

ПРОГРАМА РОЗПІЗНАВАННЯ ЗОБРАЖЕННЯ ТА ПОШУК АНАЛОГІВ У БАЗІ ДАНИХ

Місюра Олександр Сергійович, студент групи 41 П

Науковий керівник:

Баклан Ігор Всеволодович, кандидат технічних наук

Анотація. Метою дослідження є застосування методів та систем розпізнавання зображень. Головна задача розпізнавання образів — це задача віднесення вихідних даних до певного класу за допомогою виділення істотних ознак, які характеризують ці дані, із загальної маси несуттєвих даних.

Ключові слова: комп'ютерний зір, аналіз зображень, образ, контекстна інформація, шум.

Вступ. Більшість вчених і практиків з сфери ІТ усвідомлюють роль, яку відіграє комп'ютерний зір, але мало хто знає про всі області, в яких він використовується. Наприклад, більшість людей знає, що комп'ютерний зір використовується в системах спостереження, а також при роботі з зображеннями та відео в інтернеті. Мало хто знає, що аерозйомка і вуличні карти широко використовують калібрування камери і технологію зшивання зображень. Широкого поширення набув комп'ютерний зір і у виробництві: практично все масове виробництво в певний момент технологічного процесу потребує використання комп'ютерного зору.

Ось ще деякі проблеми які вирішує комп'ютерний зір: вирівнювання відсканованих зображень, зменшення шуму на медичних знімках, аналіз об'єктів, системи виявлення вторгнення, автоматичний моніторинг, системи контролю, калібрування камери, додатки для військових, безпілотників, наземні і підводні апарати. Деякі витонченими рішеннями є розпізнавання звуків і музики, за допомогою аналізу спектрограм.

Постановка задачі. Комп'ютерне зір - процес перетворення даних, отриманих з фотоапаратів і відеокамер, в нову виставу. Всі ці перетворення виконуються для досягнення якоїсь конкретної мети. Вхідні дані можуть містити деяку контекстну інформацію, таку як "камера встановлена в машині" або "датчик глибини визначив об'єкт в радіусі 1 метра". Рішеннями може бути "є людина в цій сцені" або "є 14 пухлинних клітин на знімку". Процес перетворення кольорового зображення в чорно-біле або усунення ефекту руху камери з послідовності зображень.

Мета написання дослідження. Метою дослідження є застосування методів та систем розпізнавання зображень. Аналіз проблем комп'ютерного зору.

Основна частина. Не варто себе думати, що завдання комп'ютерного зору такі вже й легкі. Яким тяжким може бути пошук, скажімо, автомобіля, коли є тільки його образ? Інтуїція може бути вельми оманливою. Людський мозок розділяє сигнал надійшов від зору на безліч каналів, які згодом передають різного виду інформацію в ваш мозок. Мозок влаштований таким чином, що увага концентрується тільки на важливих ділянках зображення, виключаючи при цьому інші. Існує маса відповідних сигналів, що переміщуються в зоровому каналі, які поки погано вивчені. Сигнали, що надходять на асоціативні входи, які прийшли від датчиків контролю м'язів і всіх інших почуттів, дозволяють мозку спиратися на перехресні асоціації, накопичені за роки прожитих років. За рахунок зворотного зв'язку в мозку, процес повторюється знову і знову і включає в себе датчик (очей), який механічно управляє освітленням за допомогою райдужної оболонки і наполягає прийом на поверхні сітківки.

Однак, в системі комп'ютерного зору все що отримує комп'ютер це сітку з числами від камери або з диска. Здебільшого, немає ні вбудованого розпізнавання образів, що не автоматичного управління фокусом і діафрагмою, ні перехресних асоціацій з багаторічним досвідом роботи. Здебільшого, системи комп'ютерного зору все ще досить примітивні.

Наприклад на малюнку людина бачить бічне дзеркало з боку водія. Комп'ютер же "бачить" тільки сітку з числами. Будь-який номер з цієї сітки має досить таки велику шумову складову і сам по собі дає нам мало інформації. Наше завдання зводиться до перетворення сітки чисел в результат: "бічне дзеркало".

Насправді, проблему, про яку йдеться мова, формально неможливо вирішити. Маючи тільки двовимірне подання тривимірного світу, неможливо однозначно відновити 3D сигнал. Формально, такого роду некоректна завдання не має однозначного або остаточного рішення. Двовимірне зображення можна уявити будь-яким з нескінченних поєднань 3D-сцен, навіть якщо дані були досконалі. До того ж, як згадувалося раніше, дані схильні до шуму і спотворень. Такого роду ушкодження пов'язані зі змінами в світі (погода, освітлення, відображення, переміщення), недоліками в об'єктиві і механічних установках, кінцевим часом інтеграції на датчику (розмиття при русі), електричними шумами в датчику або іншу електроніку, артефактами стиснення після захоплення зображення.

При проектуванні практичної системи, додаткові знання контексту часто можуть бути використані, щоб обійти обмеження, що накладаються візуальними датчиками. Розглянемо наступний приклад: мобільний робот повинен знайти і підібрати степлер. Робот може використовувати наступний факт: стіл - це об'єкт, який знаходиться всередині офісу і найчастіше степлер можна знайти на столах. Це дає неявну послання про розмір; степлер повинен поміститися на столі. Це також допомагає усунути помилкові "розпізнавання" розташування степлера в неможливих місцях (наприклад, на стелі або вікні). В інших випадках, коли доводиться витягувати зображення степлера з бази даних, може виявитися так, що на цих зображеннях степлер буде величезних розмірів або ж мати незвичайну форму, що потягне за собою виключення цих зображень з розгляду.

Контекстна інформація також може бути змодельована явно методами машинного навчання. Приховані змінні, такі як розмір, орієнтація тяжіння і т.д. можливо співвіднести з їх значеннями в навальному середовищі. В альтернативі, можна спробувати виміряти приховані змінні зсуву за допомогою додаткових датчиків. Використовуючи датчик глибини, можна точно виміряти розмір об'єкта.

Наступна проблема комп'ютерного зору - це шум. Як правило, ми маємо справу з шумом, коли використовуємо статистичні методи. Наприклад, може бути неможливо виявити контур в зображенні простим порівнянням сусідніх точок. Але якщо ми зберемо статистику по локальній області, то задача виявлення контуру стає легше. Контур повинен з'явитися у вигляді рядка безпосередніх відповідей по локальній області, кожен з яких орієнтований відповідно до його сусідами. Так само можливо компенсувати шум, приймаючи статистичні дані протягом довгого часу. Проте інші методи обліку шуму або спотворення шляхом створення чітких моделей дозволяють розпізнавати безпосередньо з наявних даних (характеристики об'єктиву).

Ті чи інші дії або рішення, що приймаються в комп'ютерному зорі, зроблені на основі даних отриманих з камери в контексті конкретної мети або завдання. Можна видалити шум або пошкодження в зображенні так, що система безпеки буде видавати попередження, якщо хтось спробує залізти на паркан або ж система моніторингу буде підраховувати, скільки людей переступило через певну область в парку атракціонів. ПЗ з використанням комп'ютерного зору для роботів, які пересуваються по офісу будуть використовувати інші стратегії, ніж стаціонарні камери відеоспостереження, оскільки ці дві системи мають суттєво різні контексти і завдання. Як правило, контекст задач комп'ютерного зору обмежений і тому, розраховуючи на ці обмеження, завдання можна спростити і, як наслідок, кінцеве рішення буде більш надійним.

Висновки. Комп'ютерний зір має широку сферу застосування в різних галузях. Нажаль сучасні методи аналізу та розпізнавання зображень дають змогу знайти лише часткове рішення відповідно до поставлених задач. Перспективою даної теми можна віднести застосування нейронних мереж, які дадуть можливість отримувати більш точні результати при розпізнаванні зображень.

Список використаних джерел:

1. A practical introduction to computer vision. Kenneth Dawson-Howe, 2014
2. Learning OpenCV. Gary Bradski and Adrian Kaehler, 2008
3. <http://www.racurs.ru/>
4. <http://wiki.technicalvision.ru/>