ПОСТРОЕНИЕ КОДЕРА ВИДЕОДАННЫХ, ОРИЕНТИРОВАННОГО НА ПРИМЕНЕНИЕ НА БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТАХ

METHOD OF BUILDING VIDEO DATA CODER, UAV-ORIENTED

K. Mudrak

Summary. Review of method of building video data coder, exploiting special features, given by shooting video from UAV. Formalized basic principles for building one. Founded conditions for calculating the matrices of the primary projective transformation. Adaptation of the inverse Lucas-Canade algorithm for pyramidal projection transformation is carried out in terms of exact transformation. Combined the results of the actions of algorithms, proposed the way to further decomposition and compression of the frame

Keywords: video compressing, unmanned aerial vehicle, projection transformation.

Мудрак Константин Русланович
Аспирант, ФГОБУ ВПО «Санкт-Петербургский государственный университет телекоммуникаций им. проф. М. А. Бонч-Бруевича»
zzzmurdoczzz@bk.ru

Аннотация. Рассмотрен способ построения кодера видеоданных, использующий особенности съемки с беспилотного летательного аппарата (БЛА). Предложены основные принципы построения такого кодера. Проведена работа по нахождению условий для вычисления матриц первичного проективного преобразования. Проведена адаптация инверсного алгоритма Лукаса-Канаде для решения задач точного пирамидального проекционного преобразования. Совмещены результаты действий алгоритмов, предложен путь к дальнейшей декомпозиции и сжатию кадра.

Ключевые слова: видеосжатие, беспилотный летательный аппарат, проективные искажения.

Введение

В силу инерционности человеческого зрения [1], достаточно отображать с частотой более 15–20 Гц статичные изображения для создания ощущения движения на отображающем устройстве. Мозг человека сам интерполирует полученные данные и «достраивает» движение. Собственно, на этом принципе основано телевидение, от него же унаследовало подход к хранению информации и цифровое телевидение. То есть видеоданные представляют собой набор статичных изображений, а значит занимает огромный объем памяти. С целью снижения объема памяти, занимаемого видеоконтентом, необходимо применение схем сжатия. Вне зависимости от применяемого алгоритма кодирования, сами видеоданные высокоэнтропийны, а значит слабо сжимаются. Поэтому для передачи и хранения видеоизображений применяется сжатие с потерями. Для того, чтобы потери информации минимально сказывались на субъективном качестве восприятия изображений, необходимо установить, какие данные являются важными, а какие нет. Внутри одного кадра избыточными, согласно JPEG [2], являются данные высокой пространственной частоты и данные о полной цветности каждого пикселя. Этот тип избыточности носит название пространственная избыточность. Между кадрами, согласно MPEG [3], избыточными данными являются блоки изображения, присутствующие в предыдущем кадре. Классические межкадровые кодеры строятся из предположений, что между двумя соседними кадрами разница в перемещении контекстной информации невелика. Согласно этому постулату, можно разбить изображение предыдущего кадра на блоки пикселей, найти соответствующие им блоки на текущем кадре, и вместо передачи блоков текущего изображения, передавать только вектор перемещения блока. Этот тип избыточности называется временная избыточность. Из-за возникающих в процессе реконструкции изображения только из блоков из предыдущего кадра неточностей, передается не только таблица векторов, но и сигнал ошибки, сформированный абсолютной разностью между текущим кадром и кадром, сформированным на основе сдвига блоков предыдущего кадра.

Применение видеокодирования к БЛА

В случае применения межкадрового кодирования к видеоряду с борта БЛА самолетного типа, несложно заметить, что:

1. Вектора всех блоков должны быть сонаправленны. Это происходит из-за установки видеокамеры ортогонально фюзеляжу БЛА. То есть в ходе полета, информация из предыдущего кадра (Кпред) смещается «вниз» текущего кадра (Ктек).
2. В силу неравномерности полета БЛА, а именно, наличия крена и тангажа в ходе съемки, вектора перемещения недостаточно для полного описания изменения частей изображения. В общем случае блоки подвергаются проекционному преобразованию.

Исходя из вышесказанного, и общего принципа ведения съемки с БЛА, можно сформулировать следующие постулаты адаптации принципов межкадрового кодирования, применительно к видеоинформации с борта БЛА:

1. Можно, не разбивая изображения на блоки, вычислить смещение Кпред относительно Ктек целиком и передавать не таблицу векторов, а одну матрицу преобразования кадра.

2. Благодаря наличию телеметрической информации, можно вычислить изначальное смещение Ктек относительно Кпред. Кроме смещения, эта информация также даст необходимые коэффициенты матрицы проективного преобразования.

3. Из-за ошибок телеметрии, задержки поступления телеметрии относительно времени экспозиции кадра, необходимо процесс нахождения матрицы преобразования кадра разбить на два этапа: Первичное преобразование, на основе телеметрической информации, и вторичное преобразование на основе контекстуальной информации внутри кадров.

Рассмотрим каждый из постулатов. Информация между кадрами, снятыми с БЛА, является разновидностью панорамной съемки, при которой происходит смещение камеры относительно объекта съемки. При этом в каждом новом кадре содержится ранее не отснятый регион подстилающей поверхности. Регион из Кпред смещается в нижнюю часть кадра. Такое перемещение можно описать как сдвиг в евклидовом пространстве и осуществить с помощью матрицы A размерностью 2 на 2 элемента.

$$A = \begin{pmatrix} \Delta x & 0 \\ 0 & \Delta y \end{pmatrix} \tag{1}$$

Следующим образом:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \Delta x & 0 \\ 0 & \Delta y \end{pmatrix} * \begin{pmatrix} x \\ y \end{pmatrix} \tag{2}$$

Для прямого полета БЛА без крена, тангажа и поворотов, такой матрицы достаточно для описания необходимого преобразования одного кадра к другому, с целью образования разностного кадра, в котором отсутствует временная избыточность. Но, к сожалению, в ходе полета летательного аппарата неизбежно происходят крены, тангажи и повороты из-за маневров, связанных с корректировкой траектории. В общем случае кадры могут быть искривлены друг относительно друга проективно. Тогда необходимо формирование матрицы A размерностью 3 на 3 элемента, а также переход к однородным координатам. Переход к однородным координатам на двумерном пространстве достаточно прост, необходимо к исходным координатам добавить просто еще один единичный элемент [4]. Тогда уравнение (2) примет вид:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{3}$$

Для проецирования координат, полученных в (3) на ту же плоскость, в которой лежат исходные координаты, необходимо:

$$x'' = \frac{x'}{z'},\ y'' = \frac{y'}{z'} \tag{4}$$

Таким образом для устранения временной избыточности необходимо сформировать матрицу преобразования A для Кпред к Ктек, на основе преобразованного предыдущего кадра и текущего вычислить разностный кадр, применить к нему любой из необходимых алгоритмов устранения пространственной избыточности и сжатия. Разностный кадр таким образом содержит в себе и новую информацию, вошедшую в кадр и сигнал ошибки преобразования.

Отдельной задачей стоит формирование преобразованного Кпред. Очевидно матрица A содержит дробные значения, а значит вычисленные координаты будут иметь также дробные значения. Изображение же, в своей сущности состоит из целых значений координат пикселей. Для того чтобы вычислить значения элементов изображения, лежащих по целым значениям координат, можно применить следующее преобразование к (3):

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}^{-1} * \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} \tag{5}$$

Тогда, для вычисления значения элемента с целыми координатами в преобразованном изображении, необходимо вычислить координаты этого элемента в исходном изображении и, в случае, если эти координаты дробные, вычислить элемент с помощью суммы взвешенных соседних элементов. Собственно, единственной проблемой теперь стоит нахождение матрицы A.

Первичное искажение кадра

Для вычисления матрицы A необходимо воспользоваться дополнительной информацией, поступающей к видеокамере с САУ БЛА. Эти данные называются телеметрическими. В данных содержатся поля, в которых отображаются текущие координаты фюзеляжа, углы крена (α), тангажа (β), магнитного склонения (γ). На основе разности телеметрии, зафиксированными на момент экспозиции Кпред и Ктек можно вычислить приращение координат и углов за время между кадрами. На основе этого приращения можно вычислить относительные координаты углов снимка [5]:

$$(x' \quad y' \quad h') = (\Delta_x \quad \Delta_y \quad \Delta_h) + \\ + \left(\frac{h}{\langle\langle(0 \quad 0 \quad -1),(M*R)\rangle*M*R\rangle}\right) \tag{6}$$

Где Δ_x, Δ_y, Δ_z — приращения координат между двумя кадрами, h — высота снимка над подстилающей поверхностью. M — матрица поворота, сформированная на основе данных об углах наклона фюзеляжа БЛА:

$$M = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix} * \\ * \begin{pmatrix} \cos(\beta) & 0 & \sin(\beta) \\ 0 & 1 & 0 \\ -\sin(\beta) & 0 & \cos(\beta) \end{pmatrix} * \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix} \tag{7}$$

R — вектор содержащий данные об углах, относительно нормали, луч наклоненный на которые при пересечении с подстилающей поверхностью даст координаты точки.

$$(a_1 \quad a_2 \quad a_3) = (\sin(\theta) \quad -\cos(\varphi) * \\ * \cos(\theta) \quad \sin(\varphi) * \cos(\theta)) \tag{8}$$

Где θ, φ — половина угла обзора объектива камеры. Так как ноль системы координат лежит на оптической оси объектива, для вычисления координат углов снимка необходимо лишь сменить знаки углов в (8). Тогда, согласно [6], нужно найти такую матрицу А, чтобы:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = A * \begin{pmatrix} x \\ y \\ w \end{pmatrix} \tag{9}$$

При этом, вектора координат кадров однородные, что значит, что формально вектор с результирующими координатами не равен преобразованному вектору исходных. Эти векторы сонаправлены, но могут отличаться на ненулевой масштабный коэффициент. Для ликвидации этого эффекта, уравнение (9) можно выразить в форме векторного произведения:

$$\left\langle \begin{pmatrix} x' \\ y' \\ w' \end{pmatrix}, A * \begin{pmatrix} x \\ y \\ w \end{pmatrix} \right\rangle = 0 \tag{10}$$

Раскрывая операцию векторного произведения, результирующее уравнение принимает вид (11).

В данной системе уравнений, только первые два уравнения линейно независимы. Согласно сложившейся практике, третье уравнение отбрасывается. Исходных точек изображения четыре, результирующих также четыре, а исходя из (11), коэффициентов в векторе девять. Очевидно, что решение системы, где девятый коэффициент равен нулю не интересен, поэтому элемент a9 приравнивается единице [6]. Данное приравнивание имеет смысл только тогда, когда, суммарный угол крена или тангажа с половиной угла обзора камеры превышает девяносто градусов к нормали, то есть когда в кадре появляется небо, для которого понятия геокоординат не определены. Тогда (11) примет вид (12).

Далее необходимо сформировать систему уравнений, в которой преобразованным и исходным координатам соответствуют результаты (6) и решить ее для А. При добавлении единичного элемента в вектор А, с помощью которого и было найдено решение и перекомпоновке в матрицу и получается искомая матрица первичного преобразования изображения:

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & 1 \end{pmatrix} \tag{13}$$

Уточнение искажения кадра

В случае идеальной фиксации телеметрии и идеальной точности, этих преобразований достаточ-

$$\begin{pmatrix} 0 & 0 & 0 & -w'x & -w'y & -w'w & y'x & y'y & y'w \\ w'x & w'y & w'w & 0 & 0 & 0 & -x'x & -x'y & -x'w \\ -y'x & -y'y & -y'w & x'x & x'y & x'w & 0 & 0 & 0 \end{pmatrix} * \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \\ a_7 \\ a_8 \\ a_9 \end{pmatrix} = 0 \tag{11}$$

$$\begin{pmatrix} 0 & 0 & 0 & -w'x & -w'y & -w'w & y'x & y'y \\ w'x & w'y & w'w & 0 & 0 & 0 & -x'x & -x'y \end{pmatrix} * \tilde{A} = \begin{pmatrix} -y'w \\ x'w \end{pmatrix} \tag{12}$$

но, чтобы в области пересечения соседних кадров разность их равнялась нулю. Но, к сожалению, такая ситуация крайне маловероятна, а значит необходима корректировка наложений кадров. Выполнить ее можно с помощью инверсного алгоритма Лукаса-Канаде. Применяется инверсный алгоритм, так как в нем единожды, на одну переменную, считается матрица Гессе. Но так как алгоритм Лукаса-Канаде является частным случаем применения итерационного численного метода Ньютона, количество итераций, а значит, и вычислений может оказаться настолько большим, что не успеет выполниться за время, в течении которого кадр может находиться в оперативной памяти обработчика. Поэтому предлагается применить пирамидальное разложение преобразуемого кадра на изображения меньшего разрешения и первые, наибольшие, с точки зрения инкремента коэффициентов, шаги, производить с изображением малого разрешения. Для этого к Кпред и Ктек применяются матрицы следующего вида:

$$S = \begin{pmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (14)$$

Где s — константа, обратная коэффициенту уменьшения изображения. В результате получаются изображения меньшего разрешения. От Ктек, с помощью оператора Собеля [6] берется частный градиент по горизонтальной составляющей и вертикальной (gradx, grady). Оператор Собеля в данном случае применяется как оператор взятия производной по направлению.

После этого действия для каждого пикселя изображения формируется матрица Якоби для всех переменных, по которым осуществляется поиск решения и для всех составляющих, по которым выполнена функция градиента. В случае текущего применения, матрица Якоби обладает размерностью 8x2 и сформирована следующим образом:

$$J = \begin{pmatrix} y & x & 1 & 0 & 0 & 0 & -yy & -yx \\ 0 & 0 & 0 & y & x & 1 & -xy & -xx \end{pmatrix} \quad (15)$$

Где y, x — координаты пикселя по вертикали и горизонтали соответственно. Далее с каждым же пикселем формируется восемь изображений, элементами которого становятся суммы произведений элементов градиентов, на соответствующие элементы матрицы Якоби:

$$P(x,y,z) = grad_x(x,y) * J[1,z] +$$

$$+ grad_y(x,y) * J[2,z] \quad (16)$$

Теперь, на основе изображений P, можно построить матрицу Гессе, состоящую из вторых частных производных Ктек и вычисляемую как:

$$Hess = \begin{pmatrix} P_1P_1 & \cdots & P_1P_8 \\ \vdots & \ddots & \vdots \\ P_8P_1 & \cdots & P_8P_8 \end{pmatrix} \quad (17)$$

Где P_z — сумма всех пикселей изображения $P(x, y, z)$. На этом подготовительный этап работы с изображениями завершен. Далее происходят итерации, в ходе которых метод оптимизации находит минимум разности пересекающихся участков изображений и, соответственно формирует матрицу параметров B. Для этого в области пересечения считаются параметры ошибки E_z:

$$E_z = \sum_{1,1}^{x,y} (\text{Кпред}(x,y) - \text{Ктек}(x,y)) * P_z \quad (18)$$

А коэффициенты матрицы B:

$$B = Hess^{-1} * E \quad (19)$$

К B — в свою очередь, добавляется элемент b_{33}, равный единице, применяется перекомпоновка, и получается матрица, аналогичная полученной в (13). После применения этой матрицы к Кпред, происходит очередное формирование параметров ошибки. Итерации происходят до того момента, пока среднее значение параметра ошибки не понизится ниже предела, определяемого эмпирически.

В уравнении (19) получилась матрица B, которая при применении ее к уменьшенной копии $K_{пред}$ преобразовывает его к уменьшенной копии $K_{тек}$. Теперь необходимо произвести переход к полноразмерным кадрам. Для этого необходимо проследить осуществленные преобразования:

$$\text{Кпред} \xrightarrow{S} \text{Кпред мал} \xrightarrow{B} \text{Кпреобр} \xleftarrow{S} \text{Ктек}$$

Согласно вышесказанному, для того, чтобы получить матрицу перехода от $K_{пред}$ к $K_{тек}$ необходимо:

$$\text{Ктек} = G * \text{Кпред} \quad (20)$$

Где:

$$G = S * B * S^{-1} \quad (21)$$

Таким образом матрица преобразования, совмещающая в себе преобразования, полученные как за первичный шаг, так и за последующую оптимизацию примет вид:

$$F = A * S * B * S^{-1} \quad (22)$$

Заключение

Применяя это преобразование к предыдущему кадру видеопоследовательности, получается кадр, вычитание которого из текущего приведет к обнулению пересекающейся зоны, за исключением объектов, переместившихся за время между кадрами на величины более одного пикселя. В неперекрывающейся зоне будет присутствовать новая информация. В таком случае в потоке данных нет необходимости использовать ключевые кадры, в каждом кадре есть новый участок данных. Для компрессии результирующих данных, автор рекомендует использовать кодеры, основанные на вейвлет-декомпозиции, в силу их особенности к делению пространственных частот. В таком случае даже если оба алгоритма совершили ошибки, и пересекающиеся участки не совпали полностью, информация, содержащая ошибки несовпадения, окажутся лишь в одном сегменте декомпозиции, что позволит добиться существенного сжатия всех остальных сегментов.

ЛИТЕРАТУРА

1. Браммер Ю. А. Радиотехника Издание 2. М.: Высшая школа, 1969
2. William B. Pennebaker; Joan L. Mitchell. JPEG still image data compression standard. Springer, 1993.
3. ISO. ISO/IEC11172–2:1993 — Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s — Part 2: Video. 1993
4. Р. Хартсхорн. Основы проективной геометрии. М.: Мир, 1970.
5. С.М. Вишняков, Н. В. Фильченко. Оценка точности определения координат наземных объектов средствами оптико-электронного контроля на беспилотных летательных аппаратах малой дальности //Успехи современной радиоэлектроники № 7 за 2015
6. Hartley R. Multiple View Geometry in Computer Vision. Cambridge University Press 2000.