

УДК 004.932

ОБРАБОТКА ОБЛАСТЕЙ НАЛОЖЕНИЙ ОБЪЕКТОВ ПРИ ПОСТРОЕНИИ КАРТ ГЛУБИНЫ ДЛЯ ВИДЕО

*Матюнин С.Б., аспирант факультета вычислительной математики и кибернетики МГУ
им. М. В. Ломоносова, e-mail: smatyunin@graphics.cs.msu.ru;*

*Ватолин Д.С., к. ф.-м. н., с. н. с. факультета вычислительной математики и кибернетики МГУ
им. М. В. Ломоносова, e-mail: dmitriy@graphics.cs.msu.ru.*

Ключевые слова: карта глубины, видео, обработка изображений, оптический поток, сжатие стереовидео.

Введение

В данной статье рассматривается задача полуавтоматического построения карты глубины для видео. Под термином «карта глубины» для некоторого кадра будем понимать изображение, содержащее вместо яркостной или цветовой информации расстояния до соответствующих видимых точек сцены. Карты глубины, полученные с помощью прямого измерения специализированными приборами [1] или посредством анализа стереопары [2], применяются для трехмерного моделирования. В киноиндустрии карты глубины применяются как промежуточный этап преобразования обычного кино (будем его называть 2D кино или 2D видео) в стереоформат. Процесс создания карт глубины для видео подразумевает создание карты расстояний для каждого кадра. Данная задача в общем случае не имеет решения. На практике применяются ручная разметка и различные эвристические методы [3, 4, 5]. При ручной обработке трудозатраты можно уменьшить с помощью компьютерных инструментов. Это могут быть как инструменты обработки изображений общего назначения: кисти, заливка, сегментация и т.п., а также специализированные: фильтры карты глубины, алгоритмы распространения ручной разметки, методы автоматического построения карты глубины для видео

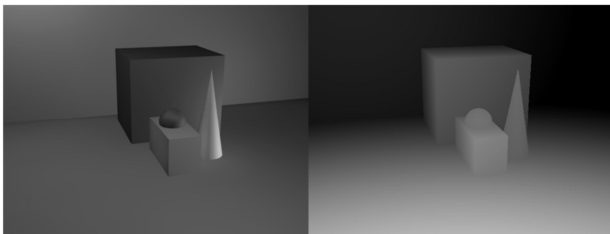


Рис. 1. Пример трехмерной сцены (слева) и карты глубины для нее (справа)

Один из распространенных полуавтоматических подходов [6, 7, 8, 9] состоит в том, что пользователь вручную создает карту глубины для некоторых кадров видео. Будем называть эти кадры ключевыми. Объекты между соседними кадрами, как правило, не сильно изменяют свою позицию, а следовательно, и глубину. Поэтому карту глубины для остальных кадров можно стро-

Рассматривается задача полуавтоматического построения карты глубины для видео. Предложен метод обработки областей наложений объектов для повышения качества карт глубины. Информация о таких областях аккумулируется при последовательной обработке кадров видео, а затем используется для маскирования ненадежных интерполированных областей. Приведено сравнение алгоритмов поиска областей наложений объектов и способов их использования в рассматриваемой задаче.

ить автоматически – интерполируя ее для необработанных кадров. Для повышения качества интерполированной карты глубины предлагается отдельно обрабатывать области наложения объектов в видео.

Пример трехмерной сцены (слева) и карты глубины для нее (справа) представлен на рис. 1. Яркость пикселей карты глубины соответствует расстоянию от точки наблюдения до поверхности объекта. Большим значениям яркости соответствует более близкие поверхности.

Обзор методов построения карты глубины

В работах [6, 8, 10] предлагаются методы полуавтоматической разметки карты глубины для видео, основанные на использовании пользовательского ввода для некоторых кадров последовательности (ключевых кадров). Исходная разметка распространяется с ключевых кадров на остальные с учетом информации о движении, цвете и других характеристиках, которые можно извлечь из видео.

В [10] для интерполяции карты глубины между ключевыми кадрами составляется система уравнений, ограничивающих разность по глубине между соседними пикселями в кадре и соответствующими пикселями в соседних кадрах. Системой также задается глубина для размеченных пикселей и глубина пикселей, найденная с помощью классификатора по цвету. Полученная система решается методом наименьших квадратов. Решение соответствует карте глубины для всей последовательности. Такой подход требует решения системы уравнений большой размерности. Это приводит к необходимости уменьшать размерность за счет разрешения искомой карты глубины. Аналогичным образом карта глубины строится в работе [11]. Для каждого пикселя p видео-последовательности задается уравнение вида

$$d(p) = \sum_{q \in N(p)} \omega_{pq} d(q), \quad (1)$$

где $d(p)$ – глубина пикселя p , $N(p)$ – множество пик-

селей, состоящее из 8 соседних пикселей на текущем кадре и соответствующего соседнего пикселя с предыдущего кадра, ω_{pq} – вес соседнего пикселя, обратно пропорциональный разности цветов пикселей p и q . Соседний пиксель с предыдущего кадра определяется с учетом векторов оптического потока. Из полученных уравнений составляется линейная система $Ax = 0$. Для нахождения карты глубины решается задача квадратичного программирования: $\min_x \|Ax\|^2$ при ограничениях $Ex \geq f$, заданных пользовательским вводом.

Чаще используется подход, который позволяет последовательно вычислять карту глубины для соседнего кадра. В работе [8] глубина для каждого следующего кадра вычисляется с помощью билатерального фильтра, веса которого зависят от геометрического и фотометрического расстояния между пикселями. Возможные ошибки исправляются с помощью компенсации движения. За счет применения компенсации движения только на этапе постобработки данный подход не применим для видеопоследовательностей с сильным движением.

В работе [12] межкадровая интерполяция используется для увеличения частоты кадров карты глубины. Интерполяция осуществляется за счет использования векторов движения, полученных по соответствующему 2D-видеопотоку.

В задаче сжатия карты глубины также применяются аналогичные подходы. Например, в методе [11] компенсация движения используется при предсказании карты глубины для очередного кодируемого кадра.

Подход, аналогичный описанному в [8, 12, 13, 14], мы использовали как базовый для построения карты глубины. Он позволяет последовательно получать карту глубины для каждого следующего кадра. Размерность решаемой задачи при этом значительно меньше, чем при построения карты глубины для всей последовательности сразу.

Построение областей открытия для пары кадров

При обработке сцен, содержащих движущиеся объекты, обычно применяется оценка движения. При поиске векторов движения критичной является корректная обработка областей открытия, то есть тех областей, которые были не видны на одном кадре, но на другом из-за движения объектов стали видимы. В данной работе мы используем существующие методы построения областей открытия для пары кадров. На их основе предлагается строить маску «доверия» для интерполированной карты глубины, а далее использовать ее для улучшения качества результата.

Задача построения областей открытия по видео в общем случае не имеет точного решения. Например, для двух последовательных кадров можно определить области открытия несколькими способами, в зависимости от реального направления движения (см. рис. 2). Поэтому на практике используются различные эвристические подходы.

Один из наиболее часто используемых подходов [17] основан на анализе ошибки компенсации движения между двумя кадрами. Область кадра, где ошибка компен-

сации высока в обоих направлениях, рассматривается как область наложения. В силу простоты критерия данный подход имеет высокую скорость работы, однако он неустойчив к шуму, дает большое число ложных срабатываний при малых значениях порога Θ и наоборот, малое число верных срабатываний при больших Θ . Аналогичный подход (LRC) основан на анализе разности векторов движения от одного кадра к следующему и обратно. В работе [17] также предложен геометрический подход, основанный на анализе плотности концов векторов движения на опорном кадре. Так как на кадре t не существуют соответствия для точек из области закрытия кадра $t - 1$, поэтому в таких областях функция плотности будет иметь малые значения. Данный подход обладает хорошей устойчивостью, так как опирается на надежные вектора движения в областях закрытия вместо некорректно определенных векторов в областях открытия.

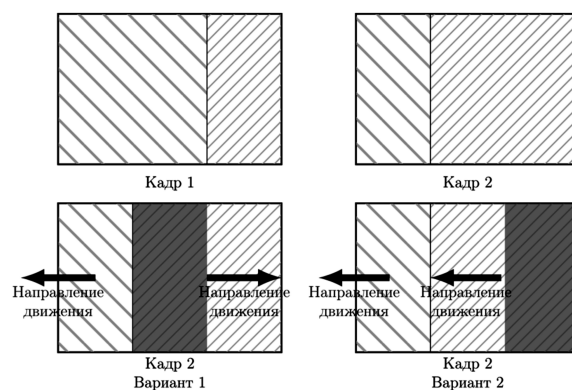


Рис. 2. Неоднозначность определения области открытия по двум кадрам. Возможная область открытия отмечена серым цветом, объекты отмечены штриховкой

Обозначения

Пусть размер кадра равен $W \times H$. Обозначим через $I_t(p)$ яркость в точке с координатами $p = (p_x, p_y)$, $p \in [0, W - 1] \times [0, H - 1]$ на кадре в момент времени t . При этом запись I_t будет обозначать весь кадр в момент времени t . Оценка движения – поиск соответствия между двумя кадрами, один из которых называется базовым, а другой – опорным. Для каждой точки x_{t_0} базового кадра I_{t_0} алгоритм оценки движения находит соответствующую ей точку x_{t_1} опорного кадра I_{t_1} . Разность между координатами этих точек называется вектором движения $V_{t_0, t_1}(x_{t_0})$, вычисленным в точке x_{t_0} базового кадра I_{t_0} : $V_{t_0, t_1}(x_{t_0}) = x_{t_1} - x_{t_0}$. Совокупность V_{t_0, t_1} всех векторов движения $V_{t_0, t_1}(x_{t_0})$ базового кадра называется полем векторов движения для кадров I_{t_0} и I_{t_1} .

При движении объектов на видео возникают области их взаимного наложения. Можно выделить два типа наложений на заданной паре кадров [15]:

1. **Открытие** – область базового кадра, отсутствующая на опорном кадре;
2. **Закрытие** – область опорного кадра, отсутствующая на базовом кадре.

Компенсация движения – процесс построения скомпенсированных кадров \tilde{I}_{t_0} на основе опорного кадра I_{t_1} с использованием информации о движении между исходными кадрами I_{t_0} и I_{t_1} :

$$\tilde{I}_{t_0}(x) = I_{t_1}(x + V_{t_0,t_1}(x_{t_0})). \quad (2)$$

Для распространения карты глубины D с кадра $t-1$ на кадр t используются векторы движения $V_{t-1,t}$, найденные по исходному 2D видео с помощью преобразования компенсации движения MC :

$$MC_{t,t-1} * D_i(x) \stackrel{def}{=} D_i(x + V_{t-1,t}(x)). \quad (3)$$

Если известна карта глубины D_{t-1} для предыдущего кадра, то можно экстраполировать ее для кадра t :

$$\hat{D}_t(x) = MC_{t,t-1} * D_{t-1}(x). \quad (4)$$

Для краткости обозначим преобразование для всего кадра $\hat{D}_t = MC_{t,t-1} * D_{t-1}$. Если значение $x + V_{t-1,t}(x)$ выходит за допустимые пределы, оно приводится к ближайшему допустимому. Преобразование можно применять последовательно, начиная с некоторого заданного кадра k , распространяя карту глубины на всю видео-последовательность:

$$\hat{D}_t = MC_{t,t-1} * MC_{t-1,t-2} * \dots * MC_{k+1,k} * D_k. \quad (5)$$

Положим $\hat{D}_k \stackrel{def}{=} D_k$. При таком подходе качество полученной карты глубины напрямую зависит от алгоритма оценки движения. Основным недостатком такого подхода является то, что в областях открытия, которые присутствуют на базовом кадре и отсутствуют на опорном, векторы движения не определены. Большинство алгоритмов оценки движения опирается на скомпенсированную межкадровую разность и поэтому в областях открытия они дают некоторые неадекватные (зависящие от алгоритма оценки движения и входного изображения) значения векторов движения.

В областях открытия векторы движения определяются некорректно, поэтому «копирование» карты глубины для области фона может происходить из области объекта (рис. 3).

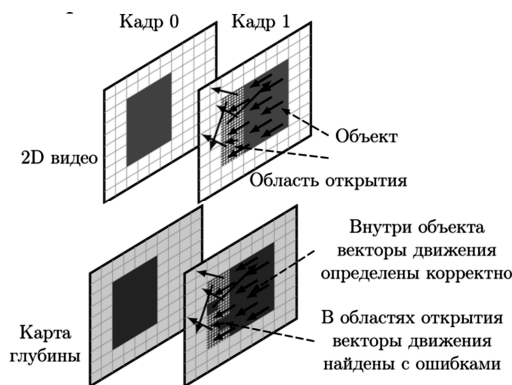


Рис. 3. Причина возникновения ошибки интерполяции карты глубины при использовании компенсации движения.

Обработка областей открытия

Компенсация карты глубины (4) в областях открытия приводит тому, что за движущимся объектом образуется «шлейф» неправильной глубины (см. рис. 4). Перед объектом карта глубины получена с использованием более надежных векторов движения, поэтому результат в этой области корректный. При распространении с разных направлений корректные значения карты глубины получают по разные стороны от объекта. Это можно использовать для улучшения результата обработки.

Рассмотрим последовательность кадров длины n (сцену длины n): I_0, \dots, I_{n-1} . Допустим, что известна карта глубины D_0 и D_{n-1} для первого и последнего кадров соответственно (рис. 4). Рассмотрим объект движущийся поступательно слева направо. При распространении от 0-го кадра до некоторого кадра $k \in [0, n-1]$ «шлейф» находится слева от объекта (рис. 4 е), а при распространении от последнего кадра – справа от объекта (рис. 4д).

Результат работы алгоритма может быть улучшен, путем слияния результатов распространения с нескольких направлений. Предположим, имеется m размеченных ключевых кадров с индексами k_0, \dots, k_{m-1} . Для каждого кадра последовательности можно построить m вариантов экстраполированной карты глубины $\hat{D}^0, \dots, \hat{D}^{m-1}$. Можно использовать простое усреднение результатов распространения с разных направлений:

$$\hat{D}_t^{avg}(x) = \frac{1}{m} \sum_{i=0}^{m-1} \hat{D}_t^i(x) \quad (6)$$

где \hat{D}_t^{avg} – итоговая карта глубины, m – количество направлений распространения (для экспериментов использовалось распространение с первого и с последнего кадра, т.е. $m = 2$). Такой подход не дает удовлетворительного результата, т. к. «шлейф» остается с обеих сторон объекта (рис. 8 а, б). Другой подход состоит в том, что результаты с двух направлений смешиваются с весами, зависящими от расстояния до соответствующего ключевого кадра, например:

$$\hat{D}_t^w(x) = \frac{\sum_{i=0}^{m-1} w_t^i \hat{D}_t^i(x)}{\sum_{i=0}^{m-1} w_t^i} \quad (7)$$

$w_t^i = 1$, при $t = k_i$ и убывает до нуля при приближении t к соседнему ключевому кадру k_{i-1} или k_{i+1} (рис. 5).

Таким образом, для ключевого кадра берется только исходная карта глубины, а на промежуточных кадрах, расположенных между размеченными, используется среднее значение из результатов распространения с двух ближайших ключевых кадров. Когда область открытия образуется близко к ключевому кадру, веса ошибочных пикселей будут большими. Это увеличивает заметность ошибок распространения. Для более эффективного удаления шлейфа необходимо производить смешивание направлений с учетом положения областей открытия.

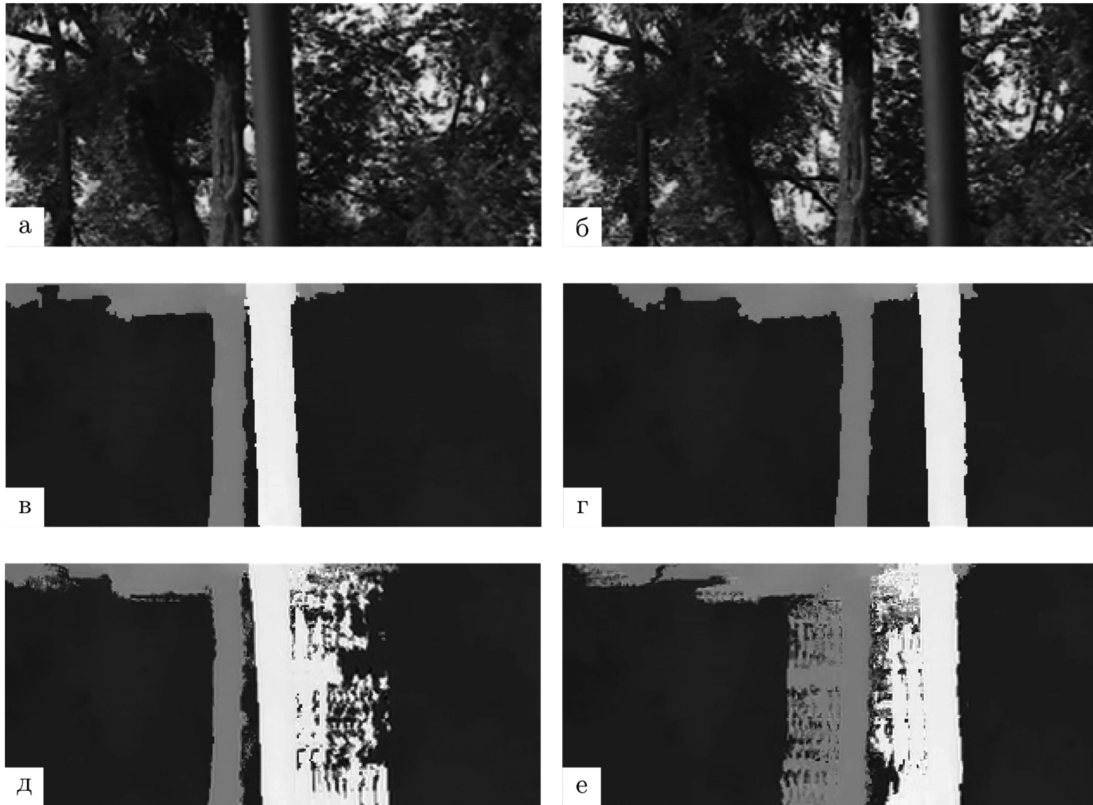


Рис. 4. Возникновение затеканий при покадровом распространении карты глубины при использовании базового подхода. Фрагмент видеопоследовательности «Road» [16]. а) Кадр №1; б) Кадр №7; в) Эталонная карта глубины для кадра №1; г) Эталонная карта глубины для кадра №7; д) Результат покадровой интерполяции с 7-го до 1-го кадра; е) Результат покадровой интерполяции с 1-го до 7-го кадра

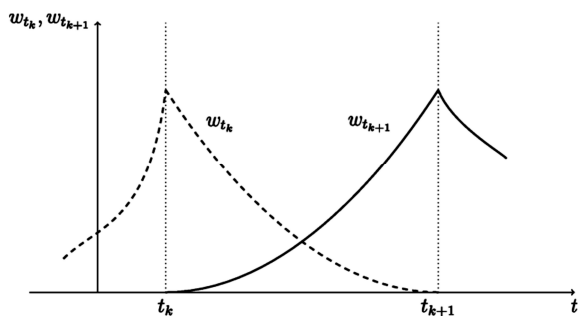


Рис. 5. Пример весовой функции для объединения результатов интерполяции для ключевых кадров t_{k-1} и t_{k+1}

Построение меры доверия

Покадровую маску областей открытия без дополнительной обработки трудно применить для улучшения качества интерполированной карты глубины. Одним из возможных подходов к заполнению областей открытий является синтез изображения в этих областях (inpainting, например [18]). Эта задача имеет множество решений для изображений. Для карт глубины она фактически представляет собой задачу построения карты глубины по одному изображению. Чтобы избежать синтеза карты глубины, и максимально использовать достаточно надежную информацию о глубине на ключевых кадрах и о движении объектов между кадрами, мы предлагаем использовать интерполированную карту глубины с нескольких ключевых кадров и построенные области открытия.

Мы предлагаем аккумулировать информацию об областях открытия при переходе от кадра к кадру. При этом области, помеченные как открытие на текущем кадре, в общем случае могут двигаться на последующих кадрах. Пусть $\Omega(x)$ – карта областей открытия для кадра t :

$$\Omega(x) = \begin{cases} 1, & \text{если } x \text{ – точка, не принадлежащая} \\ & \text{области открытия,} \\ 0, & \text{иначе.} \end{cases} \quad (8)$$

Тогда аккумулированная маска «доверия» $\bar{\Omega}^j$ при проходе от ключевого кадра k_j будет строиться следующим образом:

$$\hat{\Omega}_{k_j}^j = \Omega_{k_j} \quad (9)$$

$$\hat{\Omega}_t^j(x) = \min(MC_{t,t-1} * \hat{\Omega}_{t-1}^{k_j}, \Omega_t) \quad (10)$$

Процесс построения маски областей открытия схематично показан на рис. 6. Если при компенсации движения координаты выходят за допустимую область, то значение $MC_{t,t-1} * \hat{\Omega}_{t-1}^{k_j}$ принимается равным нулю, т.к. фактически это означает движения из-за границы кадра. Пиксель кадра t с координатой x_t помечается как область открытия в двух случаях: 1) если он лежит в области открытия на текущем кадре; 2) если соответствующий ему с учетом движения пиксель x_{t-1} на кадре $t-1$ в аккумулированной маске $\hat{\Omega}_{t-1}^j$ принадлежит области открытия.

Использование меры доверия

Построенную на основе областей открытия маску доверия предлагается использовать для повышения качества итоговой карты глубины. Из предварительно построенных вариантов интерполяции карты глубины больший приоритет в каждом пикселе имеют варианты с большей мерой доверия. За счет этого уменьшается вклад в итоговый результат областей, в которых интерполяция производилась по некорректным векторам движения. Мы разработали и протестировали несколько способов использования построенной меры доверия.

Наиболее простым вариантом является выбор в каждой точке того направления распространения s_t^Ω , которое соответствует большему значению меры доверия:

$$s_t^\Omega = \arg \max(\hat{\Omega}_t^i) \quad (11)$$

Результирующим значением глубины будет

$$D_t^\Omega = D_t^{s_t^\Omega} \quad (12)$$

При таком подходе результат для областей видео, не являющихся областями открытия, определяется порядком обхода $\hat{\Omega}_t^i$. Если D_t^i значительно отличаются для разных i , и на некотором кадре t s_t^Ω меняется для значительной части кадра, то результат будет содержать резкий переход по глубине. Обычно это отрицательно сказывается на качестве видео, обработанного с использованием карты глубины, например при переводе видео из 2D формата в 3D.

Чтобы уменьшить количество резких изменений глубины, можно в областях, где $\hat{\Omega}_t^i = 1$, применить дополнительный критерий, например, расстояние до ключевого кадра. Среди направлений распространения можно выбирать те, которые имеют максимальное значение меры доверия:

$$J_t^\Omega \left\{ j : \hat{\Omega}_t^j = \hat{\Omega}_t^{s_t^\Omega} \right\} \quad (13)$$

Из них выбираем направление, которое имеет наименьшее расстояние до ключевого кадра:

$$s_t^{\Omega Dist}(x) = \arg \min_{j \in J_t^\Omega} (|j - k_j|) \quad (14)$$

Найденное направление используется в качестве итогового для данной точки:

$$D_t^{\Omega Dist} = D_t^{s_t^{\Omega Dist}(x)} \quad (15)$$

При таком подходе резкий скачок может быть только на середине отрезка между двумя ключевыми кадрами, поэтому можно усреднить глубину для направлений с близкими к максимальному значениями меры доверия.

$$J_t^{\Omega T} \left\{ j : \left| \hat{\Omega}_t^j - \hat{\Omega}_t^{s_t^\Omega} \right| < T \right\} \quad (16)$$

где T – заданный порог. Итоговое значение получаем усреднением выбранных кандидатов:

$$D_t^{\Omega T} = \frac{1}{|J_t^{\Omega T}|} \sum_{j \in J_t^{\Omega T}} D_t^j \quad (17)$$

Усреднение может производиться с весами, убывающими при возрастании расстояния до ключевого кадра:

$$D_t^{\Omega \omega T}(x) = \frac{1}{\mu} \sum_{j \in J_t^{\Omega T}} \omega(|j - k_j|) D_t^j \quad (18)$$

Здесь μ – нормировочный коэффициент. Более значимыми при близких значениях меры доверия считаются кандидаты, которые расположены ближе к соответствующим ключевым кадрам. Далее будет приведено сравнение описанных подходов.

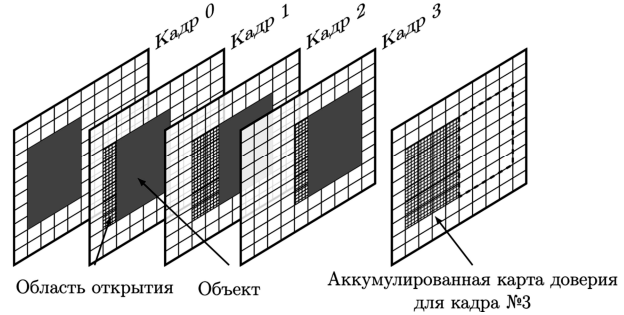


Рис. 6. Построение аккумулятивной карты областей открытия для нескольких кадров

Результаты

Разработанный метод использования областей открытия был протестирован на наборе видеопоследовательностей с построенной картой глубины [16]. Для тестирования качества интерполяции карты глубины разработанному алгоритму на вход подавались карты глубины для двух ключевых кадров и отрезков видео между ними. Полученная карта глубины сравнивалась с эталонной. Для этого измерялось пиковое отношение сигнал/шум (PSNR). Результаты представлены на рис. 7. Для оценки поля векторов движения использовался алгоритм построения оптического потока [19].



Рис. 7. Сравнение методов построения областей открытия по качеству распространения карты глубины с учетом областей открытия

В качестве опорного метода было взято усреднение кандидатов без учета областей открытия (\hat{D}^{avg}). Второй вариант алгоритма помечал как ненадежные те области, которые были скомпенсированы по движению из-за границ кадра. Это позволило повысить качество построенной карты глубины более чем на 10 дБ. Дальнейшее улучшение качества было достигнуто за счет обработки областей открытия внутри кадра. Наилучшие результаты по PSNR были достигнуты с использованием подхода

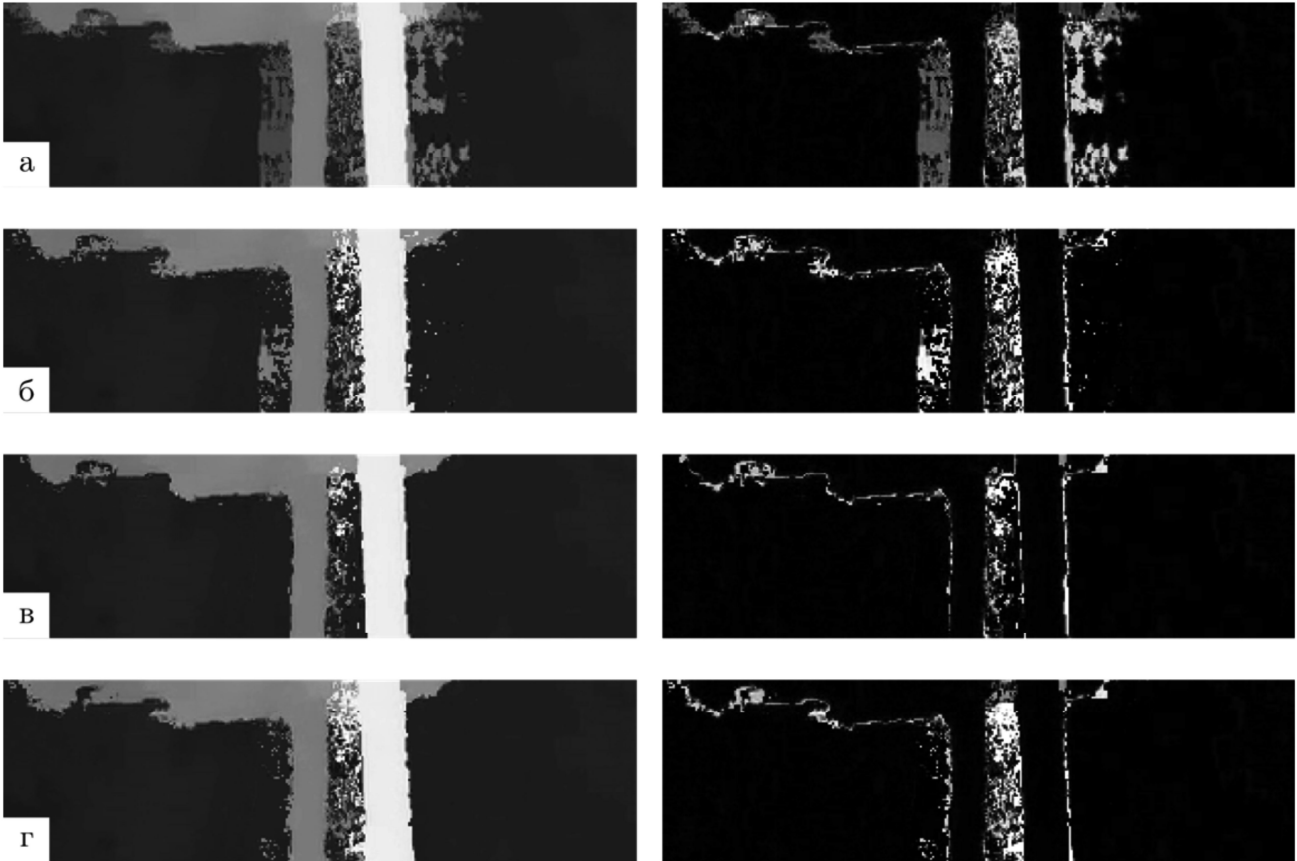


Рис. 8. Карта глубины, построенная с помощью различных методов обработки областей открытия (слева) и визуализация отличий по сравнению с эталоном (справа)

LRC, но этот метод порождает большое количество выбросов вблизи границ объектов (рис. 8 г). Геометрический подход имеет близкие показатели PSNR и продемонстрировал хорошую устойчивость вблизи границ (рис. 8 в). Среди протестированных подходов объединения результатов распространения карты глубины с нескольких кадров наилучшим оказался подход, использующий усреднение кандидатов с близкими к максимальному значениями меры доверия ($D_t^{\Omega_{OT}}$).

Дополнительно было протестировано влияние обработки областей открытия на результаты работы алгоритма сжатия стереовидео. Использованный подход более подробно описан в [14]. Его суть в следующем. Стереовидео представляется в формате 2D+глубина. 2D видео кодируется с использованием классического кодека общего назначения. В данном случае использовался кодек x264. Для эффективного сжатия карты глубины используется ее согласованность с 2D видео, за счет чего можно повышать качество декодированной карты глубины. Карты глубины прореживаются как по времени, так и по пространственным измерениям и сжимаются. На этапе декодирования разжатое 2D видео используется для восстановления исходного разрешения и прореженных кадров карты глубины.

Декодированная карта глубины сравнивалась с исходной по пиковому соотношению сигнал/шум (PSNR). Для оценки влияния обработки областей открытий на алгоритм сжатия были построены RD-кривые. Исполь-

зование детектора областей открытия при декодировании карты глубины позволило повысить качество декодированной карты глубины до 2 дБ без изменения размера сжатых данных (рис. 9).

Заключение

В данной работе был проведен сравнительный анализ трех методов определения областей открытия для видео в применении к задаче межкадровой интерполяции карты глубины. Наиболее перспективным оказался геометрический метод анализа векторов движения из работы [17]. Этот метод показал близкие к наилучшим значения PSNR, и в то же время показал лучшие результаты на границах объектов – меньше артефактов чем у метода LRC, что важно для карт глубины.

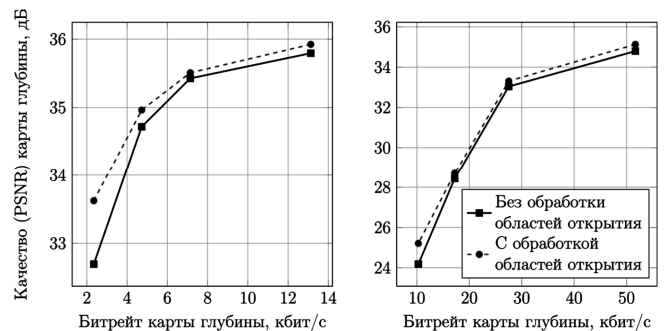


Рис. 9. Сравнение алгоритмов сжатия карты глубины на основе межкадрового распространения с обработкой областей открытия и без нее

Было предложено построение маски доверия для интерполированной карты глубины на основе пок кадровых областей открытия и несколько вариантов её использования. По результатам тестов наилучшим оказался метод с использованием расстояния до ключевого кадра со взвешенным усреднением. На тестовом наборе он показал улучшение показателя качества (PSNR) интерполированной карты глубины более чем на 1 дБ.

Обработка областей открытия позволила повысить качество карты глубины (до 2 дБ при низких значениях битрейта). Соотношение качество/битрейт регулировалось количеством использованных ключевых кадров карты глубины.

На визуализации (рис. 8) более светлые области соответствуют большей ошибке. Здесь: а) карта глубины, полученная простым усреднением; б) результат с обработкой областей открытия, полученных по ошибке компенсации движения; в) результат с обработкой областей открытия, полученных с помощью геометрического метода; г) результат с обработкой областей открытия, полученных по критерию LRC.

Литература

1. Blais F. Review of 20 years of range sensor development. // *Journal of Electronic Imaging*. – 2004. – Vol. 13, no. 1. – Pp. 231–243.
2. Ogale A. S., Aloimonos Y. Shape and the Stereo Correspondence Problem // *International Journal of Computer Vision*. – 2005. – Vol. 65, no. 3. – Pp. 147–162.
3. Zhuo S., Sim T. On the Recovery of Depth from a Single Defocused Image // *Computer Analysis of Images and Patterns* / Ed. by Xiaoyi Jiang, Nicolai Petkov. – Springer Berlin Heidelberg, 2009. – Vol. 5702 of *Lecture Notes in Computer Science*. – Pp. 889–897.
4. Battiato S., Curti S., La Cascia M. et al. Depth map generation by image classification. – 2004. <http://dx.doi.org/10.1117/12.526634>.
5. Saxena A., Ng A., Chung S. Learning Depth from Single Monocular Images // *IEEE Neural Information Processing Systems*. – 2005. – Vol. 18.
6. Li Z., Cao X., Dai Q. A novel method for 2D-to-3D video conversion using bi-directional motion estimation // *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. – 2012. – Pp. 1429–1432.
7. Rapid 2D-to-3D conversion / P. V. Harman, J. Flack, S. Fox, M. Dowley / Ed. by A. J. Woods, J. O. Merritt, S. A. Benton, M. T. Bolas. – Vol. 4660. – SPIE, 2002. – Pp. 78–86. <http://link.aip.org/link/?PSI/4660/78/1>.
8. Varekamp C., Barenbrug B. Improved depth propagation for 2D to 3D video conversion using key-frames // *IET Conference Publications*. – 2007. – Vol. 2007, no. CP534. – Pp. 29–29. <http://link.aip.org/link/abstract/IEECPS/v2007/iCP534/p29/s1>.
9. Practical temporal consistency for image-based graphics applications / M. Lang, O. Wang, T. Aydin et al. – Vol. 31. – New York, NY, USA: ACM, 2012. – Pp. 34:1–34:8. <http://doi.acm.org/http://doi.acm.org/10.1145/2185520.2185530>.
10. Guttman M., Wolf L., Cohen-Or D. Semi-automatic stereo extraction from video footage. // *ICCV'09*. – 2009. – Pp. 136–142.
11. Video stereolization: Combining motion analysis with user interaction / M. Liao, J. Gao, R. Yang, M. Gong // *Visualization and Computer Graphics, IEEE Transactions on*. – 2012. – Vol. 18, no. 7. – Pp. 1079–1088.
12. Choi J., Min D., Sohn K. 2D-plus-depth based resolution and frame-rate up-conversion technique for depth video // *Consumer Electronics, IEEE Transactions on*. – 2010. – November. – Vol. 56, no. 4. – Pp. 2489–2497.
13. De Silva D.V.S.X., Fernando W. A C, Yasakethu S. L P. Object based coding of the depth maps for 3D video coding // *Consumer Electronics, IEEE Transactions on*. – 2009. – August. – Vol. 55, no. 3. – Pp. 1699–1706.
14. Matyunin S., Vatolin D. 3D Video Compression Using Depth Map Propagation // *Multimedia Communications, Services and Security* / Ed. by A. Dziech, A. Czyżewski. – Springer Berlin Heidelberg, 2013. – Vol. 368 of *Communications in Computer and Information Science*. – Pp. 153–166.
15. Гришин С.В. Программная система для преобразования частоты кадров цифровых видео сигналов: Дис. канд. физ.-мат. наук: 05.13.11 / МГУ. – М., 2009.
16. Consistent depth maps recovery from a video sequence / Guofeng Zhang, Jiaya Jia, Tien-Tsin Wong, Hujun Bao // *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. – 2009. – Vol. 31, no. 6. – Pp. 974–988.
17. Ince S., Konrad J. Geometry-based estimation of occlusions from video frame pairs // *Acoustics, Speech, and Signal Processing (ICASSP'05). IEEE International Conference on / IEEE*. – Vol. 2. – 2005. – Pp. ii–933.
18. Telea A. An image inpainting technique based on the fast marching method // *Journal of graphics tools*. – 2004. – Vol. 9, no. 1. – Pp. 23–34.
19. Ayvaci A., Raptis M., Soatto S. Sparse Occlusion Detection with Optical Flow // *International Journal of Computer Vision*. – 2012. – Vol. 97, no. 3. – Pp. 322–338. <http://dx.doi.org/10.1007/s11263-011-0490-7>.

OCCLUSION PROCESSING FOR DEPTH MAP ESTIMATION

Matyunin S., Vatolin D.

We develop a system for semi-automatic converting 2D video into 3D stereoscopic content. Traditional pipeline of conversion includes mainly manual depth map mark-up for video and automatic views generation. Instead of manual depth map creation, we use only limited user input for several frames in video and automatically propagate it on other frames.

Most video processing algorithms fail in occlusion area where motion estimation is unreliable. We propose a technique for occlusion processing. We evaluated several occlusion area detection and usage strategies. The proposed method was applied to depth map creation and depth map compression.