

ИНФОРМАТИКА

УДК 621.391

ОБЪЕКТНОЕ ПРЕДСКАЗАНИЕ ИЗМЕНЕНИЯ ВИДЕОДАНЫХ ПРИ ГОРИЗОНТАЛЬНОМ ПЕРЕМЕЩЕНИИ КАМЕРЫ

Т.М. АЛЬ-ДЖУБУРИ, В.Ю. ЦВЕТКОВ, В.К. КОНОПЕЛЬКО

*Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6 Минск 220013, Беларусь**Поступила в редакцию 5 октября 2009*

Предложен метод объектного предсказания изменения видеоданных на основе модели движения камеры. Суть метода состоит в объектной декомпозиции опорного кадра видеоданных, классификации объектов по дальности, расслоении опорного кадра по дальности и формировании нового кадра на основе модели движения камеры, учитывающей смещения объектов навстречу вектору движения камеры. Метод обеспечивает уменьшение ошибки предсказания изменения видеоданных при движении камеры за счет использования информации о характеристиках, условиях установки и параметрах движения камеры.

Ключевые слова: модель движения камеры, объектная декомпозиция изображения, предсказание изменения видеоданных.

Введение

В связи с ростом объемов передаваемых и хранимых видеоданных актуальна проблема их компактного представления. Эффективными способами решения данной проблемы являются сжатие видеоданных с устранением межкадровой избыточности на основе компенсации движения и частичное или полное замещение естественных видеоданных псевдореалистическими, синтезированными на основе естественных, но требующими существенно меньших ресурсов при передаче и хранении. Эти подходы используются в стандартах сжатия видео MPEG-2, MPEG-4 и H.264 [1], технологии MPEG-7 [2], а также методах формирования трехмерных панорамных [3] и псевдостереоскопических [4] изображений. Эффективность данных подходов ограничена отсутствием предсказания изменения видеоданных при перемещении камеры, для реализации которого необходима информация о характеристиках, условиях установки и параметрах движения камеры. В общем случае получение данной информации проблематично. Однако, существует ряд систем, для которых данная проблема может быть решена. Примерами могут служить системы видеонаблюдения с несколькими стационарно установленными камерами, а также одной или несколькими перемещающимися камерами. В этих случаях возможно предсказание изменения видеоданных при перемещении камеры или при переключении между стационарными камерами с перекрывающимися областями видимости по известным значениям угла видимости, высоты установки, угла наклона и направления перемещения камеры.

Цель данной работы – разработать метод предсказания изменения видеоданных на основе модели движения камеры.

Модель движения камеры

В общем случае модель M_C движения видеокамеры может быть задана следующим набором параметров:

$$M_C(t) = \|D_Y, D_U, D_V, Y, X, \alpha_Y(t), \alpha_X(t), A(t), H(t), S_X(t), S_Z(t), \beta_Z(t), \beta_Y(t), \beta_X(t)\|, \quad (1)$$

где D_Y, D_U, D_V – битовая глубина формируемого камерой цветного изображения для яркостной и двух цветоразностных компонент; Y, X – размеры формируемого камерой изображения по вертикали и горизонтали; $t = \overline{0, T-1}$ – дискреты времени, совпадающие с номерами кадров, формируемых камерой; T – период дискретного времени моделирования движения камеры; $\alpha_Y(t), \alpha_X(t)$ – углы обзора камеры по вертикали и горизонтали; $A(t) = \|a(y, x, t)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$ – матрица корректирующих коэффициентов, учитывающих аберрации оптической системы камеры для каждого пикселя формируемого изображения; $H(t)$ – высота установки камеры; $S_X(t), S_Z(t)$ – смещения камеры в горизонтальной плоскости перпендикулярно и вдоль оптической оси относительно предыдущего положения камеры ($S_X(0) = 0, S_Z(0) = 0$); $\beta_Z(t)$ – угол наклона нижней границы изображения к поверхности (угол крена камеры); $\beta_Y(t)$ – угол поворота оптической оси камеры в горизонтальной плоскости относительно предыдущего положения камеры ($\beta_Y(0) = 0$); $\beta_X(t)$ – угол наклона оптической оси камеры к поверхности ($\beta_X(t) > 0$, если оптическая направлена вверх и $\beta_X(t) < 0$, если оптическая ось направлена вниз).

Параметры $H(t), S_X(t), S_Z(t), \beta_Z(t), \beta_Y(t), \beta_X(t)$ при $t = \overline{0, T-1}$ определяют траекторию движения камеры.

С учетом конкретных условий применения модель движения камеры может быть существенно упрощена. Частным, но важным случаем является горизонтальное движение ($S_Z(t) = 0, \beta_Y(t) = 0, H(t) = const$) камеры, формирующей полутонное изображение ($D_Y = 8, D_U = 0, D_V = 0$). Если предсказание основано на одном исходном кадре и распространяется только на один последующий кадр, то $T = 2$. Учитывая особенности построения ряда систем видеонаблюдения, при $T = 2$ модель подвижной камеры может быть упрощена за счет следующих дополнительных ограничений: $\alpha_Y(t) = const, \alpha_X(t) = const$ (постоянное фокусное расстояние); $a(y, x, t) = 1$ (оптическая система описывается моделью тонкой линзы); $\beta_X(t) = const$ (постоянный угол наклона оптической оси камеры); $\beta_Z(t) = 0$ (отсутствие крена). В результате упрощенная модель подвижной камеры представляется следующим выражением:

$$M_C = \|Y, X, \alpha_Y, \alpha_X, H, S_X, \beta_X\|. \quad (2)$$

В данной модели траектория движения камеры задается единственным параметром S_X .

Описание метода

Предлагается метод объектного предсказания изменения видеоданных при перемещении камеры на основе модели $M_C(t)$ движения камеры, использующий опорный кадр $C_G = \|c_G(y, x)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1})}$ видеоданных, соответствующий исходному положению камеры, а также информацию о характеристиках, параметрах установки и перемещения камеры для формирования кадров $\hat{C}_G(t) = \|\hat{c}_G(y, x, t)\|_{(y=\overline{0, Y-1}, x=\overline{0, X-1}, t=\overline{1, T-1})}$, соответствующих положениям камеры в дискретные моменты времени на периоде T . Метод описывается следующим выражением

$$\hat{C}_G(t) = f_{OP}(C_G, M_C(t)), \quad (3)$$

где f_{OP} – функция объектного предсказания изменения видеоданных.

Основными шагами метода являются объектная декомпозиция опорного кадра видеоданных (формируется информация о количестве и расположении объектов опорного кадра), классификация объектов по дальности (каждый объект причисляется к некоторому классу объектов по дальности в зависимости от его удаленности от камеры), расслоение опорного кадра видеоданных по дальности (определяются расстояния до равноудаленных от камеры объектов, группируемых в слои изображения по дальности), формирование нового кадра видеоданных (объекты и фон опорного кадра смещаются с учетом вектора движения камеры и интерполируются значения образующихся неопределенных пикселей, заслоненных объектами на опорном кадре).

Для эффективного использования метода объекты опорного кадра должны иметь высокий контраст по отношению к фону и располагаться на ровной поверхности, достаточно хорошо аппроксимирующей плоскостью. Несмотря на данные ограничения, сужающие область применения предлагаемого метода, он может достаточно эффективно использоваться для сжатия видеоданных с предсказанием движения камеры в системах видеонаблюдения помещений, открытых участков равнинной местности и поверхностей водоемов. Кроме того, возможно использование данного метода в этих условиях для формирования второго изображения псевдостереопары, а также последовательности кадров синтетического видео.

В случае простой структуры опорного кадра, содержащего высококонтрастные по отношению к однородному фону и низкоконтрастные по отношению друг к другу объекты, не заслоняющие частично или полностью другие объекты, отдельные шаги метода могут быть существенно упрощены. При использовании в этом случае упрощенной модели движения камеры, представленной выражением (2), метод описывается соотношением

$$\hat{C}_G = f_{OP}(C_G, M_C) \quad (4)$$

и сводится к следующим четырем шагам.

1) Объектная декомпозиция опорного кадра C_G видеоданных. На данном шаге осуществляется бинаризация опорного кадра видеоданных и формируется матрица объектной декомпозиции, содержащая информацию о количестве и расположении изолированных объектов опорного кадра, которые представляют собой совокупности единичных пикселей, разделенные друг от друга нулями. Результатом выполнения данного шага являются количество N_O изолированных объектов в опорном кадре видеоданных и матрица

$M_D = \|m_D(y, x)\|_{(y=0, \overline{Y-1}, x=0, \overline{X-1})}$ объектной декомпозиции опорного кадра видеоданных.

2) Классификация объектов опорного кадра видеоданных по дальности. Формируется вектор классификации объектов по дальности, с помощью которого каждый изолированный объект, выделенный в матрице объектно-ориентированной декомпозиции, причисляется к некоторому классу объектов по дальности в зависимости от минимального значения y -координат пикселей, образующих данный объект. Результатом выполнения данного шага является вектор $D_Y = \|d_Y(i)\|_{(i=0, \overline{N_O-1})}$ классификации объектов по дальности, значение каждого i -го элемента $d_Y(i) \in [0, Y-1]$ которого определяет класс соответствующего i -го объекта по дальности. Класс характеризует удаленность объекта от камеры. Чем меньше значение $d_Y(i)$, тем ближе к камере расположен i -й объект. Максимальное число различных классов равно $Y+1$ – на один больше числа пикселей в кадре по вертикали. К классу Y принадлежат пиксели, не принадлежащие ни одному из N_O объектов.

3) Расслоение опорного кадра видеоданных по дальности. Оно основано на использовании параметров Y , α_Y , H и β_X модели движения камеры для представления опорного кадра видеоданных в виде бинарной проекции трехмерной сцены, расстояние до каждого объекта которой оценивается по смещению проекции точки местоположения объекта от нижней границы кадра. В результате выполнения данного расслоения формируется вектор $D_Z = \|d_Z(y)\|_{(y=0, \overline{Y-1})}$ расстояний от камеры до объектов, значение каждого y -го элемента $d_Z(y) \in [0, \infty)$ которого определяет расстояние от камеры до объектов y -го слоя, смещенных на y пикселей от нижнего края опорного кадра видеоданных. Расстояния определяются в тех единицах измерения, которые использовались для определения параметра H модели движения камеры.

4) Формирование нового кадра $\hat{C}_G$ видеоданных. Объекты каждого слоя опорного кадра видеоданных сдвигаются на некоторое число пикселей в зависимости от удаленности слоя от камеры. Значения пикселей, заслоненных объектами на опорном кадре, устанавливается равным среднему значению яркости фона.

Формирование нового кадра видеоданных

Предлагается алгоритм формирования по опорному кадру видеоданных, соответствующему исходному положению камеры, нового кадра, соответствующего горизонтальному смещению камеры относительно исходного положения. Алгоритм основан на модели движения камеры и предположении, что при перемещении камеры расстояния от камеры до объектов не изменяются (или, по крайней мере, эти изменения не могут быть отображены на новом кадре видеоданных из-за недостаточной разрешающей способности фотоприемной матрицы камеры). Такое предположение возможно, если выполняется следующее условие

$$\sqrt{d_Z(0)^2 + S_X^2} - d_Z(0) < d_Z(1) - d_Z(0). \quad (5)$$

Условие (5) имеет место во многих практических случаях, в том числе в системах видеонаблюдения.

Алгоритм формирования нового кадра видеоданных состоит из следующих шагов.

- 1) Инициализация вектора $S_D = \|s_D(y)\|_{(y=0, Y-1)}$ смещений слоев опорного кадра

$$s_D(y) = 0 \quad (6)$$

при $y = \overline{0, Y-1}$.

- 2) Инициализация матрицы $M_S = \|m_S(y, x)\|_{(y=0, Y-1, x=0, X-1)}$ смещенных пикселей опорного кадра видеоданных

$$m_S(y, x) = 0 \quad (7)$$

при $y = \overline{0, Y-1}$ и $x = \overline{0, X-1}$.

- 3) Определение горизонтального виртуального фокуса F_X модели оптической системы камеры

$$F_X = (X/2) / \operatorname{tg}(\alpha_X/2). \quad (8)$$

- 4) Определение средней яркости $\bar{c}_G$ фона опорного кадра видеоданных

$$\bar{c}_G = \left(\sum_{y=0}^{Y-1} \sum_{x=0}^{X-1} (c_G(y, x) \cdot (1 - c_B(y, x))) \right) / \left(Y \cdot X - \sum_{y=0}^{Y-1} \sum_{x=0}^{X-1} c_B(y, x) \right). \quad (9)$$

- 5) Цикл определения величин смещений слоев опорного кадра видеоданных. Для каждого элемента вектора D_Z расстояний до объектов рассчитывается значение соответствующего элемента вектора S_D с помощью выражения

$$s_D(y) = \left[F_X \cdot \frac{S_X}{\sqrt{d_Z(y)^2 - S_X^2}} \right] \quad (10)$$

при $y = \overline{0, Y-1}$,

где $[\cdot]$ – операция округления до ближайшего целого.

- 6) Цикл смещения слоев опорного кадра видеоданных. Операция смещения слоев применяется к строкам опорного кадра видеоданных в направлении сверху вниз. При смещении камеры вправо ($S_X > 0$) операция смещения применяется к пикселям каждой строки в направлении слева направо. При смещении камеры влево ($S_X < 0$) операция смещения применяется к пикселям каждой строки в направлении справа налево.

Для пикселей $c_G(y, x)$ опорного кадра C_G видеоданных, не отмеченных в матрице M_D объектной декомпозиции ($m_D(y, x) = 0$ – т.е. пиксель не принадлежит ни одному из N_O объектов), смещение осуществляется на величину $s_D(y)$ при условии, что пиксель $\hat{c}_G(y, x - s_D(y))$ формируемого кадра $\hat{C}_G$ видеоданных является неопределенным и находится в пределах кадра ($0 \leq (x - s_D(y)) < X$)

$$(m_S(y, x') = 0) \rightarrow (\hat{c}_G(y, x') = c_G(y, x)) \quad (11)$$

при $y = \overline{0, Y-1}$ и $x = \overline{0, X-1}$,

где $x' = x - s_D(y)$ – координата пикселя формируемого кадра видеоданных.

Для пикселей $c_G(y, x)$ опорного кадра C_G видеоданных, отмеченных в матрице M_D объектной декомпозиции ($m_D(y, x) \neq 0$ – т.е. пиксель принадлежит одному из N_O объектов), смещение осуществляется на величину $s_D(d_Y(m_D(y, x)))$ при условии, что пиксель $\hat{c}_G(y, x - s_D(d_Y(m_D(y, x))))$ формируемого кадра $\hat{C}_G$ видеоданных не принадлежит более приближенному к камере объекту и находится в пределах кадра ($0 \leq (x - s_D(d_Y(m_D(y, x)))) < X$)

$$\begin{cases} (m_S(y, x'') = 0) \rightarrow ((\hat{c}_G(y, x'') = c_G(y, x)) \wedge (m_S(y, x'') = m_D(y, x))) \\ (m_S(y, x'') \neq 0) \rightarrow \left((d_Y(m_D(y, x)) < d_Y(m_S(y, x''))) \rightarrow \right. \\ \left. \rightarrow ((\hat{c}_G(y, x'') = c_G(y, x)) \wedge (m_S(y, x'') = m_D(y, x))) \right) \end{cases} \quad (12)$$

при $y = \overline{0, Y-1}$ и $x = \overline{0, X-1}$,

где $x'' = x - s_D(d_Y(m_D(y, x)))$ – координата пикселя формируемого кадра видеоданных.

7) Цикл поиска и определения значений пикселей формируемого кадра видеоданных, заслоненных объектами на опорном кадре. Значения пикселей, заслоненных объектами на опорном кадре, устанавливаются равным среднему значению яркости фона в соответствии с выражением

$$(m_S(y, x) = 0) \rightarrow (\hat{c}_G = \bar{c}_G) \quad (13)$$

при $y = \overline{0, Y-1}$ и $x = \overline{0, X-1}$.

По окончании цикла поиска и определения значений пикселей формируемого кадра видеоданных осуществляется выход из алгоритма.

В результате выполнения данного алгоритма формируется кадр $\hat{C}_G$ видеоданных, соответствующий новому положению камеры при ее горизонтальном перемещении на S_X относительно исходного положения в плоскости, параллельной плоскости изображения.

Оценка эффективности метода объектного предсказания изменения видеоданных

Для оценки эффективности разработанного метода объектного предсказания изменения видеоданных получены 5 полутоновых изображений тестовой трехмерной сцены – исходный видеокادر (рис. 1) и 4 видеокадра (рис. 2, а – г), соответствующие горизонтальному смещению камеры вправо на 1, 2, 3 и 4 см. Изображения получены с помощью цифровой камеры Sony Cyber-shot с отключенной автофокусировкой, установленной на высоте $H = 10$ см и обеспечивающей размеры изображения по вертикали $Y = 480$ и горизонтали $X = 640$, углы видимости по вертикали $\alpha_Y = 22.5^\circ$ и горизонтали $\alpha_X = 50.2^\circ$, угол наклона оптической оси камеры к поверхности $\beta_X = -6^\circ$.

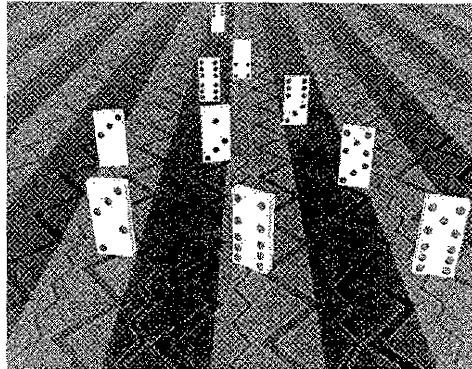


Рис. 1. Исходный видеокادر

На рис. 3 представлены видеокadres при смещении камеры вправо по горизонтали на $S_X=1\text{см}$, $S_X=2\text{см}$, $S_X=3\text{см}$ и $S_X=4\text{см}$, синтезированные с помощью разработанного метода объектного предсказания изменения видеоданных и форматированные по горизонтали для удаления области неопределенности, возникающей справа при смещении камеры. На рис. 4 представлены образы ошибок предсказания, соответствующие этим синтезированным видеокadres.

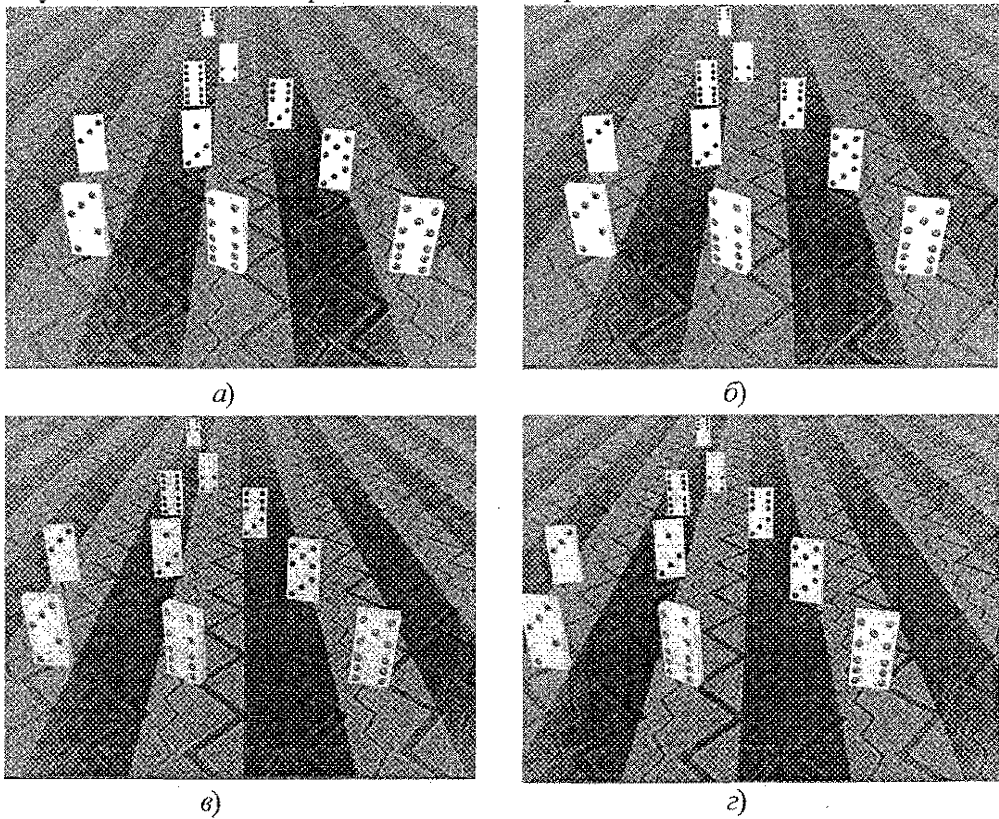


Рис. 2. Полутоновые изображения тестовой трехмерной сцены:
а – видеокادر при $S_X = 1\text{ см}$; б – видеокادر при $S_X = 2\text{ см}$;
в – видеокادر при $S_X = 3\text{ см}$; г – видеокادر при $S_X = 4\text{ см}$

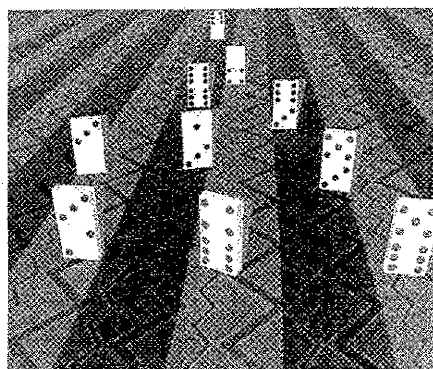
Для количественной оценки ошибки предсказания изменения видеоданных при движении камеры использована среднеквадратическая ошибка MSE, вычисляемая с помощью выражения

$$MSE = \frac{\sum_{y=0}^{Y-1} \sum_{x=0}^{X'-1} (c_G(y,x) - \hat{c}_G(y,x))^2}{Y X'}, \tag{14}$$

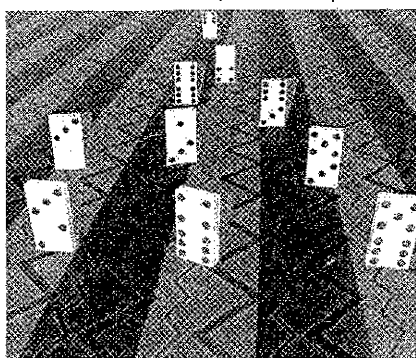
где X' – размер форматированного видеокадра по горизонтали с учетом удаления области неопределенности, возникающей при смещении камеры.



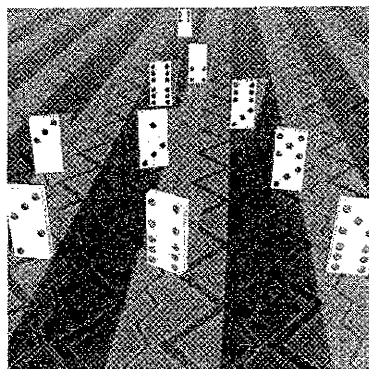
a)



б)



в)

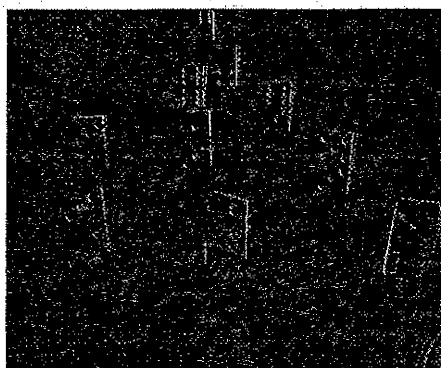


г)

Рис. 3. Синтезированные полутоновые изображения тестовой трехмерной сцены:

а – видеокадр при $S_X = 1$ см; б – видеокадр при $S_X = 2$ см;

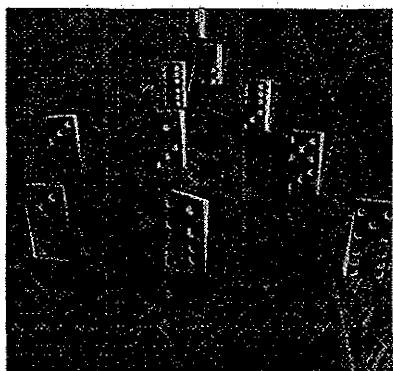
в – видеокадр при $S_X = 3$ см; г – видеокадр при $S_X = 4$ см



a)



б)



в)



г)

Рис. 4. Образы ошибок синтезированных полутоновых изображений тестовой трехмерной сцены:

а – образ ошибки при $S_X = 1$ см; б – образ ошибки при $S_X = 2$ см;

в – образ ошибки при $S_X = 3$ см; г – образ ошибки при $S_X = 4$ см

В таблице приведены значения среднеквадратической ошибки MSE предсказания для трех способов устранения межкадровой избыточности видеоданных, используемых в методах Cinepak (кадровая разность), MPEG-2, MPEG-4 (блочная компенсация движения) и разработанном методе объектного предсказания изменения видеоданных (объектная компенсация движения). Из таблицы видно, что метод объектного предсказания изменения видеоданных обеспечивает уменьшение ошибки предсказания примерно в 6 раз по сравнению с методом Cinepak и примерно в 4,4 раза по сравнению с методами MPEG-2, MPEG-4.

Среднеквадратическая ошибка предсказания

Величина горизонтального смещения камеры	Среднеквадратическая ошибка MSE предсказания		
	Кадровая разность (Cinepak)	Блочная компенсация движения (MPEG-2, MPEG-4)	Объектное предсказание изменений видеоданных
1 см	2362	1294	215,364
2 см	3604	1868	615,367
3 см	3713	2202	623,401
4 см	3801	2802	624,219

Заключение

Предложен метод объектного предсказания изменения видеоданных при перемещении камеры на основе модели ее движения. Суть метода состоит в объектной декомпозиции опорного кадра видеоданных, классификации объектов по дальности, расслоении опорного кадра по дальности и формировании нового кадра на основе модели движения камеры, учитывающей смещения объектов навстречу вектору движения камеры. Метод обеспечивает уменьшение ошибки предсказания изменения видеоданных при движении камеры в 6 раз по сравнению с методом Cinepak и примерно в 4,4 раза по сравнению с методами MPEG-2, MPEG-4.

OBJECT PREDICTION CHANGES FOR VIDEO DATA TRANSFORMATION WITH HORIZONTAL CAMERA MOVMENT

T.M. AL-JUBOORI , V.YU. TSVIATKOU, V.K. KONOPELKO

Abstract

This is a suggested method for predicting the changes in the object in video data by using a model of camera motion. The method contains with, object decomposition of the reference frame of video data, the classification of objects in range, the start position of the reference frame for the range and the formation a new frame-based on the model of camera motion, taking in consideration, the displacement vector of the motion of objects towards the camera. The method provides a reduction of prediction error for video data changes when camera moves, through the use of information regarding characteristics of camera, installation conditions and parameters of camera motion.

Литература

1. Iain R. H // The Robert Gordon University. Aberdeen. UK. John Wiley & Sons Ltd. The Atrium. Southern Gate. Chichester. 2003. P. 306.
2. A.B. Benitez, S. Chang // Signals, Systems and Computers. Conference Record of the Thirty-Seventh Asilomar Conference. 2003. Vol.1 P. 92–96.
3. Zhigang Z., Allen R. // Eighth International Conference on Computer Vision (ICCV'01). 2001. Vol. 2. P. 723.
4. Ianir I., Leonid P., Barak F. // Real-time Image Processing Jornal. 2007. Vol. 2. P. 3–9.