

УДК 004.932.4

МАТИРОВАНИЕ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ С ИСПОЛЬЗОВАНИЕМ ВОССТАНОВЛЕННОГО ФОНА

Ерофеев М.В., аспирант Московского государственного университета им. М.В. Ломоносова, факультет вычислительной математики и кибернетики, e-mail: merofeev@graphics.cs.msu.ru; Ватолин Д.С., к.ф.-м.н., с.н.с. Московского государственного университета им. М.В. Ломоносова, e-mail: dmitriy@graphics.cs.msu.ru.

VIDEO MATTING WITH AID OF RECONSTRUCTED BACKGROUND

Erofeev M.V., Vatolin D.S.

Formally, matting is a problem of image decomposition into foreground image, background image and foreground transparency map. This problem is extremely important for such video and image editing problems as: background replacement, applying transform to background or foreground only, stereoscopic image generation. In this paper we propose video matting method based on Learning Based Matting method. We describe modification to base method which enabled us to use reconstructed background sequence as additional input. We also propose iterative method for spatio-temporal transparency map smoothing. Finally, we show that proposed approach outperforms 11 image and video matting methods in objective comparison.

Key words: matting, video processing, background reconstruction, background inpainting, matting Laplacian.

Ключевые слова: матирование, обработка видео, восстановление фона, матирующий Лапласиан.

Введение

Одной из ключевых задач, возникающих при редактировании изображений и видеопоследовательностей, является построение карты прозрачности (матирования) объекта переднего плана для последующей замены фона или элементов фона, или изменения положения объекта относительно фона.

Формальная постановка задачи матирования имеет следующий вид: разделить данное изображение I на изображение объекта переднего плана F , изображение заднего плана B и карту прозрачности переднего плана α таким образом, чтобы было верно следующее уравнение:

$$I = \alpha F + (1 - \alpha) B \quad (1)$$

Задача матирования видео является обобщением задачи матирования изображений на видеопоследовательности. Такое обобщение порождает дополнительные требования к алгоритмам матирования видео:

Результирующая карта прозрачности должна быть стабильна во времени, ввиду высокой чувствительности зрительной системы человека к различиям между соседними кадрами.

Алгоритм должен иметь адекватную объемам видеоданных вычислительную сложность.

В данной работе предлагается алгоритм матирования видеопоследовательностей, в основу которого положен алгоритм матирования изображений на основе обучения (Learning Based Matting [1]). Отличительной особенностью предложенного алгоритма является использование изображений фона, содержащих фрагменты, скрытые объектом переднего плана на исходной

Задача матирования состоит в разделении данного изображения или видео на передний план, задний план и карту прозрачности переднего плана. Построение такого разделения важно для решения многих задач редактирования и обработки видео, таких как: замена фона, применение спецэффекта только к фону либо только к переднему плану, построение стереоизображения. Предлагается метод матирования видеопоследовательностей, основанный на методе матирования изображений Learning Based Matting. Описывается способ модификации базового метода, позволивший использовать информацию о фрагментах фона, скрытых объектом переднего плана, в качестве дополнительных данных; а также способ итеративной фильтрации последовательности карт прозрачности. Приводятся результаты сравнения предложенного алгоритма с 11 аналогами, иллюстрирующие преимущества предложенного подхода.

видеопоследовательности, в качестве дополнительных входных данных. Для построения таких изображений фона применяется алгоритм, описанный в [2], выполняющий восстановление фрагментов фона, скрытых на данном кадре, путем копирования этих же фрагментов из других кадров видеопоследовательности, где они являлись видимыми. Предлагаемый подход также выполняют два последовательных прохода по видеопоследовательности с целью обеспечения большей стабильности во времени результата в сравнении с результатом независимого применения алгоритма [1] к кадрам видеопоследовательности. Подробное описание алгоритма содержится в разделе «Предлагаемый алгоритм».

Предложенный алгоритм был протестирован на наборе из 10 видеопоследовательностей с известными эталонными картами прозрачности (данный набор был подготовлен в рамках работы [3] и сравнен с 11 другими методами матирования изображений и видео). Описание проведенной экспериментальной оценки содержится далее в разделе: «Экспериментальная оценка».

Обзор области исследований

Исследования в области матирования видео, в сравнении с областью матирования изображений, затруднены большим объемом входных данных, требуемых для работы алгоритмов матирования видео, и необходимостью обеспечения стабильности во времени результирующей карты прозрачности. Ввиду этого исследования в области матирования видео всегда отставали от прогресса в области матирования изображений. Несмотря на это, на сегодняшний день существуют 12 оригинальных методов матирования видео, а также объективная методика сравнения алгоритмов матирования видео, описанная в [3].

Большинство методов матирования видео основаны на одном из существующих методов матирования изображений. При этом основное внимание их авторы уделяли повышению временной стабильности результирующей последовательности карт прозрачности и уменьшению объема дополнительных данных, требуемых от пользователя.

Наиболее широко используемым в литературе способом сокращения объема дополнительных данных является распространение тернарных масок (карт, задающих три области: гарантированный передний план, гарантированный задний план и неизвестная область, для которой требуется решить задачу матирования) ключевых кадров на все кадры видеопоследовательности. Данный метод позволяет пользователю предоставить тернарные маски лишь для части кадров, затем алгоритм вычисляет вектора оптического потока между кадрами исходной видеопоследовательности и использует их для переноса масок на остальные кадры. Такой метод применяется в алгоритмах матирования видео [4, 5, 6, 7]. Схожим образом задачу сокращения объема ввода пользователя решает алгоритм [8], требуя в качестве дополнительных входных данных лишь карту прозрачности первого плана. Следует также отметить алгоритм [9], который, аналогично алгоритму сегментации изображений Grab Cut [10], лежащему в его основе, не требует никаких дополнительных данных от пользователя. Однако получаемые с его помощью карты прозрачности имеют крайне низкое качество.

Для обеспечения стабильности во времени результирующей последовательности карт прозрачности существует множество разнообразных подходов. Первый из них был описан в работе [5], ее авторы внедрили в оптимизируемую алгоритмом функцию стоимости слагаемое, моделирующее совместное распределение градиента исходной видеопоследовательности и последовательности карт прозрачности.

Большая группа существующих на сегодня алгоритмов матирования выполняет поиск итоговой карты прозрачности как минимума квадратичной функции стоимости, аналогичной описанной в [11]. Работы [12, 13, 14] обеспечивают стабильность результата во времени, внедряя в эту функцию слагаемое, требующее постоянства карты прозрачности соседних кадров вдоль векторов движения. В работах [6, 8] пространственный матирующий лапласиан с окнами 3×3 , описанный в [11], за-

менен пространственно-временным матирующим лапласианом с окнами $3 \times 3 \times 3$. Эта идея была развита в работе [15], заменившей матирующий лапласиан с окнами $3 \times 3 \times 3$ нелокальным матирующим лапласианом – лапласианом, задающим связи между пикселями, не обязательно находящимися по соседству. В работе [16] пространственно-временной лапласиан был заменен быстрым методом фильтрации Guided Filter [17], позволившим уменьшить время обработки одного кадра.

Другой подход к обеспечению стабильности карт прозрачности во времени состоит в создании моделей переднего плана и фона для групп кадров или всего видео вместо моделей, построенных для каждого кадра независимо (как в алгоритмах матирования изображений). Примерами этого подхода являются: метод [12], переносящий параметры классификаторов между соседними кадрами; построение тернарной маски с использованием модели цвета переднего плана текущего и предыдущего кадра в работе [8]; пространственно-временное сэмплирование образцов пикселей, примененное в [13]; использование глобального словаря цветов переднего плана и фона в методе [18].

Последним из рассматриваемых в данном обзоре методов обеспечения стабильности карты прозрачности во времени является темпоральная фильтрация карт прозрачности. В [7] авторы применяют алгоритм фильтрации на основе дистантного преобразования [19] для обеспечения стабильности во времени последовательности карт прозрачности, полученных с помощью алгоритма матирования изображений [20].

В заключение обзора области следует также упомянуть работу [3], в рамках которой было подготовлено 10 видеопоследовательностей с эталонными картами прозрачности и тернарными масками, предложены методы объективной оценки качества карт прозрачности, а также было проведено объективное сравнение ряда существующих алгоритмов матирования изображений и видео.

Предлагаемый алгоритм

Как уже отмечалось во введении, отличительной особенностью предлагаемого алгоритма является использование восстановленного фона. Очевидно, что знание цвета пикселей фона B в неизвестной области тернарной маски уменьшает множество возможных решений уравнения (1). Однако в случае задания изображения можно лишь примерно оценить их цвет (что и делают многие алгоритмы матирования изображений). Ситуация кардинально меняется в случае матирования видеопоследовательности: фрагменты фона, скрытые объектом переднего плана на одном кадре, зачастую видимы на других кадрах этой видеопоследовательности. Этот факт позволяет вычислить цвет пикселей фона неизвестной области путем копирования данных с соседних кадров.

Несмотря на то, что различными авторами ранее предлагались алгоритмы матирования видеопоследовательностей, использующие восстановленный фон, данная работа имеет новизну, благодаря:

Использованию алгоритма восстановления фона, применимого к более широкому классу видеопоследо-

вательностей, чем алгоритмы ранних работ.

Предлагаемому обобщению алгоритма матирования, основанного на матирующем лапласиане, используемому восстановленный фон в качестве дополнительных входных данных (ранее в литературе восстановленный фон использовался только в методах, основанных на сэмпировании (последние методы существенно отличаются от первых)).

Методу вычисления доверия к векторам оптического потока, оптимизированному для задачи матирования видео.

Входными данными предлагаемого в данной работе алгоритма являются: исходная видеопоследовательность, последовательность тернарных масок. Алгоритм выполняет следующие основные шаги:

Применить к исходной видеопоследовательности алгоритм восстановления фрагментов фона, скрытых объектами переднего плана.

Применить к каждому кадру исходной последовательности модифицированную версию алгоритма матирования изображений на основе обучения, использующую восстановленный фон в качестве дополнительных входных данных.

Применить итеративный алгоритм пространственно-временного сглаживания последовательности карт прозрачности.

Следует отдельно отметить использование итеративного подхода к сглаживанию карт прозрачности. Формально это может быть записано следующим образом: пусть α_t^0 – кадр t последовательности карт прозрачности, вычисленной изначально; $S(\cdot)$ – оператор сглаживания, α_t^i – кадр t результата фильтрации на итерации i , тогда

$$\alpha_t^i = S(\alpha_{t-1}^{i-1}, \alpha_t^{i-1}, \alpha_{t+1}^{i-1}). \quad (2)$$

Доминирующим подходом в литературе на сегодняшний день является метод последовательной фильтрации, результат которой (α_t^i) может быть записан как

$$\alpha_t^1 = \alpha_t^0 \quad (3)$$

$$\alpha_t^i = S(\alpha_t^{i-1} - 1, \alpha_t^0, \alpha_t^0 + 1). \quad (4)$$

Согласно проведенным экспериментам, последовательная фильтрация склонна накапливать ошибку, то есть расстояние между вычисленной картой прозрачности и эталоном увеличивается с номером кадра. Итеративный подход лишен данного недостатка, а также допускает параллельное вычисление карт прозрачности нескольких кадров. Далее мы рассмотрим основные шаги алгоритма более детально.

Восстановление фрагментов фона

Для восстановления фрагментов фона, скрытых передним планом, был использован алгоритм, описанный в [2]. Входными данными алгоритма являются видеопоследовательность и набор масок для каждого ее кадра, отмечающих области, подлежащие восстановлению. Результатом работы алгоритма является видеопоследовательность, совпадающая с исходной в областях, не отмеченных на масках, и содержащая фрагменты вос-

становленного фона в остальных областях. Дополнительными выходными данными алгоритма является набор масок областей, восстановление которых не было выполнено. Принцип работы алгоритма состоит в экстраполяции поля векторов движения в область восстановления и последующем копировании фрагментов фона других кадров согласно экстраполированным векторам. В данной работе в качестве маски восстановления используется объединение областей тернарной маски, отвечающих переднему плану и неизвестной области.

Базовый метод матирования изображений

Для лучшего понимания дальнейшего описания предлагаемого алгоритма матирования видео рассмотрим краткое описание алгоритма матирования изображений [1], лежащего в его основе. Входными данными алгоритма являются изображение в цветовом пространстве RGB и тернарная маска. Алгоритм выполняет обход изображения с использованием скользящего окна 3×3 пиксела. Далее для каждого окна ω , содержащего хотя бы один неизвестный пиксел, выразим прозрачность каждого его пиксела α' как линейную комбинацию w' прозрачностей его соседей α_ω :

$$\alpha' = w' \alpha_\omega. \quad (5)$$

Для получения w' воспользуемся приемом, схожим с линейной регрессией. Пусть, зная цвет пиксела $x' = [R; G; B; 1]$ и всех его соседей X_ω (каждая строка матрицы состоит из цвета пиксела в системе RGB и единицы), требуется предсказать его прозрачность α' . Дополнительно предположим, что известны прозрачности всех его соседей в текущем окне α_ω , тогда:

$$\alpha' = x'^T \beta, \quad (6)$$

где β – столбец из четырех элементов, таких, что:

$$\beta = \arg \min_{\beta} \| X_\omega \beta - \alpha_\omega \|^2 + \varepsilon \| \beta \|^2, \quad (7)$$

где $\varepsilon = 10^{-6}$ – параметр регуляризации. Последнее выражение содержит задачу минимизации квадратичного многочлена с положительно определенным гессианом, а значит, может быть сведено к:

$$\beta = (X_\omega^T X_\omega + \varepsilon I)^{-1} X_\omega^T \alpha_\omega. \quad (8)$$

Подставляя правую часть (8) в уравнение (6) и сопоставляя с уравнением (5), получим:

$$w' = x'^T (X_\omega^T X_\omega + \varepsilon I)^{-1} X_\omega^T. \quad (9)$$

Вектор-строка w вычисляется для каждого пиксела в каждом окне, содержащем хотя бы один пиксел неизвестной области. Для удобства записи дополним полученные вектора нулями, чтобы было верно:

$$\alpha' = w' \alpha, \quad (10)$$

где α – коэффициенты прозрачности всех пикселов изображения. Матрица W образуется из дополненных нулями векторов. Теперь задача вычисления карты прозрачности сводится к следующей задаче оптимизации:

$$\alpha = \arg \min_{\alpha} \| \alpha - W \alpha \|^2 + (\alpha - T)^T D (\alpha - T), \quad (11)$$

где T – тернарная маска, D – диагональная матрица, у которой элементы главной диагонали, отвечающие пикселам неизвестной области, равны нулю, а осталь-

ные – 10^6 . Решение последней задачи может быть сведено к решению разреженной системы линейных уравнений. Ее решение и есть результат работы базового метода матирования изображений.

Использование восстановленного фона

Рассмотрим, как базовый метод может быть модифицирован для использования восстановленного фона в качестве дополнительных данных. Важно, чтобы данная модификация была выполнена с учетом следующих фактов:

- результаты работы метода восстановления фона могут содержать неверно восстановленные фрагменты.
- восстановленный фон зачастую смещен на несколько пикселей из-за накопления ошибки восстановления.

Для тех окон, которые не содержат восстановленного фона, базовый метод применяется без модификаций. Для каждого окна ω , содержащего хотя бы один пиксел восстановленного фона, дополним матрицу X_ω (см. уравнение (7)) цветами пикселей восстановленного фона. Использование соответствий окон фона окнам изображений вместо прямых соответствий между пикселями обеспечивает устойчивость метода к сдвигам фона до 3 пикселей. Для обеспечения устойчивости метода к ошибкам восстановления фона модифицируем уравнение (7), взвесив образцы пикселей в матрице X_ω на доверие к ним:

$$\beta = \arg \min_{\beta} \|C(X_\omega \beta - \alpha_\omega)\|^2 + \varepsilon \|\beta\|^2, \quad (12)$$

где C – диагональная матрица, элементы главной диагонали которой определяют доверие к соответствующим строкам матрицы X_ω . В предлагаемом методе доверие ко всем образцам фона устанавливается равным 0,08 и 1 для остальных пикселей. Задача (12), аналогично предыдущему разделу, может быть переписана в виде:

$$\beta = (X_\omega^T C X_\omega + \varepsilon I)^{-1} X_\omega^T C \alpha_\omega, \quad (13)$$

тогда (аналогично уравнению (9)):

$$w' = x'^T (X_\omega^T C X_\omega + \varepsilon I)^{-1} X_\omega^T C. \quad (14)$$

Так как прозрачность пикселей восстановленного фона равна нулю, в уравнении (5) подставим нули вместо соответствующих им элементов вектора α_ω , эта операция эквивалентна удалению из вектора w' элементов, отвечающих пикселям восстановленного фона. Аналогично базовому методу получаем задачу вида (11), которая сводится к решению системы линейных уравнений. Таким образом, полученный метод расширяет каждое окно базового метода образцами пикселей восстановленного фона, прозрачность которых заведомо равна нулю, и учитывает их с низким доверием для обеспечения устойчивости к неточностям восстановленного фона.

Обеспечение стабильности результата во времени

Как уже упоминалось ранее, предложенный метод выполняет итеративную фильтрацию последовательно-сти карт прозрачности по схеме, описанной в уравнении (2). Здесь и далее запись v_k^n обозначает поле векторов

движения от кадра n к кадру k . Рассмотрим, какие операции выполняются при фильтрации кадра t :

1. Вычисление полей векторов движения $v_{t-1}^t, v_{t+1}^t, v_{t-1}^{t-1}$ и v_{t+1}^{t-1} с помощью алгоритма, описанного в [21] (алгоритм [21] вычисляет поле векторов движения v_k^n , выполняя поиск блоков кадра k , наиболее похожих на блоки кадра n ; смещения между блоками кадра n и найденными блоками кадра k образуют искомое поле векторов).

2. Вычисление карт доверия к полям векторов движения K_{t-1}^t и K_{t+1}^t .

3. Применение базового метода матирования, учитывая восстановленный фон, с дополнительными требованиями стабильности результата во времени.

Для вычисления доверия к векторам движения, полученным на первом шаге, используются две вспомогательные величины: **FBC** и E . **FBC** – несоответствие между полями векторов движения для прямого и обратного проходов. Значение **FBC** в точке с координатой x может быть выражено формулой:

$$\mathbf{FBC}_k^n(x) = \|v_k^n(x) - v_n^k(x + v_k^n(x))\|_2^2. \quad (15)$$

$E_k^n(x) - L_1$ расстояние между блоком 8x8 пикселей с центром в точке x кадра n исходной видеопоследовательности и аналогичным блоком кадра k с центром в точке $x + v_k^n(x)$.

С использованием введенных величин доверие к векторам движения вычисляется по формуле:

$$K_{t-1}^t = \exp\left(-\frac{\mathbf{FBC}_{t-1}^t}{\sigma_1} - \frac{E_{t-1}^t}{\sigma_2}\right), \quad (16)$$

где σ_1, σ_2 – константы, выбранные с использованием двух видеопоследовательностей с известной эталонной картой прозрачности из набора [3]. Для этого были вычислены вектора движения для этих последовательностей. Затем исключены из рассмотрения вектора, принадлежащие областям с малой дисперсией значений эталонной прозрачности. Для оценки качества вектора $v_k^n(x)$ используется пиковое отношение сигнала к шуму $P_k^n(x)$ между блоком 8x8 пикселей с центром в точке x эталонной карты прозрачности кадра n и аналогичным блоком кадра k с центром в точке $x + v_k^n(x)$. Путем визуального контроля было установлено, что значения P менее 12 децибел соответствуют ошибочно найденным векторам. σ_1, σ_2 были выбраны как:

$$\langle \sigma_1, \sigma_2 \rangle = \arg \min_{\sigma_1, \sigma_2} \sum_{P_k^n(x)} \exp\left(-\frac{\mathbf{FBC}_{t-1}^t}{\sigma_1} - \frac{E_{t-1}^t}{\sigma_2}\right) (12 - P_k^n(x)). \quad (17)$$

Для решения последней задачи был использован метод внутренней точки. Найденные значения: $\sigma_1 = 3,9$; $\sigma_2 = 29$.

Рассмотрим шаг матирования. Под операцией компенсации кадра k к кадру n будем понимать отображение R :

$$\Gamma = R(I, v_k^n), \tag{18}$$

такое, что

$$\Gamma(x + v_k^n(x)) = \Gamma(x). \tag{19}$$

Компенсируем предыдущий кадр, следующий кадр и их карты прозрачности (вычисленные на предыдущей итерации) к текущему кадру, будем обозначать их I_B , I_F , α_B и α_F соответственно. Аналогично описанному ранее способу расширим для каждого окна ω матрицу X_ω пикселями из I_B и I_F . Коэффициенты доверия к этим образцам вычисляются по формуле (16), а затем умножаются на 0,2 (экспериментально выбранное значение). В качестве значений прозрачности добавленных пикселей используются значения прозрачности в окне ω карт α_B и α_F . Добавим в уравнение (11) новые слагаемые:

$$\begin{aligned} &\psi[(\alpha - \alpha_B)^T \text{Diag}(K_{t-1}^t)(\alpha - \alpha_B) + \\ &+(\alpha - \alpha_F)^T \text{Diag}(K_{t+1}^t)(\alpha - \alpha_F)], \end{aligned} \tag{20}$$

где $\text{Diag}(K_{t-1}^t)$ – диагональная матрица, образованная значениями карты доверия K_{t-1}^t ; $\psi = 0,1$.

Предложенный метод выполняет две итерации сглаживания для вычисления итоговой карты прозрачности.

Экспериментальная оценка

Описанный метод был сравнен с другими методами матирования изображений и видео на 10 тестовых видео с известными эталонными картами прозрачности из набора, предложенного в [3] (примеры двух видеопоследовательностей из набора изображены рис. 1). Критериями оценки выступали: L_2 расстояние до эталона и **meSSDdt** (метод сравнения, предложенный в работе

[22], имеющий высокую корреляцию с результатами экспертной оценки). Величина **meSSDdt** выражает среднее различие между ошибкой метода на текущем кадре и скомпенсированной, согласно векторам движения ошибкой на следующем кадре. Иными словами, **meSSDdt** штрафует изменения ошибки между соседними кадрами и тем самым позволяет оценить стабильность результата во времени. Результаты сравнения предложенного метода с 10 ранее существовавшими приведены в табл. 1 и 2. Для каждой тестовой последовательности методы были упорядочены согласно критерию сравнения, а затем было вычислено среднее место метода во всех тестах. Как видно из таблиц, предложенный метод превзошел аналоги по данному показателю. Рассмотрим также сравнение предложенного метода с алгоритмом матирования видео, предложенным в 2015 году [18]. Сравнение было выполнено на устаревшей версии набора тестовых видео (к сожалению, авторы метода не предоставили результатов тестирования своего метода на обновленном наборе). Результаты представлены на рис. 1, согласно которому предложенный метод превзошел аналог.

В табл. 1 представлены результаты сравнения предложенного алгоритма с 10 аналогами на 10 тестовых видеопоследовательностях из набора, описанного в [3]. Пространственные ошибки методов были вычислены как L_2 расстояния между результатами работы алгоритмов и эталоном (приведены в столбцах 3-12). Затем методы были упорядочены по близости результата к эталону (надстрочные числа в столбцах 3-12). Для каждого метода было вычислено среднее место во всех тестах (второй столбец). Меньшие значения соответствуют более высоким местам в рейтинге. Предложенный метод превзошел аналоги по среднему месту во всех тестах.



а) Исходная видеопоследовательность

б) Эталонная карта прозрачности

Рис. 1. Пример двух видеопоследовательностей (а) с эталонными картами прозрачности (б) из набора тестовых данных [3]

Таблица 1.

Метод	Ср. место	Тестовая последовательность									
		№1	№2	№3	№4	№5	№6	№7	№8	№9	№10
Предложенный	1,9	29,39 ¹	27,79 ²	8,83 ²	122,22 ⁵	13,96 ¹	21,01 ¹	31,83 ³	27,66 ¹	49,26 ¹	59,67 ²
Learning Based [1]	3,1	29,98 ²	33,08 ⁵	10,84 ⁶	134,82 ⁷	15,23 ²	23,35 ²	31,46 ²	30,20 ²	58,38 ²	56,60 ¹
Closed Form [11]	4,3	31,34 ³	31,80 ⁴	10,96 ⁷	149,48 ⁹	15,86 ³	23,67 ³	30,15 ¹	30,60 ³	69,94 ⁴	82,13 ⁶
Comprehensive Sampling [23]	4,3	34,48 ⁴	30,18 ³	12,10 ⁸	79,60 ¹	16,17 ⁴	25,02 ⁴	35,25 ⁵	41,05 ⁵	77,72 ⁵	70,18 ⁴
Robust Matting [20]	4,6	37,54 ⁶	26,15 ¹	9,61 ³	88,15 ³	18,52 ⁵	33,31 ⁶	34,15 ⁴	46,91 ⁷	78,56 ⁶	80,41 ⁵
Shared Matting [24]	4,9	36,79 ⁵	46,72 ⁷	10,30 ⁴	87,64 ²	24,41 ⁶	35,18 ⁷	35,26 ⁶	45,23 ⁶	68,52 ³	67,17 ³
Refine Edge [7]	5,8	37,78 ⁷	38,48 ⁶	8,64 ¹	129,56 ⁶	27,47 ⁷	31,36 ⁵	36,65 ⁷	38,75 ⁴	102,02 ⁸	94,50 ⁷
Bayesian Matting [25]	7,9	58,76 ⁹	51,66 ⁸	10,79 ⁵	144,53 ⁸	30,36 ⁸	53,28 ⁸	43,70 ⁸	70,50 ⁹	96,69 ⁷	140,12 ⁹
KNN Matting [26]	8,8	67,83 ¹⁰	92,22 ⁹	64,73 ¹¹	112,46 ⁴	32,96 ⁹	61,50 ⁹	50,87 ⁹	79,20 ¹⁰	113,41 ⁹	123,63 ⁸
Nonlocal matting [27]	9,7	54,21 ⁸	92,37 ¹⁰	33,86 ¹⁰	243,47 ¹¹	45,13 ¹⁰	64,05 ¹⁰	61,75 ¹⁰	67,30 ⁸	135,08 ¹⁰	171,79 ¹⁰
Spectral Matting [28]	10,7	86,32 ¹¹	105,24 ¹¹	13,15 ⁹	210,04 ¹⁰	48,76 ¹¹	65,72 ¹¹	68,83 ¹¹	109,44 ¹¹	157,18 ¹¹	182,82 ¹¹

Таблица 2.

Метод	Ср. место	Тестовая последовательность									
		№1	№2	№3	№4	№5	№6	№7	№8	№9	№10
Предложенный	1,3	0,44 ¹	0,941 ²	0,07 ²	2,65 ¹	0,15 ¹	0,25 ¹	0,66 ²	0,51 ¹	1,60 ¹	2,20 ¹
Closed Form [11]	3,1	0,49 ²	1,35 ⁴	0,09 ⁶	2,88 ²	0,19 ³	0,35 ³	0,61 ¹	0,61 ²	2,75 ³	3,38 ⁵
Learning Based [1]	3,2	0,50 ³	1,42 ⁵	0,09 ⁷	2,88 ³	0,18 ²	0,34 ²	0,68 ³	0,64 ³	2,37 ²	2,43 ²
Refine Edge [29]	4,6	0,55 ⁴	1,48 ⁶	0,06 ¹	3,99 ⁶	0,34 ⁶	0,54 ⁵	0,70 ⁴	0,81 ⁴	3,80 ⁶	3,16 ⁴
Comprehensive Sampling [23]	5,1	0,74 ⁵	1,06 ³	0,19 ⁹	3,15 ⁴	0,22 ⁴	0,39 ⁴	0,86 ⁶	1,24 ⁵	3,65 ⁵	3,40 ⁶
Robust Matting [20]	5,6	0,89 ⁷	0,936 ¹	0,07 ³	4,82 ⁸	0,25 ⁵	0,88 ⁶	0,77 ⁵	1,58 ⁷	3,89 ⁷	3,92 ⁷
Shared Matting [24]	5,9	0,79 ⁶	2,73 ⁷	0,09 ⁵	4,50 ⁷	0,35 ⁷	1,00 ⁷	0,91 ⁷	1,43 ⁶	2,94 ⁴	2,94 ³
KNN Matting [26]	8,4	2,28 ¹⁰	6,40 ⁹	1,22 ¹¹	3,47 ⁵	0,61 ⁸	1,02 ⁸	1,73 ⁹	1,94 ⁸	6,42 ⁸	7,73 ⁸
Bayesian Matting [25]	9,1	1,85 ⁹	3,58 ⁸	0,12 ⁸	13,33 ⁹	0,74 ⁹	2,31 ¹¹	1,47 ⁸	4,51 ¹⁰	8,01 ⁹	16,23 ¹⁰
Spectral Matting [28]	9,8	7,03 ¹¹	13,25 ¹¹	0,08 ⁴	15,36 ¹⁰	2,09 ¹¹	1,84 ⁹	2,45 ¹⁰	6,59 ¹¹	8,71 ¹⁰	19,97 ¹¹
Nonlocal matting [27]	9,9	1,56 ⁸	8,17 ¹⁰	0,70 ¹⁰	20,97 ¹¹	1,11 ¹⁰	1,88 ¹⁰	2,51 ¹¹	2,35 ⁹	11,17 ¹¹	13,37 ⁹

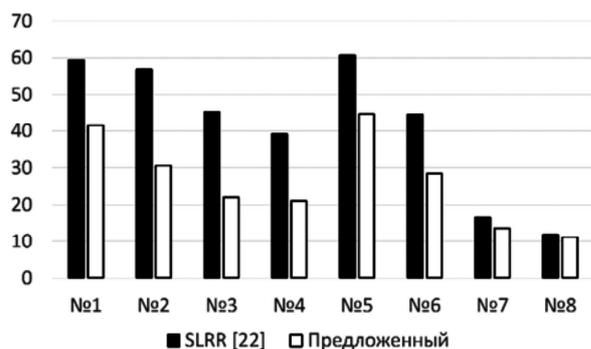
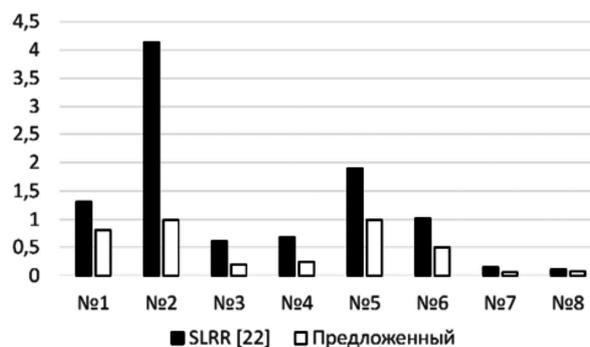
а) L_2 расстояние до эталонаб) $meSSDdt$ (предложена в [22])

Рис. 2. Сравнение предложенного алгоритма с алгоритмом, описанным в [18], на 8 тестовых видео последовательностях (каждая пара столбцов соответствует результату для одной тестовой последовательности)

с использованием L_2 расстояния до эталона (а) для оценки пространственной ошибки и $meSSDdt$ (б) для оценки временной стабильности результата

В табл. 2 представлены результаты сравнения предложенного алгоритма с 10 аналогами на 10 тестовых видеопоследовательностях из набора, описанного в [3]. Для каждой видеопоследовательности были вычислены значения квазиметрики $meSSDdt$ (в работе [22] показана высокая корреляция данного способа оценки с результатами экспертной оценки) между результатами работы алгоритмов и эталоном (приведены в столбцах

3-12) для оценки стабильности во времени результатов. Затем методы были упорядочены по близости результата к эталону (надстрочные числа в столбцах 3-12). Для каждого метода было вычислено среднее место во всех тестах (второй столбец). Меньшие значения соответствуют более высоким местам в рейтинге. Предложенный метод превзошел аналоги по среднему месту во всех тестах.

Заключение

В работе был предложен алгоритм матирования видеопоследовательностей, использующий восстановленный фон в качестве дополнительных данных. К основным особенностям можно отнести:

- модификацию базового алгоритма матирования, позволившую использовать фон в качестве дополнительных данных;

- итеративный подход к пространственно-временному сглаживанию карт прозрачности, позволивший избежать накопления ошибки и вычислять несколько кадров параллельно;

- способ вычисления доверия к векторам движения, выбранный путем анализа двух видео с эталонной картой прозрачности.

Преимущества предложенного метода были продемонстрированы путем объективного сравнения результатов его работы с 11 аналогами.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта 15-01-08632 а.

Данная статья поступила в редакцию журнала «Цифровая обработка сигналов» 22.03.2016 г.

Литература

1. Zheng Yuanjie, Kambhamettu C. Learning based digital matting // International Conference on Computer Vision (ICCV). – 2009. – P. 889-896.

2. Зачесов А. А., Ерофеев М. В., Ватолин Д. С. Использование карт глубины при восстановлении фона в видеопоследовательностях // Новые информационные технологии в автоматизированных системах: материалы научно-практического семинара. – М.: МИЭМ НИУ ВШЭ, 2015.

3. Методика объективного сравнения алгоритмов матирования видео/М. В. Ерофеев, Ю. А. Гитман, Д. С. Ватолин, А. А. Федоров // Цифровая обработка сигналов. – 2015. – № 3. – С. 53–59.

4. Video matting of complex scenes / Yung-Yu Chuang, Aseem Agarwala, Brian Curless et al.//ACM Transactions on Graphics (TOG). – 2002. – Vol. 21, no. 3. – P. 243-248.

5. Apostoloff N., Fitzgibbon A. Bayesian video matting using learnt image priors // Computer Vision Pattern Recognition (CVPR). – Vol. 1. – 2004. – P. 1-407-1-414.

6. Lee Sun-Young, Yoon Jong-Chul, Lee In-Kwon. Temporally coherent video matting // Graphical Models. – 2010. – Vol. 72, no. 3. – P. 25–33.

7. Bai Xue, Wang Jue, Simons David. Towards temporally coherent video matting // International Conference on Computer Vision (ICCV). – 2011. – P. 63-74.

8. Video matting via opacity propagation / Z. Tang, Z. Miao, Y. Wan, D. Zhang // The Visual Computer. – 2012. – Vol. 28, no. 1. – P. 47-61.

9. Corrigan D., Robinson S, Kokaram A. Video matting using motion extended grabcut // European Conference on Visual Media Production (CVMP). – 2008. – P. 3–3(1).

10. Rother C., Kolmogorov V., Blake A. Grabcut: Interactive foreground extraction using iterated graph cuts // ACM transactions on graphics (TOG) / ACM. – Vol. 23. – 2004. – P. 309–314.

11. Levin A., Lischinski D., Weiss Y. A closed-form solution to natural image matting // IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). – 2008. – Vol. 30, no. 2. – P. 228–242.

12. Video snapcut: Robust video object cutout using localized classifiers / X. Bai, J. Wang, D. Simons, G. Sapiro // ACM Transactions on Graphics (TOG). – 2009. – Vol. 28, no. 3. – P. 70:1–70:11.

13. Temporally coherent and spatially accurate video matting / Ehsan Shahrian, Brian Price, Scott Cohen, Deepu Rajan // Computer Graphics Forum. – 2014. – Vol. 33, no. 2. – P. 381–390.

14. Sindeev M., Konushin A., Rother C. Alpha-flow for video matting // Computer Vision-ACCV 2012. – Springer, 2012. – P. 438-452.

15. Choi I., Lee M., Tai Y.-W. Video matting using multi-frame nonlocal matting laplacian // European Conference on Computer Vision (ECCV). – 2012. – P. 540-553.

16. Spatio-temporally coherent interactive video object segmentation via efficient filtering / N. Brosch, A. Hosni, C. Rhemann, M. Gelautz // Pattern Recognition. – Vol. 7476. – 2012. – P. 418-427.

17. He Kaiming, Sun Jian, Tang Xiaoou. Guided image filtering // European Conference on Computer Vision (ECCV). – 2010. – P. 1–14.

18. Video matting via sparse and low-rank representation / D. Zou, X. Chen, G. Cao, X. Wang // Proceedings of the IEEE International Conference on Computer Vision. – 2015. – P. 1564-1572.

19. Linear time euclidean distance transform algorithms / H. Breu, J. Gil, D. Kirkpatrick, M. Werman // Pattern Analysis and Machine Intelligence, IEEE Transactions on. – 1995. – Vol. 17, no. 5. – P. 529–533.

20. Wang J., Cohen M. F. Optimized color sampling for robust matting // Computer Vision Pattern Recognition (CVPR). – 2007. – P. 1–8.

21. Fast video super-resolution via classification / K. Simonyan, S. Grishin, D. Vatolin, D. Popov // International Conference on Image Processing (ICIP). – 2008. – P. 349–352.

22. Perceptually motivated benchmark for video matting / M. Erofeev, Y. Gitman, D. Vatolin et al. // Proceedings of the British Machine Vision Conference (BMVC). – BMVA Press, 2015. – September. – P. 99.1–99.12. – URL: <https://dx.doi.org/10.5244/C.29.99>.

23. Improving image matting using comprehensive sampling sets / E. Shahrian, D. Rajan, B. Price, S. Cohen // Computer Vision Pattern Recognition (CVPR). – 2013. – P. 636–643.

24. Gastal Eduardo S.L., Oliveira Manuel M. Shared sampling for real-time alpha matting // Computer Graphics Forum. – 2010. – Vol. 29, no. 2. – P. 575–584.

25. A bayesian approach to digital matting / Y.-Y. Chuang, B. Curless, D. H. Salesin, R. Szeliski // Computer Vision Pattern Recognition (CVPR). – Vol. 2. – 2001. – P. 11-264-11-271.

26. Chen Qifeng, Li Dingzeyu, Tang Chi-Keung. KNN matting // IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). – 2013. – Vol. 35, no. 9. – P. 2175-2188.

27. Lee Philip, Wu Ying. Nonlocal matting // Computer Vision Pattern Recognition (CVPR). – 2011. – P. 2193–2200.

28. Levin A., Rav Acha A., Lischinski D. Spectral matting // IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI). – 2008. – Vol. 30, no. 10. – P. 1699-1712.

29. <http://www.adobe.com/en/products/aftereffects.html>.