

УДК 004.932.2

ОТСЛЕЖИВАНИЕ НЕСКОЛЬКИХ РЕГИОНОВ ИНТЕРЕСА В ОДНОМ КАДРЕ

*Борисов А.С., Гридин Д.А., Самощенко А.В.
Донецкий национальный технический университет
Кафедра компьютерной инженерии
reterete@mail.ru*

Рассматривается решение задачи отслеживания необходимого числа ROI (регионов интереса) на протяжении всей видеопоследовательности. Приведены практические результаты отслеживания нескольких регионов интереса в одном видеокадре.

Регион интереса

Передача больших объемов видеоданных высокого качества в реальном времени до сих пор является проблематичной. Снижение влияния ошибок достигается путем использования канала с большой пропускной способностью. Но пропускная способность - это дорогой ресурс. Для снижения битрейта применимо выделение регионов интереса. При этом областям с большей информационной ценностью обеспечивается лучшее качество. Обычно кодировки регионов интереса применяют для видеоконференций, чтобы придать большей качества регионам, содержащие лица, при трансляциях спортивных соревнований, удержании внимания зрителя на важных моментах в кино и т.д.



Рисунок 1. Пример назначения регионов интереса в видеокадре

Задачи слежения

Регион интереса – это область в видеокадре, которая характеризуется несколькими атрибутами, такими как цвет, структура, текстура, яркость и прочее. Отслеживание ROI (регион интереса) включает в себя выявление таких областей в последовательности кадров с использованием их места размещения на первом кадре. Сложность заключается в том, что регион интереса может подвергаться незначительным изменениям во внешнем виде, в освещении, превращение сцены и прочее. Задача для системы отслеживания регионов интереса заключается в применении метода, который бы выявлял ROI в присутствии таких отклонений. Отслеживание регионов интереса может быть реализовано с помощью оценки состояния целевого блока (ROI) или сравнении его со следующими блоками кандидатами.

Отслеживание возможно рассматривать как задачу оценивания. Указанная задача оценки может быть реализована с использованием алгоритма EM (Expectation-Maximization, англ. Ожидания максимизации (EM) алгоритм) - алгоритм, используемый в математической статистике для нахождения оценок максимального правдоподобия параметров вероятностных моделей, в случае, когда модель зависит от некоторых скрытых переменных.

Отслеживание также возможно рассматривать как проблему соответствия. Задача состоит в поиске блока кандидата в текущем кадре, что соответствует целевому блоку предыдущего кадра. Меры сходства используются для определения степени соответствия. Измеряются подобию геометрических особенностей, такие как форма, положение, структура, а так же негеометрические подобию - цвет и освещение. Блок кандидат, который является наиболее похож, становится новым регионом интереса.

Гистограммы цветов используют для отслеживания [2]. Гистограммы цветов целевого региона и блока кандидата рассчитываются, их сходство измеряется с помощью коэффициентов Бхаттачария (1). Низкие показатели свидетельствуют о хорошем совпадении, а высокие – о плохом. Идеальное совпадение соответствует нулю 0, а расхождение – единице.

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

Ковариация. Ковариационные матрицы

Критерии подобию – безразмерная величина, составленная из размерных параметров, определяющих рассматриваемые свойства. Одной из основных проблем является отсутствие критериев подобию, объединяющих пространственные и статистические свойства, то есть большинство подходов или зависят только от распределения цвета, или от распределения структурных моделей. В этом плане ковариационная матрицы являются преобладающими [3]. Они предоставляют отличный способ для объединения свойств и хорошо масштабируются, поскольку размер дескриптора (матрицы ковариаций) зависит только от числа свойств d . Они имеют только $(d^2 + d) / 2$ различных значений. Необходимо детально рассмотреть матрицы ковариаций и вычисления расстояния между ними.

Ковариация – в теории вероятностей и математической статистике, числовая характеристика зависимости случайных величин. Сущность ковариации заключается в том, что она возникает вследствие неопределенности результата перемножения двух совокупностей чисел.

Ковариация двух случайных величин X, Y обозначается как $Cov(X, Y)$:

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E(X, Y) - \mu_X \mu_Y = \mu_{XY} - \mu_X \mu_Y \quad (2)$$

где E - оператор математического ожидания; μ_X - среднее значение величины X ; μ_Y - среднее значение величины Y ; $E(X, Y)$ - математическое ожидание произведения величин X, Y ; μ_{XY} - среднее значение произведения этих величин.

Случайные величины, имеющие нулевую ковариацию, называются некоррелируемые. Независимые случайные величины всегда некоррелируемые, но не наоборот [4].

Ковариационная матрица (или матрица ковариаций) в теории вероятностей - это квадратная матрица, составленная из попарных ковариаций и дисперсий двух или более случайных величин.

Если $X: \Omega \rightarrow R^n, Y: \Omega \rightarrow R^m$ - два случайных вектора размерностью n и m соответственно. Пусть также случайные величины $X_i, Y_i, i = 1, \dots, n, j = 1, \dots, m$ имеют конечный второй момент. Тогда матрицей ковариаций X, Y называется:

$$\Sigma_{ij} = cov(X_i, Y_j) = E[(X_i - \mu_i)(X_j - \mu_j)], \quad (3)$$

где $\mu_i = E(X_i)$.

$$C = \begin{vmatrix} cov(x,x) & cov(x,y) & cov(x,z) \\ cov(y,x) & cov(y,y) & cov(y,z) \\ cov(z,x) & cov(z,y) & cov(z,z) \end{vmatrix}$$

Рисунок 2. Матрица ковариации

Алгоритм отслеживания регионов интереса

Для начала работы алгоритма отслеживания необходима априорная информация относительно внешнего вида объекта. Одним из ее вариантов могут быть координаты расположения региона интереса на первом кадре видеопоследовательности. После того, как получена матрица ковариаций и координаты на первом кадре, начинается отслеживание регионов интереса на всех последующих кадрах. Для всех полученных матриц ковариаций вычисляется расстояние до эталонной матрицы. Регион, который имеет наименьшее расстояние, устанавливается новым ROI. Запоминается его позиция и матрица ковариаций.

Алгоритм повторяется для всех последующих кадров. Все описанные действия можно разбить на несколько шагов:

1. Получить местоположение регионов интереса на первом кадре. Вычислить матрицы ковариаций.
2. Перейти на следующий кадр. Вычислить матрицы ковариаций для позиций расположения ROI в предыдущем кадре и для всех соседних позиций.
3. Вычислить расстояние между всеми матрицами ковариаций регионов кандидатов с эталонными матрицами.
4. Выбрать матрицу с минимальным расстоянием. Соответствующий регион устанавливается новым ROI. Запомнить позицию и матрицу ковариаций.
5. Повторить шаги с 1 по 4 для всех кадров видеопоследовательности.

Блок-схема работы алгоритма изображена на рис. 3.

Отслеживание нескольких регионов интереса в одном видеокадре

Для отслеживания нескольких регионов интереса на одном видеокадре были подобраны различные видео. Наиболее подходящие для этого видео – концертные записи, видеоклипы, спортивные состязания или кино. Регионы интереса на выбранном кадре могут быть различного размера. Подбор параметров поиска ROI на последующих кадрах позволяет отлично работать с такими регионами интереса.

Для большего сосредоточения на необходимых регионах интереса и снижения размера полученного видео был применен blur-эффект для оставшегося фона. В некоторых видео с большим количеством артефактов можно применить fuzzy-фильтрацию [5] ROI для улучшения изображения регионов интереса.

Выводы

Экспериментально показана возможность отслеживания нескольких регионов интереса на одном видеокадре. Работа с ROI с различными параметрами высоты и ширины. Применение blur-эффекта для снижения размера выходного видеофайла.

Литература

- [1] Karlsson L.S. Spatio-temporal pre-processing methods for region-of-interest video coding / Linda S. Karlsson – Department of Information Technology and Media Mid Sweden University, 2007.
- [2] Comaniciu D. Real-Time Tracking of Non-Rigid Objects using Mean Shift / D. Comaniciu, V. Ramesh, P. Meer - IEEE Transactions on Computer Vision and Pattern Recognition, 2009.
- [3] Porikli F. Covariance Tracking using Model Update Based on Lie Algebra / F.Porikli, P.Meer, O.Tuzel - IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006.
- [4] Forstner W. A metric for covariance matrices / W. Forstner and B. Moonen – Technical report, Dept. of Geodesy and Geoinformatics, Stuttgart University, 1999.

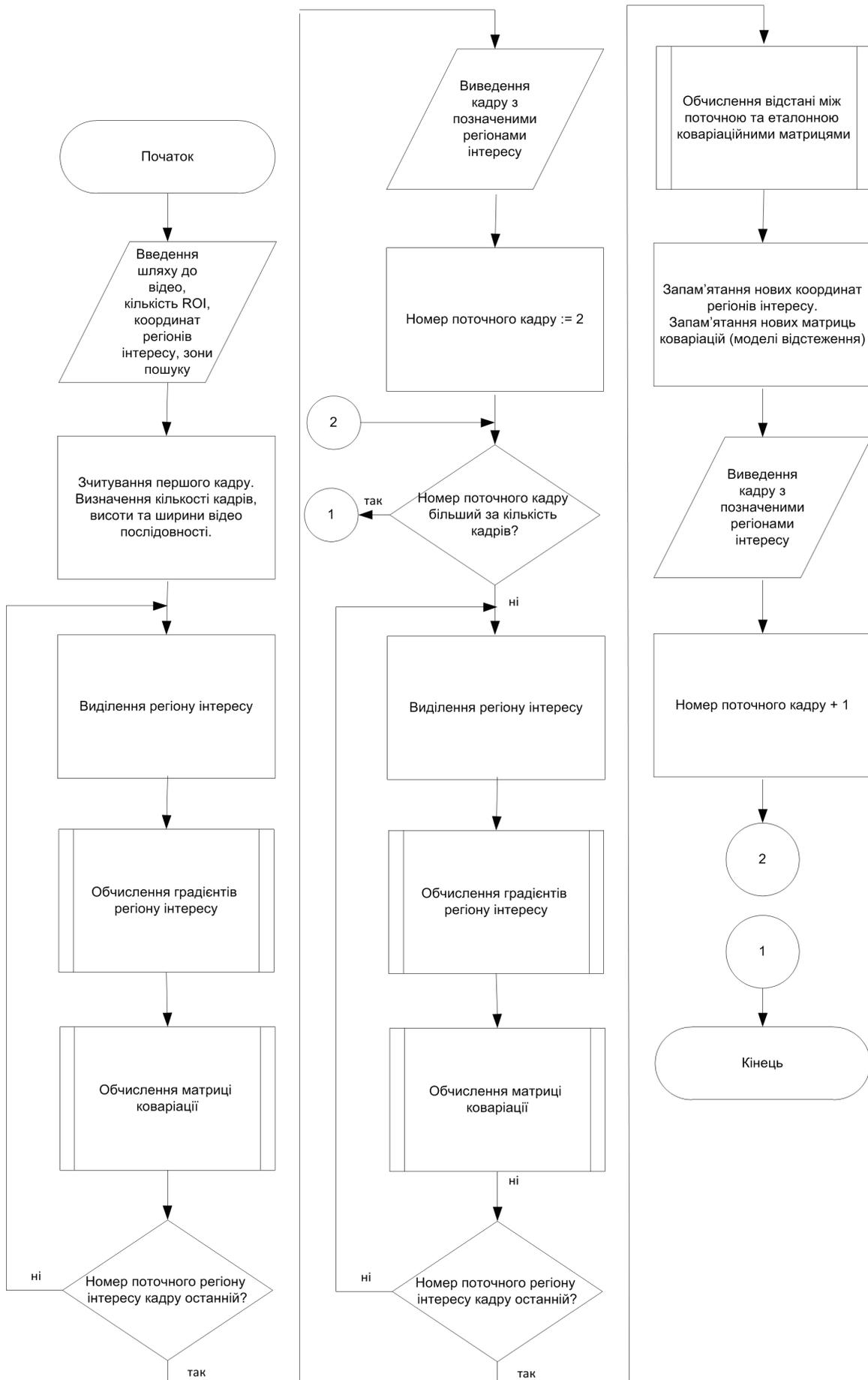


Рисунок 3. Блок-схема алгоритма отслідження



а)

б)

Рисунок 4. Результаты отслеживания двух регионов интереса:

а) первый кадр видеопоследовательности, б) десятый кадр видеопоследовательности



а)

б)

в)

Рисунок 5. Результаты отслеживания трех регионов интереса:

а) первый кадр видеопоследовательности, б) десятый кадр видеопоследовательности;

в) двадцать третий кадр видеопоследовательности

[5] Dung T. Vo, Truong Q. Nguyen, Sehoon Yea, Anthony Vetro, Adaptive Fuzzy Filtering for Artifact Reduction in Compressed Images and Videos // IEEE Transactions on Image Processing – v.18 – no.1 – 2009 – Pp. 1166-1178.

[6] Gopalan R. Exploiting Region of Interest for Improved Video Coding / R. Gopalan – Ohio, 2009.