УДК 004.932.2

Геометрический анализ видеоданных

Geometrical video data analysis

Зуев С.В., Кабалянц П.С., Поляков В.М.

**Аннотация**. Работа посвящена анализу видеоданных. Предполагается, что видеоданные могут содержать как валидные объекты (находившиеся на сцене), так и невалидные объекты – объекты, внедренные в видеоданные уже после съемки. Известно, что при внедрении объекта в видеоданные, учитываются цветовые и яркостные характеристики изображения. Однако, даже если внедрённый объект является реальным, например, снятым в другом месте и времени, но не валидным – он будет иметь отличные от валидных объектов характеристики движения и трансформации. Проекция на пространство изображений трехмерного объекта при движении этого объекта испытывает определенные преобразования – конформные, повороты, трансляции и растяжения вдоль выделенной оси. Характеристики этих преобразований для невалидного объекта в сравнении с валидными дают возможность сделать заключение о его нахождении на сцене. Работа посвящена предварительной подготовке к такому анализу объектов. Описан способ параметризации целевых объектов и представления их в виде структур данных. Описаны алгоритмы представления объектов и их преобразований в пространстве изображений и в реальном пространстве. Представлен алгоритм определения валидности объекта, удовлетворяющего некоторым естественным требованиям. Результаты проведённого вычислительного эксперимента соответствуют положениям, предлагаемым в работе.

**Ключевые слова**: видеоданные, распознавание образов, проверка подлинности объекта, детектирование внедрений.

**Abstract**. This paper is devoted to video data analysis. The considered objects are those which have doubted validity, i.e. they are possible embedded into the video. It is well-known that when the object has been embedded, the color and brightness characteristics are taken into account. Suppose that the embedded object really exists, but it has been shot in other place, other time. And let this object had been embedded and hence invalid for the considered video. In this case it will be described by some characteristics different from the valid objects. The projection of the 3D-object on the image space transforms in definite way when the very object moves in the space. These transformations are conformal one, rotation, translations and stretch transformation along an axe. The characteristics of the transformations give us possibility to analyze the object and make the conclusion about validity. The article presents preliminarily considerations about such analysis. The parametrization of the object is described at the first and the corresponding data structures are proposed. Then the algorithms of the object presentation in the image space and in real space are presented. Finally, the algorithm of the object validity determination is proposed under some natural assumptions. The computational experiment has been made and it shows that the propositions given in the paper are really take place.

**Keywords**: video data, pattern recognition, object validation check, embeddings detection.

**Введение**

Что такое видеоряд: кино, цифровое видео, кодировка, стрим

Видеоряд – последовательность видеоизображений, каждое из которых является кадром

Видеоряды несут в себе информацию, которую можно не только наблюдать, но и анализировать

Работа над анализом видеоданных: анализ видеоданных, компьютерное зрение ( чем занимаются люди и чего добились в общем)

Форензика – расследование компьютерных преступлений(?)

Анализ видеоданных в контексте расследования преступлений

Валидность - ?

Вопрос о валидности объектов в видеоданных возникает, чаще всего, в следующих случаях:

- ведется расследование или судебное разбирательство и требуется подтвердить или опровергнуть подлинность улик на видеозаписи;

- имеются основания опасаться внедрения объектов в потоковые видеоданные системы наблюдения;

- есть подозрения на активность с целью влияния на пользователя посредством внешне нейтрального видео, но с опасными внедренными объектами.

В первом случае речь идет об отложенном анализе видеоданных. Во втором и третьем случаях необходим либо анализ в реальном времени, либо отложенный анализ с коротким временем ожидания результата. В любом случае, скорость обработки и принятия решения является важным параметром в рассматриваемой задаче.

Анализ цифровых изображений был в центре исследований еще в 60-е годы прошлого века [1]. Постепенно выделяется отдельная область исследований – компьютерное зрение. Основными задачами в этой области являются задачи выделения объектов на цифровом изображении и их классификация. Классическим алгоритмом для выделения контуров объектов является детектор Кэнни [2]. В его основе лежит анализ градиентов составляющих RGB-модели. Он реализован в виде отдельной функции популярной библиотеки OpenCV.

С другой стороны, достоверность выделенных объектов, их валидность, стала изучаться существенно позже. Целый ряд работ появляется в 90-е годы 20 века, а в 2001 году в Нью-Йорке состоялся семинар по компьютерно-технической судебной экспертизе (the First Digital Forensic Research Workshop (DFRWS, http:/dfrws.org)), на котором была выработана дорожная карта исследований по валидности в судебной экспертизе (A Road Map for Digital Forensic Research). Позже появляется целый ряд исследований, в которых обсуждается именно валидность выделенных объектов [3]. Однако большинство работ основано на анализе именно RGB-модели без учета соотношения двумерного плоского изображения и исходной трехмерной сцены.

В последние годы появилось значительное число работ, связанных с анализом валидности объектов видеоданных при проведении судебной экспертизы (см., например, [4-6]). Особенно важной эта задача считается в странах, в которых большая часть разбирательств проводится судом присяжных. Однако, важность таких разработок подчеркивается во всем мире.

В настоящей работе предлагается построить процесс анализа наличия невалидных объектов в видеоданных с использованием существующих инструментов выделения и трассировки опорных точек объектов. Полигоны опорных точек объектов являются плоскими геометрическими фигурами, которые от кадра к кадру испытывают преобразования: трансляции, повороты, конформные растяжения и сжатия, а также растяжения и сжатия вдоль какой-то оси. Эти преобразования являются результатами движений трехмерных объектов в реальном пространстве: трансляций и вращений. Для большинства приложений можно полагать объект приближенно сохраняющим форму: если объектом является активное живое существо, механизм, жидкость и т.п., для которых, вообще говоря, это предположение не выполнено, то требуются дополнительные меры для их исследования: например, выделение из всего объекта части, сохраняющей форму, и указания опорных точек на ней.

Реальное пространство 3-мерно и объект в нем тоже. В 2017 году в работе [7] был предложен алгоритм, строящий 3D модель объекта по видеоданным. Однако, этот алгоритм требует специальной съемки объекта – камера должна двигаться вокруг него. С точки зрения расследования зачастую нельзя потребовать специальной съемки, но можно узнать характеристики камеры и расстояние до объекта в какой-то момент (по относительному расположению других объектов). Сохраняющий форму объект будет совершать в 3-мерном пространстве только *движения*, то есть трансляции и вращения. Эти движения, спроецированные на матрицу камеры, становятся 2-мерными преобразованиями *образа* объекта, данные о которых могут быть получены из видеоряда. А именно, трансляции объекта приводят к трансляциям и конформным преобразованиям образа, а вращения объекта приводят для образа к вращениям и растяжениям вдоль некоторой оси. Эта ось не обязательно совпадает с какой-то координатной осью — это может быть любая прямая в пространстве изображений. В литературе образ реального объекта в пространстве изображений называется обычно *объектом видеоданных*. При совместном употреблении с реальным объектом, объект видеоданных будем также называть *образом*. Набор опорных точек объекта видеоданных задается соответствующим полигоном. С помощью специально подобранной параметризации, мы представим полигон в виде набора векторов, имеющих в полярных координатах углы где . Такое представление поможет нам отфильтровать трансляции по осям, параллельным плоскости изображения, оставив в представлении образа зависимости только от вращений реального объекта, от его трансляций по оси , перпендикулярной плоскости пространства изображений, а также от времени.

**Параметризация объекта и его образа**

Предположим, что опорные точки уже выделены и осуществляется их трассировка

Для определения валидности объекта используются различные методы:

1. Статистический анализ цветовых и яркостных характеристик объекта и фона
2. Объемный анализ цветовых характеристик (подобно контрольной сумме файлов)
3. Метод метаданных (очень много по нему работ в США). Основан на том, что каждый видеофайл или картинка имеет свои атрибуты – метаданные – которые при внедрении объекта в другой файл, «переходят» в другой файл
4. Метод, для которого используются результаты проекта: обратный рендеринг, заключающийся в восстановлении трёхмерного объекта из двумерного изображения

Для обратного рендеринга очень важна параметризация объекта видеоряда – она должна быть такой, чтобы процедура рендеринга происходила максимально экономично( по ресурсам), так как она довольно сложна.

Пусть имеется некоторый выделенный в каждом кадре видеоряда объект и набор его опорных точек представлен в виде некоторого полигона. Будем предполагать, что в каждом рассматриваемом кадре образ не попадает на второй план (обычно внедренный объект не располагается на втором плане).

Выберем целевой объект и обозначим пары координат опорных точек его образа через , где – число опорных точек. Тогда координаты центра прямоугольника, построенного на опорных точках, будут иметь вид

|  |  |
| --- | --- |
|  | (1) |

В пространстве изображений перейдем в систему координат с осями, параллельными изначально заданным, и центром в точке (*X*,*Y*), координаты которой определяются (1). Образ объекта в новой системе координат будет задаваться с помощью следующих пар координат:

Таким образом, с одной стороны, у нас есть система координат с центром на пересечении диагоналей прямоугольника в двумерном пространстве изображений и, с другой стороны, система координат с тем же началом отсчета, но в реальном трехмерном пространстве. Между ними имеется соотношение

|  |  |
| --- | --- |
|  | (2) |

построенное по формуле тонкой линзы с фокусным расстоянием .

Пусть объект в реальном пространстве задан координатами его опорных точек, представляющими собой некоторые постоянно присутствующие в видеоряде точки образа, координаты которых известны для любого кадра. При этом координаты опорных точек объекта задаются в виде начальных условий только для первого кадра. Индексы номеров кадров в дальнейшем не пишем для краткости изложения.

В заключение этого пункта отметим, что количество опорных точек может быть довольно малым, а для вычисления характеристик преобразования нам нужно будет много точек. Чтобы преодолеть эту трудность, введем параметризацию полигона опорных точек специально построенными векторами – выходя из начала координат, каждый вектор будет иметь вершину на отрезке, соединяющем соседние по полярным углам опорные точки. Полярные углы самих векторов задаются соотношением

Тем самым, теперь объект будет характеризоваться не точками, а любым наперед заданным количеством точек и можно установить , ограничиваясь только вычислительными ресурсами.

**Преобразования в реальном пространстве и пространстве изображений**

Если тело сохраняет форму, то преобразования в реальном пространстве между соседними кадрами состоят из трансляций по осям координат, вращения вокруг оси на угол и вращения вокруг оси на угол . Других преобразований нет. Трансляция по оси , имеющая вид с учетом (2), дает следующее

где , а – перемещение в пространстве изображений в результате сдвига на в трехмерном пространстве. Отсюда легко видеть, что трансляция по оси не влияет на отношение и аналогично по оси Эти трансляции мы используем чтобы перейти в новую систему координат с началом отсчета, которое будем считать неподвижным при вращении. Центр ее пусть будет иметь координаты , определяемые аналогично (1). В текущей системе для вектора с углом :

и эти координаты будут вычисляться аналогично в каждом кадре.

В реальном пространстве, в момент первого кадра, мы перейдем в систему координат с центром в точке где

Получим такие новые координаты:

Так как и , можем записать

Трансляция по оси в реальном пространстве определяется следующим образом:

а в системе координат, связанной с объектом,

В пространстве изображений будем иметь соотношения

|  |  |
| --- | --- |
|  | (3) |

где - масштабный фактор для перехода к следующему кадру и важная величина для вычисления расстояния до объекта в следующем кадре:

|  |  |
| --- | --- |
|  | (4) |

Вращения вокруг оси с неподвижной точкой на угол , при сохранении формы объекта (то есть длины радиус-векторов точек), осуществляются по правилу

Перепишем это для пространства изображений:

отсюда находим

|  |  |
| --- | --- |
|  | (5) |

Рассмотрим теперь вращения вокруг оси на угол . Имеем в реальном пространстве:

В пространстве изображений, соответственно,

Величину найдем из (3) и, подставим сюда, получим

|  |  |
| --- | --- |
|  | (6) |

**Алгоритм расчета параметров объекта**

Рассмотрим приведенные выше формулы в качестве основы для расчета параметров объекта в следующем кадре, если известны соответствующие его параметры в текущем кадре. Если объект реально был на сцене, то его преобразования приведут к естественным изменениям параметров. Если же он был внедрен, то при монтаже были применены специальные инструменты, приводящие объект к нужному масштабу так, чтобы он казался естественным в видеоряде. При этом, однако, параметры его трансформаций будут изменяться негладким образом, чем он и будет отличаться от валидного объекта.

Параметрами объекта выберем его удаление от объектива и разброс (стандартное отклонение) координат его опорных точек. Покажем способ расчета этих параметров и продемонстрируем то, каким образом нарушается гладкость при вложении объекта в видеоряд.

Итак, у нас имеются величины координат образа объекта в пространстве изображений в начальном кадре – в принятых выше обозначениях это , а также в следующем кадре – после всех преобразований это Кроме этого, мы можем указать начальные значения координат опорных точек по оси : . В результате имеем набор координат для контура, натянутого на опорные точки объекта в данном кадре, а также координаты в следующем кадре. Кроме того, для подстановки в формулы следует учесть, что так как последнее вращение производится вокруг оси то .

Величины найдем из уравнения поворота вокруг оси минимизацией функции

Подставим найденные величины , соответствующие локальному минимуму функции , в первое уравнение из (5) и найдем из него величины .

Далее рассмотрим поворот вокруг оси . При этом расстояние до объекта, равное , не претерпевает изменений и определяется формулой (4), а величина может быть найдена из первого уравнения (6) минимизацией функции:

где величины определяются следующим образом:

и в них определены все характеристики, кроме угла .

Далее, по найденной величине , соответствующей локальному минимуму функции , находим -координаты объекта в следующем кадре из второго уравнения (6):

Получено рекуррентное соотношение для расстояния до объекта в нужном кадре при известных величинах первого кадра, а также последовательность стандартных отклонений от среднего от кадра к кадру. То есть, характеристиками объекта при таком рассмотрении могут быть последовательности

– расстояния и стандартные отклонения для каждого кадра. Для валидных объектов последовательность не должна превышать реальных размеров объекта и перекрывать другие объекты на сцене. Расстояние также должно быть сопоставляемо с другими объектами видеоряда, не претерпевать резких скачков и быть согласованным с видимым поведением объекта на сцене. Такое поведение пар ( должно отличать валидные объекты от девиантного поведения невалидных объектов [8]. Это подтверждают и результаты эксперимента.

**Эксперимент: описание движений объекта видеоданных, представленного несколькими опорными точками**

Рассмотрим видеоряд из 24 кадров (около 1 секунды продолжительности) с реальным живым объектом. Кадры 1, 5, 10 и 15 представлены на рис. 1.

Изображение выглядит как животное, кот, млекопитающее, забор

Автоматически созданное описание Изображение выглядит как животное, рептилия, забор, здание

Автоматически созданное описание Изображение выглядит как животное, млекопитающее, кенгуру, сидит

Автоматически созданное описание Изображение выглядит как животное, млекопитающее, кот, сидит

Автоматически созданное описание

Рис. 1. Фрагменты видеоряда.

В качестве опорных точек для моделирования 3D-объекта были взяты правое ухо, глаз, нос и верхняя губа животного. Рассматривается движение только головы. На видеоряде голова перемещалась по всем направлениям. Полигон, моделирующий объект, в пространстве изображений выглядит так, как показано на рис. 2 (для тех же кадров, но в системе координат, связанной с центром головы).

Изображение выглядит как текст, карта

Автоматически созданное описание

Рис. 2. Параметризация целевого объекта.

Начальные координаты опорных точек были восстановлены по масштабам изображения на матрице камеры (в 1 см – 2500 пикселей) и реальным размерам животного, находящегося примерно на расстоянии 10 см от объектива, а см.

Расчеты дают последовательности, представленные на графиках на рис. 3 и 4.

Изображение выглядит как текст, карта

Автоматически созданное описание

Рис. 3. Расстояние до объекта (в пикселях) при различных значениях фокусного расстояния и начального удаления от объектива.

Изображение выглядит как текст, карта

Автоматически созданное описание

Рис. 4. Стандартное отклонение координаты объекта для расчетов с разными фокусными расстояниями и начальными удалениями от объектива.

На графиках видно, что при несоответствующим реальным значениям фокусного расстояния, а особенно начального расстояния при съемке объекта, эволюция дистанции до объекта протекает со значительным отличием от видимых значений. При этом расчетное стандартное отклонение (то есть размеры объекта по оси ) зависит только от фокусного расстояния.

Таким образом, если объект снят с другого расстояния, а затем вставлен в видеоряд, его параметры будут нуждаться в постоянной подстройке, что обязательно отразится на графике расстояний до объекта.

**Выводы и дальнейшие исследования**

Предложенная в работе идея восстановления расчетной 3D-модели объекта видеоданных при сравнительно бедной входящей информации может стать основой для определения валидности объектов видеоданных, для распознавания угроз информационной безопасности в потоке видеоданных, а также для других целей. Чтобы сделать возможным применение этой методики, требуется провести больше экспериментов, оптимизировать алгоритмы и, возможно, сделать вычисления параллельными. Кроме того, для эффективного множественного анализа данных по совокупности объектов, среди которых фиксированы реальные объекты и выделены объекты для анализа, скорее всего, потребуется применение нейросетевых технологий.

***Работа выполнена при поддержке гранта РФФИ, проект № 19-29-09056 мк***

БИБЛИОГРАФИЯ

[1] Rosenfeld A., Pfaltz J.L. Sequential operations in digital picture processing. J. Assoc. Comput. Machinery. – 1966., v.13. – P. 471-494.

[2] Canny J. A computational approach to edge detection. IEEE Trans. Pattern Anal. And Machine Intelligence. – 1986, v. 8(6). – P. 679-698.

[3] Beckett, J., Slay, J.: Digital Forensics: Validation and Verification in a Dynamic Work Environment. In: Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS 2007), p. 266a (2007).

[4] Anderson Rocha, Walter Scheirer, Terrance Boult, Siome Goldenstein. Vision of the unseen: Current trends and challenges in digital image and video forensics. ACM Computing Surveys (CSUR). – 2011, v.43, Issue 4. – P.1-42.

[5] Daniel Moreira, Sandra Avila, Mauricio Perez, Daniel Moraes, Vanessa Testoni, Eduardo Valle, Siome Goldenstein, Anderson Rocha. Multimodal data fusion for sensitive scene localization. - Information Fusion, 2019. – V.45. – P.307-323.

[6] Joseph Roth, Andrew Carriveau, Xiaoming Liu, and Anil K. Jain. Learning-based Ballistic Breech Face Impression Image Matching, IEEE 7th Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS 2015), Arlington, VA.

[7] Буханов Д.Г. Алгоритм построения 3D модели объекта по видеопотоку / А.М. Макаров, Д.Г. Буханов // В сборнике: Оптико-электронные приборы и устройства в системах распознавания образов, обработки изображений и символьной информации. Распознавание - 2017 сборник материалов XIII Международной научно-технической конференции. Курск, 2017. С. 227-230.

[8] Shapiro L.G., Stokman G.C. Computer Vision. – Prentice Hall, 2001. – 608 p. или Шапиро Л., Стокман Дж. Компьютерное зрение. – М.: БИНОМ, 2015. – 763 с.

**Сведения об авторах**

**Зуев Сергей Валентинович**; Белгородский государственный технологический университет им. В.Г. Шухова, к.ф.-м.н., доцент кафедры программного обеспечения вычислительной техники и автоматизированных систем, 308012 Белгород, ул. Костюкова, 46, тел. (4722) 309901, доб. 16-83, sergey.zuev@bk.ru

**Кабалянц Петр Степанович**; Белгородский государственный технологический университет им. В.Г. Шухова, к.т.н., доцент кафедры программного обеспечения вычислительной техники и автоматизированных систем, 308012 Белгород, ул. Костюкова, 46, тел. (4722) 309901, доб. 16-83, kabalyants@gmail.com

**Поляков Владимир Михайлович**; Белгородский государственный технологический университет им. В.Г. Шухова, к.т.н., доцент, заведующий кафедрой программного обеспечения вычислительной техники и автоматизированных систем, 308012 Белгород, ул. Костюкова, 46, тел. (4722) 54-95-09, p\_v\_m@mail.ru

**Zuev Sergei Valentinovich**; Belgorod Shoukhov State University of Technology, PhD, Associated Professor, 46 Kostyukova str., Belgorod 308012 Russia, Phone. +7 (4722) 309901, add. 16-83, sergey.zuev@bk.ru

**Kabalyants Petr Stepanovich**; Belgorod Shoukhov State University of Technology, PhD, Associated Professor, 46 Kostyukova str., Belgorod 308012 Russia, Phone. +7 (4722) 309901, add. 16-83, kabalyants@gmail.com

**Polyakov Vladimir Mikhailovich**; Belgorod Shoukhov State University of Technology, PhD, Head of Department, 46 Kostyukova str., Belgorod 308012 Russia, Phone. +7 (4722) 54-95-09, [p\_v\_m@mail.ru](mailto:p_v_m@mail.ru)

Задачи на эту неделю:

1. Написание введения по всем тем тезисам ( всё красное) – 2-4 страницы
2. Разобраться в коде