

УДК 004.932.2

ВІДСТЕЖЕННЯ РЕГІОНІВ ІНТЕРЕСУ В ПОСЛІДОВНОСТІ ВІДЕОКАДРІВ

Савченко Д.А., Самощенко О. В.

Донецький національний технічний університет

У статті розглядається задача відстеження регіонів інтересу у відео послідовності. За допомогою цієї технології є можливість кодувати частини відео з різним бітрейтом. Виконано аналіз переваг і недоліків різних методів при вирішенні завдання відстеження. Методика застосовується для об'єктів з постійною формою і для рухомих об'єктів. Розглядається вибір характеристик регіонів інтересу, які будуть використовуватися при відстеженні. Обговорюються методи вирішення виникаючих проблем при виконанні поставленого завдання.

Відстеження регіонів інтересу (ROI - Region of Interest) містить у собі виявлення регіонів інтересу в кожному кадрі за допомогою апіорної інформації, що описує об'єкти пошуку. Це дуже важливо для гарного відео кодування, тому що коректне прогнозування і виявлення ROI має вирішальне значення. Хибно виявлені регіони інтересу призводять до зниження якості сприйняття відео порівняно з початковим відео.

Відстеження ROI може виконуватися для досягнення особливих цілей (відстеження обличчя для програм відео конференцій), або може базуватися на моделях системи зору людини. Також можливий такий варіант, коли користувач визначає регіон інтересу на першому кадрі, а система відстежує об'єкт на усіх подальших кадрах. Таким чином, це включає виявлення об'єкту/регіону у послідовності відеокадрів відштовхуючись від його розташування на першому кадрі. Іншими словами, якщо V відео, що складається з n кадрів $F_1..F_n$ і B_1 – це контур регіону інтересу в кадрі F_1 ; проблема відстеження полягає у визначенні

контурів $B2..Bn$ з $B1$. Такий вид відстеження може бути реалізовано не тільки за допомогою моделей просторового подання, але й з використанням алгоритмів компенсації руху. Відстеження регіонів інтересу доречно застосовувати при телевізійному спостереженні (кримінальні телепрограми, спортивна статистика або репортажі), при пост опрацюванні фільмів (застосування особливих кольорових ефектів стосовно об'єктів), при передачі відео за допомогою мереж (відео конференції) та інше. Не зважаючи на те, що значна кількість робіт, пов'язаних з цією проблемою, була зроблена, мало робіт по виявленню тіл не чіткої форми, що швидко рухаються. У цій роботі лише вивчаються методи, що комбінують просторову, структурну та інші атрибути, необхідні для виявлення ROI.

Регіони інтересу – це область відеокадрі, яка характеризується декількома атрибутами, такими як колір, структура, текстура, яскравість та інше. Відстеження ROI містить у собі виявлення таких областей у послідовності кадрів з використанням їхнього місця розміщення на першому кадрі. Складність полягає у тому, що регіон інтересу може зазнавати незначних змін у зовнішньому вигляді, у освітленні, перетворення сцени та інше. Завдання для системи відстеження регіонів інтересу полягає у застосуванні методу, який би виявляв ROI у присутності таких відхилень. Відстеження регіонів інтересу може бути реалізовано за допомогою двох шляхів.

1. Порівняння цільового блоку (ROI) з блоками кандидатами. Найбільш співпадаючий блок – це наступний цільовий блок.
2. Оцінка стану (цільового блоку) з урахуванням усіх вимірів до цього моменту, або розробка еквівалентної функції ймовірності щільності розташування об'єкту.

Відстеження можливо розглядати як проблему відповідності. Задача полягає у пошуку блоку кандидату в поточному кадрі, що відповідає цільовому блоку попереднього кадру. Міри подібності використовуються для визначення кількості відповідності. Вимірюються подібності геометричних особливостей, таких як форма, положення, структура та негеометричних особливостей, та-

ких як колір та освітлення. Блок кандидат, що є найбільш подібний, визначається новим регіоном інтересу.

2 За допомогою гістограм кольорів відстеження здійснюється як проблема відповідності. Гістограми кольорів цільового регіону та блоку кандидата обчислюються, їх схожість вимірюється за допомогою коефіцієнтів Bhattacharya. Блок кандидат з найбільшим коефіцієнтом встановлюється новим цільовим регіоном. Необхідно детально описати цей метод. Цільовий регіон опрацьовується за своїм розподілом кольору. Якщо x_i для $i=1,2,\dots,n$ це позиції пікселів цільової моделі, що відцентровані стосовно нуля; необхідно визначити функцію $b: R^2 \rightarrow 1..m$, що асоціює піксель з індексом гістограми щодо кольору пікселів. Ймовірність кольору u цільової моделі, що позначається як $q(u)$, присвоюється шляхом опуклого і монотонного віддалення від центру. Менша вага призначається тім місцям, які більш віддалені від центру цільової моделі. Ця ймовірність обчислюється для усіх кольорів $u=1..m$. Починаючи з оціночної позиції, поточний кадр розщеплюється на блоки такого ж розміру, як блоки цільової моделі. Гістограми кольору цих регіонів знаходяться та їх подібність вимірюється. Вираз для обчислення:

$$\rho(y) = \rho[p(y), q] = \sum_{i=1}^n \sqrt{p_u(y)q_u(y)} \quad (1.1)$$

де $p(y) = p_u(y)$ для $u=1..m$ з $\sum_{i=1}^n p(u) = 1$ – це гістограма з m стовпчиками кольору регіону кандидата у позиції y , та $q(y) = q_u(y)$ для $u=1..m$ з $\sum_{i=1}^n q(u) = 1$ – це гістограма з m стовпчиками кольору цільового регіону.

Схожість між двома гістограмами обчислюється з використанням наступної формули:

$$d(y) = \sqrt{1 - \rho[p(y), q]} \quad (1.2)$$

Мінімальна відстань позначає максимальну схожість. Коефіцієнт Bhattacharya $\rho[p(y), q]$ проходить процедуру максимізації для пошуку мінімальної відстані. Ітерації зсуву середньої величини використовуються для максимізації. Ітерації починаються з тієї ж самої позиції, де починається ROI у попередньому кадрі.

Переваги:

1. Метод надійний при обертанні об'єкту, зсуву камери, часткової оклюзії та інше.
2. Алгоритм показує гарні результати, якщо регіон інтересу специфічний стосовно кольорів. Гістограми кольору варто використовувати для відстеження облич, оскільки колір обличчя можливо описати у двох просторових компонент кольору.

Недоліки:

1. Метод не варто використовувати, коли дискримінаційна здатність між гістограмами не є великою. Таке трапляється, коли регіон інтересу важко характеризувати кольором.

Цей метод не є надійним для відстеження регіонів інтересу, оскільки він занадто сильно залежить від кольорової специфікації об'єктів. Також існують об'єднані методи відповідності гістограм, які опрацьовують декілька особливостей ROI, але їх використання значно підвищить розмірність при опрацюванні декількох властивостей.

Відстеження також можливо розглядати як проблему оцінювання. Воно також називається ймовірнісне виявлення/відстеження. У цьому методі регіон інтересу кадру Y_t моделюється як Z_t з параметрами Q_t (як показано у виразі 1.3), також застосовується модель простору станів (описується переходами станів та моделями спостережень) для розміщення декількох кадрів, як показано у виразах 1.4 і 1.5.

Регіон інтересу

$$Z_t = T[Y_t, \theta_{t-1}] \quad (1.3)$$

Модель переходів станів

$$\theta_t = F_t(\theta_{t-1}) \quad (1.4)$$

Модель спостереження

$$\theta_t = Y_t(\theta_t) \quad (1.5)$$

Задача вирішення проблеми оцінювання полягає у пошуку $p(\theta(t) | Y(1:t))$ з урахуванням ймовірності моделі спостереження

$p(Y(t)|\theta(t))$ та ймовірності переходу станів $p(\theta(t)|\theta(t-1))$.

Зазначена проблема оцінки може бути реалізована за допомогою моделі спостереження (моделі зовнішнього вигляду), що є складовою методу фільтру часток. Також можливе використання алгоритму EM (Expectation-Maximization). Ці методи не будуть описані, але зроблені висновки щодо переваг і недоліків.

Одназголовних переваг при вирішенні відстеження як проблеми оцінювання полягає у тому, що знаходиться не лише найкращий локальний збіг, а глобальний. Але ця перевага дається за рахунок великої кількості обчислень. Алгоритм дуже трудомісткий, тому що він стохастичний, що потребує аналізу випадкових явищ. Аналіз випадкових явищ майже завжди потребує затрат при обчисленні. Також цей алгоритм є дуже складним, тому його використання не бажане.

Регіон інтересу повинен бути добре описаний/змодельований для виконання відстеження ROI. Модель, що описана з використанням різних властивостей регіону, називається дескриптором регіону. Властивості, такі як колір, контур (градієнт), інтенсивність та інші, які є особливими для кожного регіону, формують параметри моделі. У результаті, вибір властивостей є одним з найважливіших кроків для проблем відстеження та класифікації. Гарні властивості повинні бути характерними та легко обчислюватися. Такі властивості як колір, градієнт, відклик фільтру – це найпростіший вибір. Їх використовували протягом багатьох років для вирішення завдань комп'ютерного бачення. Гістограми – це природне розширення таких властивостей, які не є параметричними оцінювачами, ефективні при стеженні за не жорстко закріпленими об'єктами.

Однією з основних проблем є брак компетентних критеріїв подібності, що фіксує просторові та статистичні властивості, тобто більшість підходів або залежать тільки від розподілу кольору, або від розподілу структурних моделей. Є можливість використання гістограм з об'єднаними властивостями, але це занадто ускладнить обчислення. Кількість невідомих дорівнює b^d , де b – це кількість стовпчиків гістограми, d – це кількість властивостей. Навіть

моделі, де регіон представлено необробленими значеннями, не приводяться до більшої розмірності, оскільки необхідна розмірність $n \cdot d$, де n – кількість пікселів. У цьому плані коваріаційні матриці є переважаючими. Вони надають чудовий засіб для об'єднання властивостей та добре масштабуються, оскільки розмір дескриптора (матриці коваріації) залежить тільки від числа властивостей d . Вони мають лише $(d^2 + d)/2$ різних значень. Необхідно детально розглянути матриці коваріації.

Нехай I – одновимірне полутонове зображення або три вимірне кольорове зображення. Нехай F – властивість зображення отримана з I . Кількість вимірів $F = W \cdot H \cdot d$, де W і H – розмір ROI, і d – це кількість властивостей, що використовуються для характеризування регіону інтересу. Використовується наступний вираз:

$$F(x, y) = \phi(I, x, y) \quad (1.7)$$

де ϕ може будь-якою властивістю зображення: інтенсивність, колір, градієнти, відклики фільтрів та інше. Треба зазначити, що властивість зображення/вектору може складатися з використанням двох типів атрибутів: просторових та структурних.

$$f_k = [x, y, I(x, y), I_x(x, y)..] \quad (1.8)$$

де \mathbf{k} – піксель, \mathbf{x} та \mathbf{y} координати пікселю, $\mathbf{I}(\mathbf{x}, \mathbf{y})$ – інтенсивність пікселю \mathbf{k} , $\mathbf{I}_x(\mathbf{x}, \mathbf{y})$ – градієнт інтенсивності за напрямком \mathbf{x} . Після того як було отримано властивість F регіону R , створюється матриця коваріації:

$$C_R = \frac{1}{n-1} \sum_{k=1}^n (z_k - \mu)(z_k - \mu)^T \quad (1.9)$$

де z_k для $\mathbf{k}=1..n$ точки властивості у \mathbf{d} вимірах всередині регіону інтересу R та μ – середнє значення точок.

Можна побачити з вище приведених виразів, що матриця коваріації – це ефективний шлях для об'єднання декількох властивостей. Діагональні значення відповідають за розбіжність кожної властивості, не діагональні значення представляють кореляції. Оскільки при коваріації береться середнє значення та обчислюється його коваріація, то шум, який є присутній на зображенні, у більшості

випадків є відфільтрованим.

Є декілька переваги при використанні матриць коваріації як дескрипторів регіону. Деякі з них:

1. Це самий зрозумілий та простий шлях для об'єднання структурних та статистичних властивостей.
2. Матриці є низько вимірними при порівнянні з іншими дескрипторами. Вони мають тільки $(d^2 + d)/2$ різних значень, оскільки розмір матриці коваріації не залежить від розмірів регіону, а залежить лише від кількості використаних властивостей.
3. Можливе досягнення інваріантності при масштабуванні або обертанні в залежності від того, як описані властивості.

Коваріація не залежить від змін середніх значень, як це відбувається при зсуві значень кольору, оскільки при коваріації різниця середніх значень береться з дійсних значень. У результаті, можливо відстежувати об'єкти при освітленні, що змінюється.

Для обчислення відстані між коваріаційними матрицями необхідно використовувати спеціальний підхід. Просте віднімання матриць не спрацює. Використовується інша метрика, що включає об'єднані характеристичні значення матриць. Відстань між значеннями обчислюється за допомогою інваріантної ріманової метрики для роботи з симетричними додатними матрицями.

$$\rho(C_1, C_2) = \sqrt{\sum_{i=1}^d [\ln^2 \lambda_i(C_1, C_2)]} \quad (1.10)$$

де $\lambda_i(C_1, C_2)$, для $i=1..d$ це узагальнені характеристичні значення, які обчислюються з $\lambda_i C_1 x_i - C_2 x_i = 0$.

Алгоритм відстеження припускає, що регіон інтересу задається на першому кадрі. Кроки для реалізації алгоритму:

1. Обчислення матриці коваріації у тій області, яка описує ROI попереднього кадру.
2. Повторити крок 1 для сусідніх регіонів. Розмір околиці визначається користувачем, завданням відстеження.
3. Обчислити відстань між усіма матрицями коваріації, що були

отримані у попередніх двох кроках, з матрицею коваріації регіону інтересу попереднього кадру.

4. Визначити мінімальну відстань та призначити новим ROI регіон, з якого була отримана матриця з мінімальною відстанню.
5. Повторити кроки 1-4 для усіх відеокадрів.

На рисунку 1 наведено приклад відстеження. При моделюванні обиралися такі властивості: позиція (x, y), атрибути зовнішнього вигляду (кольорові компоненти RGB), структура області зображення (градієнт).

Матриці коваріації – це зручний інструмент для відстеження регіонів інтересу. Головні переваги – простота реалізації, можливість легко об'єднувати властивості зображення, кількість яких обирається у залежності від поставленого завдання. Це робить



Рисунок 1 – Відстеження човна на послідовності відеокадрів

2 алгоритм відстеження гнучким. Треба зазначити, що невелика кількість властивостей негативно позначиться на успішності відстеження. Розмірність матриць коваріації не залежить від розміру фрагменту зображення, тому нема ніяких перешкод для додання підтримки масштабування. Алгоритм справляється із завданням відстеження об'єктів як постійної, так і змінної форми.

Література

- [1] Covariance Tracking using Model Update Based on Lie Algebra [Електронний ресурс] / Fatih Porikli, Oncel Tuzel – 2006. – Режим доступу до статті: <http://www.merl.com/papers/docs/TR2005-127.pdf>
- [2] Exploiting Region Of Interest For Improved Video Coding [Електронний ресурс] / Ramya Gopalan – 2009. – Режим доступу до статті: <http://etd.ohiolink.edu/send-pdf.cgi/Gopalan%20Ramya.pdf?osu1250622014>
- [3] A Metric For Covariance Matrices [Електронний ресурс] / Wolfgang Forstner, Boudewijn Moonen – Режим доступу до статті: http://www.uni-stuttgart.de/gi/research/schriftenreihe/quo_vadis/pdf/foerstner.pdf