

УДК 681.007.05

О.В. Ліпанов, М.В. Фесенко

Харківський національний університет радіоелектроніки, Харків

АНАЛІЗ МЕТОДІВ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ В СИСТЕМАХ АНАЛІЗУ ВІЗУАЛЬНОЇ ІНФОРМАЦІЇ

У статті розглядаються наступні підходи до розпізнавання об'єктів у відеопотоці: визначення особливостей кадру, визначення точок інтересу, кореляційний аналіз, методи машинного навчання. В результаті їх аналізу був проведений вибір методу для реалізації бібліотеки аналізу відеопотоку. За допомогою такої бібліотеки можуть бути побудовані різноманітні системи аналізу відеопотоку, які можуть бути використані для відеоспостереження, біометричної ідентифікації, індексації та пошуку. Робота бібліотеки була перевірена для різних умов якравості, контрастності та рівня впливу шумів.

Ключові слова: розпізнавання об'єктів, відеопотік, точка інтересу, детектор, кореляція, машинне навчання, бібліотека аналізу відеопотоку.

Вступ

Побудова систем комп'ютерного зору та аналізу відеопотоку – актуальна задача сьогодення. Рішення у цій області знаходять широке застосування у найрізноманітніших сферах життя людини: відстеження появи машин на стоянці, забутого багажу в приміщеннях аеропортів та вокзалів, забезпечення захисту від несанкціонованого проникнення на певні об'єкти, розпізнавання облич людей, автомобільних номерів, написів, тощо.

Розпізнавання об'єктів – одна з найважливіших задач в області комп'ютерного зору. Задача, з якою людина легко вправляється кожного дня, є неймовірно складною для обчислювальної техніки. Пошук відповідності між двома зображеннями деякої сцени чи об'єкту є складовою частиною багатьох систем аналізу візуальної інформації, зокрема відеопотоків. Крім того розпізнавання об'єктів у відеопотоці ускладнюється наступними факторами [9]:

- втрата інформації в результаті проектування трьохмірного світу на двомірне зображення;
- зашумленість кадрів;
- часткове або повне перекриття об'єктів елементами сцени або іншими об'єктами;
- зміни у освітленні сцени;
- аналіз відеопотоків від різних камер;
- вимоги до аналізу в режимі реального часу.

При розв'язанні задачі розпізнавання об'єкту у відеопотоці однією з основних проблем є пошук компромісу між якістю розпізнавання та його швидкістю.

Системи аналізу відеопотоку часто реалізуються у вигляді цілісного програмно-апаратного комплексу. Наприклад, компанія DVTel пропонує IP камери та кодери іoimage, що дозволяють проводити аналіз відеопотоку в режимі реального часу. Аналіз відеопотоку проводиться безпосередньо у самому пристрої «на льоту», що знижує технічні вимоги до сервера, зменшує навантаження на мережу, підви-

щує ефективність аналізу. Інтелектуальне відеоспостереження дозволяє відстежувати такі події як вторгнення на певну територію, забуті речі, несанкціонована зупинка автомобіля, зникнення предмету з поля зору, може проводитися автоматичне супроводження об'єкта за допомогою рухомої камери. Також у модулях, що пропонуються, є вбудовані механізми корекції та покращення якості зображення для зменшення кількості помилкових спрацьовувань. Також відбувається сповіщення про відсутність сигналу від камери або спроби зменшити її огляд.

Крім того відбувається постійне розширення областей застосування систем відеоаналізу та засобів, які вони використовують. Наприклад, компанія DVTel представила додаток для мобільних телефонів TruWitness, який перетворює мобільний телефон у повноцінну IP-камеру відео-спостереження, яка направляє відеопотік у спеціальну систему управління відео.

Цікавим застосування методів розпізнавання об'єктів є біометричні системи. BioID Web Services – це онлайн сервіс, що пропонує розробникам будь-яких сервісів або додатків потужну біометричну технологію. Прикладом такого сервісу є система MyBioID, що пропонує простий та зручний спосіб авторизації користувача на різноманітних інтернет ресурсах. Після реєстрації у системі користувач має надати кілька своїх фотографій та зразків звучання свого голосу у якості навчальної вибірки. При спробі авторизації на сайті, що приймає MyBioID, користувач перенаправляється на сторінку MyBioID, що призначена для проведення процесу розпізнавання. Користувач надає фотографію та звукозапис свого голосу, отриманий за допомогою веб-камери, і якщо система його впізнала – то перенаправляється на початковий ресурс. Також розв'язання задач в області аналізу відеопотоку можуть бути використані для автоматичної індексації відео, створення автоматичних систем управління механізмами, контролю за безпекою на об'єктах, контролю за дорожньо-транспортним рухом, тощо.

1. Методи розпізнавання образів

У роботі вивчаються методи розпізнавання образів з метою вибору методу для використання при реалізації бібліотеки аналізу відеопотоку.

1.1. Виділення особливостей зображень. Зображення може бути описане глобальним дескриптором, що по суті є вектором ознак. При цьому кожна точка зображення робить певний внесок у значення дескриптора.

Однією з найважливіших характеристик зображення для людини є колір. Для опису кольорових особливостей зображення часто використовуються гістограми розподілу кольорів. При цьому колірний простір розбивають на інтервали, для кожного з яких обчислюється частота потрапляння пікселів. Основною проблемою побудови такого представлення зображень і механізмів розпізнавання на його основі є те, що важко побудувати розбиття, при якому кольори з одного інтервали сприймалися людиною як однакові, а з різних – як різні. Можна розбити зображення на блоки фіксованої величини та обчислювати гістограму для кожного блоку окремо. Форма гістограми розподілу частот дає багато інформації про властивості зображення. Наприклад, вузька гістограма вказує на низький контраст.

Крім цього колір можна представити через колірні моменти. При такому представленні пропонується замість повного кольорового розподілу зберігати домінантні особливості. У [6] пропонується розглядати розподіл кольору зображення як розподіл вірогідностей і використовувати три перших моменти розподілу для унікального опису кожного кольорового каналу. Авторами цього методу було показано, що він перевершує гістограмний метод як у швидкості обробки, так і у стійкості результатів.

Також при описі зображення велике значення має текстура, за допомогою якої задається структура об'єктів. Зазвичай під структурою розуміють повторення певних базових примітивів, що мають різну орієнтацію у просторі. Прикладами натуральних текстур є зображення крони дерев, кущів.

При статистичному описанні структурні характеристики зображення можуть бути представлені за допомогою матриці суміжності. Матриця суміжності показує як часто кожна пара різних рівнів сірого кольору знаходиться на певній відстані d вздовж певного напрямку, що задається кутим α . За допомогою цієї матриці можна визначити такі показники як контраст, ступінь однорідності, ентропія.

Структурний опис текстур опирається на те, що текстура складається з регулярно або майже регулярно повторення елементів. Тому структуру можна описати за допомогою таких елементів та правил їх розміщення. Для опису текстур широко застосовуються ознаки Тамура, які були виділені в результаті психологічних експериментів [4]. До них відносять-

ся контраст, зернистість, спрямованість, лінійність, регулярність, грубість.

У [1] був запропонований алгоритм розпізнавання текстур за допомогою моментних ознак так нейронної мережі. Початкове зображення розбивається на області, і для кожної області обчислюються центральні моменти першого порядку для кожного каналу. Ці значення формують вектор сигналів, що подається на вхід нейронної мережі. Оскільки моментні ознаки не дозволяють описати взаємне розташування елементів, то для подолання цього недоліку була введена однозначна відповідність між номером підобласті зображення та номером синаптичного входу.

Також текстуру можна подати у вигляді фракталу і застосовувати методи фрактальної геометрії для її опису. Це пов'язано з тим, що текстурам притаманна властивість самоподібності, яка є характерною для фракталів. При такому поданні текстури для аналізу зображення у різних масштабах використовується вейвлетний аналіз.

1.2. Визначення точок інтересу. Точка інтересу – така точка сцени, зображення околу якої можна відрізнити від зображень околів всіх інших точок сцени. Важливою властивістю точки інтересу об'єкта є її інваріантність до змін освітленості або точки спостереження камери. Для кожного об'єкта на зображенні можна знайти його точки інтересу і описувати ними сам об'єкт. Отримавши опис об'єкта з певного тренувального зображення його можна використовувати для подальшого пошуку цього об'єкта на інших зображеннях. Для надійності такого алгоритму розпізнавання важливим є те, що точки інтересу об'єкта мають визначитися незалежно від масштабу зображення, впливу шумів та зміни яскравості. Крім того відносне розташування між точками інтересу об'єкта має залишатися незмінним. Такі точки зазвичай знаходяться на межах об'єктів.

При стеженні за об'єктами у відеопотоці з використанням точок інтересу спочатку необхідно визначити точки інтересу у першому кадрі відеопослідовності. Потім необхідно визначити якість отриманих точок та відібрати лише ті, якість яких задовольняє певному порогу. Після цього для кожного наступного кадру визначається нове положення точок інтересу, оновлюється інформація про їх якість та проводиться перевірка на задоволення порогу якості. Якщо кількість точок інтересу стає меншою за певний поріг, то визначаються нові точки з поточного кадру і додаються до існуючих точок.

За останні десятиріччя було розроблено величезну кількість різних детекторів точкових особливостей зображень. Всі вони використовують різні підходи до формування функції оцінки пікселів для знаходження особливостей. У нашому випадку, в системах стеження за особливостями, визначальним параметром якості детектора буде якість подальшого відстеження.

Одним з найперших детекторів особливих точок зображення був запропонований оператор Моравека. Він визначає зміни інтенсивності зображення у блоках 4x4 пікселі у горизонтальному, вертикальному, діагональному та протидіагональному напрямках та обирає найменше з чотирьох відхилень як характерне значення для вікна. Точка позначається як точка інтересу, якщо відхилення інтенсивності для неї є локальним максимумом у блоці 12x12. Цей оператор є одним із найстаріших і дозволяє визначити кути зображення. Але зазвичай об'єкти, за якими відбувається спостереження є доволі складними, і просте визначення його кутів не є достатнім для опису образу та його подальшої ідентифікації.

Детектор Гарріса розглядає зображення як функцію двох змінних $I(x,y)$. Для кожного пікселя зображення обчислюється значення особливої функції відгуку кута (corner response function), що оцінює ступінь подібності зображення околиці точки до кута. Для цього спочатку розраховується матриця Гарріса, задана такою формулою:

$$M = \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \\ \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) & \left(\frac{\partial I}{\partial y}\right)^2 \end{bmatrix}, \quad (1)$$

де $I(x,y)$ – яскравість зображення в точці (x,y) .

Функція відгуку кута задається формулою:

$$R = \det M - k(\text{tr } M)^2, \quad (2)$$

де k – параметр, що зазвичай дорівнює 0,04.

Точками інтересу вважаються точки зображення, що відповідають локальним максимумам цієї функції. У багатьох випадках детектор Гарріса знаходить занадто велику кількість кутів, що ускладнює їх подальше відстеження. Для уникнення цієї проблеми вводять обмеження на мінімальну відстань між знайденими особливостями, і всі зайві відкидаються.

Детектор SIFT (Scale-invariant feature transform), запропонований Девідом Лоу у [5], дозволяє ефективно ідентифікувати об'єкти навіть на зашумлених зображеннях і при частковому перекритті об'єктів, тому що дескриптор особливостей SIFT інваріантний до рівномірного масштабування та зміни положення у просторі, частково інваріантний до операції афінного скосу та змін освітленості. SIFT дозволяє визначити «точки інтересу» та сформувати дескриптори околиць цих точок. На першому кроці алгоритму SIFT конструюється масштабний простір шляхом застосування Гаусового фільтра для зображення у різних масштабах. З отриманих зображень отримують різницю між гаусіанами. Точки-кандидати обираються в точках максимуму та мінімуму зображень з різницями гаусіан. На наступному кроці оновлюється положення кожної точки-кандидата шляхом інтреполяції кольору по сусіднім пікселям. Далі відкидаються точки-кандидати, що недостатньо контрастні або зна-

ходяться на краях зображення. Для точок, що залишилися, визначається напрям за градієнтом у невеликому околі. Цей метод знаходить значно більшу кількість точок інтересу порівняно з іншими детекторами, і емпірично розробники цього алгоритму показали, що він стійкіший до змін у зображенні.

Найбільшим недоліком SIFT є те, що він працює лише з зображеннями у градаціях сірого, втрачаючи при цьому важливу інформацію про зображення, яку несе в собі колір. У [8] була запропонована модифікація алгоритму SIFT, що дозволяє враховувати кольорові особливості зображення. На першому кроці алгоритму CGSIFT проводиться квантування кольорового зображення з метою представлення його невеликою кількістю кольорів. На другому кроці застосовується детектор SIFT для виділення особливостей у квантованому зображенні.

Дескриптор SURF (Speeded Up Robust Feature) був запропонований у [7]. Він схожий на SIFT, але у кілька разів швидший за нього. Цей дескриптор інваріантний до масштабу та повороту об'єкта. Пошук точок інтересу у алгоритмі SURF базується на обчисленні гессіана. Матриця Гессе – це матриця, утворена другими частковими похідними функції та має вигляд, заданий формулою:

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}. \quad (3)$$

Гессіаном називається детермінант цієї матриці. Якщо розглядати зображення як функцію, то у точках максимальної зміни яскравості гессіан цієї функції досягає екстремумів. Після знаходження ключових точок за допомогою гессіана у алгоритмі SURF формуються їх дескриптори. Кожен дескриптор є набором чисел, що відображає зміни градієнту функції поблизу ключової точки. Для визначення матриці Гессе використовується метод бінаризованої апроксимації лапласіану гаусіан.

1.3. Кореляційне розпізнавання зображень.

Кореляція – статистичний взаємозв'язок декількох випадкових величин, при якому зміни значень однієї з величин супроводжується систематичними змінами інших величин. Метою кореляційного аналізу є виявлення кореляційного зв'язку. Для того, щоб чисельно охарактеризувати взаємозв'язок між змінними, вводиться коефіцієнт кореляції.

Кореляційне розпізнавання зображень характеризується високою надійністю та надійною працездатністю у широкому діапазоні зовнішніх умов [2, 3]. Кореляційними вважаються підходи, що базуються на побудові міри подібності $\xi(B_0, B)$ між зо-

браженням B та еталоном B_0 . Обсяг обчислень при використанні кореляційних алгоритмів залежить від області визначення функції $\xi(B_0, B)$. Кореляційні методи широко використовуються при виявленні та розпізнаванні зображень у системах навігації, стеження та промислових роботах.

1.4. Методи машинного навчання. Для розпізнавання зображень широко застосовуються штучні нейронні мережі (наприклад, мережі Кохонена, Хопфілда). Нейромережеві методи забезпечують швидке та надійне розпізнавання зображень, але з ними виникають проблеми при розпізнаванні тривимірних об'єктів, пов'язані з просторовим поворотом та змінами освітленості. Також є приклади використання методів генетичного програмування для обробки зображень (здебільшого для військових потреб). Зокрема еволюційним шляхом були отримані алгоритми пошуку танків на інфрачервоних зображеннях, визначення автомобілів по фотографіях повітряної розвідки, кораблів з зображень від радара, розташованого на супутнику. Також генетичне програмування використовувалось для попередньої обробки зображень людських облич для пошуку областей інтересу, необхідних для подальшого аналізу.

2. Бібліотека аналізу відеопотоку

При реалізації бібліотеки аналізу відеопотоку необхідно було забезпечити можливість розпізнавання об'єктів у відеопотоці. Після аналізу методів розпізнавання було вирішено реалізувати цю можливість на основі методів визначення точок інтересу, зокрема детектору SIFT.

Розроблена бібліотека дозволяє виконувати захоплення відеопотоку з файлу або відеокамери, задавати об'єкти для розпізнавання за допомогою їх зображень. Якщо в процесі аналізу відеопотоку було виявлено один з заданих об'єктів, то виконується оповіщення про цю подію через інтерфейс бібліотеки. Розроблена бібліотека може бути використана як для побудови настільних додатків, так і для використання у різноманітних веб-сервісах.

Якість роботи розробленої бібліотеки перевірялася при різних умовах яскравості, контрастності та рівня впливу шумів. Результати аналізу наведені на рис. 1 – 3. Також при перевірці роботи бібліотеки було виявлено, що розмір області кадру, що аналізується, та кількість об'єктів, що розпізнаються, помітно сповільнюють роботу системи.

Висновки

В ході досліджень, приведених у статті, були вивчені методи розпізнавання образів. В результаті досліджень був обраний метод визначення точок інтересу, зокрема детектор

SIFT. На основі цього методу була реалізована бібліотека аналізу відеопотоку, яка може бути використана для побудови різноманітних систем аналізу візуальної інформації. Якість роботи розробленої бібліотеки була перевірена для різних умов яскравості, контрастності та рівня зашумленості кадрів. Подальші дослідження будуть направлені на розробку алгоритмів шумозаглушення, які дозволять покращити стійкість розпізнавання об'єктів у відеопотоці, та застосування методів паралельної обробки для забезпечення можливості виконання аналізу в режимі реального часу.

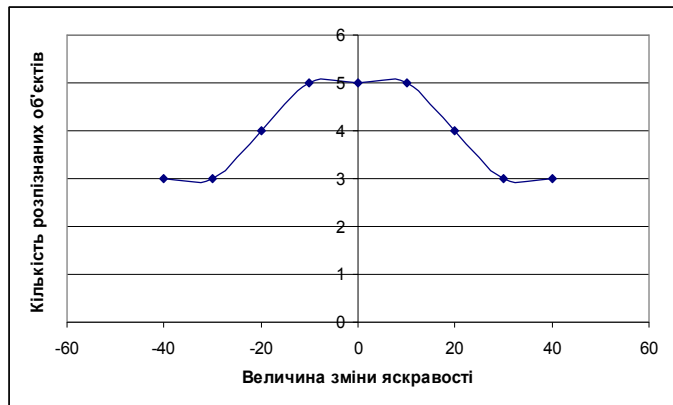


Рис. 1. Діаграма залежності кількості вдало розпізнаних об'єктів від ступеня зміни яскравості зображення

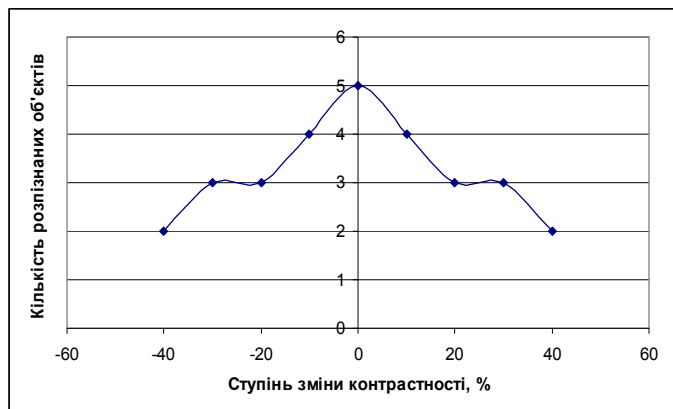


Рис. 2. Діаграма залежності кількості вдало розпізнаних об'єктів від ступеня зміни контрастності зображення

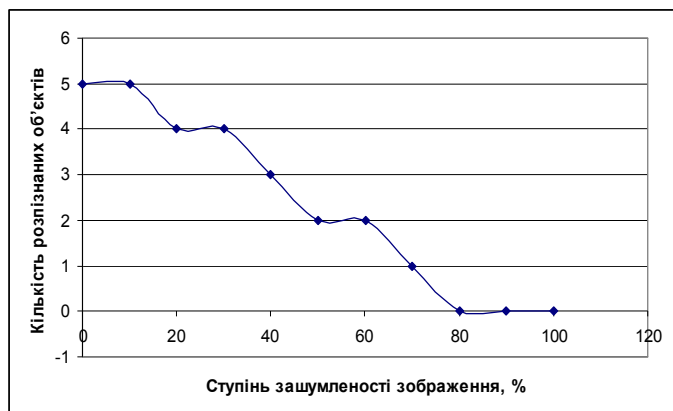


Рис. 3. Діаграма залежності кількості вдало розпізнаних об'єктів від рівня впливу шумів

Список літератури

1. Липанов А.В. Алгоритм распознавания изображений текстур с использованием моментных признаков и методов нейронных сетей / А.В. Липанов, А.Ю. Михайлов // Системи обробки інформації: зб. наук. праць. – Х.: ХУПС, 2007. – Вып. 3 (61). – С. 49-52.
2. Методы корреляционного обнаружения объектов / А.В. Гуренко, В.В. Ляшенко, В.П. Машиталир, Е.П. Пуятин. – Х.: АО "БизнесИнформ", 1996. – 112 с.
3. Пуятин Е.П. Нормализация и распознавание изображений [Электронный ресурс] / Е.П. Пуятин. – Режим доступа до ресурсу: <http://sumschool.sumdu.edu.ua/is-02/rus/lectures/pyyatin/pyyatin.htm>.
4. CBIR: Texture Features [Электронный ресурс]. – Режим доступа до ресурсу: <http://www.cs.auckland.ac.nz/compsci708slc/lectures/Glect.html/topic4c708FSC.htm>.
5. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints / D.G. Lowe // International Journal of Computer Vision. – 2004. – Vol. 60, № 2. – P. 91-110.
6. Stricker M. Similarity of color images / M. Stricker, M. Orengo // SPIE Conf. – 1995. – P. 381-392.
7. SURF: Speeded Up Robust Features / H. Bay, A. Ess, T. Tuytelaars, L. Van Gool // Computer Vision and Image Understanding. – 2008. – Vol. 110, № 3. – P. 346-359.
8. Unsupervised Video Shot Detection Using Clustering Ensemble with a Color Global Scale-Invariant Feature Transform Descriptor [Электронный ресурс] / Y. Chang, D.J. Lee, Y. Hong, J. Archibald. – Режим доступа до ресурсу: http://jivp.eurasipjournals.com/content/pdf/1687-5281-200886_0743.pdf.
9. Yilmaz A. Object Tracking: A Survey / A. Yilmaz, O. Javed, M. Shah // ACM Computing Surveys. – 2006. – Vol. 38, № 4. – P. 13-45.

Надійшла до редколегії 12.03.2012

Рецензент: д-р техн. наук, проф. Є.П. Пуятин, Харківський національний університет радіоелектроніки, Харків.

АНАЛИЗ МЕТОДОВ РАСПОЗНАВАНИЯ ОБЪЕКТОВ В СИСТЕМАХ АНАЛИЗА ВИЗУАЛЬНОЙ ИНФОРМАЦИИ

А.В. Липанов, М.В. Фесенко

В статье рассматриваются следующие подходы к распознаванию объектов в видеопотоке: определение особенностей кадра, определения точек интереса, корреляционный анализ, методы машинной учебы. В результате их анализа был проведен выбор метода для реализации библиотеки анализа видеопотока. С помощью такой библиотеки могут быть построены разнообразные системы анализа видеопотока, которые могут быть использованы для видеонаблюдения, биометрической идентификации, индексации и поиска. Работа библиотеки была проверена для разных условий яркости, контрастности и уровня влияния шумов.

Ключевые слова: распознавание объектов, видеопоток, точка интереса, детектор, корреляция, машинная учеба, библиотека анализа видеопотока.

ANALYSIS OF METHODS FOR OBJECT RECOGNITION IN SYSTEMS FOR VISUAL INFORMATION ANALYSIS

O.V. Lipanov, M.V. Fesenko

This paper contains overview of following approaches for object recognition in videostream: feature detection, points of interest detection, correlation analysis, machine learning. This overview helped to select method to build the library for videostream analysis. Using this library different systems for videostream analysis can be built. These systems can be used for video-surveillance, biometric identification, indexation and search. The library was tested for different brightness and contrast changes, for different noise levels.

Keywords: objects recognition, videostream, point of insterest, detector, corelation, machine learning, library for videostream analysis.