 Аерофотознімки: (a) – (e) Зображення 1–6. Зліва: оригінальні зображення на 8 біт/с, праворуч: класифіковані зображення, позначені вручну. Білий: штучний, сірий: натуральний

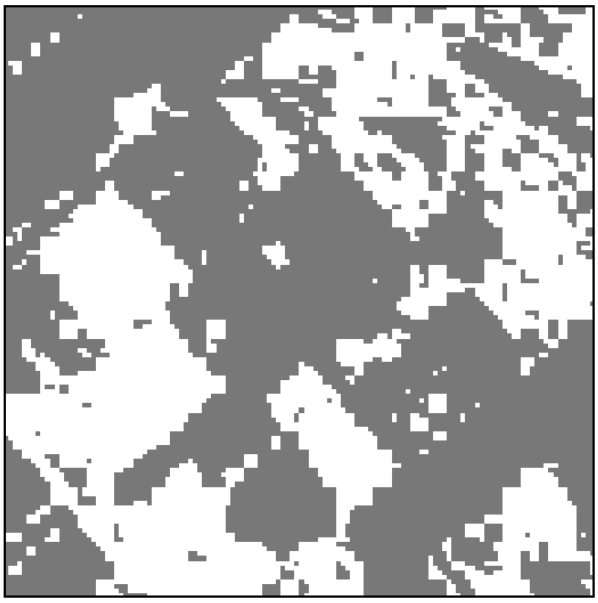


Рис. 3.3 Порівняння результатів класифікації з використанням двовимірної ПММ для аерофотознімка.

3.2 Сегментація зображення документа

Друге застосування двовимірного алгоритму HMM полягає у сегментації зображень документа на текст та фотографію.

Фотографія відноситься до зображень з безперервним тоном, таким як скановані зображення; а текст відноситься до тексту, таблиць і штучним графіками, створеним комп'ютерним програмним забезпеченням. Ми називаємо звичайний текст текстом для простоти, якщо сенс зрозумілий з контексту. Зображення, доповнені напівтоновими зображеннями, складають 8 біт на піксель. Приклад зображення і його сегментоване зображення показані на «рис. 3.4». Цей тип класифікації корисний в процесі друку для роздільного візуалізації різних типів локальних зображень. Це також інструмент для ефективного вилучення даних з баз даних зображень.

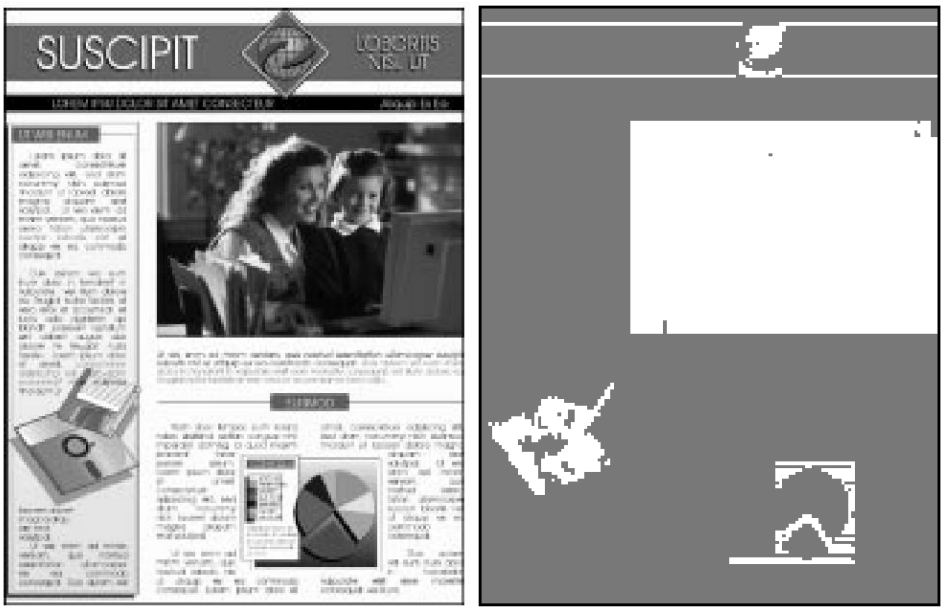


Рис. 3.4 Зображення тестового документа: зліва – оригінальне зображення, зправа – результат двовимірної класифікації ПММ. Білий: фотографія, сірий: текст

Використовувані нами функції містять дві функції. Перша особливість – це міра випадковості збігу між емпіричним розподілом вейвлет-коефіцієнтів у високочастотних діапазонах і розподілом Лапласа [35]. Він визначається як x^2

статистика, нормалізована за розміром вибірки. Друга функція вимірює вірогідність того, що вейвлет-коефіцієнти в високочастотних діапазонах складаються з висококонцентрованих значень. Ми також використовуємо просторові похідні від середніх значень інтенсивності блоків в якості характеристик, як і в попередньому додатку. Використаний розмір блоку – 8x8. ПММ має 5 станів для кожного класу. Експерименти показують, що моделі з 2-5 станами для кожного класу дають аналогічні результати.

Набір зображень – це кольорові зображення розміром близько 1600x1300. Кожна колірна складова становить 8 біт на піксель. В експериментах використовувалася тільки складова яскравості (тобто напівтонові зображення). Для більшості протестованих зображень алгоритм забезпечує дуже низький рівень помилок класифікації, в середньому близько 2%. Результати класифікації двомірного алгоритму ПММ показані на рис. 3.4. Ми бачимо, що результат з використанням ПММ чистий, особливо в областях фотографій. Це очікувано, оскільки класифікація, заснована на ПММ, враховує контекст. В результаті деякі гладкі блоки в областях фотографії, які локально нагадують текстові блоки, можна правильно ідентифікувати як фотографію.

ВИСНОВКИ

Мета та задачі, поставлені в цій роботі, були виконані. А саме:

1) Провести аналіз методів обробки зображень;

2) Розробити математичне формулювання основних припущень двовимірної прихованої марковської моделі;

3) Обгрунтувати обчислювальну складність;

4) Розробити алгоритм класифікації зображення за допомогою двовимірної прихованої марковської моделі;

5) Провести тестування розробленого алгоритму класифікації.

У дипломній роботі було запропоновано двовимірну приховану марковскую модель для класифікації зображень. Двовимірна модель забезпечує структурований спосіб включення контекстної інформації в класифікацію. Оскільки модель є двомірної, обчислювальна складність є важливою проблемою. Розроблено швидкий алгоритм для ефективної оцінки моделі і виконання класифікації на основі моделі. Застосування цього алгоритму до кількох завданням показує кращу продуктивність, ніж у декількох популярних алгоритмів блокової статистичної класифікації.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] KMM Rao, Medical Image Processing, Proc. Of workshop on Medical Image Processing and Applications, 8th October 1995 @ NRSA, Hyderabad-37.

[2] KMM Rao, Image Processing for Medical Applications, Proc. of 14th world conference on NDT, 8th – 13 th Dec 1996.

[3] Ramanjaneyulu M, KMM Rao , A Novel technique to Resample High Resolution Remote Sensing Satellite Images, Proc. of IGRASS-02, Colorado.

[4] KMM et al., Design and Fabrication of Color Scanner, Indian Journal of Technology, Vol 15, Apr 1997.

[5] Fundamentals Of Digital Image Processing – Anil K. Jain, Prentice-Hall, 1989.

[6] Computer Image Processing And Recognition - Ernest L.Hal, Academic Press, 1979.

[7] Digital Image Processing - A Remote Sensing Perspective, Jhon R. Jenson, 3rd Edition, Prentice – Hall, 2003.

[8] Digital Image Processing - Chellappa, 2nd Edition, IEEE Computer Society Press, 1992.

[9] P. A. Devijver, “Probabilistic labeling in a hidden second order Markov mesh” Pattern Recognition in Practice II, pp. 113-123, Amsterdam, Holland, 1985.

[10] P. A. Devijver, “Segmentation of binary images using third order Markov mesh image models," Proc. 8th Int. Conf. Pattern Recognition, pp. 259-261, Paris, Oct. 1986.

[11] P. A. Devijver, «Modeling of digital images using hidden Markov mesh random fields," Signal Pro-cessing IV: Theories and Applications (Proc. EUSIPCO-88), pp. 23-28, 1988.

[12] P. A. Devijver, “Real-time modeling of image sequences based on hidden Markov mesh random field models," Proc. 10th Int. Conf. Pattern Recognition, vol. 2, pp. 194-199, Los Alamitos, California, 1990.

[13] E. Levin and R. Pieraccini, “Dynamic planar warping for optical character recognition" Int. Conf. Acoust., Speech and Signal Processing, vol. 3, pp. 149-152, San Francisco, CA, March 1992.

[14] M. Park and D. J. Miller, “Image decoding over noisy channels using minimum mean-squared estimation and a Markov mesh," Proc. Int. Conf. Image Processing, vol. 3, pp. 594-597, Santa Barbara, CA, Oct. 1997.

[15] S. S. Kuo and O. E. Agazzi, “Machine vision for keyword spotting using pseudo 2D hidden Markov models” Proc. Int. Conf. Acoust., Speech and Signal Processing, vol. 5, pp. 81-84, 1993.

[16] J. Li and R. M. Gray, “Context based multiscale classification of images” Proc. Int. Conf. Image Processing, Chicago, Oct. 1998.

[17] J. Li and R. M. Gray, “Text and picture segmentation by the distribution analysis of wavelet coeficients” Proc. Int. Conf. Image Processing, Chicago, Oct. 1998.

[18] A. A. Markov, “An example of statistical investigation in the text of `Eugene Onyegin' illustrating coupling of `tests' in chains” Proc. Acad. Sci. St., Petersburg, VI Series 7, pp. 153.

[19] N. J. Nilsson, Learning Machines: Foundations of Trainable Pattern-Classifying Systems, McGraw-Hill, NY.

[20] T. Kohonen, J. Hynninen, J. Kangas, J. Laaksonen, and K. Torkkola, “LVQ PAK: The learning vector quantization program package (version 3.1)” Technical Report, Helsinki University of Technology, Lab oratory of Computer and Information Science, Finland, April, 1995.

[21] R. Kindermann and J. L. Snell, Markov Random Fields and Their Applications, American Mathematical So ciety, 1980.

[22] T. Kohonen, G. Barna, and R. Chrisley, “Statistical pattern recognition with Neural Networks: benchmarking studies” IEEE Int. Conf. Neural Networks, pp. I-61-68, July 1988.

[23] K. Abend, T. J. Harley, and L. N. Kanal, “Classification of binary random patterns” IEEE Trans. Inform. Theory, vol. IT-11, no. 4, pp. 538-544.

[24] L. N. Kanal, “Markov mesh models” Image Modeling, pp. 239-243, New York: Academic.

[25] S. Young, J. Jansen, J. Odell, D. Ollason, and P. Woodland, HTK - Hidden Markov Model Toolkit, Cambridge University, 1995.

[26] K. O. Perlmutter, S. M. Perlmutter, R. M. Gray, R. A. Olshen, and K. L. Oehler, “Bayes risk weighted vector quantization with posterior estimation for image compression and classification” IEEE Trans. Image Processing, vol. 5, no. 2, pp. 347-360, Feb. 1996.

[27] D. K. Pickard, “A curious binary lattice process” J. Appl. Prob., vol. 14, pp. 717-731.

[28] F. Jelinek and J. B. Anderson, “Instrumentable tree encoding of information sources” IEEE Trans. Inform. Theory, vol. IT-17, pp. 118-119.

[29] C. F. J. Wu, “On the convergence properties of the EM algorithm” Annals of Statistics, vol. 11, no. 1, pp. 95-103.

[30] C. C. Yen and S. S. Kuo, “Degraded documents recognition using pseudo 2d hidden Markov models in gray-scale images” Proc. SPIE, vol. 2277, pp. 180-191, 1994.

[31] A. J. Viterbi and J. K. Omura, “Trellis enco ding of memoryless discrete-time sources with a fidelity criterion” IEEE Trans. Inform. Theory, vol. IT-20, pp. 325-332.

[32] P. S. Williams and M. D. Alder, “Generic texture analysis applied to newspaper segmentation” Proc. Int. Conf. Neural Networks, vol. 3, pp. 1664-1669, Washington, DC, June 1996.

[33] K. L. Oehler, “Image compression and classification using vector quantization” Ph.D thesis, Stanford University, 1993.

[34] M. Stone, “Cross-validation: a review” Math. Operationforsch. Statist. Ser. Statist., no. 9, pp. 127-139.

[35] C. E. Shannon, «A mathematical theory of communication” Bell System Technical Journal, vol. 27, pp. 379-423.