

МЕТОДЫ КОНВЕРТАЦИИ 2D-ИЗОБРАЖЕНИЙ И ВИДЕО В СТЕРЕОСКОПИЧЕСКИЙ ФОРМАТ

Н. Н. Красильников^а, доктор техн. наук, профессор

О. И. Красильникова^а, канд. техн. наук, доцент

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Санкт-Петербург, РФ

Постановка проблемы: известные методы, обеспечивающие качественную конвертацию 2D-изображений и видео в стереоскопический формат, имеют общий недостаток — ручную реализацию ряда этапов и, следовательно, являются дорогостоящими. **Цель:** разработка полуавтоматических и потому более экономичных методов конвертации 2D-изображений и видео в стереоскопический формат. **Результаты:** предложенные шесть методов конвертации основаны, во-первых, на положениях выдвинутой ранее и экспериментально подтвержденной гипотезы, согласно которой ощущение глубины пространства при наблюдении 3D-сцен и стереоскопических изображений создается благодаря «конвергенционно-дивергенционным» движениям глаз, а во-вторых, на следующем из нее выводе, что при генерации стереоизображений важны лишь области карты глубины, примыкающие к резким световым границам. Первый метод заключается в аппроксимации поверхностей объектов, представленных на конвертируемом изображении, небольшим количеством плоскостей. Второй метод основан на синтезе карты глубины небольшим набором примитивных поверхностей. Особенностью третьего метода конвертации является синтез карты глубины из ахроматического компонента самого конвертируемого изображения. Четвертый метод конвертации отличается от предыдущего тем, что карта глубины формируется из контуров ахроматического компонента конвертируемого изображения. В пятом и шестом методах при формировании карты глубины использован фактор движения объектов. **Практическая значимость:** разработанные методы обладают большим быстродействием и требуют меньших затрат ручного труда, поэтому их можно эффективно применять на практике для конвертации как отдельных изображений, так и видео.

Ключевые слова — 3D-изображение, стереоскопическое изображение, конвертация, 2D-изображение, видео, 3D-формат.

Введение

В настоящее время значительное внимание уделяется конвертации 2D-изображений и фильмов в так называемый «3D»-формат. Мы заключаем в кавычки термин «3D», поскольку на самом деле речь идет о конвертации 2D-изображений и фильмов в стереоскопический формат, а использование термина «3D»-формат носит, скорее, рекламный характер. Повышенный интерес к конвертации обусловлен отсутствием достаточного количества качественного контента в данном формате. Несмотря на то что в Интернете имеется множество предложений по конвертации 2D-изображений и видео в «3D»-формат, а также бесплатных приложений конвертации и приложений конвертации, предоставляемых за весьма умеренную плату, описания самих методов конвертации практически отсутствуют. В Интернете имеются упоминания лишь о двух технологиях конвертации, которые считаются основными [1, 2].

Первая технология заключается в том, что для конвертируемого 2D-изображения сначала рисуется карта глубины, т. е. по существу для каждого объекта, представленного на этом изображении, рисуется силуэт, залитый серыми тонами разной светлоты. Затем, с использованием конвертируемого 2D-изображения и карты гл-

бины, рассчитываются левое и правое изображения стереопары.

При использовании второй технологии в 3D-среде 3D-графического редактора, например 3Ds Max, создается виртуальная модель сцены, повторяющая сцену, представленную на 2D-изображении. Затем на эту виртуальную модель проецируется 2D-изображение, и результат фотографируется двумя виртуальными камерами, смещенными по горизонтали на величину базы. Полученные таким образом изображения образуют стереопару.

Недостатком этих технологий является то, что их реализация требует больших затрат ручного труда. Действительно, вначале оператор вручную разделяет сцену конвертируемого изображения на составляющие ее объекты, т. е. выполняет сегментацию исходного 2D-изображения. Эта сегментация является семантической, т. е. смысловой, поскольку при этом определяется, какие пиксели изображения к какому объекту относятся. Последующие операции каждой из двух технологий также выполняются в ручном режиме. При использовании первой технологии после семантической сегментации области, занимаемые каждым из объектов на 2D-изображении, заполняются серым цветом, интенсивность которого (или его отдельных областей, если объект протяженный в глубину сцены) определяется оператором в соответствии с удаленностью этого объекта от

камеры. Применяя вторую технологию, после выполнения семантической сегментации виртуальную модель сцены, содержащей все объекты, оператор создает вручную.

Для автоматизации процесса конвертации необходимо решить две задачи, относящиеся к области искусственного интеллекта, которые до настоящего времени не решены. Это задача семантической сегментации изображений и задача автоматического определения расстояния от объекта до камеры на основании перспективных трансформаций изображений сегментированных объектов. В результате того, что описанные выше технологии базируются на ручном труде оператора, они оказались чрезвычайно дорогими. По словам режиссера Джеймса Кэмерона, создателя фильма «Титаник», стоимость конвертации этого фильма с использованием технологии рисования карт глубины составила около \$18 млн и потребовала работы 300 специалистов в течение 60 недель [2].

Стремление автоматизировать конвертацию 2D-изображений и видео в «3D»-формат привело к появлению ряда упрощенных методов, возможность разработки которых положили исследования, опубликованные в работах [3–6]. Некоторые из этих методов рассмотрены в настоящей статье.

Метод конвертации, основанный на аппроксимации 2D-изображения набором плоскостей

В основе метода, рассматриваемого здесь, лежат результаты наших исследований [9], по итогам которых было показано, что:

— ощущение глубины пространства создается благодаря тому, что при наблюдении 3D-сцен и стереоскопических изображений глаза зрителя при сканировании постоянно совершают «конвергенционно-дивергенционные» движения, одновременно фиксируясь вблизи резких световых границ составляющих их объектов;

— стереоскопические изображения допускают большие искажения параллакса без возникновения заметных зрителю искажений, из чего

следует допустимость больших погрешностей при создании карт глубины.

Отличием разработанного нами метода от метода, в котором применяется виртуальная модель сцены, представленной на 2D-изображении, является то, что при ее создании выполняется весьма грубая аппроксимация поверхностей составляющих ее объектов малым количеством плоскостей [7, 8]. Благодаря этому существенно упрощается процедура моделирования и в то же время возникающие погрешности параллакса для зрителя остаются незаметными в силу отмеченного выше свойства зрения. В качестве примера на рис. 1, а в черно-белом варианте приведено 2D-изображение, на котором белыми линиями показаны границы аппроксимирующих плоскостей, которыми аппроксимируются стены, пол и потолок помещения, а на рис. 1, б — карта глубины, которую представляют эти пять плоскостей.

Светлые области карты глубины соответствуют частям сцены, близко расположенным к зрителю, а темные области — удаленным частям сцены.

Другим отличием данного метода является способ получения стереопары, которая рассчитывается на основе полученной карты глубины. Вначале вычисляются значения параллакса p по известной формуле

$$p = \frac{b(d_2 - d_1)}{d_2},$$

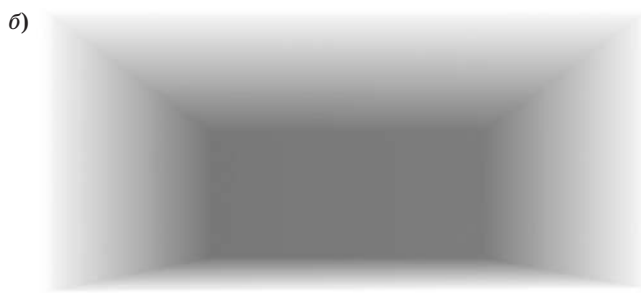
где b — база (расстояние между оптическими осями); d_1 — расстояние от зрителя до экрана, на котором наблюдается стереоскопическое изображение; d_2 — расстояние от зрителя до наблюдаемого виртуального объекта (берется из карты глубины). После этого рассчитываются компоненты изображений стереопары:

— левого

$$L_{LR}(x, y) = L_R(x - p/2, y);$$

$$L_{LG}(x, y) = L_G(x - p/2, y);$$

$$L_{LB}(x, y) = L_B(x - p/2, y);$$



■ Рис. 1. Исходное изображение, подлежащее конвертации (а), и его карта глубины (б)

— правого

$$L_{RR}(x, y) = L_R(x + p/2, y);$$

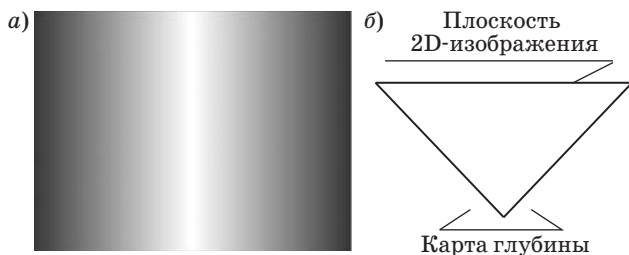
$$L_{RG}(x, y) = L_G(x + p/2, y);$$

$$L_{RB}(x, y) = L_B(x + p/2, y).$$

Здесь $L_{LR}(x, y)$, $L_{LG}(x, y)$, $L_{LB}(x, y)$ — зависимость интенсивности излучения от координат x и y для левого компонента в красном, зеленом и синем цветах; $L_{RR}(x, y)$, $L_{RG}(x, y)$, $L_{RB}(x, y)$ — то же самое для правого компонента. В приведенных формулах для простоты изложения изображения рассматривались как непрерывные функции координат x и y . Однако реальные изображения, представленные набором пикселей, дискретны в пространстве. Вследствие пространственных преобразований, обусловленных зависимостью параллакса от пространственных координат, в одних зонах стереоскопических изображений имеют место растяжения в горизонтальном направлении, а в других — сжатия. В результате этого на изображениях возникают места, в которых интенсивность пикселей не определена. Эта неопределенность устраняется методом интерполяции в горизонтальном направлении. В работе [10] на рис. П1 и П2 приведены исходное цветное 2D-изображение и результат его конвертации в «3D»-формат описанным методом. Следует особо отметить, что для изображений сцен, в которых отсутствуют объекты, заслоняющие собой какие-либо части изображенной сцены, полученная карта глубины достаточно точно отражает реальную ситуацию, что позволяет при расчете стереоскопического изображения изменять ракурс наблюдения в довольно широких пределах. Рис. П3 и П4 [10] демонстрируют это.

Метод конвертации 2D-изображения в «3D»-изображение, основанный на задании примитивной карты глубины

Следующий метод, предложенный нами, основан на задании в качестве карты глубины произвольной поверхности, в простейшем случае — двух плоскостей, ориентированных перпендикулярно плоскости горизонта (рис. 2, а, б).



■ Рис. 2. Примитивная карта глубины (а) в виде двух плоскостей, ориентированных перпендикулярно плоскости горизонта (б)

Карта глубины, заданная таким образом, обеспечивает лишь разные значения параллаксов для пикселей, имеющих различное горизонтальное расположение. Но благодаря тому, что при просмотре такого изображения имеют место «конвергенционно-дивергенционные» движения зрительных осей, будет возникать ощущение глубины пространства [9]. Естественно, вследствие того, что карта глубины задана произвольно, значения параллаксов стереоскопического изображения оказываются сильно искаженными, следовательно, расстояния до объектов, которые зрительная система «вычисляет» через значения параллаксов, могут существенно отличаться от расстояний, которые она вычисляет путем сопоставления масштабов объектов, расположенных на различных расстояниях, тем не менее в большинстве случаев эти расхождения остаются незамеченными зрителем. В этом легко убедиться, обратившись к рисункам сайта [11], на которых приведены исходные 2D-изображения и стереоскопические изображения, представляющие результат их конвертации описанным способом. В качестве карты глубины могут быть использованы также фрагменты цилиндрических, гауссовых и других поверхностей. В отличие от метода, описанного выше, при расчете стереоскопического изображения попытка изменить ракурс наблюдения даже в небольших пределах в данном случае приводит к недопустимо большим геометрическим искажениям.

Метод конвертации, основанный на использовании карты глубины, сформированной из исходного изображения

Недостатком метода, рассмотренного в предыдущем разделе, является то, что карта глубины формируется вне какой-либо связи с объектами сцены. В результате этого возникает упорядоченное изменение параллакса в направлении, которое определяется только положением задаваемой произвольной поверхности, используемой как карта глубины, и которое не связано с положением объектов в пространстве. Однако для ряда изображений удается избежать этого недостатка, формируя карту глубины из самого конвертируемого изображения. Примером таких изображений являются изображения, на которых сильно выражен эффект воздушной перспективы. Используя то обстоятельство, что близко расположенные объекты имеют больший контраст, чем удаленные, и формируя из ахроматических компонентов таких изображений путем их небольшого размытия карты глубины, удается получать на сформированных таким методом картах сильное различие по интенсивности между областями,

которые соответствуют близко расположенным объектам, и областями, которые соответствуют удаленным объектам, а следовательно, разнести их по дальности в правильном направлении. В результате многочисленных экспериментов было установлено, что наилучший результат достигается при размытии ахроматического компонента конвертируемого изображения в горизонтальном направлении путем его свертки с гауссовой функцией

$$L_d(n, k) = \sum_{n'=-\frac{N_n-1}{2}}^{\frac{N_n-1}{2}} L_{ca}(n+n', k)h(n'),$$

где $L_{ca}(n, k)$ — ахроматический компонент конвертируемого изображения; k — номер строки; n — номер столбца; N_n — протяженность импульсной функции;

$$h(n) = k_1 \exp(-n/n_0).$$

Здесь k_1 — нормирующий множитель; n_0 — параметр, определяющий ширину импульсной функции. В качестве примера, иллюстрирующего эффективность описанного метода, на рис. П5 [10] приведено изображение, которое затем было конвертировано в анаглифные стереоскопические изображения методом, основанным на задании примитивной карты глубины (рис. П6, [10]), и методом, в котором карта глубины формировалась из исходного изображения в соответствии с приведенным описанием (рис. П7, [10]). Сопоставление изображений, представленных на рис. П6 и П7 [10], показывает, что использование дополнительной информации, которая заключена в исходном изображении в виде различия контраста, позволило существенно усилить эффект объема стереоскопического изображения. Так же, как и в случае метода конвертации, основанного на задании примитивной карты глубины, при использовании данного метода попытка изменить ракурс наблюдения при расчете стереоскопического изображения приводит к недопустимо большому геометрическим искажениям.

Метод конвертации, основанный на использовании карты глубины, сформированной из контуров исходного изображения

При выполнении экспериментов, описанных в предыдущем разделе, было обращено внимание на то, что карта глубины в своих внутренних областях, соответствующих областям объекта с плавным изменением яркости и цвета и отстоящих от световых границ объекта на расстояние, превышающее значение величины погрешности фиксации зрительных осей относительно вертикальных световых границ объекта, может иметь произвольные значения интенсивности (яркости), и это не сказывается на формируемом стереоскопическом изображении. Кроме того, если карту глубины конкретного объекта расширить за его пределы, но так, чтобы она не вторглась в зону световых границ другого объекта, это также не скажется на формируемом стереоскопическом изображении. Эти факты не являются неожиданными и объясняются тем, что интенсивность (яркость) карты глубины важна лишь вблизи световых границ объекта, поскольку она в конечном итоге определяет, насколько должна быть смещена эта граница при расчете компонентов стереоскопического изображения вследствие параллакса. Это полностью согласуется с выдвинутой нами гипотезой [9], которая объясняет механизм создания у зрителя ощущения глубины наблюдаемой сцены. Данное свойство карты глубины может быть использовано на практике при конвертации 2D-изображений в стереоскопические изображения в процессе редактирования карты глубины. Получение контурного препарата из конвертируемого изображения — процедура не сложная, а редактировать карту глубины в виде контурного изображения значительно проще, чем редактировать карту глубины, представленную в виде полутонового изображения. Поясним сказанное на примере конвертации изображения, ахроматическая версия которого приведена на рис. 3, а, с дополнительным требованием сильного разделения переднего и заднего планов.



■ Рис. 3. Исходное изображение (а) и карта глубины (б), сформированная из контуров исходного изображения

То есть при конвертации требуется усилить разделение по глубине сцены переднего и заднего планов.

Поставленная задача рассматриваемым методом решается следующим образом. Вначале из исходного изображения получается контурный препарат одним из известных методов. Лучше всего для этой цели использовать метод Канны, поскольку при этом на контурном препарате получается минимальное количество разрывов и отдельных штрихов и точек. После этого в данном конкретном случае следует устранить контуры, созданные объектами заднего плана. Работая с контурами, это сделать легко. Затем отредактированный таким образом препарат нужно подвергнуть гауссовой фильтрации в горизонтальном и вертикальном направлениях с целью устранить остаточные разрывы контуров и расширения контурных линий, а затем после двустороннего ограничения присвоить ему соответствующий уровень яркости. Полученный результат можно использовать как карту глубины (рис. 3, б). На рис. П8 и П9 [10] приведены исходное изображение и результат его конвертации в анаглифное стереоскопическое изображение. Из рис. П9 [10] видно, что полученный результат впечатляет.

Метод конвертации, основанный на использовании фактора движения при формировании карты глубины

В ряде случаев дополнительным признаком, отличающим задний план сцены от переднего плана, на котором сосредоточены объекты, является движение объектов. Это может быть использовано при формировании карты глубины. Переходя к рассмотрению предлагаемого метода конвертации, основанного на использовании данного признака, будем считать видеокамеру, посредством которой производится съемка, неподвижной. В качестве исходного препарата, из которого синтезируется карта глубины в этом методе, используются абсолютные значения разности интенсивности (яркости) ахроматических компонентов конвертируемого кадра и кадра, который ему предшествует:

$$L_{\Pi}(k, n, i) = |L_c(k, n, i) - L_c(k, n, i - 1)|,$$

где $L_{\Pi}(k, n, i)$ — исходный препарат; $L_c(k, n, i)$ — распределение яркости в конвертируемом кадре изображения; $L_c(k, n, i - 1)$ — распределение яркости в кадре, предшествующем конвертируемому кадру; k — номер строки; n — номер столбца (пикселя в строке); i — номер кадра. Исходный препарат не может быть непосредственно использован в качестве карты глубины по двум причинам. Первой причиной является наличие флуктуационного шума, а также небольшого разност-

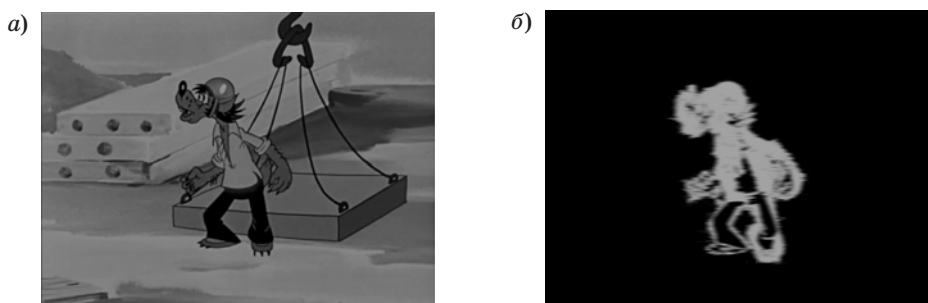
ного сигнала, обусловленного небольшими движениями видеокамеры, особенно при съемке с рук, которые в конечном итоге будут создавать ложные, изменяющиеся во времени приращения параллакса. Для того чтобы избавиться от этого эффекта, исходный препарат $L_{\Pi}(k, n, i)$ необходимо подвергнуть ограничению снизу, в результате чего получаем первый промежуточный препарат $L_{\Pi1}(k, n, i)$:

$$L_{\Pi1}(k, n, i) = \begin{cases} L_{\Pi}(k, n, i) & \text{при } L_{\Pi}(k, n, i) \geq P_{\Pi} \\ 0 & \text{при } L_{\Pi}(k, n, i) < P_{\Pi} \end{cases}, \quad (1)$$

где P_{Π} — нижний порог. Вторая причина, по которой, даже избавившись от флуктуационного шума и разностного сигнала, обусловленного небольшими движениями видеокамеры, нельзя использовать препарат $L_{\Pi1}(k, n, i)$ в качестве карты глубины, является сильная зависимость интенсивности $L_{\Pi1}(k, n, i)$ не только от яркости, но и от скорости движения объекта и его отдельных частей, которая непрерывно изменяется внутри одной сцены. Примером может служить фигура человека, который, будучи сам неподвижным, лишь двигает рукой. В этом случае, рассматривая препарат $L_{\Pi1}(k, n, i)$ как карту глубины, обнаружим на нем только нечто вроде силуэта руки, так как лишь она движется, а все остальные части изображения неподвижны. А поскольку от скорости движения руки зависит интенсивность этой части препарата (яркость), то величина рассчитанного по нему параллакса также будет определяться этой скоростью. В результате кажущаяся дальность руки будет изменяться вместе с изменением скорости ее движения, при этом она может на стереоскопическом изображении сильно отдаляться от своего владельца. Выход из этого положения был найден в сильном ограничении интенсивности препарата сверху путем его сравнения с порогом, который устанавливается на малом уровне. Таким образом, следующая вычислительная операция при формировании карты глубины, которая делает независимой ее интенсивность от скорости движения объекта, заключается в ограничении первого промежуточного препарата сверху:

$$L_{\Pi2}(k, n, i) = \begin{cases} L_{\Pi1}(k, n, i) & \text{при } L_{\Pi1}(k, n, i) \leq P_{\Pi} \\ P_{\Pi} & \text{при } L_{\Pi1}(k, n, i) > P_{\Pi} \end{cases}, \quad (2)$$

где $L_{\Pi2}(k, n, i)$ — второй промежуточный препарат; P_{Π} — верхний порог. И заключительной вычислительной операцией формирования карты глубины является размытие по горизонтали и вертикали второго промежуточного препарата подобно тому, как это сделано в методе кон-



■ **Рис. 4.** Исходное изображение (а) и карта глубины (б) в методе конвертации, основанном на использовании фактора движения при формировании карты глубины

вертации, основанном на использовании карты глубины, сформированной из исходного изображения, и присвоение ему соответствующего уровня яркости. При этом значение интенсивности (яркости) карты глубины, которая в конечном итоге определяет расположение объекта по дальности, устанавливается оператором в диалоговом режиме. На рис. 4, а в полутоновом черно-белом формате приведено исходное изображение, а на рис. 4, б — соответствующая ему карта глубины, полученная описанным выше способом. В приложении работы [10] на рис. П10 показан цветной вариант изображения, приведенного на рис. 4, а на рис. П11 — его стереоскопическое изображение в анаглифном формате.

Формирование карты глубины в случае движения камеры

Продолжая рассмотрение метода конвертации, основанного на использовании признака движения объекта, будем считать, что видеокamera, посредством которой производится съемка, также движется, но с другой скоростью, нежели объект. В этом случае для того, чтобы выделить на карте глубины только движущийся объект, необходимо скомпенсировать смещение от кадра к кадру изображения снимаемой сцены за исключением объекта, который мы хотим выделить. Другими словами, необходимо в изображениях двух следующих друг за другом кадров фоновую часть сделать неподвижной (разумеется, за исключением небольших периферийных областей, которые в соседних кадрах не совпадают вследствие движения камеры). При горизонтальном движении камеры (наиболее часто встречающийся случай) в изображениях кадров, следующих друг за другом, необходимо выделить по одной строке с одинаковым номером k_0 , которые не пересекают выделяемый движущийся объект (или объекты, если их несколько), и путем сравнения их содержимого определить взаимное смещение фоновой части изображений в этом направлении. Аналогичным образом определяется величина

взаимного смещения и при движении камеры в любом другом направлении. Все вычислительные операции целесообразно выполнять с ахроматическими компонентами изображений. Для конкретности дальнейшего изложения будем считать, что видеокamera при съемке смещается в горизонтальном направлении.

Вначале вычисляются суммы абсолютных значений разностей яркостей пикселей в строках k_0 сравниваемых изображений $L_c(k_0, n, i)$ и $L_c(k_0, n, i - 1)$ при различных величинах их взаимного смещения на r пикселей в направлении движения камеры, которые записываются в массив $M(r)$:

$$M(r) = \sum_{n=r_m}^{N-r_m} |L_c(k_0, n, i) - L_c(k_0, n+r, i-1)|,$$

где N — число пикселей (столбцов) в изображении; k_0 — номер строки в сравниваемых изображениях; r_m — максимальный ожидаемый интервал взаимного смещения фоновой части изображений в ту или иную сторону в пикселях. При этом $-r_m \leq r \leq r_m$. Затем путем перебора элементов массива $M(r)$ находится элемент массива с минимальным значением, индекс которого r_0 представляет собой величину взаимного смещения изображений текущего и предыдущего кадров, обусловленного движением снимающей видеокamera. После этого, как и в предыдущем разделе, по формуле

$$L_{\Pi}(k, n, i) = |L_c(k, n, i) - L_c(k, n+r_0, i-1)|$$

определяется исходный препарат $L_{\Pi}(k, n, i)$ для области изображения, занимаемой выделяемым движущимся объектом, из которого затем по формулам (1) и (2) синтезируются первый $L_{\Pi 1}(k, n, i)$ и второй $L_{\Pi 2}(k, n, i)$ промежуточные препараты, как это описано в разделе «Метод конвертации, основанный на использовании фактора движения при формировании карты глубины». Заключительным этапом формирования карты глубины является размытие по горизонтали и вертикали второго промежуточного препарата



■ **Рис. 5.** Текущий (а) и предшествующий (б) кадры и карта глубины (в) в методе конвертации, основанном на использовании фактора движения при формировании карты глубины в случае движения камеры

и присвоения соответствующего уровня яркости. Значение интенсивности (яркости) карты глубины, которая в конечном итоге определяет расположение объекта по дальности, устанавливается оператором в диалоговом режиме. На рис. 5, а в полутоновом черно-белом формате приведено исходное изображение текущего кадра, на рис. 5, б — предшествующего ему кадра, а на рис. 5, в — соответствующая ему карта глубины, полученная описанным выше способом. В приложении [10] на рис. П12 показан цветной вариант изображения, приведенного на рис. 5, а, а на рис. П13 — его стереоскопическое изображение в анаглифном формате.

Заключение

Последнее время характеризуется повышенным интересом к так называемым 3D-кино и телевидению, но отсутствие качественного контента сдерживает развитие этих областей медиаиндустрии. Вместе с тем съемка фильмов в этом формате по сравнению с обычными, в формате 2D, требует значительных финансовых затрат. Не многим дешевле обходится конвертация уже снятых 2D-фильмов в 3D-формат, поскольку существующие технологии требуют больших за-

трат ручного труда. Выходом из сложившейся ситуации является разработка методов конвертации, базирующихся на различных особенностях конкретного конвертируемого контента: факторе движения; различии в резкости световых границ объектов, составляющих сцену; наличии в сцене объектов, поверхности которых могут быть аппроксимированы плоскостями; наличии характерных контуров; явно выраженной перспективе в изображении конвертируемой сцены и т. д. Имея пакет компьютерных приложений, в основу которых положен набор таких методов, роль оператора можно свести только к выбору надлежащего приложения для конвертации конкретного клипа (фрагмента видео, содержащего определенный тип 2D-изображений) и подбору необходимых параметров конвертации. Опыт работы с приложениями, основанными на описанных методах, показал реальность подобного решения. Отметим также, что в наших экспериментах с описанными в статье приложениями время конвертации одного кадра видео размером 1920×1080 не превышало 1 с для любого из них. На сайте [12] в качестве примера приведено видео, при конвертации которого был применен метод, основанный на использовании карты глубины, сформированной из исходного изображения.

Литература

1. Конвертация 2d в 3d. Что такое псевдостерео? <http://www.televizor-3d.ru/konvertaciya-2-d-v-3d.html> (дата обращения: 10.10.2014).
2. Почему «Титаник» стоит смотреть в стерео 3D. <http://total3d.ru/diy/92707/> (дата обращения: 10.10.2014).
3. Красильников Н. Н. Метод получения 3D-изображений, основанный на диффузном отражении света сканируемыми объектами // Информационно-управляющие системы. 2009. № 6(43). С. 7–11.
4. Красильников Н. Н., Красильникова О. И. Получение трехмерного изображения объекта путем измерения интенсивности диффузного отражения света различными точками его поверхности // Оптический журнал. 2010. Т. 77. № 6. С. 19–24.
5. Красильников Н. Н., Красильникова О. И. Определение координаты глубины по 2D изображению // Оптический журнал. 2011. Т. 78. № 12. С. 30–33.
6. Красильников Н. Н., Красильникова О. И. Исследование погрешностей определения координаты глубины при 3D-сканировании методом, основанным на диффузном отражении света // Информационно-управляющие системы. 2012. № 3(58). С. 2–8.
7. Красильников Н. Н. Метод формирования 3D-изображения сцены по одной фотографии // Информационно-управляющие системы. 2013. № 3. С. 2–7.
8. Красильников Н. Н., Красильникова О. И. Метод конвертации 2D-изображения в стереоскопическое

3D-изображение // Оптический журнал. 2014. Т. 81. № 2. С. 20–28.

9. Красильников Н. Н. Механизмы и погрешности восприятия 3D-пространства зрительной системой // Информационно-управляющие системы. 2014. № 6(73). С. 13–19.
10. Приложение к статье «Методы конвертации 2D изображений и видео в стереоскопический формат». http://cv.guap.ru/krasilnikov_nn/files/pril006 (дата обращения: 10.07.2014).

11. Приложение к исследованиям нового метода конвертации 2D-изображений в 3D-изображения. http://cv.guap.ru/krasilnikov_nn/files/pril003 (дата обращения: 01.07.2015).

12. Приложение к исследованиям нового метода конвертации 2D-изображений в 3D-изображения. http://cv.guap.ru/krasilnikov_nn/files/pril004 (дата обращения: 01.07.2015).

UDC 612.843.721

doi:10.15217/issn1684-8853.2015.5.18

Methods of Converting 2D Images and Videos into Stereoscopic Format

N. N. Krasilnikov^a, Dr. Sc., Tech., Professor, NNKrasilnikov@yandex.ru

O. I. Krasilnikova^a, PhD, Tech., Associate Professor, OIKrasilnikova@yandex.ru

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Purpose: The known methods of providing high-quality conversion of 2D images and videos into stereoscopic format have a common drawback: some stages require manual operations which make this conversion costly. The aim of this work is developing more economical semi-automatic conversion methods. **Results:** The six proposed conversion methods are based on the hypothesis (put forward by us and experimentally confirmed) that space depth perception when observing 3D scenes and stereoscopic images is provided by "converging-diverging" eye movements. This hypothesis has a corollary that when generating stereo images, only the fields of the depth map adjacent to the sharp light borders are important. The first method is approximating the surfaces of the objects represented on the converted image by a small number of planes. The second method is synthesizing the depth map by a small set of primitive surfaces. The main feature of the third conversion method is synthesizing the depth map from the achromatic component of the converted image. The fourth method differs from the previous one by generating the depth map out of the contours of the achromatic component of the converted image. The fifth and sixth methods use the objects movement factor to form the depth map. **Practical relevance:** The proposed methods, as compared to the conventional ones, have greater performance and require less manual labor. They can be effectively used in practice, to convert either single images or videos.

Keywords — 3D Image, Stereoscopic Image, Conversion, 2D Image, Video, 3D Format.

References

1. *Konvertatsiya 2d v 3d. Chto takoe pseudostereo?* [Converting 2D to 3D. What is Pseudostereo?]. Available at: <http://www.televizor-3d.ru/konvertatsiya-2-d-v-3d.html> (accessed 10 October 2014).
2. *Pochemu «Titanik» stoit smotret' v stereo 3D* [Why «Titanic» is Worth Watching in Stereo 3D]. Available at: <http://total3d.ru/diy/92707/> (accessed 10 October 2014).
3. Krasilnikov N. N. A Method of 3D Scanning Based on Scanned Objects' Light Radiosity. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2009, no. 6(43), pp. 7–11 (In Russian).
4. Krasilnikov N. N., Krasilnikova O. I. Obtaining a Three Dimensional Image of an Object by Measuring the Intensity of the Diffuse Reflection of Light by Various Points of its Surface. *Opticheskii zhurnal*, 2010, vol. 77, no. 6, pp. 19–24 (In Russian).
5. Krasilnikov N. N., Krasilnikova O. I. Determining the Depth Coordinate from a 2D Image. *Opticheskii zhurnal*, 2011, vol. 78, no. 12, pp. 30–33 (In Russian).
6. Krasilnikov N. N., Krasilnikova O. I. The Investigation of Measurement Errors of Depth Coordinate during 3D Scanning by the Method Based on Diffuse Reflection of Light. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2012, no. 3(58), pp. 2–8 (In Russian).
7. Krasilnikov N. N. The Method of Generating a 3D-Image Scene Based on a Single Photograph. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2013, no. 3(64), pp. 2–7 (In Russian).
8. Krasilnikov N. N., Krasilnikova O. I. The Method of Converting a 2D Image into a Stereoscopic 3D Image. *Opticheskii zhurnal*, 2014, vol. 81, no. 2, pp. 20–28 (In Russian).
9. Krasilnikov N. N. Mechanisms and Inaccuracies of 3D Spatial Perception by Visual System. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2014, no. 6(73), pp. 13–19 (In Russian).
10. *Prilozhenie k stat'e "Metody konvertatsii 2D izobrazhenii i video v ste-reoskopicheskii format"* [The Appendix to the Article "Methods of Converting a 2D Image and Video into a Stereoscopic Format"]. Available at: http://cv.guap.ru/krasilnikov_nn/files/pril006 (accessed 10 April 2014).
11. *Prilozhenie k issledovaniyam novogo metoda konvertatsii 2D-izobrazhenii v 3D-izobrazheniia* [Appendix to the Research of a New Method of Converting 2D Images into 3D Images]. Available at: http://cv.guap.ru/krasilnikov_nn/files/pril003 (accessed 01 July 2015).
12. *Prilozhenie k issledovaniyam novogo metoda konvertatsii 2D-izobrazhenii v 3D-izobrazheniia* [Appendix to the Research of a New Method of Converting 2D Images into 3D Images]. Available at: http://cv.guap.ru/krasilnikov_nn/files/pril004 (accessed 01 July 2015).