Розділ 2.

Розпізнавання маркерів.

В якості маркерів використаємо QR(Quick Responsive) код. QR код може містити інформацію про те, який віртуальний об’єкт відображати в доповненій реальності. Положення QR коду визначає положення віртуального об’єкту.

Розпізнавання QR-кодів, як і розпізнавання будь яких інших образів, на зображенні складається з декількох етапів. На кожному етапі виконується робота, результат якої використовується в наступних етапах. На рис. 2.1 наведено етапи розпізнавання QR-коду.

C:\Users\Ivan\Desktop\qr_recognition.png

Рис. 2.1. Блок-схема розпізнавання QR-коду на зображенні

Далі розглянемо детально роботу кожного етапу окремо.

Порогова бінаризація зображення.

За допомогою порогової бінаризації з зображення можна отримати бінарне зображення. При виконанні порогової бінаризації частина пікселів обирається в якості пікселів переднього плану, що представляє об’єкт інтересу, інша частина – в якості фонових пікселів. Беручи до уваги розподіл яскравості на зображенні, деякі значення можливо обрати в якості порогів, що розділяють пікселі на групи. П найпростішому випадку вибирається одне порогове значення t. Всі пікселі з яскравістю більше або рівні t стають пікселями переднього плану, всі інші – фоновими пікселями. Описана операція називається верхньою пороговою бінаризацією. Існують й інші види порогової бінаризації. Нижня порогова бінаризація пікселями переднього плану робить ті пікселі, які менші або рівні t. При пороговій бінаризації по діапазону вказується два порогових значення – верхнє та нижнє. Пікселі з яскравістю між пороговими значеннями обираються в якості пікселів переднього плану. Також можливо виконати зворотну операцію – пікселі більше рівні значенню верхнього порогу або менше рівні значенню нижнього порогу стають пікселями переднього плану.

Вище описані способи бінаризації зображень відносно порогу t. Необхідно виконати задачу – визначити поріг t. Розрахунок порогу має займати мінімум розрахунків, щоб встигати розпізнавати потік зображень. З іншого боку треба виконати бінаризіцію, яка виділить необхідні для аналізу компоненти на зобраенні.

Обрано алгоритм автоматичного вибору порогу бінаризіції за методом Оцу(Otsu). Даний алгоритм підходить для зображень на гістограмах яких є дві виділені моди. Рекурентна формула, що використовується для знаходження порогу зменшує кількість обчислень. Розглянемо зображення монет на рис. 2.2. Припустимо необхідно виконати розрахунок кількості монет, для цього треба виконати бінаризацію зображення таким чином, щоб монети стали об’єктами переднього плану, все інше фоном.



Рис. 2.2. Зображення монет, над якого необхідно знайти найкращий поріг бінаризації

Для визначення порогу t методом Оцу необхідно скласти гістограму інтенсивності, оскільки зображення напівтонове та має значення пікселів в діапазоні 0-255. Гістограму можна представити масивом, індекс масива відповідає інтенсивності , значення елемента масива – кількості пікселів х даною інтенсивністю. Гістограма інтенсивності до зображення на рис. 2.2 наведена на рис. 2.3.

Рис. 2.3. Гістограма інтенсивності

На гістограмі рис. 2.3 є виділені дві моди значень, перша мода відповідає фону на зображенні, друга мода – монетам на зображенні. Оскільки гістограма бімодальна, то доцільно застосувати метод Оцу для пошуку порога бінаризації.

Вибір порогу в методі Оцу заснований за мінімізації внутрішньо групової дисперсії двох груп пікселів, які розділені порогом t. При розгляді методу будемо вважати, що гістограма задана у вигляді дискретної функції розподілу ймовірності P. Значення P(0),…,P(I) являють собою ймовірності значень яскравості 0,…,I. P(i) виражається через відношення кількості пікселів яскравістю i на зображенні до загальної кількості пікселів на зображенні. Якщо гістограма P бімодальна, то задача вибору порогу бінаризації закладається в визначенні найкращого порогу t, який відділяє дві гістограми одна від одної. Кожному пороговому значенню t відповідає дисперсія групи значень, менших або рівних t, та дисперсія групи значень, більших від t. Оцу запропонував визначення найкращого порогу бінаризації таким, при якому зважена сума внутрішньо групових дисперсій мінімальна. Ваги рівні сумарним ймовірностям відповідних груп.

Мірою однорідності груп є дисперсія. Чим менше розкид значень всередині групи, тим менше значення дисперсії. Можливим способом вибору критерія для розділення груп може бути вибір такого значення, при якому зважена сума внутрішньо групових дисперсій буде мінімальною. В даному критерії передбачається існування однорідних груп, які складаються з приблизно однакових значень. Другий спосіб закладається в виборі значення, при якому квадрат різниці між середніми значеннями отриманих груп матиме максимальне значення. Дана різниця зв’язана с між груповою дисперсією. Обидва критерії призводять до аналогічних результатів, так як сума внутрішньо групових та між групових дисперсій постійна. В роботі обрано перший спосіб.

Позначимо через зважену суму внутрішньо групових дисперсій. Дана сума називається загальною внутрішньо груповою дисперсією. Через позначимо дисперсію першої групи, яка складається зі значень, менших або рівних t. Через позначимо дисперсію другої групи, яка складається зі значень, більших від t. Позначимо через сумарну ймовірність першої групи і через – сумарну ймовірність другої групи. Середні значення першої та другої групи та відповідно. Пояснивши позначення, визначимо через них загальну внутрішньо групову дисперсію :

|  |  |
| --- | --- |
|  | (2.1) |

, де

|  |  |
| --- | --- |
|  |  |
|  | (2.2) |
|  |  |
|  | (2.3) |
|  |  |
|  | (2.4) |

Найкраще значення порогу можна визначити простим перебором всіх можливих значень t для пошуку значення t, яке мінімізує . В більшості випадків регіон перебору можливо зменшити до проміжку між двома модами гістограми. Однак затрати на виявлення мод рівні з пошуком значення t, що розділяє гістограму в окрузі цих мод.

Між внутрішньо груповою дисперсією та повною дисперсією існує взаємозв’язок, який не залежить від порогового значення. Повна дисперсія дорівнює:

Зв'язок повної та внутрішньо групової дисперсій дозволяє скоротити обчислювальні затрати на визначення найкращого порога. Переписавши вираз для , отримаємо:

Оскільки

Та враховуючи формулу (2.2) отримаємо наступну форму повної дисперсії:

|  |  |
| --- | --- |
|  | (2.5) |

Перший доданок у дужках дорівнює внутрішньо груповій дисперсії . Це сума зважених дисперсій кожної групи. Другий доданок в дужках називається між груповою дисперсією . Це сума зважених квадратів відстаней між середніми значеннями кожної групи і середнім значенням всього розподілу. Вираз міжгрупової дисперсії можна ще спростити. Зауважимо, що середнє значення всього розподілу можна записати у вигляді:

|  |  |
| --- | --- |
|  | (2.6) |

За допомогою (2.6) виключимо із (2.5) і поставимо замість . Після простих перетворень отримаємо:

Так як повна дисперсія не залежить від t, то значенням t, мінімізуючим , буде таке значення t, при якому досягається максимум міжгрупової дисперсії ,

|  |  |
| --- | --- |
|  | (2.7) |

Для визначення t, що максимізує , всі необхідні величини мають бути визначені відповідно (2.2) і (2.3). Обчислення цих величин не обов'язково незалежно виконувати для кожного значення t. Існує залежність між значеннями цих величин для t і для наступного значення t+1. Безпосередньо із (2.2) отримаємо рекурентне співвідношення:

|  |  |
| --- | --- |
|  | (2.8) |

з початковим значенням .

Із (2.3) отримуємо рекурентне співвідношення:

|  |  |
| --- | --- |
|  | (2.9) |

з початковим значенням . В результаті, із (2.6) отримаємо:

|  |  |
| --- | --- |
|  | (2.10) |

Спробуємо застосувати метод Оцу до зображення на рис. 2.2. Як зазначалося раніше дане зображення підходить для пошуку порогу бінаризіції методом Оцу, адже має бімодальну гістограму, зображену на рис. 2.3. Описаний вище алгоритм пошуку порогу t по максимальному значенню міжгрупової дисперсії з застосуванням рекурентної формули було програмно реалізовано. Програма опрацювала зображення на рис. 2.2 та розрахувала поріг бінаризації t=126, якому відповідає максимальне значення міжгрупової дисперсії =2863.60784679457. Після визначення порогу проведено верхню порогову бінаризацію. Результат роботи програми наведено на рис. 2.4. На рис. 2.4 монети відповідають області переднього плану. Кожна монета, окрім однієї, практично не мають внутрішніх отворів, тому їх легко порахувати на зображені. Для монети, яка має отвори всередині, можливо застосувати морфологічну операцію замикання структурним елементом диском, який має розмір з монету.

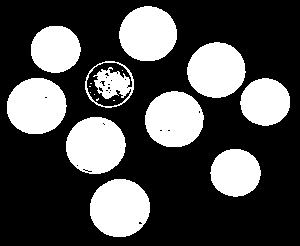
**

Рис. 2.4. Результат роботи бінаризації методом Оцу

При проектуванні алгоритмів автоматичного визначення порога бінаризації зазвичай робляться деякі припущення відносно форми розподілу значення яскравості пікселів зображення. Тому такі алгоритми добре працюють на зображеннях, які задовольняють прийнятим в алгоритмах припущенням. В методі Оцу розподіл значення яскравості вважається бімодальним. Якщо зображення приблизно задовольняє це обмеження, то алгоритм матиме хороші результати. Якщо гістограма зображення далека від бімодальної, то результати можуть виявитися практично не корисними.