

*Войтко Вікторія Володимирівна,
к.т.н., доцент кафедри програмного забезпечення,
Вінницький національний технічний університет, Україна,
Ракитянська Ганна Борисівна,
к.т.н., доцент кафедри програмного забезпечення,
Вінницький національний технічний університет, Україна,
Мокану Пінтілій Пентелійович, студент групи ІПЗ-15мн,
факультет інформаційних технологій і комп'ютерної інженерії,
Вінницький національний технічний університет, Україна*

АВТОМАТИЗОВАНА СИСТЕМА ФОРМУВАННЯ ГОЛОСОВИХ ПОВІДОМЛЕНЬ ДЛЯ ДОПОМОГИ ЛЮДЯМ З ОБМЕЖЕНИМИ МОЖЛИВОСТЯМИ ЗОРУ

Розглянуто можливості Google Cloud для створення програм, які допоможуть незрячим людям.

Ключові слова: Web API, Google Cloud, Node.js, Google Visio, допомога незрячим.

The possibilities of Google cloud to create programs that help not seeing people.

Keywords: Web API, Google cloud, Node.js, Google Visio, help the blind.

Вступ. Як прогнозує всесвітня організація охорони здоров'я, до 2020 року у світі буде близько 75 мільйонів незрячих людей [1]. Сьогодні сліпих людей нараховують близько 34 мільйонів та 124 мільйони людей мають поганий зір [1]. Тому актуальним завданням є розробка сучасних систем, призначених для формування голосових повідомлень з описом предметів навколишнього світу, що допоможе людям з обмеженим зором комфортніше почуватися та орієнтуватися в просторі, керуючись звуковими повідомленнями.

Метою роботи є підвищення комфортності людей з обмеженими можливостями зору шляхом використання розробленої автоматизованої системи формування голосових повідомлень, які описують предмети навколишнього світу, ідентифіковані з допомогою камери спостереження.

Головною задачею роботи є розробка засобів автоматизованої системи формування голосових повідомлень на основі алгоритму роботи з Google API, орієнтованого на автоматичне формування описово-ідентифікаційного

повідомлення про предмети, образи яких розпізнаються з допомогою аналізу зображень, отриманих з камери спостереження.

Об'єктом дослідження є методи розпізнавання зображень та методи автоматичного створення голосових повідомлень.

Під предметом дослідження розуміємо засоби створення автоматизованої системи формування голосових повідомлень для людей з обмеженими можливостями зору, орієнтовані на ідентифікацію зображень та створення голосового повідомлення з їх описом за допомогою API сервісу Google Cloud.

Розробка автоматизованої системи формування голосових повідомлень для людей з обмеженими можливостями зору. Під автоматизованою системою допомоги людям з обмеженими можливостями зору розуміємо систему, що має допомогти незрячим людям краще орієнтуватися в просторі за допомогою системи розпізнавання об'єктів, які звичайна людина бачить перед собою. Автоматизована система складається з 2-х основних частин: розпізнавання отриманого зображення та озвучення описово-ідентифікаційного повідомлення про розпізнаний об'єкт.

Для розробки ефективного алгоритму розпізнавання образів доцільно розробити нейронну мережу, яка може вчитися на основі отриманого досвіду, що забезпечить динамічне розширення можливостей системи. Серед існуючих сервісів, які можуть надати свої послуги в розпізнанні образів, обрано відкриту систему Google Cloud та, зокрема, Google Vision. Google надав користувачам доступ до своїх сервісів через WEB API.

Google Vision дозволяє розпізнати об'єкти з допомогою отриманого зображення в графічному чи фотоформаті та отримати розширену описову інформацію про робочий об'єкт. При цьому комплексний рисунок поділяється на фрагменти з відокремленням образів його складових об'єктів. Рис. 1 ілюструє процес розділення комплексного зображення на окремі об'єкти.

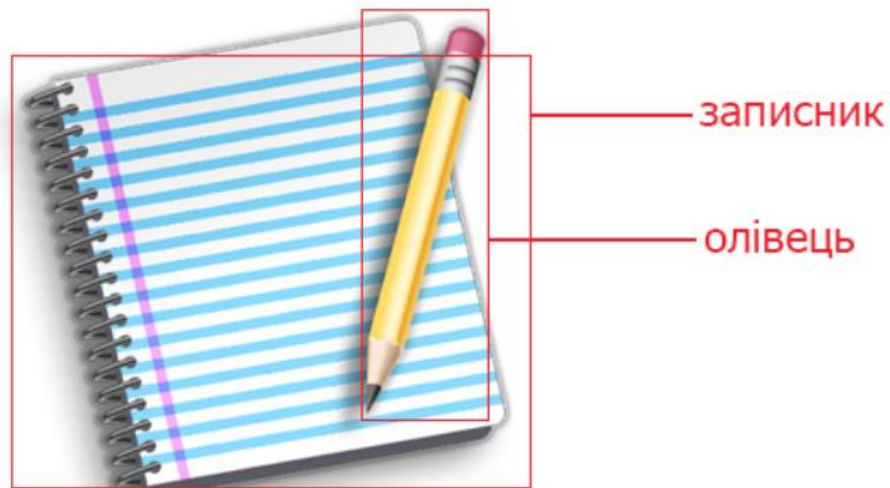


Рисунок 1 - Розділення вхідного образу на окремі об'єкти

В основі алгоритму розпізнавання образів сервісом Google Vision лежить алгоритм порівняння об'єктів. Процес порівняння здійснюється шляхом порівняння за формою робочого об'єкта з еталоном. Процес порівняння акумулює спектр етапів ідентифікації зображення. Основним етапом є перетворення координат точок контуру аналізованого об'єкта і порівняння його форми з еталоном.

Розглянемо алгоритм ідентифікації робочого об'єкта. На першому етапі проводимо сегментацію комплексного зображення на окремі образи його складових частин. Визначаємо координати початкових і кінцевих точок отриманих зображень.

Нехай координати їх початкових точок будуть відповідно (x_n, y_n) , (x_{n_em}, y_{n_em}) , а кінцевих точок – (x_k, y_k) , (x_{k_em}, y_{k_em}) . Початкова точка буде опорною для суміщення і повороту об'єкта навколо неї, а кінцева точка потрібна для визначення взаємної орієнтації об'єкта і еталона. Маючи по дві однотипні точки на Ω еталоні й досліджуваному об'єкті і знаючи їх координати, обчислюються функції $\sin \Omega$ і $\cos \Omega$ кута повороту Ω :

$$\sin \Omega = \frac{(y_{\kappa_em} - y_{n_em})(x_{\kappa} - x_n) - (x_{\kappa_em} - x_{n_em})(y_{\kappa} - y_n)}{\sqrt{(y_{\kappa_em} - y_{n_em})^2 + (x_{\kappa_em} - x_{n_em})^2} \sqrt{(y_{\kappa} - y_n)^2 + (x_{\kappa} - x_n)^2}};$$

$$\cos \Omega = \frac{(x_{\kappa_em} - x_{n_em})(x_{\kappa} - x_n) - (y_{\kappa_em} - y_{n_em})(y_{\kappa} - y_n)}{\sqrt{(y_{\kappa_em} - y_{n_em})^2 + (x_{\kappa_em} - x_{n_em})^2} \sqrt{(y_{\kappa} - y_n)^2 + (x_{\kappa} - x_n)^2}}.$$

Якщо координати точок контуру об'єкта (x_i, y_i) , а їх координати після перетворень зсуву, повороту і зміни масштабу (x_i^*, y_i^*) , маємо співвідношення:

$$x_i^* = [(x_i - x_n) \cos \Omega - (y_i - y_n) \sin \Omega] M^* + x_{n_em};$$

$$y_i^* = [(y_i - y_n) \cos \Omega + (x_i - x_n) \sin \Omega] M^* + y_{n_em}.$$

Порівняння за формою аналізованого об'єкта і еталона полягає в зчитуванні й аналізі точок сформованого кадру у порівнянні з шаблоном еталону, які мають перетворені координати точок контуру об'єкта. За розглянутим алгоритмом підраховується відношення не співпадаючих з шаблоном точок контуру об'єкта до їх загальної кількості. Якщо це відношення є найменшим, то чергове порівняння зараховується як результуюче і здійснюється перехід до наступного етапу ідентифікації зображення. Нульове значення розрахованого відношення означає, що об'єкт повністю збігається з еталоном за розмірами і формою в рамках заданого діапазону відхилень, і процедура ідентифікації зображення закінчується.

На рис. 2 зображено узагальнену модель автоматизованої системи розпізнавання образів та створення описово-ідентифікаційних голосових повідомлень, що базується на використанні сервісів Google Cloud.

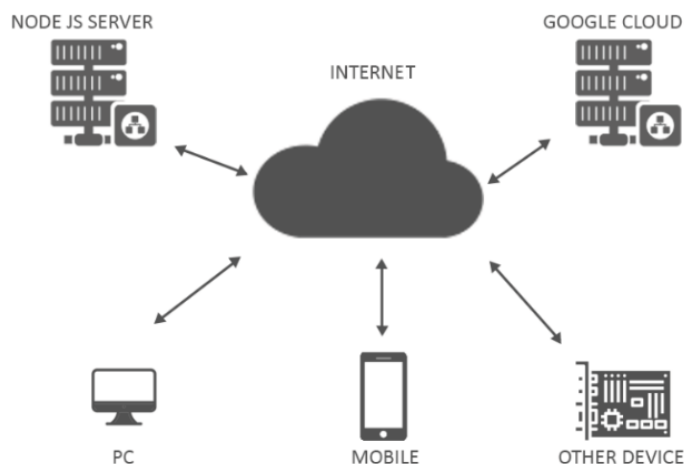


Рисунок 2 – Модель автоматизованої системи допомоги людям з обмеженими можливостями зору

Google розробив API, що дозволяє використання потужностей його серверів і забезпечує високу швидкість роботи клієнтських програм.

Висновки. Розроблена автоматизована система допомоги людям з обмеженими можливостями зору дозволяє з використанням сервісів Google Cloud розпізнавати отримані з камери спостережень зображення та формувати файл голосового повідомлення з ідентифікацією об'єкта та коротким його описом. Швидка та надійна робота програми допоможе людям з обмеженим зором комфортніше почуватися та орієнтуватися в просторі.

Список використаної літератури

1. Кожні п`ять секунд у світі сліпне одна людина [електронний ресурс] // Режим доступу: <http://health.unian.ua/country/152175-kojni-pyat-sekund-u-sviti-slipne-odna-lyudina.html> – Назва з екрану.
2. Абламейко С.В. Обработка изображений / С.В. Абламейко, Д.М.Лагуновский – Минск: Амалгея, 2000. – 206 с.
3. Фурман Я.А. Введение в контурный анализ; приложения к обработке изображений и сигналов / Я.А. Фурман, А.В. Кревецкий, А.К. Передреев и др. / Под ред. Я.А. Фурмана – М: Физматлит, 2003. – 592 с.
4. Русин В.П. Системи синтезу, обробки та розпізнавання складноструктурованих зображень / В.П. Русин – Львів: Вертикаль, 1997. – 264 с.