

## **ABSTRACT**

The thesis contains pages, figures and

Video recordings of lectures are no longer a rarity in the conditions of distance learning. Videos may be in an inconvenient format for students or contain different artifacts due to compression, camera quality, and other factors. It is useful to have a presentation of the study material, which contains only the text from the board because such a view of the material is most like the compendium.

The item of this work is creating an algorithm for obtaining panorama slides without a teacher from video lecture.

To perform the study

- SIFT method for obtaining image's keypoints;
- Homography for getting panorama slides;
- Boykov-Kolmogorov Max-flow algorithm for creating mask of moving objects;
- Denoising and binarizing operators;

IMAGES PROCESSING, INTELLECTUAL VIDEO PROCESSING, VIDEO STABILIZATION

## **РЕФЕРАТ**

Дисертація містить сторінки, ілюстрацій і бібліографію з найменувань.

Відеозаписи лекцій вже не рідкість в умовах дистанційного навчання. Відео може бути у незручному форматі для студентів або містити різні артефакти через стиснення з втратами, якість зйомки та інші фактори. Добре було б мати презентацію навчального матеріалу, яка містить лише текст з дошки, оскільки це більш наближено до формату конспекту.

Метою роботи є створення алгоритму, який з відео лекції робить панорамні слайди без викладача.

Для досягнення мети було використано

- Метод SIFT для знаходження ключових точок зображення;
- Гомографію для отримання панорамних слайдів;
- Алгоритм максимального потоку Бойкова-Колмогорова для отримання маски рухомих об'єктів;
- Оператори знешумлення та бінаризації зображень.

**ОБРОБЛЕННЯ ЗОБРАЖЕНЬ, ІНТЕЛЕКТУАЛЬНЕ ОБРОБЛЕННЯ ВІДЕО,  
СТАБІЛІЗАЦІЯ ВІДЕО**

## ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів . . . . .	9
Вступ . . . . .	10
1 Аналіз попередніх робіт. . . . .	12
1.1 Приклади . . . . .	12

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

### Стандартні позначення

$\mathbb{N}$  — множина натуральних чисел;

$\mathbb{R}$  — множина дійсних чисел;

### Позначення, введені в дисертації

$F$  — множина кадрів відео;

$T$  — довжина відео (кількість кадрів);

$P$  — множина координат пікселів зображення;

$C$  — множина інтенсивностей пікселів зображення;

$F_p^i$  — інтенсивність пікселя з координатою  $p \in P$ ;

## ВСТУП

Дана робота завдячує Кригіну Валерію Михайлович — молодшому науковому співробітнику Відділу обробки та розпізнавання образів Міжнародного науково-навчального центру інформаційних технологій і систем НАН України та МОН України.

**Актуальність роботи.** Задача створення конспекту лекцій не є новою, а з масовим переходом навчання у віддалений режим стала більш актуальною проблема створення якісних лекційних матеріалів, зокрема слайдів, які містять стислу суть занять. Було б добре мати технологію, що перетворюватиме відео лекцію у стисле слайд шоу, на якому будуть написи з дошки без викладача і до того ж всю область дошки, навіть якщо камера рухалась.

### **Мета і завдання дослідження.**

*Об'єкт дослідження* — математичні методи реєстрації зображень.

*Предмет дослідження* — автоматична обробка відеозаписів.

Метою роботи є розробка алгоритму, що перетворює відео лекції у панорамні знімки без викладача.

Завдання наступні:

- 1) вивчити методи обробки та порівняння зображень, які знадобляться для створення панорамних знімків;
- 2) вивчити та доповнити математичні методи, що допоможуть визначати рухомі об'єкти;
- 3) розробити демонстраційне програмне забезпечення.

### **Методи дослідження:**

- 1) опрацювання літератури за темою;
- 2) створення теоретичного підґрунтя алгоритму;
- 3) написання програми;
- 4) аналіз отриманих результатів.

### **Наукова новизна одержаних результатів.**

Створено алгоритм та реалізовано інформаційну технологію для одержання панорамних слайдів з відео лекції.

### **Практичне значення одержаних результатів.**

За допомогою програми викладач може надати короткий зміст відео матеріалу. Наступними кроками є автоматичне детектування дошки, розбиття її сектори та розпізнавання написаного на дошці тексту та формул.

### **Публікації.**

.... Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених «Теоретичні і прикладні проблеми фізики, математики та інформатики».



спектив за допомогою гомографії. Гарна якість оцифрування дошки досягається насамперед тим, що вона має білий колір, що в свою чергу накладає обмеження для використання технології з дошками зеленого чи чорного кольорів. У наступній своїй роботі [?] ці ж самі автори побудували систему, яка в реальному часі обробляє відеозапис, видаляє людину біля дошки за допомогою часової медіани, але в даному випадку немає панорамного склеювання знімків. Головна ідея роботи була зробити програму для телеконференцій.

### 1.1.2 Автоматичне сканування дошки

Автори даної роботи [?] створили програму яка переводить написи на білій дошці у цифрові. Вони реалізували локалізацію тексту та подальшу його обробку. Дана технологія не передбачає перекривання викладачем написів або дошку іншого кольору відмінного від білого.

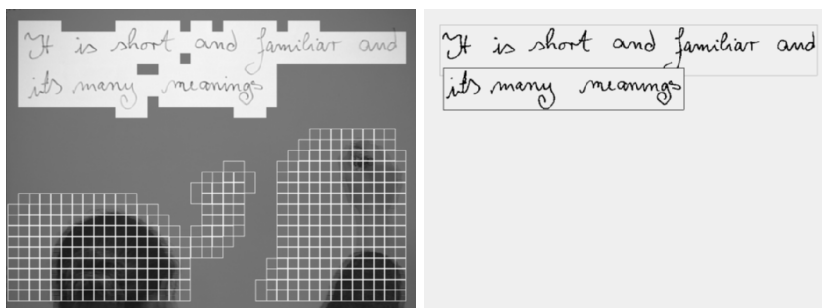


Рисунок 1.2 — Демонстрація роботи сканування дошки

Можна помітити, що як і в попередній роботі гарна якість виокремлення написів досягається тим що дошка білого кольору.

### 1.1.3 Відстежування об'єкту та віднімання фону від Стенфорду

У 2012 році науковці зі Стенфордського університету *Alex Gonzalez, Bongsoo Suh, Eun Soo Choi* представили технологію [?] локалізацію дошки (навіть такої яка



розділена на частини), відстеження викладача та його подальше прибирання. Алгоритм також може працювати з різними кольорами дошок. Для прибирання викладача і всіх рухомих об'єктів автори також використали тимчасову медіану.

Дана програма не працює в реальному часі, оскільки всі операції над кадрами відео займають тривалий час, не кажучи вже, що відео перед обробкою піддають компресії.

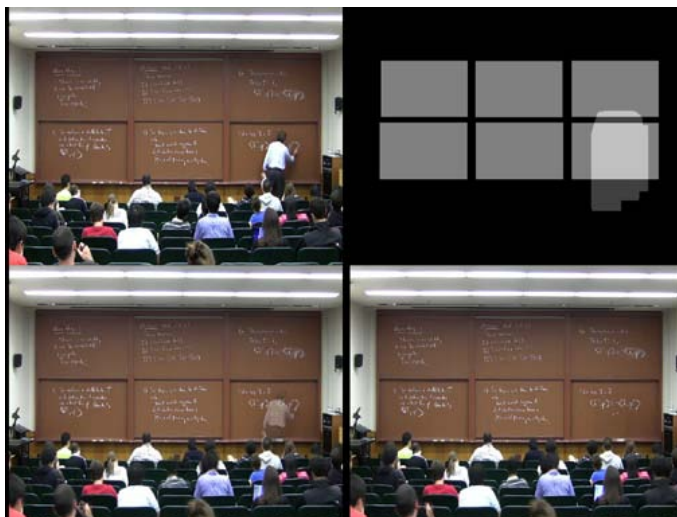


Рисунок 1.3 — Приклад роботи авторів Стенфорду

Головною особливістю даної роботи є те, що алгоритм автоматично локалізує різну кількість дошок. Однак варто відмітити, що тестування відбувалось на відео-лекціях де камера знімає всю дошку і не рухається за викладачем. Тим більше якість отриманих написів теж погана.

#### 1.1.4 Виокремлення написів дошки від Тайванського університету

У 2014 році науковці із Тайванського університету представили свій аналог [?] алгоритму по оцифровуванні дошки. Для видалення викладача автори застосували алгоритм кластеризації K-means. Для отримання бінаризованих написів з дошки використана адаптивне вирівнювання '(яке не ясно)'. Варто відмітити гарну якість власного методу знешумлення.

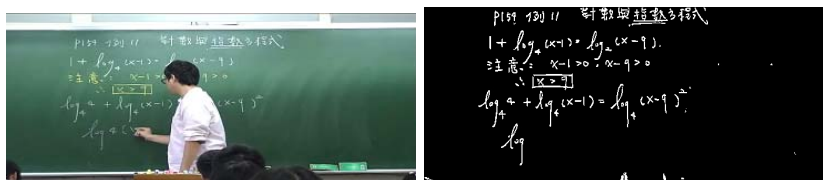


Рисунок 1.4 — Демонстрація роботи сканування дошки

Автори не представили швидкість обробки всього відео. Щоб отримати якісну сегментацію дошки та викладача потрібно, щоб кадр містив тільки викладача, причому щоб одяг викладача був сильно відмінним від кольору дошки. Це потрібно для коректної роботи K-means. Тобто нема гарантії, що якийсь рухомий об'єкт не стане дошкою під час класифікації.

### 1.1.5 Сучасна робота

На даний момент найсучаснішою повною технологією [?], яка оцифровує дошку є робота науковців з університету Рочестер. Автори *Kenny Davila* та *Richard Zanibbi* використали просторовий темпоральний індекс для викремлення написів і викладача. Відбувається видалення не самого викладача, а його контурів, після бінаризації картинки. Варто відмітити, що і тут камера має бути нерухомою.

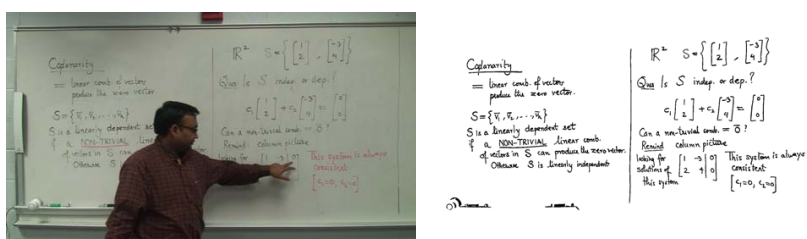


Рисунок 1.5 — Демонстрація роботи сканування дошки

Пізніше ціж автори створили повністю згорткову нейронну мережу [?] для обробки написів дошки. Дана нейронна мережа LectureNet досить добре бінаризує написи з дошки.

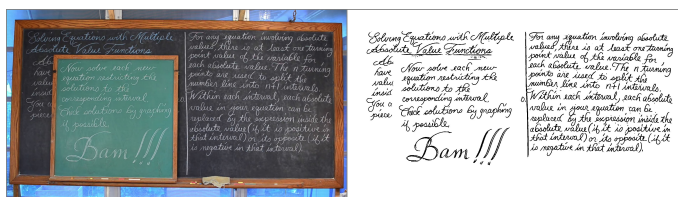


Рисунок 1.6 — Приклад роботи авторів з Рочестер