

УДК 621.377.037.3

Ю.А. Харина

МЕТОДИКА ОЦІНЮВАННЯ СИСТЕМ ВІДЕОАНАЛІТИКИ НА ПРИКЛАДІ БІБЛІОТЕКИ ОБРАЗІВ I-LIDS¹

У статті описується методика оцінювання систем відеоаналітики (далі – ВА) за допомогою бібліотеки образів на прикладі британської бібліотеки i-LIDS. Піднімається проблематика оцінювання таких систем, описується структура бібліотеки i-LIDS, її склад, а також наводяться сценарії, за якими, власне, оцінюються системи ВА, що тестуються. Розглядається п'ять рівнів сертифікації систем залежно від їх проходження чи непроходження тих чи інших сценаріїв, а також три різні набори даних у складі бібліотеки для цілей тренування та тестування систем. Наводяться ілюстровані приклади помилок спрацювання систем.

Ключові слова: відеоаналітика, i-LIDS, бібліотека образів, відхилення спрацювання.

В статье описывается методика оценивания систем видеоналитики (далее – ВА) с помощью библиотеки образов на примере британской библиотеки образов i-LIDS. Поднимается проблематика оценивания таких систем, описывается структура библиотеки i-LIDS, ее состав, а также приводятся сценарии, по которым, собственно, и оцениваются тестируемые системы ВА. Рассматриваются пять уровней сертификации систем в зависимости от их прохождения или непрохождения тех или иных сценариев, а также существование в составе библиотеки трех разных наборов данных для целей тренировки и тестирования систем. Приводятся иллюстрированные примеры ошибок срабатывания систем.

Ключевые слова: видеоналитика, i-LIDS, библиотека образов, отклонение срабатывания.

This paper discusses a method for evaluation of videoanalytic (VA) systems using British image library i-LIDS. The issues of evaluation of such systems are raised, the libraries' structure and composition are described, and scenarios for which evaluated VA-systems have been tested are stated. Five levels of certification systems, depending on their passage or failure of certain scenarios, as well as the existence in the library of three different sets of data for training and testing systems are considered. Illustrated examples of recall biases of systems are given.

Keywords: videoanalytics, i-LIDS, image library, recall bias.

На сьогодні завантаження операторів відеоспостереження постійно збільшується, і схеми відеоспостереження у виявленні злочинів не можуть повністю реалізувати свій потенціал. Системи відеоаналітики (далі – ВА) пропонують потенційне вирішення проблеми перевантаження операторів за допомогою автоматичного привертання уваги операторів у реальному часі до цікавих подій або послідовностей, що визначаються як цікаві, для прискорення післяподійного аналізу.

¹ Закінчення в наступному номері.

На разі в Україні немає впроваджених систем відеоаналітики в структурах МВС. Однак подібні системи вже давно використовуються у розвинутих країнах, і, відповідно, є сенс вивчати їх досвід з метою розпочати впровадження аналогічних систем у МВС та Національній поліції України на вже практично перевіреному рівні.

Сучасні рішення досить далеко пішли від перших прототипів, але галузь ще довго переслідували невиправдані очікування від появи тих чи інших систем. Одна з причин цього полягала у відсутності до недавнього часу зрозумілих критеріїв оцінки продукційності аналітичних систем, їх користувачам доводилось так чи інакше купувати “кота в мішку”. Так, в Україні досі немає достовірного механізму тестування систем відеоаналітики. Але за кордоном з подібною проблемою стикнулися досить давно, і тому там уже існують шляхи її вирішення. Зокрема, одне з найкращих рішень винайшли поліцейські з Департаменту наукових розробок МВС Великої Британії (CAST), які провели дуже багато роботи для систематизації практик відеоспостереження. У співробітництві з Центром захисту національної інфраструктури (CPNI) вони розробили окрему методику i-LIDS, що призначена для сертифікації систем охоронної відеоаналітики.

Абревіатура i-LIDS розшифровується як “візуальна бібліотека для систем інтелектуального детектування”, а однойменний сертифікат вручається продуктам, що задовольнили суворі критерії тесту. Експлуатаційні вимоги, на відповідність яким перевіряються рішення, були вироблені в рамках технічної специфікації на системи для державного користування.

На разі в Департаменті наукових розробок МВС Великої Британії (CAST) регулярно проводяться оцінки продукційності аналітичних систем за методикою i-LIDS. Випробування проходять за допомогою декількох сценаріїв, і система, що випробовується, “обкатується” у двох режимах – повідомлення оператору та автоматичний запис подій, що детектуються. В першому випадку від системи вимагається продемонструвати продукційність у виявленні підозрілих подій під час моніторингу оператором у реальному часі. У другому режимі переглядання записаних фрагментів відбувається тільки постфактум.

Після впровадження ініціативи i-LIDS МВС Великої Британії оцінює системи ВА для сценаріїв виявлення подій (наприклад, неправильно припаркованих транспортних засобів) та сценаріїв відстеження об’єктів (наприклад, людей в аеропорту), що є ключовими для вимог уряду Великої Британії.

Набори даних i-LIDS широко розглядаються як найбільш комплексні у своєму роді, та як такі, що досягли суттєвого визнання з моменту їх запуску в 2006 року. Попит з боку виробників був вкрай обнадійливим, а програма оцінки згодом показала, що відповідати вимогам уряду може велика кількість систем. Відеобібліотека i-LIDS забезпечує еталон для полегшення розробки та вибору систем ВА, які відповідають вимогам уряду, і вона містить відео з відеоспостереження, першопочатково в основі своїй засноване на шістьох різних сценаріях:

Виявлення подій:

- виявлення залишеного багажу;
- виявлення залишених транспортних засобів;
- спостереження за входами;
- моніторинг забороненої зони;
- нові технології (тепловізійне спостереження та інфрачервоне підсвічування).

Відстеження об'єктів:

- Стеження за допомогою відстеження багатьма камерами [1].

Для кожного сценарію були зроблені по три об'ємних комплекти тестових відеоданих. Довжина кожного такого “тестового архіву” складає близько 24 годин, і в ньому зібрано відео при всіх можливих умовах погоди та освітлення. Комплект відеоданих для сценарію з використанням декількох камер нараховує в сумі більш ніж 50 годин реального експлуатаційного відео.

Ці типові сценарії роботи бібліотеки i-LIDS є стандартом для розробки та перевірки систем, які можуть використовувати тепловізійне спостереження та інфрачервоне підсвічування – так звані сценарії “нових технологій”.

У межах кожного сценарію виявлення подій визначені певні “сигнальні події” – наприклад, наявність припаркованого транспортного засобу в певній зоні більш ніж 60 секунд. Системи ВА повинні подавати сигнал тривоги, якщо у відеоматеріалі виявляються будь-які з цих подій, з мінімальною кількістю хибних спрацювань.

За сценаріями відстеження об'єктів i-LIDS персоналії або “Цілі”, визначені системою відеоспостереження, передають до системи відстеження. Системи відстеження об'єктів потрібні для відстеження Цілі за допомогою мережі камер, доки Ціль більше не спостерігається або не задана нова Ціль.

Відео за кожним сценарієм розділяється на три набори даних, два з яких зроблені доступними для виробників ВА та академій, які допомагають розробляти підходящі системи. Той набір даних, що залишився, отримує CAST та використовує його для перевірки продукційності систем.

Технічно кожний комплект даних складається з двох чи трьох сцен, які розбиті на ролики довжиною від 30 до 60 хвилин. Для всіх сценаріїв визначені критерії оцінки подій, що виділяються: наприклад, автомобіль повинен бути присутнім у зазначеній зоні на парковці протягом не менш ніж 60 секунд. Для підготовки свого продукту до тестування за методикою i-LIDS виробник може скористатися демонстраційними наборами відеоданих для кожного зі сценаріїв. Вони надаються на платній основі.

Системи, які у випробуванні CAST демонструють значний рівень продуктивності, згодом описують у каталозі охоронного обладнання, що допущене до використання в критичній національній інфраструктурі (Critical National Infrastructure, CNI) та урядом для закупівель охоронного обладнання.

За результатами тестування аналітичний продукт може бути віднесений до однієї з п'яти категорій. Найнижча категорія не підтверджує наявності у продукта аналітичних функцій, а найвища припускає здатність системи працювати в автономному режимі, використовуючи відеоаналітику як основний, а не допоміжний засіб детектування. Тільки на такі продукти за умовами сертифікації можна нанести логотип i-LIDS. Вимоги для розміщення в каталозі охоронного обладнання Центру захисту національної інфраструктури менш суворі – в цей перелік допускаються пакети відеоаналітики двох вищих категорій.

Слід відмітити, що сертифікація та категоризація виконуються для кожного з п'яти сценаріїв окремо: продукт, наприклад, може мати сертифікат з виявлення автомобілів на парковці, але не відповідати вимогам за хибними спрацюваннями у “стерильній зоні”. На логотипі зазначений тільки той тип тесту, який продукт успішно пройшов.

Набори даних виявлення подій i-LIDS розповсюджуються на зовнішніх жорстких дисках USB2/Firewire об'ємом 500 Гб або 1 Тб. Набори даних відстеження групою камер (ВГК) розповсюджуються на зовнішніх жорстких дисках USB2/Firewire/e-SATA об'ємом 1 Тб. Можуть бути використані формати файлових систем Windows NTFS або Apple HFS+. Для користувачів Linux рекомендовано NTFS.

Відео i-LIDS рендериться до крос-платформеного файлового формату Quicktime mJPEG [1].

1. Принципи i-LIDS

1.1. Сценарії

Бібліотека i-LIDS базується на шістьох “сценаріях”, які мають вирішальне значення для вимог уряду:

- Виявлення залишеного багажу з подачею сигналу тривоги у випадку наявності сумок без нагляду на платформі на станції підземки.
- Виявлення припаркованих транспортних засобів з подачею сигналу тривоги у випадку наявності припаркованих транспортних засобів у міських умовах.
- Спостереження за входами з подачею сигналу тривоги у випадку наявності людей, які входять та виходять через дверні проходи, що моніторяться.
- Моніторинг забороненої зони з подачею сигналу тривоги у випадку наявності людей в забороненій зоні між двома парканами безпеки.
- Відстеження групою камер з цільовою подією, що складається з наявності людей (“Цілей”), що проходять через мережу камер відеоспостереження.
- “Нові технології” – відеоматеріал створюється за допомогою охолодженого та неохолодженого тепловізійного спостереження та інфрачервоного підсвічування з подачею тривоги, що містить атаки з боку пішоходів на великій площі і вздовж та атаки по воді та з боку пірсу.

1.2. Тренувальний, тестувальний та оціночний набори даних

Відповідно до академічної конвенції відеоматеріал, що виходить із кожного сценарію i-LIDS, розподіляється на три еквівалентні набори даних:

- Публічний “тренувальний” набір даних, який може бути використаний для розробки ефективних розпізнавальних алгоритмів.
- Публічний “тестувальний” набір даних, який може бути використаний для перевірки продуктивності цих алгоритмів.
- Приватний “оціночний” набір даних, що підтримується CAST та використовується для сертифікації продуктивності систем, представлених до регулярних випробувань.

1.3. Послідовності виявлення подій

Кожний набір даних виявлення подій i-LIDS порівнює приблизно 24 години “послідовностей”, записаних при різних умовах; годині дня, погоді, рівні фоновій активності тощо. Набір даних “нових технологій” для нічного спостереження містить приблизно 48 годин послідовностей для кожного способу зйомки.

Деякі з послідовностей доповнені подіями тривоги, що відбулись у відеоматеріалі. Залишкові “нетривожні” послідовності містять тільки фоновий рівень подій тривоги.

Зазвичай, кожна послідовність рендериться як поодинокий файл Quicktime в бібліотеці i-LIDS. У публічних тренувальних наборах даних, однак, “тривожні” послідовності, які містять багато подій тривоги, редаговані в “кліпи” таким чином,

щоб кожна окрема подія тривоги, що виникла, рендерилась до окремого файлу. Це пришвидшує доступ для користувачів тренувального набору даних до відеоматеріалу конкретної події тривоги. Угода про файлові імена є такою, що суміжні кліпи, які можуть бути просто ідентифіковані та суміщені послідовно, вимагають більше до- та післяподійного відеоматеріалу, наприклад, для вивчаючих алгоритмів.

Кожна тривожна послідовність містить якнайменше три хвилини відеоматеріалу до першої запланованої події тривоги так, щоб допомогти системам, які вивчаються, в пристосуванні до умов послідовності [1].

1.4. Послідовності відстеження об'єктів

Публічні набори даних відстеження групою камер i-LIDS порівнюють приблизно десять годин “сценарію” з тренувальним та тестувальним наборами. Цей сценарій відстеження групою камер (ВГК) формується з мережі, що складається з п'яти камер відеоспостереження, даючи загалом, приблизно, п'ятдесят годин відеозйомки на набір даних.

Камери можуть бути вибрані зі сценаріїв для перенакладання та ненакладання полів огляду камер або з їх комбінації.

Для точного відстеження мережею камер відеоспостереження Цілі потребуються системного групування. Цілі визначаються та структуруються як такі:

- Ціль – визначена оператором особа.
- Цільова подія – подія, в якій на відеокартинці спостереження наявна Ціль.
- Набір цільових подій – набір відеоматеріалу, зібраного з групи камер відеоспостереження (п'ять у цьому сценарії), що містить множину цільових подій.

1.5. Польова перевірка

Кожний набір даних виявлення подій i-LIDS йде в комплекті з індексом² XML, що описує на високому рівні часовий контент та події тривоги, представлені в кожному відеофайлі.

Аналогічним чином, набори даних відстеження об'єктів йдуть у комплекті з тимчасовим XML-індексом або з польовими даними. Він більш деталізований, аніж сценарій виявлення подій, і забезпечує дані про розмір Цілі та просторову інформацію, що описує їх місцезонашування на відеоматеріалі.

Як доповнення до сирого тексту індексу польової перевірки для полегшення доступу до необхідного відеоматеріалу додається фронт-енд користувацького інтерфейсу.

2. Дані i-LIDS

Кожний набір даних i-LIDS міститься на окремому жорсткому диску, містячи при цьому, окрім документації, такі файли:

- файл з визначеннями сценаріїв (наприклад, Sterile Zone.pdf) – визначає вимоги до відстеження, тривоги та інші атрибути, властиві цьому сценарію, а також приклад схеми XML, що використовується;
- текстовий індекс (index.xml) – XML-описання кожного відеофайла в наборі даних з використанням схеми, визначеної в визначенні сценарію;
- файли підтримки користувацького інтерфейсу (папка “index-files”);
- відео (папка “video”) – рендерене в форматі Quicktime mJPEG;

² Хоча забезпеченню того, щоб кожен індекс був настільки точним, наскільки це можливо, приділяється значна увага, слід врахувати, що CAST не може гарантувати цілісність даних індексу.

- кадри для калібрації – кадри в форматі .tif з кожного огляду камери, що використовується в сценарії;

- відмітки на основі кадрів (папка “annotation”) – папка доступна тільки для вибраних наборів даних. Поставляється для наших колег в Національному інституті стандартів та технологій США.

Сценарій .pdf містить таку інформацію, конкретну для сценарію.

Файл визначення подій тривоги визначає обставини, які представляють собою “подію тривоги” в цьому сценарії. Можуть бути визначені декілька різних типів події тривоги, і всі вони мають розпізнаватися системами відеоспостереження і викликати тривогу.

Кожен сценарій зазвичай містить відеоматеріал з декількох фіксованих оглядових камер. Визначення тривог містять картинку з кожної із цих “сцен” з областями, які належать до визначень подій тривоги, або які виділені за допомогою XML-розмітки.

Вимоги до відстеження описують обставини, які складаються, коли Ціль потрібно відстежувати в кожному полі огляду камери – тобто, коли системи відстеження повинні забезпечити результат, описаний в пункті 3.2.

Кожний набір подій Цілі містить відеоматеріал з п’яти фіксованих оглядів камер. Секція отримання Цілі містить рисунок з кожної із цих камер з коротким описом, коли саме особа вважається за валідну Ціль.

2.1. Відеовідмітки

Системи стеження можуть бути оцінені CAST як для “камер, що перекриваються”, так для ролі “змішаної камери”. Роль із перекриттям включає камери 2, 3 та 4, а роль змішування – всі п’ять камер. В обох ролях система повинна забезпечувати координати XY в реальному часі для потрібної Цілі, пов’язаної з певною камерою.

У випадку, коли Ціль відповідає наступним вимогам, вона помічається рамкою, і повинна відстежуватися:

Камери 1, 2, 3 та 5:

- У сцені видно 100 % зросту Цілі або повинно бути видно, якщо вона не загороджується іншими об’єктами, які не є точками виходу зі сцени³.

- Ціль дорівнює або більше ніж 10 % висоти екрану (58 пікселів).

- Обидва лікті Цілі видимі всередині сцени, або повинні бути видимі, якщо не закриваються об’єктами, які не є точками виходу зі сцени.

Камера 4:

- У сцені видно 75 % зросту Цілі або повинно бути видно, якщо не закривається іншими об’єктами, які не є точками виходу зі сцени.

- Частина Цілі, яку можна помітити в сцені, дорівнює або більше ніж 10 % висоти екрану (58 пікселів).

- Обидва лікті Цілі видимі всередині сцени або повинні бути видимі, якщо не закриваються об’єктами, які не є точками виходу зі сцени.

Ціль є єдиним відміченим об’єктом сцени. Відмітка не включає багаж, який несе або штовхає Ціль, але включає будь-що, що Ціль носить на собі, включаючи шляпи, шарфи та пальта, які можуть додавати розміру до Цілі.

³ Точки виходу зі сцени є границею поля огляду камери або деталями сцени, які закривають собою Ціль ще до того, як вони повинні досягти меж поля огляду камери.

Ціль відмічена з першого кадру, на якому Ціль відповідає вимогам, і продовжує їм відповідати для кожного п'ятого кадру.

Приклад: Ціль входить в поле огляду камери 2 з нижньої частини екрану. По-перше, коли стає видимою тільки голова Цілі, тоді Ціль не відмічена (кадр 2147). Потім стають видимими обидва лікті, але на сцені немає Цілі в повний зріст (кадр 2150). Врешті-решт, на сцені з'являється Ціль в повний зріст (кадр 2153) і відмічається. Відмічається після кожного п'ятого кадру. Згодом Ціль покидає сцену на кадрі 2296, роблячи останній кадр із позначкою на нашій п'ятикадровій послідовності для цієї події Цілі 2293.

Оточена позначка використовується, коли на сцені присутні 100 % площі Цілі, а в полі огляду камери знаходиться більш ніж 50 % Цілі.

Початкова позначка використовується для перших п'яти позначених кадрів події Цілі. Ці позначки призначені для представлення оператору Цілі для першого разу та повинні бути опрацьовані. Ці кадри подаються кожного разу, коли визначається нова Ціль. Початкова позначка в тому ж самому форматі, як і будь-яка інша позначка в наборі даних, і системи будуть потребувати інтерпретації цієї інформації в реальному часі для початку відстеження кожної Цілі.

Початково-оточена позначка використовується, коли позначка відповідає обом згаданим вище правилам.

Отже, вище була описана структура та склад бібліотеки i-LIDS, а також приклади роботи систем ВА. У наступній частині будуть описані критерії випробування систем ВА за допомогою бібліотеки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Imagery Library for Intelligent Detection Systems – i-LIDS user guide/ 10/11, v. 4.9
2. C.J. van Rijsbergen. Information Retrieval. Butterworths, London, 1979.

Отримано 10.01.2017

Рецензент Марченко О.С., к.т.н.