

# ОПИСАНИЕ ПОЛЕЗНОЙ МОДЕЛИ К ПАТЕНТУ

(12)

РЕСПУБЛИКА БЕЛАРУСЬ



НАЦИОНАЛЬНЫЙ ЦЕНТР  
ИНТЕЛЛЕКТУАЛЬНОЙ  
СОБСТВЕННОСТИ

(19) ВУ (11) 8206

(13) U

(46) 2012.04.30

(51) МПК

H 04N 1/64 (2006.01)

## (54) ВИДЕОКОДЕК С ОБЪЕКТНОЙ КОМПЕНСАЦИЕЙ ДВИЖЕНИЯ ВИДЕОКАМЕРЫ ДЛЯ СЖАТИЯ ВИДЕОДАНЫХ

(21) Номер заявки: u 20110853

(22) 2011.10.31

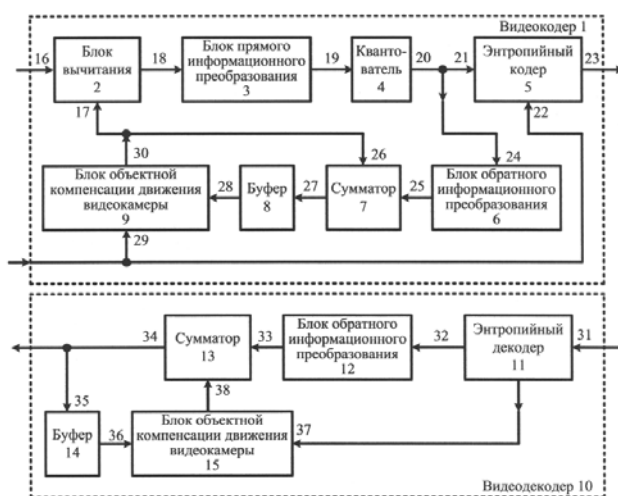
(71) Заявитель: Учреждение образования  
"Белорусский государственный уни-  
верситет информатики и радиоэлек-  
троники" (ВУ)

(72) Авторы: Конопелько Валерий Кон-  
стантинович; Цветков Виктор Юрье-  
вич; Аль-Джубури Тарик Мохаммед  
Сальман; Аль-Фурайджи Одай Джа-  
сим Мохаммед (ВУ)

(73) Патентообладатель: Учреждение образо-  
вания "Белорусский государственный  
университет информатики и радиоэ-  
лектроники" (ВУ)

(57)

Видеокодек с объектной компенсацией движения видеокамеры для сжатия видеоданных, состоящий из видеокодера, включающего блок вычитания, блок прямого информационного преобразования, квантователь, энтропийный кодер, блок обратного информационного преобразования, сумматор, буфер, блок объектной компенсации движения видеокамеры, и видеодекодера, включающего энтропийный декодер, блок обратного информационного преобразования, сумматор, буфер, блок объектной компенсации движения видеокамеры, причем первый вход блока вычитания является информационным входом видеокодера, второй вход блока объектной компенсации движения видеокамеры в видеокодере и второй вход энтропийного кодера являются управляющим входом видеокодера, выход энтропийного кодера является выходом видеокодера, вход энтропийного



Фиг. 1

ВУ 8206 U 2012.04.30

декодера является входом видеodeкодера, выход сумматора и вход буфера в видеodeкоде-ре являются выходом видеodeкодера, **отличающийся** тем, что содержит блоки объектной компенсации движения видеокамеры, первые входы которых соединены с выходами буферов, вторые входы используются для подачи информации о движении видеокамеры, а выходы соединены со вторыми входами сумматоров.

(56)

1. Gupta K., Kukarni A.V. Implementation of an Automated Single Camera Object Tracking System Using Frame Differencing and Dynamic Template Matching // Springerlink-ACISE. - 2008. - P. 245-250.

2. Iain R. H.264 and MPEG-4 Video Compression and Video Coding for Next-generation Multimedia // The Robert Gordon University. Aberdeen. UK., England. - 2003. - P. 159-162.

---

Полезная модель относится к области передачи изображений и может быть использована для эффективного кодирования видеоинформации при передаче по каналам связи. Целью полезной модели является повышение коэффициента сжатия и качества восстановления видеоинформации, формируемой движущейся видеокамерой. Данная цель достигается введением блока объектной компенсации движения видеокамеры.

Известен видеокодек, использующий кадровую разность для сжатия видеоданных и состоящий из видеокодера, включающего блок вычитания, векторный квантователь, энтропийный кодер (кодирование с переменной длиной), сумматор и буфер, и видеodeкодера, включающего энтропийный декодер, сумматор и буфер [1]. Однако данный кодек не эффективен для сжатия видеоданных, формируемых движущейся камерой, из-за отсутствия компенсации различных смещений изображений объектов на текущем кадре относительно предыдущего, возникающих при перемещении видеокамеры, зависящих от расстояния до этих объектов и обусловленных параллаксом.

Наиболее близким к предлагаемой полезной модели является видеокодек, использующий блочную компенсацию движения для сжатия видеоданных и состоящий из видеокодера, включающего блок вычитания, блоки прямого и обратного информационных преобразований, квантователь, энтропийный кодер, сумматор, буфер, блок формирования прогнозного кадра на основе блочной компенсации движения и блок оценки движения, и видеodeкодера, включающего энтропийный декодер, блок обратного информационного преобразования, сумматор, буфер, блок формирования прогнозного кадра на основе блочной компенсации движения [2]. Однако данный кодек не эффективен для сжатия видеоданных, формируемых движущейся камерой, из-за равномерного разбиения формируемого прогнозного кадра на блоки без учета глубины сцены и границ изображений объектов на прогнозном и опорном кадрах.

Техническая задача, на решение которой направлена полезная модель, - создание ко-дека для сжатия видеоданных, поступающих от движущейся камеры, с использованием объектной компенсации движения видеокамеры, учитывающей структуру сцены и параллак-с. Техническим результатом, который может быть получен при использовании данной полезной модели, является повышение коэффициента сжатия и качества восстановления видеоданных, формируемых движущейся камерой.

Видеокодек с объектной компенсацией движения видеокамеры для сжатия видеоданных, состоящий из видеокодера, включающего блок вычитания, блок прямого информационного преобразования, квантователь, энтропийный кодер, блок обратного информационного преобразования, сумматор, буфер, блок объектной компенсации движения видеокамеры, и видеodeкодера, включающего энтропийный декодер, блок обратного информационного преобразования, сумматор, буфер, блок объектной компенсации

движения видеокамеры, причем первый вход блока вычитания является информационным входом видеокодера, второй вход блока объектной компенсации движения видеокамеры в видеокодере и второй вход энтропийного кодера являются управляющим входом видеокодера, выход энтропийного кодера является выходом видеокодера, вход энтропийного декодера является входом видеодекодера, выход сумматора и вход буфера в видеокодере являются выходом видеодекодера, отличается тем, что содержит блоки объектной компенсации движения видеокамеры, первые входы которых соединены с выходами буферов, вторые входы используются для подачи информации о движении видеокамеры, а выходы соединены со вторыми входами сумматоров.

Сущность заявляемой полезной модели заключается в том, что предлагаемый видеокодек позволяет эффективно сжать и качественно восстановить видеоданные, формируемые движущейся видеокамерой, за счет использования объектной компенсации движения видеокамеры, основанной на анализе структуры трехмерной сцены и учитывающей параллакс.

Предложение иллюстрируется следующими чертежами. На фиг. 1 представлена структурная схема видеокодека с объектной компенсацией движения видеокамеры, на фиг. 2 - структура блока объектной компенсации движения видеокамеры, на фиг. 3 - структура блока объектно-ориентированного предсказания, на фиг. 4 - сегментация опорного кадра, на фиг. 5 - выравнивание сегментированного опорного кадра параллельно линии горизонта, на фиг. 6 - монокулярная оптическая модель, на фиг. 7 - объектная декомпозиция опорного кадра видеоданных, на фиг. 8 - определение расстояния до объекта по опорному кадру с помощью монокулярной оптической модели, на фиг. 9 - определение смещений пикселей опорного кадра при движении камеры с помощью бинокулярных оптических моделей, на фиг. 10 - уточнение расстояний до объектов по прогнозному кадру с помощью монокулярных оптических моделей, на фиг. 11 - неопределенности заслонения и сжатия при формировании прогнозного кадра, на фиг. 12 - возникновение неопределенного пикселя на прогнозном кадре из-за растяжения изображения объекта при движении видеокамеры, на фиг. 13 - пример возникновения неопределенности появления на прогнозном кадре, на фиг. 14 - два соседних кадра видеопоследовательности, полученной с помощью движущейся видеокамеры, на фиг. 15 - параметры установки и перемещения видеокамеры, на фиг. 16 - синтезированный прогнозный кадр, на фиг. 17 - структура потока на выходе видеокодера и входе видеодекодера, на фиг. 18 - соседние, прогнозные и разностные кадры.

Видеокодек с объектной компенсацией движения видеокамеры состоит из видеокодера 1, включающего блок вычитания 2, блок прямого информационного преобразования 3, квантователь 4, энтропийный кодер 5, блок обратного информационного преобразования 6, сумматор 7, буфер 8, блок объектной компенсации движения видеокамеры 9, и видеодекодера 10, включающего энтропийный декодер 11, блок обратного информационного преобразования 12, сумматор 13, буфер 14, блок объектной компенсации движения видеокамеры 15, причем первый вход 16 блока вычитания 2 является информационным входом видеокодера 1, второй вход 29 блока объектной компенсации движения видеокамеры 9 и второй вход 22 энтропийного кодера 5 образуют управляющий вход видеокодера 1, выход 23 энтропийного кодера 5 является выходом видеокодера 1, вход 31 энтропийного декодера 11 является входом видеодекодера 10, выход 34 сумматора 13 является выходом видеодекодера 10 (фиг. 1). Блок объектной компенсации движения видеокамеры 9 в видеокодере 1 и блок объектной компенсации движения видеокамеры 15 в видеодекодере 10 одинаковы и включают в свой состав блок сегментации 39, блок компенсации крена камеры 56, блок объектной декомпозиции 40, блок объектно-ориентированного предсказания 41, блок предсказания положения линии горизонта 42, причем вход 43 блока сегментации 39 и вход 50 блока объектно-ориентированного предсказания 41 образуют информационный вход блока объектной компенсации движения видеокамеры 9 (или бло-

ка объектной компенсации движения видеокамеры 15), второй вход 70 блока компенсации крена камеры, вход 49 блока предсказания положения линии горизонта 42 и второй вход 47 блока объектно-ориентированного предсказания 41 образуют управляющий вход блока объектной компенсации движения видеокамеры 9 (или блока объектной компенсации движения видеокамеры 15), выход 48 блока объектно-ориентированного предсказания 41 является выходом блока объектной компенсации движения видеокамеры 9 (или блока объектной компенсации движения видеокамеры 15) (фиг. 2). Блок объектно-ориентированного предсказания 41 включает в свой состав блок определения расстояний по опорному кадру 51, блок определения смещений пикселей 52, блок определения расстояний по прогнозному кадру 53, блок формирования прогнозного кадра 54, блок интерполяции пикселей прогнозного кадра 55, блок вычисления средней яркости опорного кадра 57, блок предсказания пикселей прогнозного кадра 58, причем вход 59 блока определения расстояний по опорному кадру 51 является первым информационным входом блока объектно-ориентированного предсказания 41, второй вход 60 блока определения расстояний по опорному кадру 51, второй вход 62 блока определения смещений пикселей 52 и второй вход 64 блока определения расстояний по прогнозному кадру 53 образуют управляющий вход блока объектно-ориентированного предсказания 41, первый вход 69 блока интерполяции пикселей прогнозного кадра 55 и вход 74 блока вычисления средней яркости опорного кадра 57 образуют второй информационный вход блока объектно-ориентированного предсказания 41, выход 73 блока предсказания пикселей прогнозного кадра 58 является выходом блока объектно-ориентированного предсказания 41 (фиг. 3).

В предложенном видеокодеке для достижения высокого коэффициента сжатия видеоинформации, поступающей от движущейся видеокамеры, используется объектная компенсация движения видеокамеры, основанная на отделении изображений объектов от фона на опорном кадре и объектно-ориентированном предсказании следующего кадра по управляющей информации о перемещении видеокамеры, поступающей от системы контроля движения (например, сервопривода).

Для отделения изображений объектов опорного кадра от фона используются сегментация, компенсация крена видеокамеры, определение положения линии горизонта на опорном кадре и объектная декомпозиция сегментированного опорного кадра относительно линии горизонта.

Сегментация ставит в соответствие  $t$ -му ( $t \in [0, T-1]$ ) опорному кадру  $M_F(t) = \|m_F(y, x, t)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$  сегментированный опорный кадр

$M_S(t) = \|m_S(y, x, t)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$ , значение каждого пиксела  $m_S(y, x, t) \in [1, S]$  которого

указывает на номер сегмента, к которому данный пиксел принадлежит, где  $T$  - период;  $m_F(y, x, t) \in [0, I-1]$  - пиксели опорного кадра;  $Y \times X$  - размер кадра;  $S$  - число сегментов,  $S \leq Y \times X$ ;  $I$  - число градаций яркости пикселей опорного кадра. В качестве примера на фиг. 4 представлены фрагмент опорного кадра размером  $7 \times 9$  пикселей и результат его сегментации. Для сегментации может использоваться любой известный метод, например выращения областей, разделения и слияния областей, водораздела и т.д.

Для объектной декомпозиции необходимо компенсировать возможный крен видеокамеры. Для этого сегментированный опорный кадр  $M_S(t)$  выравнивается параллельно линии горизонта за счет поворота относительно центра на угол  $\beta_Z$  крена видеокамеры, в результате чего формируется выравненный сегментированный опорный кадр  $M_C(t) = \|m_C(y_C, x_C, t)\|_{(y_C=0, \overline{Y_C-1}, x_C=0, \overline{X_C-1})}$ , где  $m_C(y_C, x_C, t)$  - пиксели выравненного сегментированного опорного кадра,  $m_C(y_C, x_C, t) = m_S(y, x, t)$ ;  $y = [r_C \cdot \sin(\beta_Z + \alpha_C)] + Y/2$ ,  $x = [r_C \cdot \cos(\beta_Z + \alpha_C)] + X/2$  - координаты пикселей сегментированного опорного кадра;  $[\cdot]$  - операция округления до ближайшего целого;  $\alpha_C = \arctg((y_C - Y_C/2)/(x_C - X_C/2))$ ,

$r_C = \sqrt{(y_C - Y_C/2)^2 + (x_C - X_C/2)^2}$  - полярный угол и полярный радиус пиксела выравненного сегментированного опорного кадра;  $Y_C \times X_C$  - размер выравненного сегментированного опорного кадра,  $Y_C = \lceil r \cdot \sin(\alpha + \beta_Z) \rceil$ ,  $X_C = \lceil r \cdot \cos(\alpha + \beta_Z) \rceil$ ;  $\lceil \cdot \rceil$  - операция округления до ближайшего целого с избытком;  $\alpha = \arctg(Y/X)$ ,  $r = \sqrt{(Y/2)^2 + (X/2)^2}$  - полярные координаты углового пиксела сегментированного опорного кадра. В качестве примера на фиг. 5 представлены фрагмент сегментированного опорного кадра с наложенной на него линией горизонта, ориентация этого фрагмента относительно линии горизонта и результат его выравнивания.

Для определения положения линии горизонта на опорном кадре используется монокулярная оптическая модель, в которой оптическая система видеокамеры представляется тонкой линзой и фотоприемной матрицей, а поверхность, на которой расположены объекты, предполагается плоской (фиг. 6). Смещение  $Y_G$  проекции линии горизонта относительно нижней границы выравненного сегментированного опорного кадра  $M_C(t)$  определяется исходя из заданного расстояния  $L_G$  до линии горизонта с помощью выражения  $Y_G = \lceil Y_C/2 - Ftg(\alpha_Z/2 - \beta) \rceil$ , где  $F = \frac{Y_C/2}{tg(\alpha_Z/2)}$  - фокусное расстояние;  $\alpha_Z$  - угол видимости камеры при крене на угол  $\beta_Z$ , соответствующий разрешению  $Y_C$  пикселей;

$(\alpha_Z/2 - \beta) = \arctg\left(\frac{H}{L_G}\right) - \beta_X$ ;  $\beta$  - угол, соответствующий смещению  $Y_G$  линии горизонта относительно нижней границы выравненного сегментированного опорного кадра;  $\beta_X$  - угол тангажа видеокамеры;  $H$  - высота установки видеокамеры.

Объектная декомпозиция ставит в соответствие  $t$ -му выравненному сегментированному опорному кадру  $M_C(t)$  объектный опорный кадр  $M_D(t) = \|m_D(y, x, t)\|_{(y=0, \overline{Y_C-1}, x=0, \overline{X_C-1})}$ , значение каждого пиксела  $m_D(y, x, t) \in [1, D+1]$  которого указывает на фон (если  $m_D(y, x, t) = 1$ ) или номер изображения объекта (если  $m_D(y, x, t) > 1$ ), к которому данный пиксел принадлежит, где  $D$  - число объектов, выделяемых на опорном кадре. Для определения значений пикселей объектного опорного кадра  $M_D(t)$  формируются векторы  $Y_{MIN} = \|y_{MIN}(S)\|_{(s=1, S)}$  и  $Y_{MAX} = \|y_{MAX}(S)\|_{(s=1, S)}$   $y$ -координат крайних нижнего и верхнего пикселей для каждого сегмента опорного кадра  $M_C(t)$ , элементы  $y_{MIN}(s)$  и  $y_{MAX}(s)$  которых вычисляются с помощью выражений

$$\forall y(y = \overline{0, Y_C - 1}) \forall x(x = \overline{0, X_C - 1}) ((m_C(y, x, t) = s) \wedge (y < y_{MIN}(S))) \rightarrow (y_{MIN}(S) = y) \text{ и}$$

$$\forall y(y = \overline{0, Y_C - 1}) \forall x(x = \overline{0, X_C - 1}) ((m_C(y, x, t) = s) \wedge (y > y_{MAX}(S))) \rightarrow (y_{MAX}(S) = y).$$

Данные векторы используются для нахождения сегментов, пересекающих проекцию линии горизонта. Такие сегменты интерпретируются как изображения объектов, а остальные - как фон. Пикселы фона определяются с помощью выражения  $\forall y(y = \overline{0, Y_C - 1}) \forall x(x = \overline{0, X_C - 1}) ((y_{MIN}(m_C(y, x, t)) > Y_G + \Delta_G) \vee (y_{MAX}(m_C(y, x, t)) < Y_G - \Delta_G)) \rightarrow (m_D(y, x, t) = 1)$ , где  $\Delta_G$  - задаваемое допустимое отклонение положения проекции линии горизонта. Для остальных пикселей объектного опорного кадра  $M_D(t)$ , относящихся к изображениям объектов, сохраняются значения соответствующих пикселей выравненного сегментированного опорного кадра  $M_C(t)$  согласно выражению  $\forall y(y = \overline{0, Y_C - 1}) \forall x(x = \overline{0, X_C - 1}) ((y_{MIN}(m_C(y, x, t)) < Y_G + \Delta_G) \wedge (y_{MAX}(m_C(y, x, t)) > Y_G - \Delta_G)) \rightarrow (m_D(y, x, t) = m_C(y, x, t))$ . Исключение составляют пикселы, относящиеся к  $s$ -м сегментам, площадь  $qC(s)$  которых превышает заданное пороговое значение  $Q_{MAX}$ . Такие пикселы определяются как фон с помощью выражения  $\forall y(y = \overline{0, Y_C - 1}) \forall x(x = \overline{0, X_C - 1}) \times$

$\times ((q_c(m_c(y, x, t)) > Q_{\max})) \rightarrow (m_D(y, x, t) = 1)$ , где  $q_c(s)$  - элемент матрицы  $Q_c = \|q_c(s)\|_{(s=\overline{1, S})}$ , определяемый с помощью выражения  $\forall s(s = \overline{1, S-1}) \forall y(y = \overline{0, Y_c-1}) \forall x(x = \overline{0, X_c-1}) \times$   
 $\times (m_c(y, x, t) = s) \rightarrow \left( q_c(s) = \sum_{y=0}^{Y_c-1} \sum_{x=0}^{X_c-1} m_c(y, x, t) \right)$ . В качестве примера на фиг. 7 представлены

результаты объектной декомпозиции опорного кадра, включающего три сегмента, два из которых (сегменты 1 и 2) пересекают проекцию линии горизонта и интерпретируется как изображения объектов, а третий сегмент не пересекает проекцию линии горизонта и интерпретируется как фон.

Объектно-ориентированное предсказание прогнозного кадра видеоданных основано на использовании результатов объектной декомпозиции опорного кадра и управляющей информации о перемещении видеокамеры, поступающей от системы контроля движения. Суть объектно-ориентированного предсказания состоит в вычислении смещения элементов изображения (сегментов и отдельных пикселей) на прогнозном кадре видеоданных относительно опорного кадра в зависимости от расстояния от видеокамеры до объектов, соответствующих этим элементам, а также параметров видеокамеры, высоты ее установки, направления, характера и скорости движения. Объектно-ориентированное предсказание включает определение расстояний до объектов с использованием опорного кадра; определение смещений пикселей опорного кадра на прогнозном кадре, обусловленных перемещением видеокамеры; уточнение расстояний до объектов с использованием прогнозного кадра; размещение пикселей опорного кадра в поле пикселей прогнозного кадра; интерполяцию неопределенных пикселей прогнозного кадра; предсказание неопределенных пикселей прогнозного кадра.

Для определения расстояния до объектов по опорному кадру используется монокулярная оптическая модель (фиг. 8), учитывающая характеристики оптической системы (угол  $\alpha_z$  видимости и разрешение  $Y_c$  по вертикали с учетом крена видеокамеры), высоту  $H$  установки, угол  $\beta_x$  тангажа видеокамеры и смещения  $Y_{\min}(s)$  сегментов относительно нижней границы опорного кадра  $M_c(t)$ . С помощью монокулярной оптической модели формируется матрица  $\tilde{D}_Z(t) = \|\tilde{d}_Z(y, x, t)\|_{(y=\overline{0, Y_c-1}, x=\overline{0, X_c-1})}$  расстояний от видеокамеры до объектов, значения элементов  $\tilde{d}_Z(y, x, t)$  которой вычисляются с помощью выражения  $\tilde{d}_Z(y, x, t) =$   
 $= H / \text{tg}(\alpha_z / 2 + \beta_x - \beta)$ , где  $\beta = \begin{cases} \alpha_z / 2 - \text{arctg}((Y_c / 2 - Y_{\min}(m_D(y, x, t))) / F) & \text{при } m_D(y, x, t) > 1, \\ \alpha_z / 2 - \text{arctg}((Y_c / 2 - y) / F) & \text{при } m_D(y, x, t) = 1 \end{cases}$ .

На основе матрицы  $\tilde{D}_Z(t)$  формируется матрица  $D_Z(t) = \|d_Z(y, x, t)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$  расстояний в результате поворота матрицы  $\tilde{D}_Z(t)$  на угол минус  $\beta_z$  крена видеокамеры. В результате пикселям опорного кадра  $M_F(t)$ , относящимся к одному объекту, ставится в соответствие одинаковое расстояние, а пикселям фона - различное расстояние, в зависимости от их смещения относительно нижней границы опорного кадра  $M_F(t)$ .

При перемещении видеокамеры происходит смещение и видоизменение изображений объектов трехмерного пространства, обусловленное явлением параллакса. Для вычисления смещений пикселей опорного кадра  $M_F(t)$  на прогнозном кадре используется бинокулярная оптическая модель, рассматривающая два положения видеокамеры: в момент времени  $t$ , соответствующий опорному кадру, и в момент времени  $(t + 1)$ , соответствующий прогнозному кадру. Бинокулярная оптическая модель учитывает два вида смещений пикселей относительно опорного кадра (вертикальное и горизонтальное), соответствующие трем видам движения видеокамеры (фиг. 9): вертикальному (вдоль оси  $y$ ), горизонтальному (вдоль оси  $x$ ) и продольному (перпендикулярно плоскости фотоприемной матрицы). С помощью бинокулярной оптической модели для каждого пикселя опорного

кадра  $M_F(t)$  рассчитываются значения элементов матриц  $S_{DY} = \|s_{DY}(y, x)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$  и  $S_{DX} = \|s_{DX}(y, x)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$  проекционного предсказания вертикального и горизонтального смещений пикселей при сдвиге и повороте видеокамеры:  $s_{DY}(y, x) = [s_{SY}(y, x) + s_{RY}(y, x)]$ ,  $s_{DX}(y, x) = [s_{SX}(y, x) + s_{RX}(y, x)]$  при  $y = \overline{0, Y-1}$ ,  $x = \overline{0, X-1}$ , где  $s_{SY}(y, x) = [s_{SYX}(y, x) + s_{SYZ}(y, x)]$ ,  $s_{SX}(y, x) = [s_{SXX}(y, x) + s_{SXZ}(y, x)]$  - составляющие, обусловленные сдвигом камеры;  $s_{RY}(y, x) = [s_{RYX} + s_{RYZ}(y, x)]$ ,  $s_{RX}(y, x) = [s_{RXY} + s_{RXZ}(y, x)]$  - составляющие, обусловленные поворотом камеры;  $s_{SYX}(y, x) = F_Y S_Y / (d_Z(y, x) \cos \gamma_X)$  - прогнозируемое вертикальное смещение при вертикальном сдвиге камеры;  $s_{SXX}(y, x) = F_X S_X / (d_Z(y, x) \cos \gamma_X)$  - прогнозируемое горизонтальное смещение при горизонтальном сдвиге камеры;  $s_{SYZ}(y, x) = y S_Z / (d_Z(y, x) \cos \gamma_X - S_Z)$ ,  $s_{SXZ}(y, x) = x S_Z / (d_Z(y, x) \cos \gamma_X - S_Z)$  - прогнозируемые вертикальное и горизонтальное смещения при продольном сдвиге камеры;  $s_{RYX} = \beta_X Y / \alpha_Y$  - прогнозируемое вертикальное смещение при тангаже камеры;  $s_{RYZ}(y, x) = (y - Y/2) - R \sin \beta_Z$ ,  $s_{RXZ}(y, x) = (x - X/2) - R \cos \beta_Z$  - прогнозируемые вертикальное и горизонтальное смещения при крене камеры;  $R = \sqrt{(x - X/2)^2 + (y - Y/2)^2}$ ;  $s_{RXY} = \beta_Y X / \alpha_X$  - прогнозируемое горизонтальное смещение при повороте камеры;  $F_Y = (Y/2) / \text{tg}(\alpha_Y / 2)$ ,  $F_X = (X/2) / \text{tg}(\alpha_X / 2)$  - вертикальный и горизонтальный виртуальные фокусы модели оптической системы;  $\gamma_X = \arctg \frac{x}{F_X}$ ;  $\alpha_Y, \alpha_X$  - углы

видимости камеры по вертикали и горизонтали.

При формировании прогнозного кадра необходимо учитывать, что продольное и горизонтальное поперечные перемещения видеокамеры приводят к изменению расстояний до объектов. Для определения расстояний до объектов, соответствующих пикселям прогнозного кадра с координатами  $(y', x')$ , используется монокулярная оптическая модель, на основе которой рассчитываются элементы матрицы  $D_C(t+1) = \|d_C(y, x, t+1)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$  дальномерной коррекции с помощью выражения  $d_C(y, x, t+1) = d_Z(y, x, t) \cos \gamma_X / \cos \gamma'_X$  при  $y = \overline{0, Y-1}$  и  $x = \overline{0, X-1}$ , где  $\gamma'_X = \arctg \frac{x'}{F_X}$ ;  $y' = y + s_{DY}(y, x)$  и  $x' = x + s_{DX}(y, x)$  - координаты пикселя прогнозного кадра, соответствующие пикселу  $(y, x)$  опорного кадра  $M_F(t)$  (фиг. 10).

При движении видеокамеры параллакс может привести к сжатию изображения объекта из-за его удаления или заслонению одних объектов другими. В этих случаях при формировании прогнозного кадра  $\hat{M}_F(t+1) = \|\hat{m}_F(y', x', t+1)\|_{(y'=0, \overline{Y-1}, x'=0, \overline{X-1})}$  не все пиксели опорного кадра  $M_F(t)$  однозначно отображаются на пиксели прогнозного кадра  $\hat{M}_F(t+1)$  и возникают неопределенности заслонения и сжатия, связанные с проецированием нескольких пикселей опорного кадра на один пиксел прогнозного кадра (фиг. 11). Для корректного размещения пикселей опорного кадра  $M_F(t)$  в поле пикселей прогнозного кадра  $\hat{M}_F(t+1)$  учитывается соотношение значений элементов матриц  $D_C(t+1)$  дальномерной коррекции и  $\hat{D}_Z(t+1) = \|\hat{d}_Z(y', x', t+1)\|_{(y'=0, \overline{Y-1}, x'=0, \overline{X-1})}$  дальномерного предсказания, которое определяется условиями  $\hat{d}_Z(y', x', t+1) > d_C(y, x, t+1)$  и  $\hat{d}_Z(y', x', t+1) = d_C(y, x, t+1)$ . Условие  $\hat{d}_Z(y', x', t+1) > d_C(y, x, t+1)$  имеет место, если пиксел  $\hat{m}_F(y', x', t+1)$  не определен (при инициализации  $\hat{d}_Z(y', x', t+1) = \infty$ ) или объект, отображенный на пиксел  $m_F(y, x, t)$  опорного кадра, заслоняет объект, отображенный на пиксел  $\hat{m}_F(y', x', t+1)$  прогнозного

кадра. В этом случае определяются значения элементов  $\hat{m}_F(y', x', t+1)$  прогнозного кадра  $\hat{M}_F(t+1)$ , матрицы  $\hat{D}_Z(t+1)$  дальномерного предсказания и матрицы  $\hat{M}_D(t+1) = \|\hat{m}_D(y', x', t+1)\|_{(y'=0, \overline{Y-1}, x'=0, \overline{X-1})}$  объектной декомпозиции прогнозного кадра с помощью выражений  $\hat{m}_F(y', x', t+1) = m_F(y, x, t)$ ,  $\hat{d}_Z(y', x', t+1) = d_C(y, x, t+1)$ ,  $\hat{m}_D(y', x', t+1) = m_D(y_C, x_C, t)$  для  $y = \overline{0, Y-1}$ ,  $x = \overline{0, X-1}$ . Условие  $\hat{d}_Z(y', x', t+1) = d_C(y, x, t+1)$  имеет место, если два или более пиксела опорного кадра проецируются на один пиксел прогнозного кадра. В этом случае определяются значения элементов  $\hat{m}_F(y', x', t+1)$  прогнозного кадра  $\hat{M}_F(t+1)$  с помощью выражения  $\hat{m}_F(y', x', t+1) = (\hat{m}_F(y', x', t+1) + \hat{m}_F(y, x, t))/2$ .

При движении видеокамеры возможно увеличение размера изображения объекта из-за его приближения (растяжение изображения). Это приводит к тому, что соседние пиксела опорного кадра не отображаются на соседние пиксела прогнозного кадра и образуются зоны неопределенности. Вычисление значений неопределенных пикселей возможно с использованием интерполяции. В качестве примера на фиг. 12 показано возникновение неопределенного пиксела на прогножном кадре из-за растяжения изображения объекта при продольном движении видеокамеры. Для интерполяции неопределенных пикселей прогнозного кадра  $\hat{M}_F(t+1)$  формируются матрицы  $D_{SY} = \|d_{SY}(y, x)\|_{(y=\overline{1, Y-1}, x=\overline{1, X-1})}$  и  $D_{SX} = \|d_{SX}(y, x)\|_{(y=\overline{1, Y-1}, x=\overline{1, X-1})}$  разностей вертикальных и горизонтальных смещений пикселей опорного кадра с использованием выражений  $d_{SY}(y, x) = [|s_{DY}(y, x) - s_{DY}(y-1, x)|]$ ,  $d_{SX}(y, x) = [|s_{DX}(y, x) - s_{DX}(y, x-1)|]$  при  $y = \overline{1, Y-1}$ ,  $x = \overline{1, X-1}$ , где  $|\cdot|$  - операция вычисления абсолютного значения. Условия интерполяции прогнозного кадра определяются тремя выражениями  $(d_{SY}(y, x) > 1) \wedge (d_{SX}(y, x) = 0)$ ,  $(d_{SY}(y, x) = 0) \wedge (d_{SX}(y, x) > 1)$ ,  $(d_{SY}(y, x) > 1) \wedge (d_{SX}(y, x) > 1)$ . Для первого из приведенных условий выполняются операции вертикальной интерполяции, определяемые выражениями  $\hat{m}_F(y' - j, x', t+1) = (m_F(y, x, t) + m_F(y-1, x, t))/2$ ,  $\hat{m}_D(y' - j, x', t+1) = m_D(y, x, t)$ ,  $\hat{d}_Z(y' - j, x') = (d_C(y, x, t) + d_C(y-1, x, t))/2$  при  $j = \overline{1, d_{SX}(y, x)}$ . Для второго условия выполняются операции горизонтальной интерполяции, определяемые выражениями  $\hat{m}_F(y', x' - i, t+1) = (m_F(y, x, t) + m_F(y, x-1, t))/2$ ,  $\hat{m}_D(y', x' - i, t+1) = m_D(y, x, t)$ ,  $\hat{d}_Z(y', x' - i) = (d_C(y, x, t) + d_C(y, x-1, t))/2$  при  $i = \overline{1, d_{SY}(y, x)}$ . Для третьего условия выполняются операции двухмерной интерполяции, определяемые выражениями  $\hat{m}_F(y' - j, x' - i) = (m_F(y, x, t) + m_F(y-1, x, t) + m_F(y, x-1, t) + m_F(y-1, x-1, t))/4$ ,  $\hat{m}_D(y' - j, x' - i, t+1) = m_D(y, x, t)$ ,  $\hat{d}_Z(y' - j, x' - i) = (d_C(y, x, t) + d_C(y-1, x, t) + d_C(y, x-1, t) + d_C(y-1, x-1, t))/4$  при  $j = \overline{1, d_{SY}(y, x)}$ ,  $i = \overline{1, d_{SX}(y, x)}$ .

При движении видеокамеры параллакс может привести к появлению областей пикселей, содержащих изображения объектов, заслоненных на опорном кадре другими объектами. Данные области представляют собой неопределенности появления, т.к. для образующих их пикселей отсутствуют соответствующие пиксела на опорном кадре. В матрице  $M_D(t+1)$  объектной декомпозиции прогнозного кадра пиксела из областей неопределенности появления соответствуют нулевым значениям  $\hat{m}_D(y', x', t+1) = 0$ . В качестве примера на фиг. 13 представлен случай заслонения объектом  $P_1$  объекта  $P_2$  при формировании опорного кадра. В результате движения видеокамеры данное заслонение перестает иметь место и значение пиксела, соответствующего объекту  $P_2$ , оказывается неопределенным (объект  $P_2$  появляется из-за объекта  $P_1$ ). Для формирования значений пикселей в областях неопределенности появления осуществляется предсказание на основе



значения средней яркости пикселей опорного кадра. В результате предсказания значения  $\hat{m}_F(y', x', t + 1)$  пикселей прогнозного кадра  $\hat{M}_F(t + 1)$ , для которых выполняется условие  $\hat{m}_D(y', x', t + 1) = 0$ , устанавливаются равными среднему значению опорного кадра  $\hat{m}_F(y', x', t + 1) = \frac{1}{YX} \sum_{y=0}^{Y-1} \sum_{x=0}^{X-1} m_F(y, x, t)$  и помечаются как пиксели фона  $\hat{m}_D(y', x', t + 1) = 1$ .

Блоки объектной компенсации движения видеокамеры 9 и 15, входящие в состав видеокодера 1 и видеодекодера 10, функционируют следующим образом.

На вход 43 блока сегментации 39 и на первый вход 50 блока объектно-ориентированного предсказания 41 с первого входа 28 блока объектной компенсации движения видеокамеры 9 в видеокодере 1 или с первого входа 36 блока объектной компенсации движения видеокамеры 15 в видеодекодере 10 поступает опорный кадр  $M_F(t)$ . В качестве примера на фиг. 14 представлены два соседних кадра видеопоследовательности, полученной от движущейся видеокамеры. На второй вход 70 блока компенсации крена камеры 56, на вход 49 блока предсказания положения линии горизонта 42 и на третий вход 47 блока объектно-ориентированного предсказания 41 со второго входа 29 блока объектной компенсации движения видеокамеры 9 в видеокодере 1 или со второго входа 37 блока объектной компенсации движения видеокамеры 15 в видеодекодере 10 поступает информация о параметрах установки и перемещения видеокамеры: значения высоты  $H$  установки, углов крена  $\beta_Z$  и тангажа  $\beta_X$  видеокамеры для каждого кадра видеоданных, а также относительные значения угла  $\beta_Y$  поворота, вертикального  $S_Y$ , горизонтального  $S_X$  и продольного  $S_Z$  перемещений видеокамеры для текущего кадра относительно предыдущего (фиг. 15). На выходе блока сегментации 39 формируется сегментированный опорный кадр  $M_S(t)$ , значение каждого пиксела которого указывает на номер сегмента, к которому данный пиксел принадлежит (фиг. 4). Сегментированный опорный кадр  $M_S(t)$  поступает на первый вход 44 блока компенсации крена камеры 56, на выходе которого формируется выравненный сегментированный опорный кадр  $M_C(t)$ , повернутый относительно центра на угол  $\beta_Z$  крена видеокамеры (фиг. 5). Выравненный сегментированный опорный кадр  $M_C(t)$  поступает на первый вход 71 блока объектной декомпозиции 40. На второй вход 45 блока объектной декомпозиции 40 поступает значение смещения  $Y_G$  проекции линии горизонта относительно нижней границы выравненного сегментированного опорного кадра  $M_C(t)$ , которое формируется в блоке предсказания положения линии горизонта 42 (фиг. 6). На выходе блока объектной декомпозиции 40 формируется объектный опорный кадр  $M_D(t)$  значение каждого пиксела которого указывает на номер соответствующего объекта или фона, и вектор  $Y_{MIN}$  у-координат крайних нижних пикселей сегментов опорного кадра  $M_C(t)$ , позволяющий вычислить расстояния до объектов (фиг. 7). Объектный опорный кадр  $M_D(t)$  и вектор  $Y_{MIN}$  поступают на второй вход 46 блока объектно-ориентированного предсказания 41, на выходе которого формируется прогнозный кадр  $\hat{M}_F(t + 1)$ . Пиксели прогнозного кадра  $\hat{M}_F(t + 1)$  формируются в результате смещения пикселей опорного кадра  $M_F(t)$  в зависимости от расстояний до соответствующих им объектов и параметров движения камеры. При формировании прогнозного кадра  $\hat{M}_F(t + 1)$  устраняются неопределенности заслонения (фиг. 11), сжатия (фиг. 11), растяжения (фиг. 12) и появления (фиг. 13) областей пикселей. В качестве примера на фиг. 16 представлен прогнозный кадр  $\hat{M}_F(t + 1)$ , синтезированный в блоке объектно-ориентированного предсказания 41 на основе опорного кадра  $M_F(t)$ , представленного на фиг. 14. Прогнозный кадр  $\hat{M}_F(t + 1)$  поступает на выход блока объектной компенсации движения видеокамеры 9 видеокодера 1 или на выход блока объектной компенсации движения видеокамеры 15 видеодекодера 10.

Блок объектно-ориентированного предсказания 41, входящий в состав блоков объектной компенсации движения видеокамеры 9 и 15, функционирует следующим образом.

На первый вход 59 блока определения расстояний по опорному кадру 51 со второго входа 46 блока объектно-ориентированного предсказания 41 поступает объектный опорный кадр  $M_D(t)$  и вектор  $Y_{\min}$  у-координат крайних нижних пикселей сегментов опорного кадра  $M_C(t)$ . На второй вход 60 блока определения расстояний по опорному кадру 51, на второй вход 62 блока определения смещений пикселей 52 и на третий вход 64 блока определения расстояний по прогнозному кадру 53 с третьего входа 47 блока объектно-ориентированного предсказания 41 поступает информация о параметрах установки и перемещения видеокамеры: значения высоты  $H$  установки, углов крена  $\beta_Z$  и тангажа  $\beta_X$  видеокамеры  $t$ -го кадра видеоданных, а также относительные значения угла  $\beta_Y$  поворота, вертикального  $S_Y$ , горизонтального  $S_X$  и продольного  $S_Z$  перемещений видеокамеры для  $(t + 1)$ -го кадра относительно  $t$ -го кадра (фиг. 15). На вход 74 блока вычисления средней яркости опорного кадра 57 и на первый вход 69 блока интерполяции пикселей прогнозного кадра 55 с первого входа 50 блока объектно-ориентированного предсказания 41 поступает опорный кадр  $M_F(t)$  (фиг. 14). На выходе блока определения расстояний по опорному кадру 51 формируется матрица  $D_Z(t)$  расстояний, которая поступает на первый вход 61 блока определения смещений пикселей 52 и на второй вход 72 блока определения расстояний по прогнозному кадру 53. На выходе блока определения смещений пикселей 52 формируются матрицы  $S_{DY}$  и  $S_{DX}$  проекционного предсказания вертикального и горизонтального смещений пикселей при сдвиге и повороте камеры (фиг. 9), которые поступают на первый вход 63 блока определения расстояний по прогнозному кадру 53 и на третий вход 68 блока интерполяции пикселей прогнозного кадра 55. На выходе блока определения расстояний по прогнозному кадру 53 формируется матрица  $D_C(t + 1)$  дальномерной коррекции (фиг. 10), которая поступает на четвертый вход 67 блока интерполяции пикселей прогнозного кадра 55 и на вход 65 блока формирования прогнозного кадра 54. На выходе блока формирования прогнозного кадра формируется прогнозный кадр  $\hat{M}_F(t + 1)$ , который поступает на второй вход 66 блока интерполяции пикселей прогнозного кадра 55. На выходе блока интерполяции пикселей прогнозного кадра 55 формируется прогнозный кадр  $\hat{M}_F(t + 1)$  с интерполированными значениями пикселей, который поступает на второй вход 76 блока предсказания пикселей прогнозного кадра 58. На выходе блока вычисления средней яркости опорного кадра 57 формируется значение средней яркости пикселей опорного кадра, которое поступает на первый вход 75 блока предсказания пикселей прогнозного кадра 58. На выходе 73 блока предсказания пикселей прогнозного кадра 58 формируется прогнозный кадр  $\hat{M}_F(t + 1)$  с предсказанными значениями неопределенных пикселей (фиг. 16), который поступает на выход 48 блока объектно-ориентированного предсказания 41.

В целом видеокодек с объектной компенсацией движения видеокамеры работает следующим образом.

При инициализации видеокодера 1 на выходе 23 энтропийного кодера 5 формируются флаг  $F_C(1) = 0$ , указывающий на кодирование без компенсации движения видеокамеры (фиг. 17). На выходе 30 блока объектной компенсации движения видеокамеры 9 формируется инициализирующий кадр  $\hat{M}_F(1)$  с нулевыми значениями пикселей, который поступает на второй вход 17 блока вычитания 2 и на второй вход 26 сумматора 7. На первый вход 16 блока вычитания 2 поступает кадр видеопоследовательности для кодирования без компенсации движения - этот кадр считается первым кадром  $F(1)$  видеопоследовательности  $\{F(t)\}$ , где  $F(t) = \|f(y, x, t)\|_{(y=0, Y-1, x=0, X-1)}$ ,  $t$  - номер кадра видеопоследовательности. На выходе блока вычитания 2 формируется кадровая разность  $\Delta F(1) = F(1) - \hat{M}_F(1) = F(1)$ , которая поступает на вход 18 блока прямого информационного преобразования 3. На выходе блока прямого информационного преобразования 3 формируется матрица  $\Delta F_T(1) =$

$=f_T(\Delta F(1))=f_T(F(1))$  коэффициентов преобразования  $f_T$  (дискретного вейвлет-преобразования или дискретного косинусного преобразования), которая поступает на вход 19 квантователя (4). После операции  $f_C$  квантования с выхода 20 квантователя 4 матрица  $\Delta F_C(1) = f_C(\Delta F_T(1)) = f_C(f_T(F(1)))$  квантованных коэффициентов преобразования поступает на первый вход 21 энтропийного кодера 5 и на вход 24 блока обратного информационного преобразования 6. На выходе блока обратного информационного преобразования 6 формируется восстановленная кадровая разность  $\Delta M_F(1) = f_T^{-1}(\Delta F_C(1)) = f_T^{-1}(f_C(f_T(F(1))))$ , которая поступает на первый вход 25 сумматора 7. На выходе сумматора 7 формируется первый опорный кадр  $M_F(1) = \hat{M}_F(1) + \Delta M_F(1) = \Delta M_F(1)$ , который поступает на вход 27 буфера 8, где запоминается. На выходе 23 энтропийного кодера 5 с помощью функции кодирования  $f_E$  формируется закодированная матрица  $F_E(1) = f_E(\Delta F_C(1)) = f_E(f_C(f_T(F(1))))$  квантованных коэффициентов преобразования, соответствующая первому кадру видеопоследовательности (фиг. 17).

При кодировании очередного кадра  $F(t+1)$  видеопоследовательности на выходе 23 энтропийного кодера 5 формируется флаг  $F_C(t+1) = 1$ , указывающий на кодирование очередного кадра видеопоследовательности с компенсацией движения видеокамеры. На второй вход 29 блока объектной компенсации движения видеокамеры 9 и на второй вход 22 энтропийного кодера 5 поступает информация о параметрах установки и перемещения видеокамеры: значения высоты  $H(t+1)$  установки, углов крена  $\beta_Z(t+1)$  и тангажа  $\beta_X(t+1)$  видеокамеры для текущего  $(t+1)$ -го кадра видеоданных, а также относительные значения угла  $\beta_Y(t+1)$  поворота, вертикального  $S_Y(t+1)$ , горизонтального  $S_X(t+1)$  и продольного  $S_Z(t+1)$  перемещений видеокамеры для текущего  $(t+1)$ -го кадра относительно предыдущего  $t$ -го кадра (фиг. 15). На выходе 23 энтропийного кодера 5 формируется закодированная с помощью функции энтропийного кодирования  $f_E$  информация  $P_E(t+1) = f_E(\{H(t+1), \beta_Z(t+1), \beta_Y(t+1), \beta_X(t+1), S_Z(t+1), S_Y(t+1), S_X(t+1)\})$  о параметрах установки и перемещения видеокамеры (см. фиг. 15, 17). На выходе буфера 8 формируется опорный кадр  $M_F(t)$ , который поступает на первый вход 28 блока объектной компенсации движения видеокамеры 9. На выходе 30 блока объектной компенсации движения видеокамеры 9 формируется прогнозный кадр  $\hat{M}_F(t+1)$ , который поступает на второй вход 17 блока вычитания 2 и на второй вход 26 сумматора 7. На первый вход 16 блока вычитания 2 поступает очередной кадр  $F(t+1)$  видеопоследовательности. На выходе блока вычитания 2 формируется кадровая разность  $\Delta F(t+1) = F(t+1) - \hat{M}_F(t+1)$ , которая поступает на вход 18 блока прямого информационного преобразования 3. На выходе блока прямого информационного преобразования 3 формируется матрица  $\Delta F_T(t+1) = f_T(\Delta F(t+1))$  коэффициентов преобразования, которая поступает на вход 19 квантователя 4. С выхода 20 квантователя 4 матрица  $\Delta F_C(t+1) = f_C(\Delta F_T(t+1)) = f_C(f_T(\Delta F(t+1)))$  квантованных коэффициентов преобразования поступает на первый вход 21 энтропийного кодера 5 и на вход 24 блока обратного информационного преобразования 6. На выходе блока обратного информационного преобразования 6 формируется восстановленная кадровая разность  $\Delta M_F(t+1) = f_T^{-1}(\Delta F_C(t+1)) = f_T^{-1}(f_C(\Delta F_T(t+1))) = f_T^{-1}(f_C(f_T(\Delta F(t+1))))$ , которая поступает на первый вход 25 сумматора 7. На выходе сумматора 7 формируется очередной опорный кадр  $M_F(t+1) = \hat{M}_F(t+1) + \Delta M_F(t+1)$ , который поступает на вход 27 буфера 8, где запоминается. На выходе 23 энтропийного кодера 5 с помощью функции кодирования  $f_E$  формируется закодированная матрица  $F_E(t+1) = f_E(\Delta F_C(t+1)) = f_E(f_C(f_T(\Delta F(t+1))))$  квантованных коэффициентов преобразования, соответствующая разности очередного кадра  $F(t+1)$  и прогнозного кадра  $\hat{M}_F(t+1)$  (фиг. 17).

Способ кодирования следующего  $(t+1)$ -го кадра видеопоследовательности зависит от величины ошибки  $E_P(t)$  предсказания, определяемой с помощью выражения

$$E_p(t) = \sum_{y=0}^{Y-1} \sum_{x=0}^{X-1} |f(y, x, t) - \hat{m}_F(y, x, t)| / (YX) \quad \text{для предыдущего } t\text{-го кадра. Если выполняется}$$

условие  $E_p(t) < E_{\text{MAX}}$ , где  $E_{\text{MAX}}$  - предельно допустимое значение ошибки предсказания, то очередной кадр видеопоследовательности кодируется с компенсацией движения видеокамеры. Иначе осуществляется инициализация видеокодера и очередной кадр видеопоследовательности условно считается первым и кодируется без компенсации движения видеокамеры.

При декодировании каждого кадра осуществляется проверка значения флага  $F_C(t)$ , указывающего на использование компенсации движения видеокамеры (фиг. 17).

Если при декодировании очередного кадра  $F_C(t+1) = 0$ , производится инициализация видеодекодера 10, декодирование осуществляется без компенсации движения и декодируемый кадр условно считается первым кадром видеопоследовательности  $F(1)$ . На выходе блока объектной компенсации движения видеокамеры 15 формируется инициализирующий кадр  $\hat{M}_F(1)$  с нулевыми значениями пикселей, который поступает на второй вход 38 сумматора 13. На вход 31 энтропийного декодера (11) поступает закодированная с помощью функции  $f_E$  матрица  $F_E(1) = f_E(\Delta F_C(1)) = f_E(f_C(f_T(F(1))))$  квантованных коэффициентов преобразования, соответствующая первому кадру видеопоследовательности. На первом выходе энтропийного декодера 11 формируется декодированная с помощью функции  $f_E^{-1}$  матрица  $\Delta F_C(1) = f_E^{-1}(F_E(1)) = f_E^{-1}(f_E(\Delta F_C(1)))$  квантованных коэффициентов преобразования, которая поступает на вход 32 блока обратного информационного преобразования 12. На выходе блока обратного информационного преобразования 12 формируется восстановленная с помощью функции  $f_T^{-1}$  кадровая разность  $\Delta M_F(1) = f_T^{-1}(\Delta F_C(1)) = f_T^{-1}(f_C(f_T(F(1))))$ , которая поступает на первый вход 33 сумматора 13. На выходе сумматора 13 формируется первый опорный кадр  $M_F(1) = \hat{M}_F(1) + \Delta M_F(1) = \Delta M_F(1)$ , который поступает на выход видеодекодера и на вход 35 буфера 14, где запоминается.

Если при декодировании очередного кадра  $F_C(t+1) = 1$  (фиг. 17), декодирование осуществляется с компенсацией движения видеокамеры. На вход 31 энтропийного декодера 11 поступает закодированная с помощью функции  $f_E$  информация о параметрах установки и перемещения видеокамеры  $P_E(t+1) = f_E(\{H(t+1), \beta_Z(t+1), \beta_Y(t+1), \beta_X(t+1), S_Z(t+1), S_Y(t+1), S_X(t+1)\})$ . На втором выходе энтропийного декодера 11 формируются декодированные с помощью функции  $f_E^{-1}$  значения высоты  $H(t+1)$  установки, углов крена  $\beta_Z(t+1)$  и тангажа  $\beta_X(t+1)$  видеокамеры для текущего  $(t+1)$ -го кадра видеоданных, а также относительные значения угла  $\beta_Y(t+1)$  поворота, вертикального  $S_Y(t+1)$ , горизонтального  $S_X(t+1)$  и продольного  $S_Z(t+1)$  перемещений видеокамеры для текущего  $(t+1)$ -го кадра относительно предыдущего  $t$ -го кадра (фиг. 15), которые поступают на второй вход 37 блока объектной компенсации движения видеокамеры 15. На выходе буфера 14 формируется опорный кадр  $M_F(t)$ , который поступает на первый вход 36 блока объектной компенсации движения видеокамеры 15. На выходе блока объектной компенсации движения видеокамеры 15 формируется прогнозный кадр  $\hat{M}_F(t+1)$ , который поступает на второй вход 38 сумматора 13. На вход 31 энтропийного декодера 11 поступает закодированная с помощью функции  $f_E$  матрица  $F_E(t+1) = f_E(\Delta F_C(t+1)) = f_E(f_C(f_T(\Delta F(t+1))))$  квантованных коэффициентов преобразования, соответствующая кадровой разности. На первом выходе энтропийного декодера (11) формируется декодированная с помощью функции  $f_E^{-1}$  матрица  $\Delta F_C(t+1) = f_E^{-1}(F_E(t+1)) = f_E^{-1}(f_E(\Delta F_C(t+1)))$  квантованных коэффициентов преобразования, которая поступает на вход 32 блока обратного информационного преобразования 12. На выходе блока обратного информационного преобразования 12 формируется восстановленная с помощью функции  $f_T^{-1}$  кадровая разность  $\Delta M_F(t+1) = f_T^{-1}(\Delta F_C(t+1)) = f_T^{-1}(f_C(f_T(\Delta F(t+1))))$ , которая поступает на первый вход 33 сумматора 13. На выходе сумматора 13 формируется очередной восстановленный кадр

$M_F(t+1) = \hat{M}_F(t+1) + \Delta M_F(t+1)$ , который поступает на выход видеodeкодера и на вход 35 буфера 14, где запоминается.

Эффективность видеокодирования определяется тремя основными факторами: коэффициентом сжатия, качеством восстановления и скоростью кодирования видеоданных. При кодировании видеoinформации, формируемой движущейся видеокамерой, с использованием компенсации движения эти параметры непосредственно определяются компактностью представления информации о движении, величиной ошибки предсказания и скоростью предсказания.

Для оценки компактности представления информации о движении целесообразно использовать коэффициент  $K_C$  компактности, определяемый с помощью выражения  $K_C = V_M/(YX)$ , имеющий размерность бит/пиксел и показывающий отношение объема координатной информации  $V_M$  о движении к размеру кадра. С учетом использования 8 бит на кодирование каждого из 7 параметров установки и движения видеокамеры ( $V_M = 8 \times 7 = 56$  бит) и размера кадра  $480 \times 640$  пикселей предложенный видеокодек обеспечивает значение коэффициента  $K_C = 1,8 \cdot 10^{-4}$  бит/пиксел. При использовании 16 бит на кодирование вектора движения каждого блока, размера кадра  $480 \times 640$  пикселей и размера блока  $8 \times 8 = 64$  пикселей ( $V_M = 16 \times 480 \times 640 / 64 = 76800$  бит) прототип, для сравнения, обеспечивает значение коэффициента  $K_C = 0,25$  бит/пиксел (примерно в 1400 раз хуже по сравнению с предложенным видеокодеком). В пересчете на коэффициент сжатия это означает, что при сокращении объема  $V_F = Y \times X \times BD$  текущего кадра за счет компенсации движения и энтропийного кодирования ошибки предсказания ( $BD$  - битовая глубина), например, в  $CR_F = 100$  раз, для каждого 8-битного ( $BD = 8$ ) канала цветного изображения или 8-битного канала яркости полутонового изображения коэффициент  $CR_B$  сжатия, обеспечиваемый прототипом, составляет  $CR_B = V_F / (V_F / CR_F + V_M) = 480 \times 640 \times 8 / (480 \times 640 \times 8 / 100 + 76800) = 24,24$  раз. В тех же условиях предложенный видеокодек с объектной компенсацией движения видеокамеры обеспечивает коэффициент сжатия  $CR_O = V_F / (V_F / CR_F + V_M) = 480 \times 640 \times 8 / (480 \times 640 \times 8 / 100 + 56) = 99,77$ . Это примерно в  $CR_O / CR_B = 99,77 / 24,24 = 4,12$  раз лучше в сравнении с прототипом. Выигрыш тем больше, чем больше размер кадра и коэффициент  $CR_F$ .

Точность предсказания целесообразно оценивать с использованием пикового отношения сигнал-шум PSNR, вычисляемого для области предсказания с помощью выражения  $PSNR = 10 \log_{10}((2^{BD} - 1)^2 / MSE)$ , где  $MSE = \frac{1}{YX} \sum_{y=0}^{Y_O-1} \sum_{x=0}^{X_O-1} (f(y, x, t) - \hat{m}_F(y, x, t))^2$  - средне-

квадратическая ошибка;  $Y_O, X_O$  - размеры области перекрытия. На фиг. 18 представлены пары соседних кадров, прогнозные кадры, сформированные с помощью предложенного видеокодека и прототипа и соответствующие прогнозным кадрам разностные кадры. В табл. 1 приведены значения пикового отношения PSNR сигнал-шум предсказания, полученные для представленных на фиг. 18 кадрах с помощью предложенного видеокодека и прототипа при среднем размере блока  $8 \times 8$  пикселей. Из табл. 1 видно, что предложенный видеокодек обеспечивает выигрыш до 1,5 дБ в пиковом отношении PSNR сигнал-шум предсказания по сравнению с прототипом.

Таблица 1

**Значения пикового отношения PSNR сигнал-шум предсказания, дБ**

Номера кадров (фиг. 18)	1-2	3-4	5-6	7-8	9-10	11-12
Предложенный видеокодек	69,70	69,14	68,40	67,15	67,36	68,55
Прототип	68,17	67,47	67,52	65,42	67,02	68,20

Скорость предсказания в предложенном видеокодеке зависит от числа выделяемых на опорном кадре сегментов. Скорость предсказания в прототипе определяются числом шагов, затрачиваемых на поиск в поле пикселей опорного кадра соответствия для каждого блока пикселей прогнозируемого кадра. Для кадров, приведенных на фиг. 18, в табл. 2 представлены результаты измерения задержки предсказания для программной модели предложенного видеокодека и программной модели прототипа, из которых видно, что скорость предсказания для обоих видеокодеков примерно одинакова.

Таблица 2

**Значения задержки предсказания, мс**

Номера кадров (фиг. 18)	1-2	3-4	5-6	7-8	9-10	11-12
Предложенный видеокодек	0,250	0,141	0,125	0,141	0,125	0,125
Прототип	0,172	0,151	0,177	0,125	0,141	0,141

Технико-экономическое преимущество предложенного видеокодека с объектной компенсацией движения видеокамеры по сравнению с прототипом заключается в повышении коэффициента сжатия видеоданных за счет увеличения компактности представления информации о движении в результате использования параметров движения для всего кадра, а не для каждого небольшого блока в отдельности, а также повышении качества восстановления видеoinформации за счет уменьшения ошибки предсказания в результате применения компенсации движения к объектам, выделяемым с учетом структуры изображения, а не к блокам, выделяемым без учета структуры изображения. Таким образом, предложенный видеокодек обеспечивает высокий коэффициент сжатия и высокое качество восстановления видеoinформации, формируемой движущейся видеокамерой.

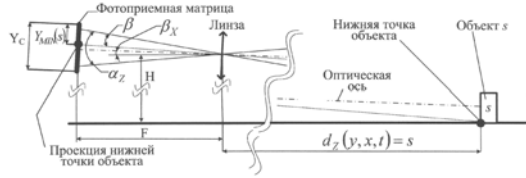


Фиг. 2

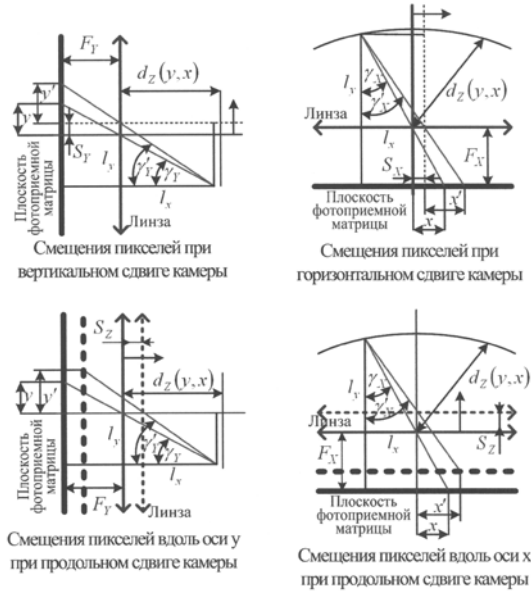


Фиг. 3

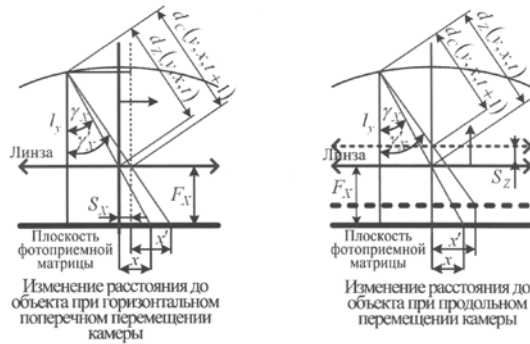




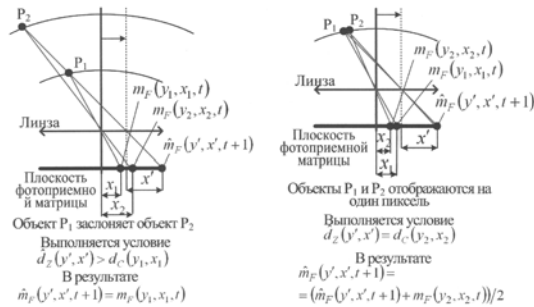
Фиг. 8



Фиг. 9

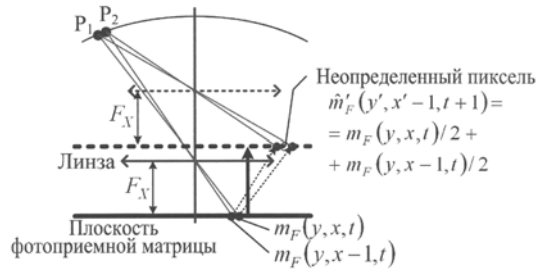


Фиг. 10

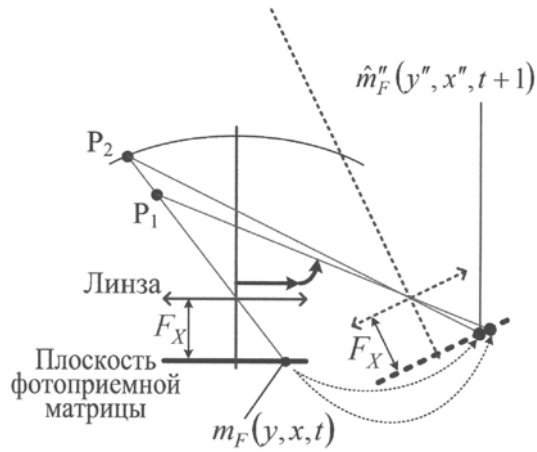


Фиг. 11

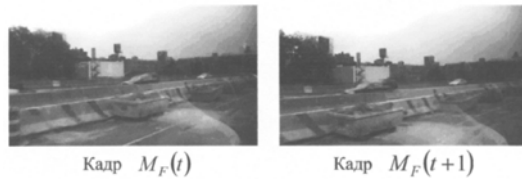




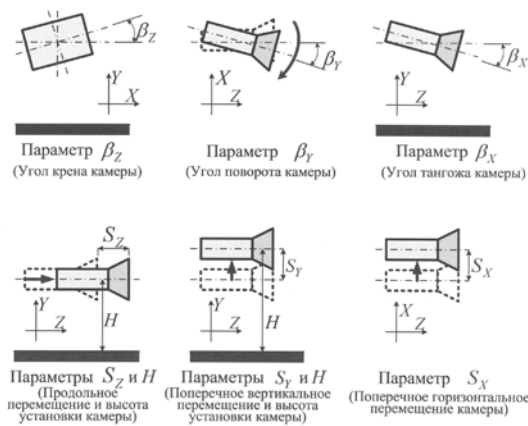
Фиг. 12



Фиг. 13



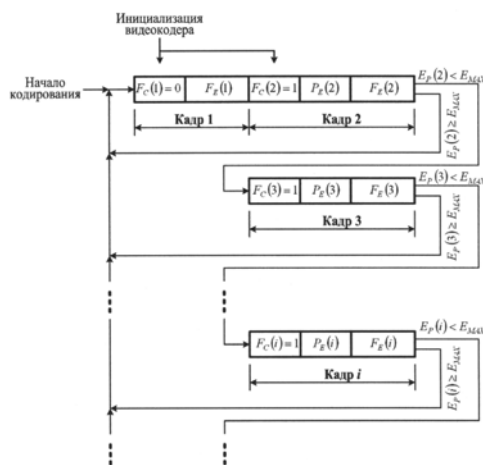
Фиг. 14



Фиг. 15



Фиг. 16



Фиг. 17

Кадр	Опорный кадр $F(i)$	Следующий кадр $F(i+1)$	Прогнозный кадр $M_p(i+1)$	Разностный кадр $\Delta F(i+1) = F(i+1) - M_p(i+1)$	Разностный кадр для прогноза
1-2					
3-4					
5-6					
7-8					
9-10					
11-12					

Фиг. 18