Міністерство освіти і науки україни

Національний авіаційний університет

Факультет кібербезпеки, комп’ютерної та програмної інженерії

Кафедра прикладної математики

ДОПУСТИТИ ДО ЗАХИСТУ

Завідувач кафедри

\_\_\_\_\_\_\_\_\_\_\_ П. О. Приставка

“\_\_\_\_\_”\_\_\_\_\_\_\_\_\_\_\_\_2019 р.

**ДИПЛОМНА РОБОТА**

**(пояснювальна записка)**

випускника освітньоГО СТУПЕНЯ

“Бакалавр”

**Тема:** Автоматизована система сегментації цифрових зображень на основі дискретних структур.

**Виконавець:** Ковдря В. Ю.

**Керівник:** к.ф.-м.н., доцент Юрчук І. А.

**Нормоконтролер:** к.ф.-м. н., доцент Кубайчук О.O.

**Київ 2019**

НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Факультет кібербезпеки, комп’ютерної та програмної інженерії

Кафедра Прикладної математики

Спеціальність 6.040301 “ Прикладна математика ”

ЗАТВЕРДЖУЮ

Завідувач кафедри

\_\_\_\_\_\_\_\_\_ П. О. Приставка

“\_\_\_\_” \_\_\_\_\_\_\_\_\_2019 р.

**ЗАВДАННЯ**

**на виконання дипломної роботи**

*Ковдрі Владислава Юрійовича*

1. **Тема дипломної роботи:** Автоматизована система сегментації цифрових зображень на основі дискретних структур, затверджена наказом ректора від "11'' квітня 2018 р. № 775/ст.
2. **Термін виконання роботи:** з «\_\_\_»\_\_\_\_\_\_\_\_\_\_\_ по «\_\_\_»\_\_\_\_\_\_\_\_\_\_\_
3. **Вихідні дані до роботи:** цифрові зображення довільного розміру, середовище розробки Visual Studio 2017.
4. **Зміст пояснювальної записки:** огляд основних теоретичних положень для виконання дипломного проекту, тобто основних методів сегментації цифрових зображень, підходу до сегментації з використанням ґраткового графу, основних аспектів обробки зображень(згортки та розмивання), кольорових моделей та їх перетворень; реалізація програмного забезпечення для сегментації цифрових зображень; аналіз та порівняння отриманих результатів.

**Календарний план-графік**

|  |  |  |  |
| --- | --- | --- | --- |
| №  пор. | Завдання | Термін виконання | Відмітка про виконання |
| 1 | Отримання завдання та огляд літератури за темою дипломної роботи. Аналіз літературних джерел |  | Виконано |
| 2 | Підготовка та написання 1-го та 2-го розділу |  | Виконано |
| 3 | Реалізація програмного забезпечення |  | Виконано |
| 4 | Оформлення дипломної роботи |  | Виконано |
| 5 | Створення презентації |  | Виконано |
| 6 | Передзахист дипломного проекту. |  | Виконано |
| 7 | Захист дипломної роботи. |  |  |

1. **Дата видачі завдання:** «\_\_\_»\_\_\_\_\_\_\_\_\_\_\_ 2019 р.

Керівник дипломної роботи Юрчук І.А.

Завдання прийняв до виконання Ковдря В. Ю.

# РЕФЕРАТ

Пояснювальна записка до дипломної роботи на тему: «Автоматизована система сегментації цифрових зображень на основі дискретних структур»: сторінки, рисунка, використаних джерел.

**Об’єкт дослідження:** процес сегментації цифрових зображень.

**Предмет дослідження:** метод сегментації за допомогою ґраткового графу.

**Програмне середовище,** що використовувалося в процесі виконання дипломної роботи Microsoft Visual Studio 2017, мова програмування - С#

**Мета роботи:** розробити інформаційну технологію для сегментації цифрового зображення на основі методу сегментації за допомогою ґраткового графу.

**Результатом** виконання дипломної роботи є інформаційна технологія для сегментації цифрових зображень за допомогою ґраткового графу, яка реалізована в програмному середовищі Microsoft Visual Studio 2017.

**Галузь застосування:** результати дипломного проекту рекомендуються для використання з метою обробки цифрових зображень.

**ЗМІСТ**

[РЕФЕРАТ 4](#_Toc11255535)

[ПОСТАНОВКА ЗАДАЧІ 6](#_Toc11255536)

[РОЗДІЛ 1 ТЕОРЕТИЧНІ ВІДОМОСТІ 8](#_Toc11255537)

[1.1. Огляд методів сегментації 8](#_Toc11255538)

[1.1.1. Методи, засновані на кластеризації 8](#_Toc11255539)

[1.1.2. Методи, засновані на стисненні 9](#_Toc11255540)

[1.1.3. Методи з використанням гістограми 10](#_Toc11255541)

[1.1.4. Методи, засновані на диференціальних рівняннях з частинними похідними 11](#_Toc11255542)

[1.2. Метод ефективної сегментації на ґратковому графі 13](#_Toc11255543)

[1.2.1. Предикат для попарного порівняння областей 13](#_Toc11255544)

[1.2.2. Алгоритм і його властивості 15](#_Toc11255545)

[1.3. Обробка зображень 20](#_Toc11255546)

[1.3.1. Згортка 20](#_Toc11255547)

[1.3.2. Обробка країв 22](#_Toc11255548)

[1.3.3. Розмивання Гауса 23](#_Toc11255549)

[1.4. Кольорові моделі 25](#_Toc11255550)

[1.4.1. RGB 26](#_Toc11255551)

[1.4.2. XYZ 27](#_Toc11255552)

[1.4.3. Lab 30](#_Toc11255553)

[РОЗДІЛ 2 Практична Частина 34](#_Toc11255554)

[2.1. Технічні засоби та структури даних 34](#_Toc11255555)

[2.1.1. Система неперетинних множин 34](#_Toc11255556)

[2.1.2. Зчитування зображень 39](#_Toc11255557)

[2.2. Опис програмного забезпечення 41](#_Toc11255558)

[Список використаних джерел 42](#_Toc11255559)

# ПОСТАНОВКА ЗАДАЧІ

Нехай задано цифрове зображення , де кожен піксель  характеризується координатами  та трьома кольоровими складовими. Необхідно визначити розбиття  на класи еквівалентності за відношенням піксель  «схожий на» піксель  (рефлексивне і симетричне).

Іншими словами, утворимо скінченний набір множин , які не перетинаються ( при  ), елементами яких є , такі, що для довільних пікселів  «схожий на»тоді і лише тоді, коли існує єдиний індекс  такий, що  і .

**Завдання:**

1. вивчити методметод ефективної сегментації на ґратковому графі;
2. розробити алгоритм сегментації цифрового зображення методом ефективної сегментації на ґратковому графі;
3. розробити обчислювальну схему для отримання сегментації цифрових зображень;
4. оцінити якість отриманої сегментації;
5. створити автоматизовану систему на основі розробленого алгоритму;
6. провести тестування інформаційної технології на реальних даних;
7. зробити порівняльний аналіз даного методу сегментації з пірамідальним методом.

ВСТУП

**Актуальність**

Задачі сегментації і групування зображень залишаються складними для комп'ютерного зору. З часів гештальтпсихології було відомо, що сприйняття форми (перцептивне сприйняття) відіграє важливу роль у зоровому сприйнятті людини. Широкий спектр задач обчислювального бачення може використовувати сегментовані зображення, якщо такі надійно і ефективно обчислюються. Наприклад, задачі бачення середнього рівня складності, такі як стереобачення та оцінка руху потребують підтримки відповідних операцій. Просторово неоднорідні області можуть бути ідентифіковані за допомогою методів сегментації. Задачі бачення вищого рівня, такі як розпізнавання та індексація зображень, також можуть використовувати результати сегментації для розпізнавання по частинах.

# РОЗДІЛ 1 ТЕОРЕТИЧНІ ВІДОМОСТІ

В комп'ютерному зорі сегментація зображення - це процес поділу цифрового зображення на кілька сегментів (наборів пікселів, також відомих як супер-пікселі). Мета сегментації - спростити і / або змінити подання зображення на щось більш значуще і більш легке для аналізу. Сегментація зображення зазвичай використовується для пошуку об'єктів і кордонів (ліній, кривих і т. д.) в образах. Точніше, сегментація зображення - це процес присвоєння мітки кожному пікселю зображення таким чином, щоб пікселі з однією і тією ж міткою мали певні характеристики[1].

Результатом сегментації зображення є набір сегментів, які в сукупності покривають все зображення, або набір контурів, витягнутих з зображення. Кожен з пікселів в області аналогічний за деякими характеристиками або обчислюваними властивостями, таких як колір, інтенсивність або текстура. Суміжні регіони істотно відрізняються один від одного по відношенню до однієї і тієї ж характеристики.

## 1.1. Огляд методів сегментації

Існують різні підходи для отримання сегментації цифрового зображення. Далі розглядаються основні з них з коротким викладенням ідеології роботи.

Побудова графу на основі зображення та подальша сегментація на його основі, що розглядаються в пункті 1.2., відповідають ідеології сімейства сегментацій на основі графів, тому їх загальна концепція не розглядається.

### 1.1.1. Методи, засновані на кластеризації

K - середніх[2] — це ітераційний метод, який використовується для того, щоб розділити зображення на K кластерів. Базовий алгоритм наведений нижче:

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

Тут за відстань зазвичай береться сума квадратів або абсолютних значень різниць між пікселем і центром кластера. Різниця зазвичай базується на кольорі, яскравості, текстурі і місцезнаходженні пікселя, або на зваженій сумі цих чинників. K може бути вибране вручну, випадково чи евристично.

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

### 1.1.2. Методи, засновані на стисненні

В методах заснованих на стисненні стверджується, що оптимальною сегментацією серед усіх можливих є та, котра використовує найменший об'єм даних для кодування результуючого зображення[3][4]. Зв'язок між сегментацією та стисненням пояснюється тим, що сегментація намагається знайти шаблони в зображенні, а будь-який взаємозв'язок в зображенні може бути використаний для його стиснення. Підхід описує кожен сегмент за його текстурою і формою контуру. Кожна складова сегменту моделюється формулою ймовірнісного розподілу, а об’єм даних для кодування обчислюється таким чином:

* 1. Кодування контурів ґрунтується на факті, що області на звичайних зображеннях стараються мати гладкий контур. Це припущення використовується в кодуванні Хаффмана для закодовування диференціального коду ланцюга контурів на зображенні. Так, наскільки контур гладкіший, на стільки ж він займає менше місця.
  2. Текстура кодується за допомогою стиснення з втратами, шляхом схожим до принципу мінімально допустимої довжини опису, але тут об'єм даних для моделі є апроксимованим кількістю вибірок ентропії моделі. Текстура в кожній області моделюється багатовимірним нормальним розподілом, ентропія котрого найбільш схожа з представленням. Цікавою властивістю даної моделі є те, що в межах оціночної ентропії міститься значення ентропії реальних даних текстури. Це тому, що серед усіх наявних функцій розподілу з відомими значеннями середнього арифметичного та кореляційного моменту, нормальний розподіл має найбільшу ентропію. З цього виходить, що реальний об'єм закодованих даних не може бути більшим ніж запропонований алгоритмом.

Для довільної сегментації дана схема обчислює кількість біт, потрібну для кодування зображення з використанням обраної сегментації. Так, серед всіх можливих сегментацій зображення, потрібно знайти сегментацію, яка представляється закодованими даними найменшої довжини. Це можна досягнути з використанням методів ієрархічної кластеризації. Спотворення в стисненні з втратами визначає похибку сегментування і її оптимальне значення може відрізнятись для окремих зображень. Цей параметр може бути оцінено евристично за показниками контрасту текстур на зображенні.

### 1.1.3. Методи з використанням гістограми

Методи з використанням гістограми дуже ефективні порівняно з іншими методами сегментації оскільки вони вимагають тільки один прохід по пікселях. У цьому методі гістограма обчислюється за всіма пікселям зображення і її мінімуми і максимуми використовуються, щоб знайти кластери на зображенні [5]. Колір або яскравість можуть бути використані при порівнянні.

Покращення цього методу — рекурсивно застосовувати його до кластерів на зображенні для того, щоб поділити їх на дрібніші кластери. Процес повторюється з усе меншими і меншими кластерами до тих пір, коли перестануть з'являтися нові кластери [5,6].

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

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

### 1.1.4. Методи, засновані на диференціальних рівняннях з частинними похідними

Використовуючи диференціальні рівняння з частинними похідними та їх розв'язки за певною числовою схемою можна одержати сегмент зображення[6]. Крива поширення є відомим підходом в цій галузі, з великими можливостями практичного застосування в виокремленні об'єктів, відслідковуванні об'єктів, відновленні просторового зображення, і т. д. Головною ідеєю є еволюція початкової кривої відповідно найменшого потенціалу функції оцінки, її визначення ж буде відображати завдання котре повинно бути виконано. Щодо зворотної задачі, мінімізації функціоналу оцінки, то вона є важкою і накладає певні обмеження щодо гладкості функції при вирішенні, що може бути виражено як геометричні обмеження на еволюційну криву.

***Параметричні методи***

Техніки відновлення многочлена Ланранжа базуються на визначенні параметра контуру відповідно до деякої стратегії вибірки, і наступної еволюції кожного елемента відповідно до зображення та внутрішніх термів. Дані техніки є швидкими та ефективними, але початкове "чисто параметричне" формулювання (за Касом Віткіним та Терзополусом в 1987 відоме як "змійки") піддане критиці через обмеження, які полягають в виборі стратегії вибірки, внутрішніх геометричних властивостей кривої, зміни топології (поділ та об’єднання кривих), проблеми адресації при вищих вимірах і т. д. На сьогодні були розроблені ефективніші "дискретизовані" формулювання для подолання даних обмежень зберігаючи ефективність. В обох випадках оптимізація геометрії виконується методом градієнтного спуску а для обчислення похідних використовується метод скінченних різниць.

***Методи встановлення рівня***

Методи встановлення рівня з самого початку використовувались для відслідковування інтерфейсів, котрі переміщувались Ошером та Сетіаном в 1988 і увійшли в інші галузі обробки зображень наприкінці 90-х. Вони можуть бути використані для ефективного вирішення задач пошуку кривих, поверхонь і т. д. в неявній формі. Головною ідеєю є представлення еволюціонуючого контуру  використовуючи знакову функцію, нульове значення якої відповідає фактичному контуру. Тоді, згідно з рівнянням руху контуру можна легко вивести схожий підхід для неявної поверхні, котра на нульовому рівні буде відображати контур об'єкта. Методи встановлення рівня мають багато переваг, вони задаються в неявній формі, не залежать від параметра, дозволяють просто оцінити геометричні властивості структури, яка еволюціонує, надає можливості зміни топології. Зао, Мерімен та Ошер в 1996 запропонували використовувати дані методи як основу в вирішенні задач оптимізації. З цього випливає, що дані методи можуть бути використані для вирішення багатьох задач в області машинного зору та аналізу медичних зображень [8]. Дослідження різноманітних структур даних для встановлення рівня дало можливість створити дуже ефективні методи вирішення цієї задачі.

## 1.2. Метод ефективної сегментації на ґратковому графі

Використаємо заснований на графі підхід до сегментації. Нехай  неорієнтований граф з вершинами (Рис. ) - набір елементів для сегментації, і ребра  відповідають парам сусідніх вершин. Кожне ребро має відповідну вагу , яка є невід'ємної мірою відмінності між сусідніми елементами  і . У разі сегментації зображення елементи з є пікселями, а вага ребра є деякою мірою відмінності між двома пікселями, пов'язаними цим ребром (наприклад, різниця в інтенсивності, кольорі, місцезнаходженні або якому-небудь іншому локальному атрибуті).

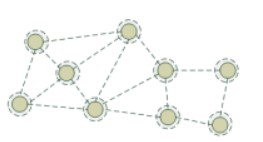


Рис. Граф для сегментації

У графовому підході сегментація  являє собою розбиття  на компоненти таким чином, що кожен компонент (або область)  відповідає з’єднаним компонентом у графі , де . Іншими словами, будь-яка сегментація є підмножиною ребер в (Рис ). Існують різні способи вимірювання якості але вцілому ми хочемо, щоб елементи в компоненті були схожі, а елементи в різних компонентах - несхожі. Це означає, що ребра між двома вершинами в одному компоненті повинні мати відносно низькі ваги, а ребра між вершинами в різних компонентах повинні мати більш високі ваги.

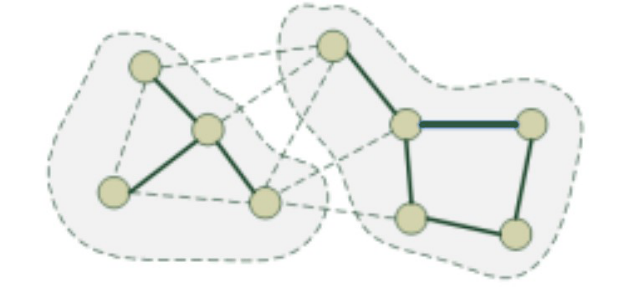


Рис. Результат сегментації на графі

### 1.2.1. Предикат для попарного порівняння областей

У цьому розділі ми визначимо предикат  для оцінки необхідності розділення, двох компонентів в сегментації (двох області зображення). Цей предикат заснований на вимірюванні відмінностей між елементами уздовж кордону двох компонентів, щодо міри відмінностей між сусідніми елементами всередині кожного з двох компонентів. Отриманий предикат порівнює міжкомпонентні відмінності з відмінностями всередині компонента і тим самим адаптується до локальних характеристик даних.

Визначимо *внутрішню різницю* компонента  як найбільшу вагу в (Рис. ), тобто



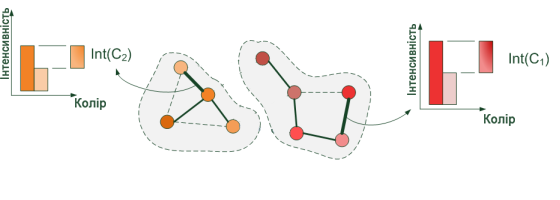


Рис. Внутрішня різниця компонента

Ідея, що лежить в основі цієї міри полягає в тому, що компонент  залишається зв’язаним доти, доки ребра, що розглядаються мають вагу меншу за .

Визначимо різницю між двома компонентами  як мінімальну вагу з ребер, що з’єднує ці компоненти(Рис.). Тобто:



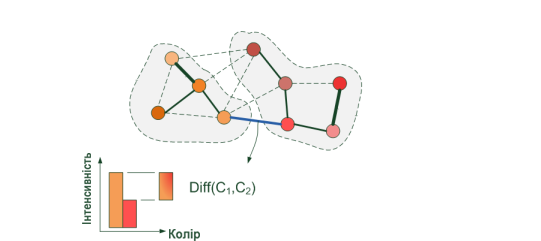


Рис. Різницю між двома компонентами

Якщо немає ребра, що з'єднує  і , то нехай . Ця міра різниці в принципі може бути проблематичною, оскільки вона відображає тільки найменшу вагу ребра між двома компонентами. Але ця міра працює досить добре, незважаючи на це очевидне обмеження. Більше того, зміна визначення для використання медіанної ваги або якого-небудь іншого квантиля, щоб зробити його більш стійким до аномалій, робить проблему пошуку хорошої сегментації NP-складною.

Предикат порівняння областей обчислює чи необхідно розділяти пару компонентів, перевіряючи, чи є різниця між компонентами великою у порівнянні до внутрішньої різниці принаймні в одному з компонентів,  і . Порогова функція використовується для керування ступенем, в якій різниця між компонентами повинна бути більше мінімальної внутрішньої різниці. Визначимо предикат попарного порівняння(Рис ) як,



де мінімальна внутрішня різниця, , визначається як,



Порогова функція керує ступенем, в якому відмінність між двома компонентами повинна бути більше за їх внутрішні відмінності, щоб між ними була межа ( повертав істинне значення). Для невеликих компонентів  не є хорошою оцінкою локальних характеристик даних. В крайньому випадку, коли . Тому ми використовуємо порогову функцію, засновану на розмірі компонента



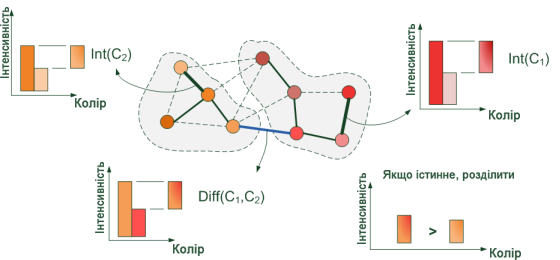


Рис. Графічний зміст предиката порівняння областей

Де  позначає розмір , а - деякий постійний параметр. Тобто для невеликих компонентів нам потрібні більш сильні докази кордону. На практиці  встановлює шкалу спостереження, в якій більше  викликає перевагу більших компонентів.

Будь-яка невід'ємна функція одного компонента може бути використана для без зміни алгоритму. Наприклад, метод сегментації може віддати перевагу компонентам певних геометричних форм, визначивши таку функцію , яка повертає великі значення для компонентів, які не відповідають деякій бажаній формі, і маленькі значення для тих, які відповідають. Це змусить алгоритм сегментації агресивно об'єднувати компоненти, які не мають бажаної форми. Зверніть увагу, що результатом цього будуть не тільки компоненти бажаної форми, проте серед будь-яких двох сусідніх компонентів один завжди буде бажаної форми

### 1.2.2. Алгоритм і його властивості

У цьому розділі ми опишемо і проаналізуємо алгоритм отримання сегментації з використанням критерію прийняття рішення , введеного вище. Ми покажемо, що сегментація, отримана за цим алгоритмом, не надто груба і не надто добра, згідно з такими визначеннями:

**Визначення 1** *Сегментація  називається «надто доброю», якшо існує пара областей  для яких немає доказів існування границі між ними.*

Щоб визначити, що значить для сегментації бути занадто грубою (мати занадто мало компонентів), ми спочатку введемо поняття *уточнення сегментації*. Для двох заданих сегментацій  і , утворених з одних початкових даних  називають уточненням , коли кожен компонент з  міститься (або еквівалентний) в деякому компоненті з . На додачу, ми кажемо, що  є *правильним уточненням* S коли . Зауважимо, що якщо  
- правильне уточнення , то  може бути отримано шляхом розщеплення однієї або декількох областей . Коли  є правильним уточненням  ми кажемо, що  більш добра за  і  більш груба за .

**Властивість 1.** Для будь-якого (скінченного) графу  існують деякі сегментації, які не є ні надто грубими, ні надто добрими.

Легко зрозуміти чому ця властивість виконується. Розглянемо сегментацію, де всі елементи знаходяться в одному компоненті. Очевидно, що ця сегментація не надто добра, тому що є тільки один компонент. Якщо сегментація також не надто груба, ми закінчили. В іншому випадку, за визначенням надто грубої сегментації, отримане правильне уточнення не є надто добрим. Обравши одне з цих удосконалень можна повторювати цю процедуру до тих пір, поки ми не отримаємо сегментацію, яка не буде надто грубою. Процедуру можна повторювати  разів, тому що всякий раз, коли ми вибираємо правильне уточнення, ми збільшуємо кількість компонентів у сегментації принаймні на один. Тоді *найкраща* сегментація, яку ми можемо отримати та, кожен елемент якої міститься у своєму компоненті.

Тепер звернемося до алгоритму сегментації, який тісно пов'язаний з алгоритмом Краскала для побудови мінімального остовного дерева графа. Його реалізація може працювати за час , де  – кількість ребер у графі.

**Алгоритм 1** *Алгоритм сегментації.*

На вхід подається граф , з  вершинами і  ребрами. Результатом є сегментація  на компоненти 

1. Сортуємо   у порядку зростання за вагами.
2. Починаємо з сегментації , де кожна вершина знаходиться у своєму компоненті.
3. Повторюємо крок 3 для .
4. Будуємо  за заданою . Нехай і  позначають вершини, з’єднані ребром  - й послідовності, тобто . Якщо і  знаходяться в системі неперетинних множин  і  мала, в порівнянні з внутрішньою різницею обох цих компонентів, то необхідно об’єднати ці два компоненти. В іншому випадку – нічого не робити. Більш формально, нехай є компонентом з , що містить  і  - компонент, що містить.Якщо  і , тоді при утворенні сегментації  компоненти об’єднуються. В іншому випадку компоненти  і переходять до сегментації без змін.
5. Повернути .

Тепер покажемо, що сегментація , отримана алгоритмом 1, підпорядковується глобальним властивостями не бути ні надто доброю, ні надто грубою при використанні предиката порівняння областей , визначеного в пункті 1.2.1. Тобто, хоча алгоритм і приймає тільки жадібні рішення, він породжує сегментацію, що задовольняє цим глобальним властивостям. Крім того, ми покажемо, що будь-які можливі сортування у порядку зростання, на 0 кроці алгоритму, породжують ту ж сегментацію.

**Лемма 1.** *Якщо на 3 кроці алгоритму, при розгляді ребра oq, компоненти, що розглядаються, не об’єднуються, то один з них обов’язково буде в кінцевій сегментації. Нехай  і  позначають два компоненти, що з’єднані ребром oq = (vi; vj), коли це ребро розглядається алгоритмом. Тоді, або , де - компонентом, що містить vi і  - компонент, що містить vj. укінцевій сегментації.*

**Доведення.** Є два випадки, які призведуть до того, що злиття не відбудеться. Розглянемо випадок . Оскільки ребра розглядаються в неспадаючому порядку за вагою , для всіх   
. Таким чином, ніяких додаткових злиттів з цим компонентом не відбудеться, тобто, . Аналогічно для .

Зверніть увагу, що у Леммі 1 мається на увазі, що ребро, яке є причиною злиття двох компонентів є саме ребром мінімальної ваги між компонентами. Таким чином, ребра, що викликають злиття, - це саме ті ребра, які будуть обрані алгоритмом Краскала для побудови мінімального остовного дерева (MST) кожного компонента.

**Теорема 1.** Сегментація, отримана за алгоритмом 1, яка використовує предикат порівняння областей , визначений в (3), не є надто доброю грубою, відповідно до визначення 1.

**Доведення.** За визначенням, для того, щоб  була надто доброю, повинна існувати деяка пара компонентів, для яких  не виконується. Між такою парою компонентів повинно бути хоча б одне ребро, яке розглядалося на кроці 3 і не викликало злиття. Нехай  перше таке ребро в послідовності. У цьому випадку алгоритм вирішить не об’єднувати з , що . За леммою 1 відомо, * або  У будь-якому випадку ми бачимо, що* . Мається на увазі, що  має місце для і , що є протиріччям.

**Теорема 2.** Сегментація, отримана за алгоритмом 1, яка використовує предикат порівняння областей , визначений в пункті 1.2.1, не є надто грубою, відповідно до визначення 2.

**Доведення.** Для того, щоб  була надто грубою, має бути якесь правильне уточнення, , яке не надто добре. Розглянемо мінімальну вагу ребра, яке є внутрішнім для компонента , але з'єднує різні компоненти . Відзначимо, що за визначенням уточнення  і .

Оскільки  не надто добра  або . Не втрачаючи загальності, скажемо, що модель вірна. За конструкцією, будь яке ребро, що з’єднує ** з іншим підкомпонентом  має масу щонайменше таку ж велику, як , яка, у свою чергу, перевищує максимальну вагу ребра в тому що . Таким чином, алгоритм, який розглядає ребра в неспадаючому порядку за вагами, повинен врахувати всі ребра з  перед іншими частинами з . Таким чином, алгоритм повинен сформувати ** перед формуванням , і при формуванні  він повинен об'єднати ** з деяким іншим підкомпонентом . Вага ребра, що викликала це злиття, повинна бути не менше . Однак алгоритм не об'єднав би ** в цьому випадку, оскільки . Протиріччя.

**Теорема 3.** *Сегментація, отримана за алгоритмом 1, не залежить від того, який неспадний порядок ребер за вагами використовується.*

***Доведення.*** Будь-який порядок може бути змінений на інший шляхом заміни сусідніх елементів з однаковою вагою. Таким чином, достатньо показати, що заміна двох сусідніх ребер однакової ваги в неспадаючому ваговому порядку не змінює результат, отриманий алгоритмом 1.

Нехай  та  – сусідні ребра в неспадній послідовності з однаковою вагою. Ясно, що якщо вони з'єднують непересічні пари компонентів, то й порядок, в якому вони розглядаються, не має значення. Єдиний випадок, який нам треба перевірити, - це коли  знаходиться між двома компонентами ** і **, а  - між одним з цих компонентів, скажімо **, та деяким іншим компонентом **.

Тепер ми покажемо, що  викличе злиття, коли розглядається після  так само як і при розгляді перед . Спершу, припустимо, що  викликає злиття при розгляді до . Це означає, що . Якби  розглядався перед , або  не викликав би злиття, і тривіально  все одно викликав би злиття, або  викликав би злиття, і в цьому випадку новий компонент  мав би  Таким чином, ми знаємо , що означає, що  все одно викличе злиття. З іншого боку, припустимо, що  не викликає злиття, якщо розглядати його до . Це означає, що. Тоді або , і в цьому випадку це все одно було б вірно, якби спочатку розглядався  (бо  не стосується **), або . У другому випадку, якщо  розглядався першим, він не міг викликати злиття, оскільки  і тому . При розгляді  після  у нас все ще є  і  не викликає злиття.

## 1.3. Обробка зображень

У цьому пункті розглядається теоретичне підґрунтя обробки цифрових зображень: описується згортка, як математична операція, та її використання на зображеннях. Також розглянуто декілька варіантів роботи на краях зображення, зокрема розширення, обгортання, віддзеркалення по краю та зріз ядра. В кінці описано розмивання зображення за допомогою фільтра на основі функції Гауса.

### 1.3.1. Згортка

**Згортка, як математична операція**

Згортка - це математична операція, що застосовується до двох функцій f і g, і породжує третю функцію, яка іноді може розглядатися як модифікована версія однієї з початкових. По суті, це особливий вид інтегрального перетворення.

Операцію згортки можна інтерпретувати як «схожість» однієї функції з відбитою і зрушеною копією іншої. Також згортка може бути описана як вага однієї функції в разі, якщо інша функція, будучи відображеною і зрушеною, є ваговою.

Нехай  — дві функції, інтегровані щодо міри Лебега у просторі . Тоді їх згорткою називається функція , визначена формулою:



Або для дискретного випадку:



**Згортка зображень**

Згортка - це операція обчислення нового значення обраного пікселя, що враховує значення оточуючих його пікселів. Для обчислення значення використовується матриця, яка називається ядром згортки. Зазвичай ядро ​​згортки є квадратною матрицею n \* n, де n - непарне, однак це не обов’язкова вимога. Під час обчислення нового значення обраного пікселя ядро ​​згортки як би «прикладається» своїм центром (саме тут важлива непарність розміру матриці) до даного пікселя. Навколишні пікселі так само накриваються ядром. Далі вираховується сума, де складовими є добутки значень пікселів на значення комірки ядра, що накрила даний піксель. Отримана сума ділиться на суму всіх елементів ядра згортки. Частка як раз і є новим значенням обраного пікселя. Якщо застосувати згортку до кожного пікселя зображення, то в результаті вийде якийсь ефект, що залежить від обраного ядра згортки.

Загальним виразом згортки є:



де — відфільтроване зображення,  — вхідне зображення,   
—ядро згортки. Розглядається кожен елемент ядра фільтра  та .

Приклад згортки для пікселя з координатами [2;2] (без нормування)



### 1.3.2. Обробка країв

При використанні операції згорки на краях зображення, стають необхідними значення пікселів за його межами. Звідси з’являються різні варіанти обробки країв:

* **Розширення (Рис 1.1)**

Найблищі до країв пікселі продовжуються настільки далеко, наскільки це необхідно для згортки. Кутові пікселі утворюють квадрати, а інші пікселі продовжуються у лінію.



Рис. 1.1 Тип обробки країв «розширення»

* **Обгортання(Рис 1.2)**

Зображення загорнуто саме в себе і значення беруться з протилежного краю або кута.



Рис. 1.2 Тип обробки країв «обгортання»

* **Віддзеркалення по краю(Рис 1. 3)**

При зверненні до пікселів за межами оригінального зображення беруться їх дзеркальні відображення відносно границь або кутів.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| e | d | d | e | f |
| b | a | a | b | c |
| b | a | a | b | c |
| e | d | d | e | f |
| h | g | g | h | i |

Рис. 1.3 Тип обробки країв «віддзеркалення по краю»

* **Зріз ядра**

Елементи, що виходять за межі зображення не враховуються. Коефіцієнт нормування коректується для компенсації.

### 1.3.3. Розмивання Гауса

Розмивання Гауса — це метод фільтрації зображення за допомогою функції Гауса, який призводить до розмивання зображення. Даний ефект широко використовується в графічних програмах, як правило, для зменшення зашумленості зображенні та зниження деталізації. Візуальний ефект цієї фільтрації розмивання аналогічний погляду на зображення крізь напівпрозорий екран.

**Принцип роботи**

Розмивання Гауса це тип фільтру розмивання зображення, що використовує функцію Гауса для розрахунку трансформації кожного пікселя у зображенні. Рівняння функції Гауса в одному вимірі:



Для двовимірного випадку:



де x — це відстань від початку координат в осі абсцис, y — це відстань від початку координат у осі ординат, а σ — стандартне відхилення розподілу Гауса.

Коли метод застосовується у двох вимірах, отримується поверхня, контури якої є концентричними колами розподілу Гауса з центральної точки (Рис 1.4). Значення з цього розподілу використовуються для створення матриці згортки(початок координат у центрі матриці). Для кожного нового значення пікселя визначається середнє зважене в околі пікселя. Значення поточного оригінального пікселя має більшу вагу (найвище значення розподілу Гауса), а сусідні пікселі отримують все меншу вагу в залежності від того наскільки далеко вони знаходяться від поточного оригінального пікселя. Це надає ефект розмитості, яка зберігає кордони та краї краще, ніж інші, аналогічні фільтри розмиття.



Рис1.4 Концентричні кола розподілу Гауса

Теоретично, в кожній точці зображення буде відмінним від нуля результат функції Гауса, це означає, що для розрахунку для кожного пікселя необхідно брати значення усіх пікселів у зображенні. На практиці, при обчисленні дискретного спрощення функції Гауса, вага пікселів на відстані більш ніж 3σ достатньо мала, щоб мати якийсь вплив на середнє зважене значення і вважається нулем. Значення пікселів, що розташовані за межами цього діапазону можуть бути проігноровані. Як правило, програмі обробки зображень необхідно тільки розрахувати матрицю з розмірами , щоб забезпечити результат, який є досить близьким до отриманого від усього розподілу Гауса.

Застосування розмивання Гауса призводить до розмивання на зображення і має ті ж наслідки, що застосування єдиного розмивання Гауса, радіус якого рівний квадратному кореню з суми квадратів радіусу розмиття, що застосовується. Наприклад, застосовуючи послідовного розмивання Гауса з радіусом 6 і 8 дає ті ж результати, як застосування єдиного розмивання радіусом 10, оскільки . Згідно з цим відношенням, час обробки не може бути зменшений шляхом імітації розмивання Гауса з послідовними процесом.

## 1.4. Кольорові моделі

Існує чимало кольорових моделей та їх варіацій. Усі вони мають переваги та недоліки в різних сферах застосування. В пункті розглядаються наступні:

* 1. **RGB,** як найпопулярніша кольорова модель для відтворення зображень на екранах комп’ютерів;
  2. **XYZ,** як проміжна модель у перетворенні RGB – LAB;
  3. **LAB,** як апаратно-незалежна розширена кольорова модель, в основі якої лежить сприйняття відмінності кольорів людиною.

Основні кольорові моделі та шляхи, що використовуються на практиці. за якими якими з однієї моделі колір може бути перетворений в іншу, наведені на наступній схемі(Рис 1.5).

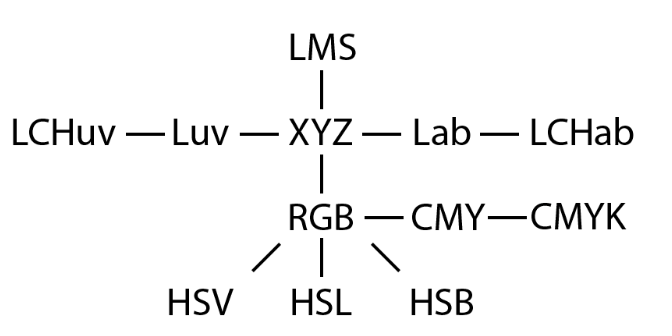


Рис 1.5 Схема перетворення кольорових моделей

### 1.4.1. RGB

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

**Кодування кольору**

У даній моделі колір кодується градаціями складових каналів (**R**ed, **G**reen, **B**lue). Тому за збільшення величини градації котрогось   
каналу — зростає його інтенсивність під час синтезу.

Кількість градацій кожного каналу залежить від розрядності бітового значення RGB. Зазвичай використовують 24-бітну модель, у котрій визначається по 8 біт на кожен канал, і тому кількість градацій дорівнює 256, що дозволяє закодувати 2563 = 16 777 216 кольорів(Рис 1.6).

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

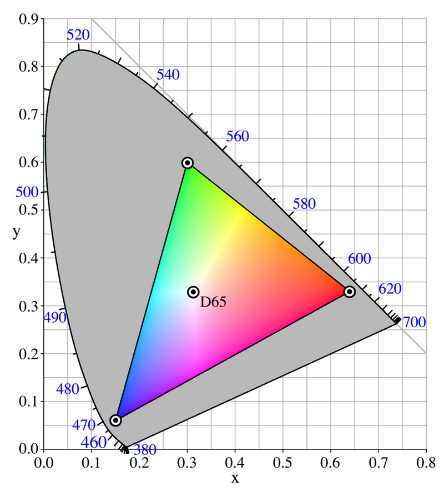


Рис 1.6 Можливості і обмеження передачі кольорів у RGB

RGB — апаратно-залежний простір кольорів.

Типовими приладами із RGB - входом є кольоровий телевізор і відеокамера, сканер і цифровий фотоапарат.

### 1.4.2. XYZ

З метою уніфікації була розроблена міжнародна стандартна колірна модель. В результаті серії експериментів міжнародна комісія з освітлення (CIE) визначила криві складання основних (червоного, зеленого і синього) кольорів. У цій системі кожному кольору, що може побачити людина, відповідає певне співвідношення основних кольорів. Використовують уявні основні кольори: X (уявний червоний), Y (уявний зелений), Z (уявний синій).

При описі кольору значення X,Y,Z називають стандартними основними збудження, а отримані на їх основі координати – стандартними кольоровими координатами. Стандартні криві складання X(λ),Y(λ),Z(λ) (див. Рис. 1.7) описують чутливість середньостатистичного спостерігача до стандартних збуджень:

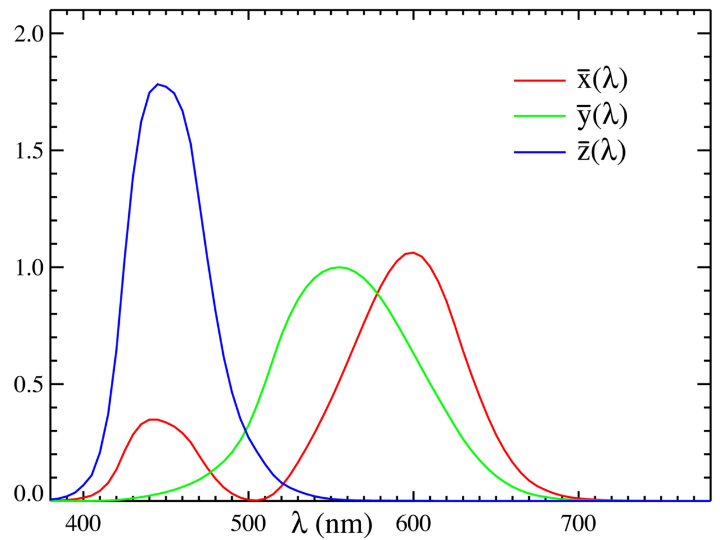


Рис. 1.7 Функції кольорового сприйняття Стандартного колориметричного спостерігача, визначені комітетом CIE у 1931 році на діапазоні довжин хвиль від 380 нм до 780 нм (з інтервалом 5 нм).

Отриманий кольоровий простір можна представити у вигляді графіку(Рис 1.8).

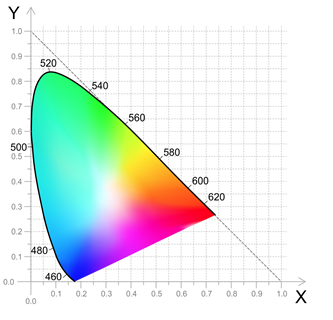


Рис 1.8 Можливості передачі кольорів в просторі XY

Множину кольорів, що задається таким способом, називають трикутником CIE. Легко помітити, що трикутник CIE описує тільки колірний тон, але ніяк не описує яскравість. Для опису яскравості вводять додаткову вісь. В результаті отримують колірне тіло CIE (див. Рис. 1.9):

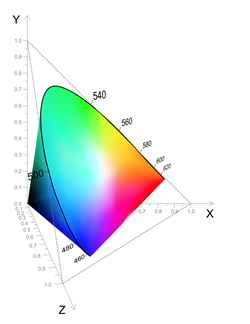


Рис 1.9 Колірне тіло СIE

Це тіло містить всі кольори, видимі середньостатистичним спостерігачем. Основним недоліком цієї системи є те, що використовуючи її, ми можемо констатувати тільки збіг чи розходження двох кольорів, але відстань між двома точками цього простору не відповідає зоровому сприйняттю відмінності кольорів.

**Конвертування RGB – XYZ**

***Перетворення RGB — XYZ***

1-й крок – гамма-коректування:

, де u —  нормалізовані компоненти RGB(поділені на 255).

2-й крок – лінійне перетворення

.

***Зворотнє перетворення***

1-й крок – лінійне перетворення

******.

2-й крок — гамма-коректування:

Ці значення не є кінцевим результатом; гамма-корекція все ще повинна застосовуватися. Наступна формула перетворює отримані значення в RGB:

, де u - .

Далі значення домножуються на 255 і округлюються до цілих чисел.

### 1.4.3. Lab

**Lab** — система задання кольорів, що використовує як параметри світлосилу, відношення зеленого до червоного та відношення синього до жовтого (Рис 1.10) . Ці три параметри утворюють тривимірний простір, точки якого відповідають певним кольорам.

Основною метою при розробці CIELAB було усунення нелінійності системи CIE XYZ з точки зору людського сприйняття. Під абревіатурою LAB зазвичай розуміється колірний простір CIE L\*a\*b\*, яке на даний момент є міжнародним стандартом.

Колірна модель L\*a\*b розроблялась як апаратно-незалежна, тобто вона задає кольори без врахування особливостей відтворення кольорів. Має три параметри для опису кольору: світлосила L (англ. *Lightness*) — рівень освітлення сцени та два хроматичні параметри. Перший (умовно позначений латинською літерою a) вказує на співвідношення зеленої і червоної складової кольору, другий (позначений літерою b) — співвідношення синьої та жовтої складової.

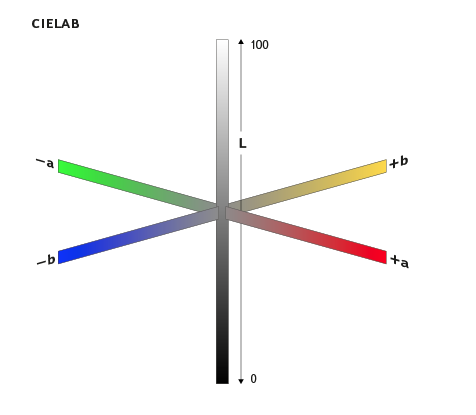


Рис 1.10 Система задання кольорів у LAB

Lab — різноконтрастний кольоровий простір, в якому відстань між кольорами відповідає мірі відчуття їх відмінності людиною.

На малюнку(Рис.1.11) представлені зрізи колірного тіла CIE L\*a\* b\* для двох значень світлини:

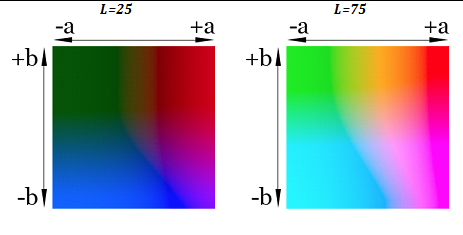


Рис 1.11 Зріз колірного тіла СIE на L = 25 і L = 75

**Конвертування XYZ – Lab**

**Перетворення XYZ -> L\*a\*b\***





, де



Значення {\displaystyle X\_{n}},{\displaystyle Y\_{n}} і {\displaystyle Z\_{n}} — це координати білої точки в значеннях CIE XYZ (буква n означає «нормалізована»).

**Зворотне перетворення L\*a\*b\* -> XYZ**

1. Задати 
2. задати 
3. задати 
4. якщо , інакше

****

1. якщо , інакше

****

1. якщо , інакше

****

**CIELAB ΔE\***

Міжнародна комісія з освітлення (CIE) задає свою метрику кольорової різниці через ΔE\*ab (також званої ΔE\*, або "дельта E"), де літера дельта - використовується для позначення різниці, а E означає Empfindung(з німецької - "відчуття").

Перцептивні неоднорідності в кольоровому просторі CIELAB призвели до того, що CIE протягом багатьох років уточнював своє визначення, що призвело до появи різних формул 1994 і 2000 років. Ці неоднорідності важливі, тому що людське око більш чутливий до певних кольорів, ніж інші. Хороша метрика повинна враховувати це, щоб поняття " помітної різниці" мало сенс. В іншому випадку певне ΔE, яке може бути незначним між двома кольорами в одній частині колірного простору, будучи значним в іншій частині.

**CIE76**

Для двох кольорів  і , заданих в кольоровому просторі CIELAB за стандартами CIE76 формула різниці визначається як

**.**

****відповідає помітній різниці в кольорах**.**

**CIE94**

Визначення 1976 року було розширено для вирішення перцептивних нерівномірностей, шляхом введення специфічних вагових коефіцієнтів.

 , де





















Деякі коефіцієнти залежать від області використання:

|  |  |  |
| --- | --- | --- |
|  | Мистецтво | Промисловість |
|  | 1 | 2 |
|  | 0.045 | 0.048 |
|  | 0.015 | 0.014 |

# РОЗДІЛ 2 Практична Частина

## 2.1. Технічні засоби та структури даних

У пункті 2.1.1 розглядається структура даних «Система неперетинних множин», яка використовується у реалізованому програмному забезпеченні для представлення ґраткового графу, на основі якого проводиться сегментація цифрового зображення.

У пункті 2.1.2 описано 2 методи зчитування цифрових зображень на мові програмування C# з тестуванням їх швидкодії.

### 2.1.1. Система неперетинних множин

Система неперетинних множин (англ. *disjoint-set-union* або **DSU**, також використовують назви англ. *union–find data structure*, англ. *merge–find set*) — структура даних, яка дозволяє відстежувати множину елементів, розбиту на неперетинні підмножини(Рис 2.1). При цьому кожній підмножині призначається її представник — елемент цієї підмножини.

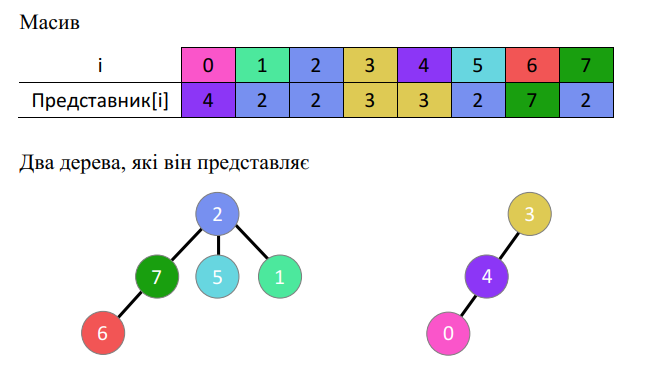


Рис 2.1 Система неперетинних множин, яка реалізована на масивах

Спочатку є кілька елементів, кожен з яких знаходиться в окремій (своїй власній) множині. За одну операцію можна об'єднати дві будь-які множини, а також можна дізнатися, в якій множині зараз знаходиться зазначений елемент. У класичному варіанті вводиться ще одна операція — створення нового елемента, який поміщається в окрему множину.

Таким чином, базовий інтерфейс даної структури даних складається всього з трьох операцій:

* *MakeSet* — додання нового елемента; розміщення його в нову множину, що складається з одного нього;
* *Union* — об'єднання двох зазначених множин;
* *Find* — повернення значення, в якій множині знаходиться зазначений елемент. Насправді при цьому повертається один з елементів множини званий представником (англ. *representative*) або лідером (*leader*). Цей представник вибирається в кожній множині самою структурою даних (і може змінюватися з плином часу). Наприклад, якщо виклик для якихось двох елементів повернув одне і те ж значення, то це означає, що ці елементи знаходяться в одній і тій же множині, а в іншому випадку — в різних множинах.

**Побудова ефективної структури даних**

Множина елементів буде зберігатися у вигляді дерев: одне дерево відповідає одній множині. Корінь дерева — це представник (лідер) множини.

При реалізації це означає, що ми заводимо масив, в якому для кожного елемента зберігаємо посилання на його предка у дереві. Для коренів дерев будемо вважати, що їх предок — вони самі (тобто посилання зациклюється в цьому місці).

**Наївна реалізація**

Наведеної вище інформації вже достатньо, щоб написати першу реалізацію системи неперетинних множин. Вона буде досить неефективною, але потім буде покращена за допомогою двох прийомів, що дадуть змогу отримати в результаті майже константний час роботи.

Отже, вся інформація про множини елементів зберігається за допомогою масиву **parrent**.

***Опис операцій***

* Щоб створити новий елемент (операція make\_set(v)), ми просто створюємо дерево з коренем у вершині, зазначаючи, що його предок — це він сам.
* Щоб об'єднати дві множини (операція union(a, b)), ми спочатку знайдемо лідерів першої і другої множини. Якщо лідери збіглися, то нічого не робимо — це означає, що множини і так вже були об'єднані. В іншому випадку можна просто вказати, що предок першої вершини дорівнює другій (або навпаки) — тим самим приєднавши одне дерево до іншого.
* Реалізація операції пошуку лідера (find(v)) проста: ми піднімаємося по предкам від вершини, поки не дійдемо до кореня. Цю операцію зручніше реалізувати рекурсивно.

void make\_set (int v) {

parent[v] = v;

}

int find (int v) {

**if** (v == parent[v])

**return** v;

**return** find\_set (parent[v]);

}

void union (int a, int b) {

a = find\_set (a);

b = find\_set (b);

**if** (a != b)

parent[b] = a;

}

Утім, така реалізація системи неперетинних множин дуже неефективна. Легко побудувати приклад, коли після кількох об'єднань множин вийде ситуація, що множина — це дерево, звиродніле в довгий ланцюжок.

Для уникнення подібних ситуацій вводяться евристики: «Евристика стиснення шляху» та «Евристика об'єднання за рангом».

**Евристика стиснення шляху**

Ця евристика призначена для прискорення роботи операції ***Find***.

Вона полягає в тому, що коли після виклику ми знайдемо шуканого представника множини, то запам'ятаємо, що у вершині *v* і всіх пройдених по шляху вершин — саме цей лідер. Найпростіше це зробити, перенаправивши їх parrent-покажчик на цю вершину(Рис 2.2).

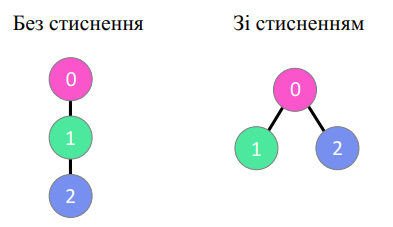


Рис 2.2 Ілюстрація стиснення шляхів

З іншого боку, зрозуміло, що не можна зробити, щоб ці покажчики parrent завжди вказували на лідера: інакше при виконанні операції довелося б оновлювати лідерів у елементів.

Таким чином, до масиву слід підходити саме як до масиву предків, можливо, частково стиснутого.

Нова реалізація операції має такий вигляд:

int find (int v) {

**if** (v == parent[v])

**return** v;

**return** parent[v] = find\_set (parent[v]);

}

Така проста реалізація робить все, що задумувалося: спочатку шляхом рекурсивних викликів знаходиться лідера множини, а потім, в процесі розкрутки стека, цей лідер присвоюється parrent посиланнями для всіх пройдених елементів.

Реалізувати цю операцію можна і не рекурсивно, але тоді   
доведеться здійснювати два проходи по дереву: перший знайде шуканого лідера, другий — проставить його всім вершин шляху. Втім, на практиці нерекурсивна реалізація не дає істотного виграшу.

**Евристика об'єднання за рангом**

Розглянемо іншу евристику, яка сама по собі здатна прискорити час роботи алгоритму, а в поєднанні з евристикою стиснення шляхів і зовсім здатна досягти практично константного часу роботи на один запит в середньому.

Ця евристика полягає в невеликій зміні роботи операції Union: якщо в наївній реалізації те, яке дерево буде приєднано до якого, визначається випадково, то тепер ми будемо це робити на основі рангів.

Є два варіанти рангової евристики: в одному варіанті рангом дерева називається кількість вершин в ньому, в іншому — глибина дерева (точніше, верхня межа на глибину дерева, оскільки при одночасному застосуванні евристики стиснення шляхів реальна глибина дерева може зменшуватися).

В обох варіантах суть евристики одна й та ж: при виконанні Union будемо приєднувати дерево з меншим рангом до дерева з більшим рангом(Рис 2.3).

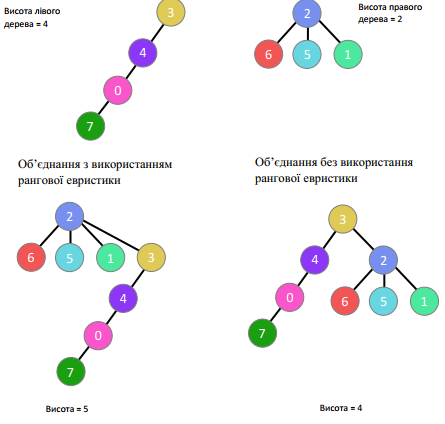


Рис 2.3 Ілюстрація евристики об'єднання за рангом

Реалізація **рангової евристики на основі глибини дерев**:

void make\_set (int v) {

parent[v] = v;

rank[v] = 0;

}

void union(int a, int b) {

a = find\_set (a);

b = find\_set (b);

**if** (a != b) {

**if** (rank[a] < rank[b])

swap (a, b);

parent[b] = a;

**if** (rank[a] == rank[b])

++rank[a];

}

}

**Об'єднання евристик: стиснення шляху та рангова евристика**

При спільному застосуванні евристик: стиснення шляху та об'єднання за рангом — час роботи на один запит виходить O(A(n)) в середньому, де A(n)— обернена функція Акермана, яка зростає дуже повільно, настільки повільно, що для всіх розумних обмежень вона не перевершує 4 (приблизно для n<=10600).

*Підсумкова реалізація системи неперетинних множин:*

void make\_set (int v) {

parent[v] = v;

rank[v] = 0;

}

int find (int v) {

**if** (v == parent[v])

**return** v;

**return** parent[v] = find\_set (parent[v]);

}

void union (int a, int b) {

a = find\_set (a);

b = find\_set (b);

**if** (a != b) {

**if** (rank[a] < rank[b])

swap (a, b);

parent[b] = a;

**if** (rank[a] == rank[b])

++rank[a];

}

}

### 2.1.2. Зчитування зображень

Перш ніж реалізовувати алгоритм сегментації на мові програмування високого рівня c#, необхідно було дослідити різні методи зчитування зображення, тобто його перетворення до матриці структур Color.

В усіх методах використовується класичний клас Bitmap (System.Drawing.Bitmap). Даний клас зручний тим, що приховує від нас деталі кодування растрових форматів, підтримуючи всі поширені формати, типу BMP, GIF, JPEG, PNG.

При тестах послідовно виконуються 100 зчитувань зображення розмірами 1000\*1000 пікселів. Процесор – Intel Core i5-8600k. При інтервальному оцінюванні, ймовірність промаху повз довірчій інтервал

**«Стандартний» метод**

Суть методу полягає у використанні методу Bitmap.GetPixel(x, y). Наведемо повністю код методу, який конвертує вміст Bitmap в тривимірний масив. При цьому перша розмірність - це колірна компонента (від 0 до 2),   
друга — позиція y, третя - позиція x.

public static double[,,] BitmapToDoubleRgbNaive(Bitmap bmp)

{

int width = bmp.Width,

height = bmp.Height;

double [,,] res = new double[3, height, width];

for (int y = 0; y < height; ++y)

{

for (int x = 0; x < width; ++x)

{

Color color = bmp.GetPixel(x, y);

res[0, y, x] = color.R;

res[1, y, x] = color.G;

res[2, y, x] = color.B;

}

}

return res;

}

**Результати**

Загальний час тесту: 72,4922 секунди.

Таблиця 2.1 Тестування швидкодії зчитування   
зображень за допомогою GetPixel

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Нижнє значення | Значення | Верхнє значення | Середньоквадратичне відхилення |
| Мат. Сподівання | 0.7247 | 0.7262 | 0.7278 | 0.0008 |
| Середньоквадратичне відхилення | 0.0070 | 0.0081 | 0.0092 | 0.0005 |

**Пряма робота з даними Bitmap через вказівники**

Клас Bitmap надає більш швидкий спосіб звернутися до своїх даних. Для цього нам необхідно скористатися посиланнями, наданими класом BitmapData та арифметикою вказівників:

public unsafe static double[,,] BitmapToDoubleRgb(Bitmap bmp)

{

int width = bmp.Width,

height = bmp.Height;

double[,,] res = new double[3, height, width];

BitmapData bd = bmp.LockBits(new Rectangle(0, 0, width, height), ImageLockMode.ReadOnly, PixelFormat.Format24bppRgb);

try

{

byte\* curpos;

fixed (double\* \_res = res)

{

double\* \_r = \_res,

\_g = \_res + width \* height,

\_b = \_res + 2 \* width \* height;

for (int h = 0; h < height; h++)

{

curpos = ((byte\*)bd.Scan0) + h \* bd.Stride;

for (int w = 0; w < width; w++)

{

\*\_b = \*(curpos++); ++\_b;

\*\_g = \*(curpos++); ++\_g;

\*\_r = \*(curpos++); ++\_r;

}

}

}

}

finally

{

bmp.UnlockBits(bd);

}

return res;

}

**Результати**

Загальний час тесту: 1,7823 секунди.

Таблиця 2.2 Тестування швидкодії зчитування   
зображень за допомогою вказівників

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Нижнє значення | Значення | Верхнє значення | Середньоквадратичне відхилення |
| Мат. Сподівання | 0.0168 | 0.0178 | 0.0189 | 0.0006 |
| Середньоквадратичне відхилення | 0.0048 | 0.0056 | 0.0063 | 0.0004 |

Незважаючи на те, що перший метод є дуже простим у застосуванні, використовувати його як модуль у сегментації зображень немає сенсу, через його низьку ефективність порівняно з другим методом.

## 2.2. Опис програмного забезпечення

Для виконання поставленої задачі було розроблено програмне забезпечення для обробки цифрових зображень.

Розглянемо основні характеристики додатку.

*Мова програмування*:C#.

*Платформа:* .NET Framework.

*Середовище розробки*: Microsoft Visual Studio 2017.

*Мінімальні системні вимоги*:

* 64 – бітна операційна система Windows;
* Microsoft .NET Framework 4.7;
* 67 КБ доступного простору на жорсткому диску.

*Вхідні дані:* цифрові зображення наступних форматів:JPEG (.jpg), BMP (.bmp), GIF (.gif), PNG (.png).

*Вихідні дані:* відсегментовані цифрові зображення. Є можливість отримати всі сегменти на вхідному зображенні, позначені різним кольором, або масив сегменів.

Після зчитування зображення з файлу, програма проводить перетворення від растрового зображення до множини тривимірних точок. Кожна точка описується:

* інтенсивністю компоненти R(red);
* інтенсивністю компоненти G(green);
* інтенсивністю компоненти B(blue);

Далі є можливість змінити яскравість і/або контрастність зображення та перевірити наскільки параметр розмиття sigma впливає на зображення.

На основі множини тривимірних точок будується ґратковий граф і на ньому проводиться сегментація. Результати можна оцінити за допомогою 3 функціоналів якості, а саме:

1. суми внутрішньокластерних дисперсій;
2. суми попарних внутрішньокластерних відстаней;
3. загальної внутрішньокластерної дисперсії.

Після отримання результатів сегментації передбачена можливість швидкого переключення до вхідних даних і назад.

При побудові ґратковий графу, для визначення ваг ребер, можуть використовуватися 3 кольорові схеми:

* відтінки сірого;
* RBG;
* Lab.

У застосунку передбачена можливість роботи з великими зображеннями. Для зручності, вони будуть пропорційно зменшені для розташування у робочій області. Слід зазначити, що всі маніпулювання(розмиття, сегментація) будуть проведені над оригінальним(не зменшеним) зображенням.

На діаграмі діяльності(Рис. 1) показані можливі сценарії роботи програми.

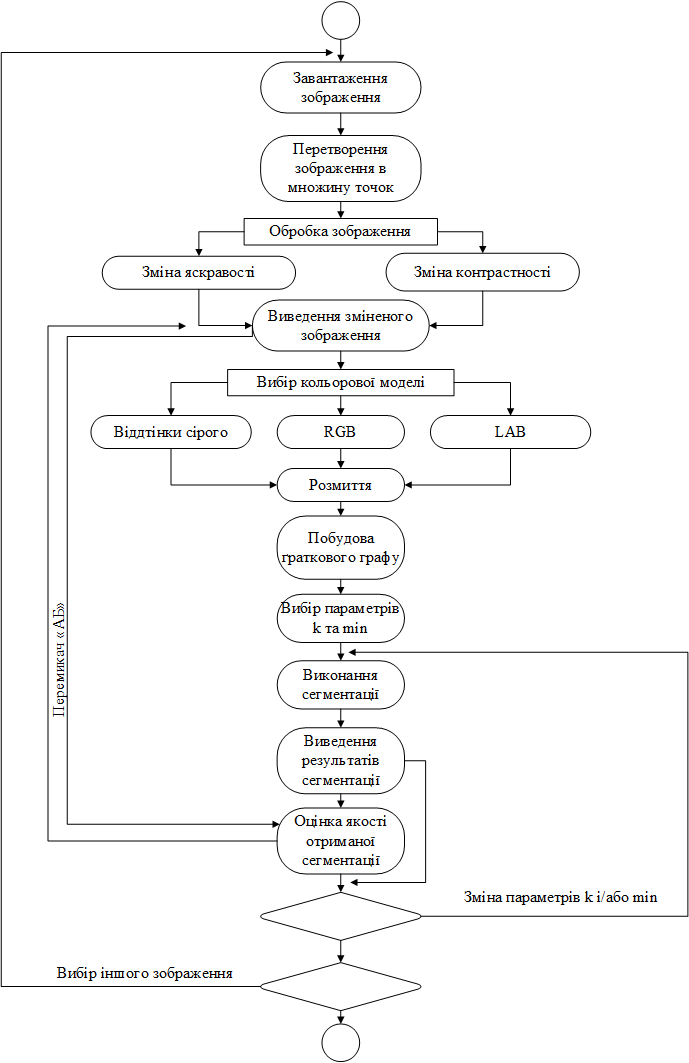


Рис. Діаграма діяльності

**Побудова ґраткового графу**

Для виконання сегментації на графі загальноприйтим є з’єднання всіх пікселів з 4 або 8 сусідами. В даній роботі сусідами вважаться не тільки ті пікселі, що стоять поруч на горизонталі та вертикалі, а й на діагоналях. Таким чином піксель має 8 сусідів.

Для того, щоб побудувати ґратковий граф для зображення за однин прохід використовується шаблон(Рис ). Результат його використання для цифрового зображення розмірами 5\*4 показано на малюнку(Рис).

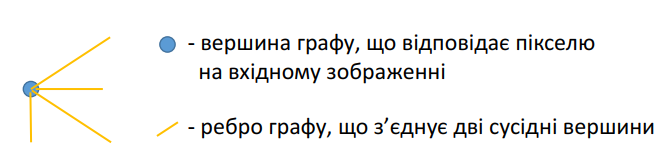


Рис. Шаблон з’єднання пікселів

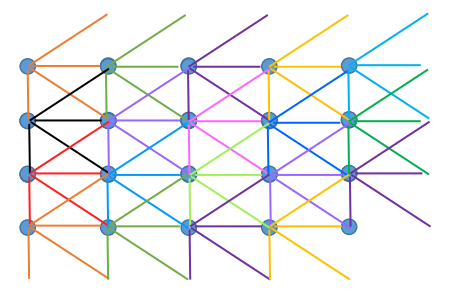


Рис. Ґратковий граф для зображення 5\*4.

Як видно з малюнку(Рис. ) шаблон, обраний для побудови графу задовольняє необхідним умовам:

* кожен некрайовий піксель має рівно по 8 сусідів;
* схема графу, побудованого з використання шаблону, що розглядається, утворює мозаїку(що характерно для ґраткового графу).

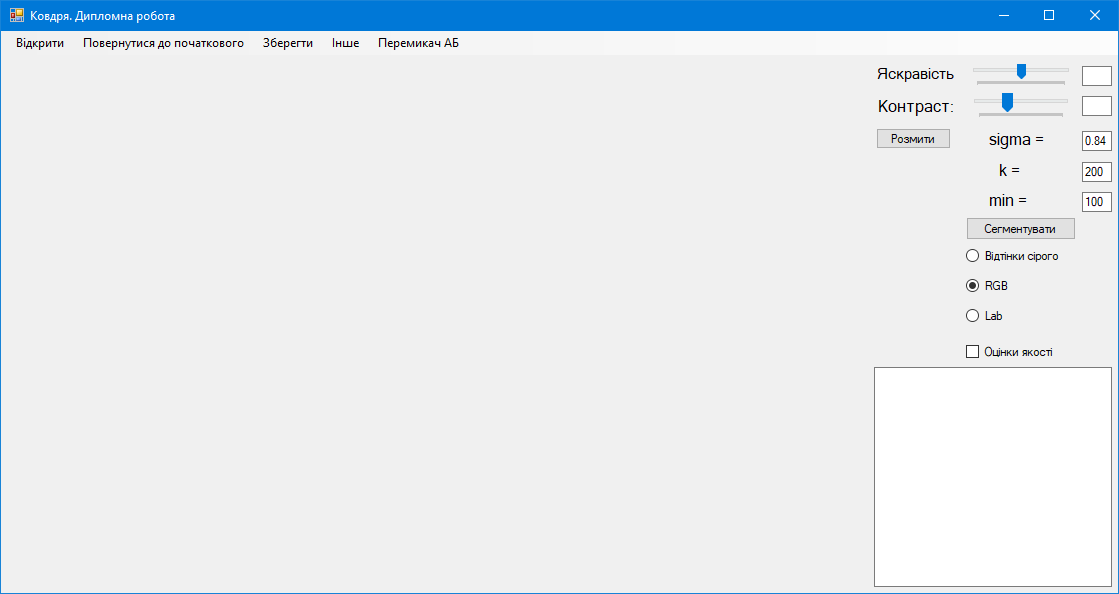
Очевидно, що для крайових пікселів не можливо побудувати усі 8 зв’язків, оскільки для цього було б необхідним існування пікселів за межами зображення. Таким чином, для останнього пікселя в зображенні, відсутня можливість побудувати нові ребра з його участю, оскільки усі можливі варіанти уже будуть існувати.

Тоді, враховуючи вищесказане, можна зробити висновок, що для довільного цифрового зображення, при побудові ґраткового графу буде створено  ребер, де

,

де  та — ширина та висота зображення відповідно.

### 2.1 Інструкція користувача

Після запуску програми на екрані відкриється початкове вікно(Рис 1).

(1)

(2)

(3)

Вікно застосунку складається з 3 основних областей:

1. меню програми;
2. область зображення;
3. область налаштувань.

Далі детально розглядаються кожна з областей.

***Меню програми***

Меню – частина інтерфейсу програми, яка маніпулює цифровими зображеннями, що подаються на вхід. З нього починається взаємодія користувача с застосунком.

В меню входять 5 пунків: «*Відкрити», «Повернутися до початкового», «Зберегти», «Інше» та «Перемикач АБ».*

Для початку роботи програми необхідно зчитати файл, що містить зображення формату BMP, GIF, JPEG, PNG, TIFF натиснувши на кнопку «Відкрити». Це призведе до відкриття діалогового вікна вибору файлу (рис. 2.6). Після вибору зображення, воно буде зчитано, перетворено на множину точок та виведено на екран в «область зображення».

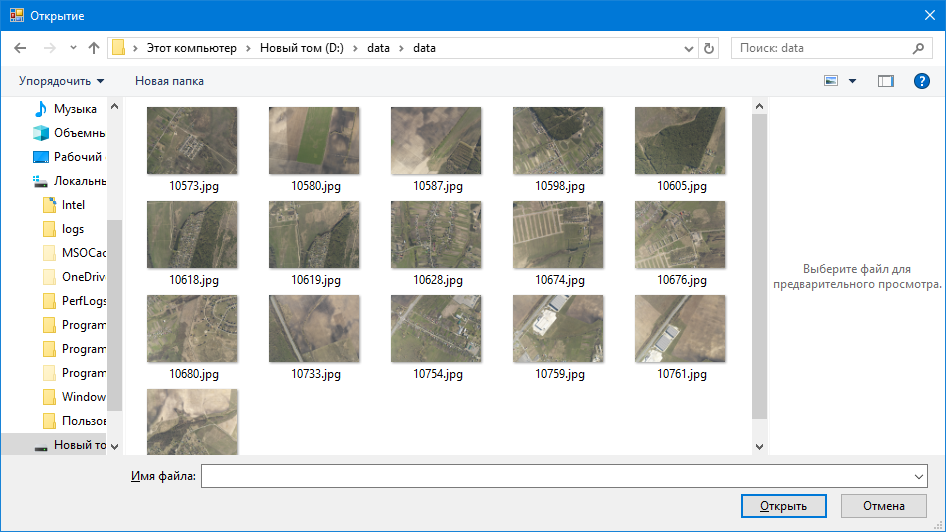


Рис . Діалогове вікно вибору зображення для роботи

Кнопка «*Повернутися до початкового*» призначена для відміни всіх перетворень зображення. Після зчитування зображення програма створює його копію, яку неможливо отримати іншим способом, окрім цієї кнопки. Якщо, за бажанням користувача, в області налаштувань буде проведена обробка зображення, яка призведе до його зміни, то для повторного отримання оригінального зображення не потрібно проводити повторне відкриття.

Кнопка «Зберегти» призначена для збереження результатів сегментації цифрового зображення. Після її натискання буде відкрито діалогове вікно збереження файлу(Рис ) для збереження останнього отриманого обробленого зображення. При збереженні користувачу буде дозволено обрати місце у його файловій системі, та ім’я файлу. За замовчуванням, застосунок пропонує встановити ім’я файлу, яке складається з початкового імені та дописаними до нього параметрами сегментації(sigma, s, min).

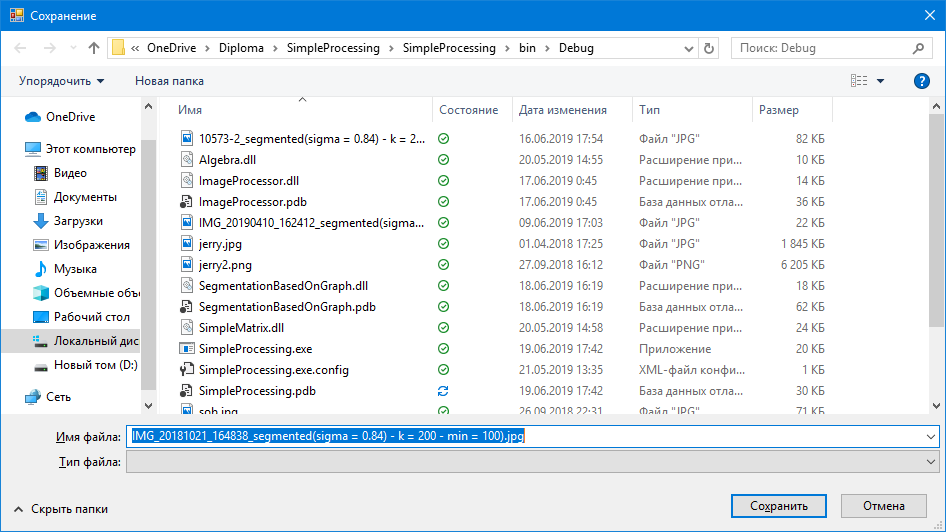


Рис. Діалогове вікно збереження результатів сегментації

В пункті «*Інше*» знаходиться процедура «Прийняти яскравість та контрастність зображення за 0». Якщо, при роботі з зображенням було змінені параметри яскравість і/або контрастність, то за допомогою цією процедури користувач зможе прийняти опрацьоване зображення за початкове і повторно змінити бажані параметри(Рис ).

Здебільшого випадний список «*Інше*» призначений для розміщення додаткових процедур, якщо такі будуть розроблені і додані до застосунку.

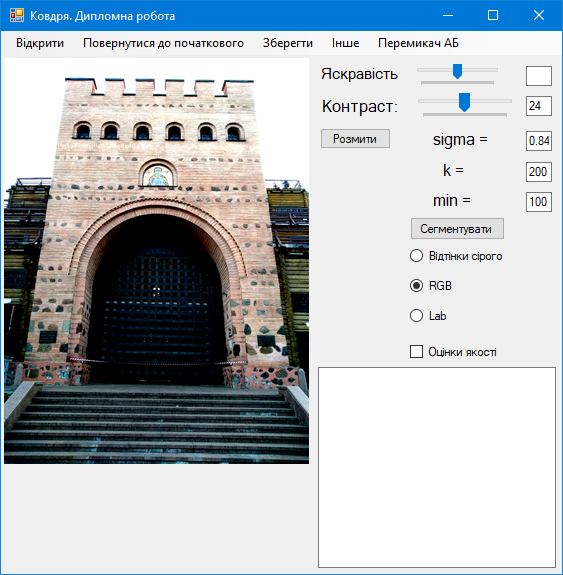
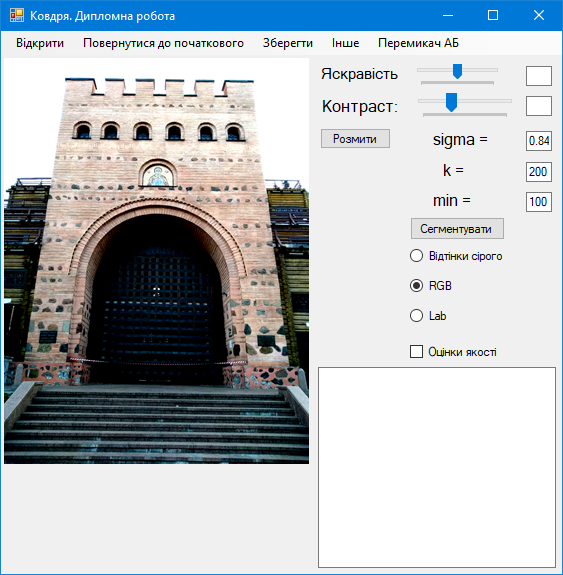
 

Рис. Зміна контрастності на початковому зображенні(зліва) та використання процедури обнулення(справа).

Кнопка «*Перемикач АБ*» призначена для порівняння результатів сегментації з початковим зображенням. Після проведення сегментації, на «область зображення» будуть виведені її результати. Якщо скористатися командою «Повернутися до початкового», то буде виведене вхідне зображення. Але щоб повернутися до результатів сегментації, в такому випадку, необхідно буде знову чекати обробку зображення, тобто виконання всього алгоритму.

Якщо ж після отримання результатів сегментації, для порівняння з початковим зображенням використати «Перемикач АБ», то вхідне зображення буде замінюватися результатами сегментації і навпаки. При чому, при виведенні результатів сегментації буде використовуватися кешована копія, тобто алгоритм заново виконуватися не буде.

**Область зображення**

Область зображення – частина інтерфейсу програми, яка призначена для виведення цифрових зображень та результатів їх сегментації.

При виведенні зображення аналізуються розміри області доступні для виведення зображення та розміри зображення, що буде виводитися(Рис ). Якщо зображення менше за «доступне місце» вікна програми, то воно буде виведено без змін. Якщо ж зображення має великі розміри, то воно буде пропорційно зменшене під час виведення.

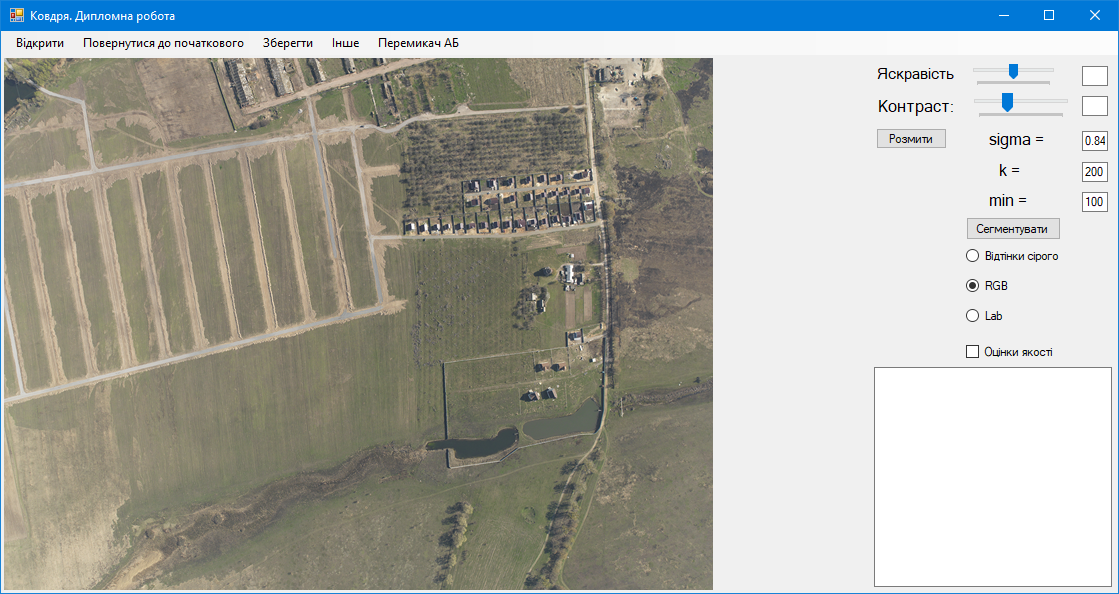


Рис. Виведення зображення розміром 2000\*1500 пікселів при розмірі робочого вікна 850\*515 пікселів.

Якщо мало місце зменшення при виведенні, то при обробці буде використовуватися початкове(велике) зображення, тобто зменшення зображення при виведенні не призводить до фактичного зменшення вхідного зображення.

**Область налаштувань**

Область налаштувань – частина інтерфейсу програми, яка відповідає за зміну яскравості та контрастності зображення, введення параметрів алгоритму сегментації(sigma, k, min), вибір кольорової схеми при виконанні алгоритму(відтінки сірого, RGB, LAB) та виведення деякої додаткової текстової інформації(кількість сегментів, та оцінки якості сегментації, якщо їх визначення було ввімкнене).

*Обробка зображення*

Для простої обробки цифрового зображення в додатку реалізовані зміни яскравості та контрастності.

**Зміна яскравості** відбувається за рахунок додавання до кожної з RGB компонент однакового значення (Рис ). За рахунок зміни значення повзунка яскравості у програмі відбувається підбір значення, яке додається до компонент.

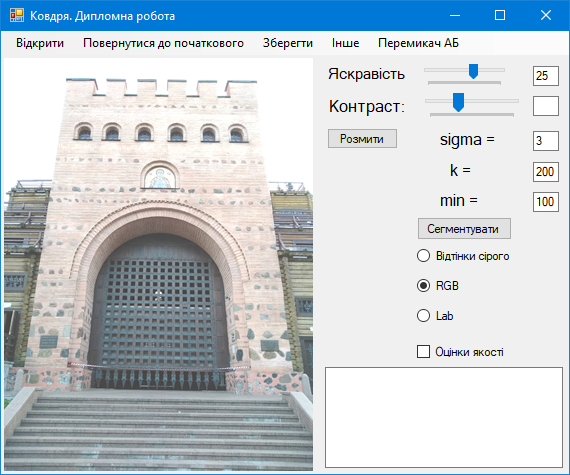
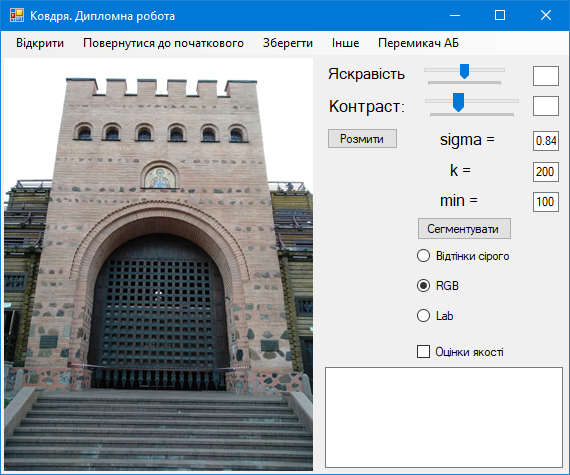


Рис. зміна яскравості

**Зміна контрастності** відбувається за рахунок збільшення різниці між конкретним пікселем та середнім значенням кольору по всьому зображенні.

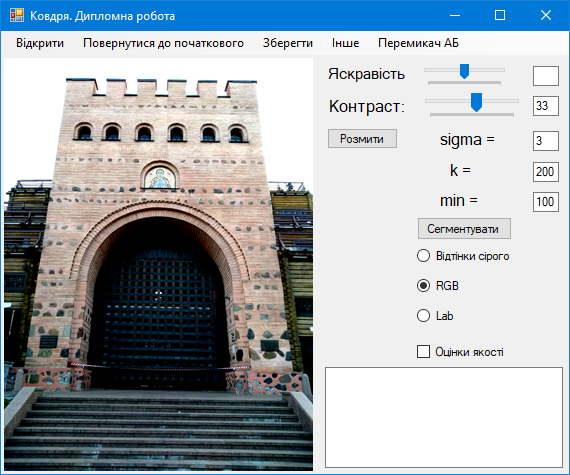
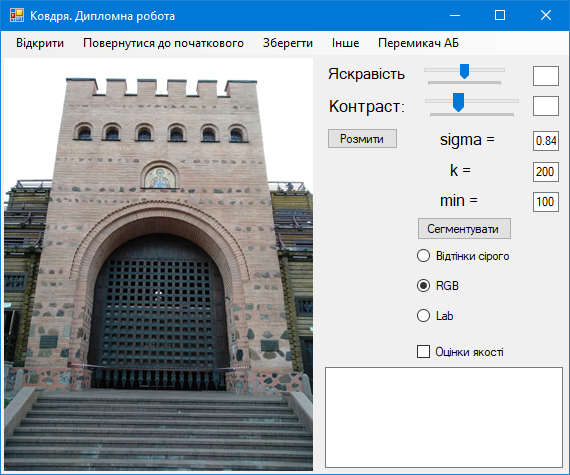


Рис. Зміна контрастності

*Параметри сегментації*

Далі задаються параметри сегментації: sigma – параметр розмиття для згладження зображення, k –параметр що відповідає за ступінь відмінності між сегментами та min – розмір мінімального сегмента(на постобробці сегментації сегменти з розмірами меншими за цей параметр будуть об’єднуватися з сусідами). Дослідження залежності результатів сегментації від цих параметрів будуть показані далі.

За бажанням, користувач може подивитися як змінюється зображення під час розмиття з заданим параметром(Рис ). Для цього поряд з параметром sigma є кнопка «Розмити». Але після підбору необхідного значення параметру, необхідно повернутися до вхідного зображення, оскільки розмиття включене в процедуру сегментації і буде виконано повторно, якщо цього не зробити.

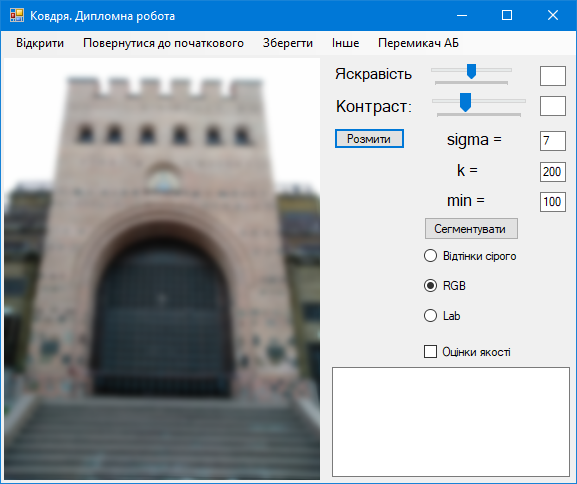
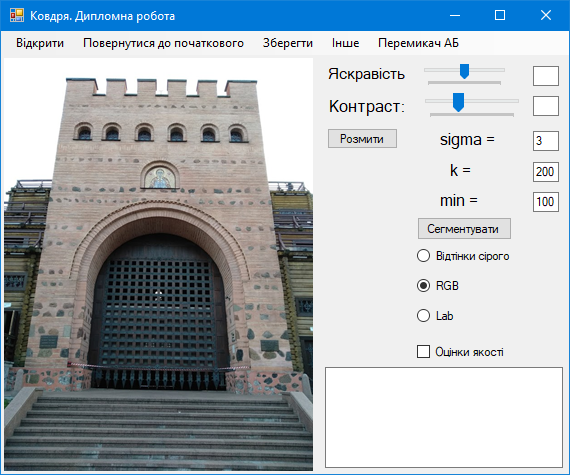


Рис. Розмиття зображення

Під кнопкою «сегментувати» надається можливість для вибору кольорового простору при виконанні сегментації. Доступні кольорові простори: відтінки сірого, RGB, LAB.

Під кольоровими просторами знаходиться перемикач для включення/виключення обрухунку оцінок якості отриманої сегментації.

Внизу області налаштувань знаходиться текстове поле для виведення кількості знайдених сегментів та оцінок якості сегментації, якщо їх обрахунок увімкнено.

Приклад виведення результатів сегментації з обрахованими оцінками якості показано на малюнку(Рис )

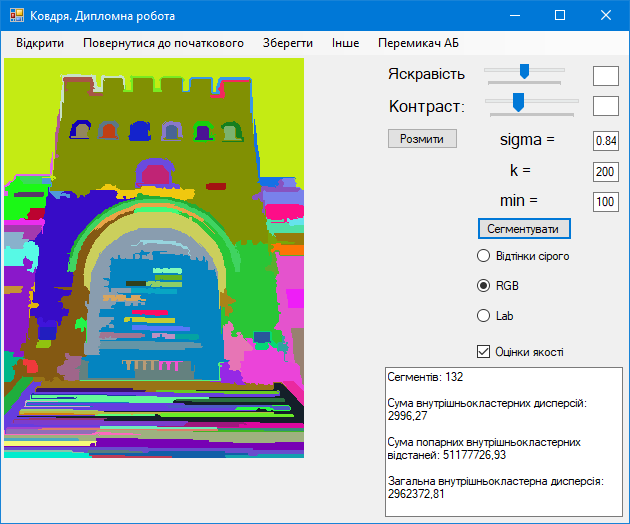


Рис. Приклад виведення результатів сегментації

# Список використаних джерел

* 1. Barghout, Lauren, and W. Lee Lawrence. Perceptual information processing system, July 11, 2003.
  2. T. Kanungo, D. M. Mount, N. S Netanyahu, C. D. Piatko, Silverman. An efficient k-means clustering algorithm: analysis and implementation, 2002.
  3. H. Mobahi, S. R. Rao, A. Y. Yang, S. S. Sastry, Y. Ma. Segmentation of natural images by texture and boundary compression. International journal of computer vision, 2011, pp. 86–98.
  4. R. Shankar, M. Hossein, Y. Allen, S. Shankar and M. Yi. Natural Image Segmentation with Adaptive Texture and Boundary Encoding, 2009 pp. 135-146.
  5. L. G. Shapiro and G. C. Stockman. Computer Vision, 2001, pp 279–325.
  6. R. Ohlander, K. Price, and D. Reddy. Picture segmentation using a recursive region splitting method, 1978, pp 313–333
  7. V. Caselles, R. Kimmel, G. Sapiro. Geodesic active contours, 1997.
  8. S. Osher and N. Paragios. Geometric Level Set Methods in Imaging Vision and Graphics, 2003.