

УДК 004.925.84

МЕТОД РЕКОНСТРУКЦИИ 3D-ОБЪЕКТОВ ПО ДАННЫМ ОДНОГО РАКУРСА ИЗОБРАЖЕНИЯ

© С. И. Вяткин, Б. С. Долговесов

*Институт автоматизи и электрометрии СО РАН,
630090, г. Новосибирск, просп. Академика Коптюга, 1
E-mail: sivser@mail.ru*

Предложен метод реконструкции по одному изображению криволинейных поверхностей. Для реконструкции поверхности используются данные силуэта объекта и многоуровневая карта высот. В отличие от известных подходов к реконструкции одного ракурса предлагаемый метод не имеет ограничений по глубине и требует на порядок меньше вычислений. 3D-модели с высоким разрешением могут быть сгенерированы в интерактивном режиме.

Ключевые слова: реконструкция по одному ракурсу, численная оптимизация, многоуровневая карта высот, взвешенная минимальная поверхность, визуализация.

DOI: 10.15372/AUT20230505

Введение. Реконструкция 3D-формы одного ракурса изображения представляет собой очень важную задачу, поскольку служит основой для формирования новых представлений об объекте, наблюдаемом на одном изображении, но с другой точки зрения, с другой текстурой и другим освещением. По сравнению с большинством задач компьютерного зрения реконструкция по одному ракурсу является крайне некорректной проблемой. Как следствие, необходимо сделать дополнительные предположения о том, какова геометрия объекта, например кусочно-плоскостная [1], форма из текстуры [2], форма из затенения [3] или форма из расфокусировки [4]. Такие подходы имеют ограничения, сложные по вычислениям, и проблематичны для использования в интерактивном режиме. Алгоритмы реконструкции на основе силуэта хорошо работают для большого набора объектов и изображений [5]. Работа для реконструкции криволинейных поверхностей описана в [6], хотя эти реконструкции ограничены формой объектов. Работа [7] посвящена реконструкции произвольных криволинейных поверхностей и обобщению класса реконструируемых объектов. Однако на практике только объекты при достаточно простой топологии могут быть реконструированы за счёт параметризованного представления поверхности. Более того, необходимый объём вводимых пользователем данных довольно большой. В [8] используют карту глубины для расчёта гладкой минимальной поверхности, которая подчиняется определённым пользователем ограничениям. Для реконструкции объём необходимого пользовательского ввода значителен. В [9] двумерная поверхность генерируется из входного силуэта, затем может быть положена в основу создания, например, шрифтов. Кроме того, существуют основанные на обучении подходы [10], которые изучают значения глубины патчей изображения из базы данных. Метод реконструкции единого вида, созданный на множествах Чигера, направлен на вычисление силуэта — последовательных минимальных поверхностей заданного пользователем объёма — и обеспечивает правдоподобные реконструкции для большого количества реальных объектов в [11]. Концепция минимальных поверхностей фиксированного объёма является интуитивно понятной и полезной парадигмой для вычисления правдоподобных одномерных реконструкций. Однако метод [11] имеет ряд преимуществ и недостатков. Объёмное представление накладывает сильные ограничения на память и время выполнения. Даже с эффективным GPU-ускорением требуется около

секунды времени вычислений для реконструкции невысокого разрешения. Вследствие этого 3D-модели с более высоким разрешением не могут быть созданы на интерактивных скоростях. Хотя метод, описанный в [11], обеспечивает точные объёмно-последовательные решения, алгоритм базируется на основе выпуклой релаксации и порогового значения. В отсутствие пороговой теоремы метод не гарантирует получения глобально минимальной поверхности заданного объёма. Кроме того, неясно, действительно ли последующее пороговое значение решения приводит к пространственно-когерентной структуре, а не к разбросанному набору вокселей? Метод вычисляет карту глубины, тем не менее в работе [11] используется полностью объёмное представление. Требуемая дискретизация возможного значения глубины накладывает ограничение на возможное разрешение в направлении Z .

Целью представленной работы является создание метода реконструкции 3D-объектов по данным одного ракурса изображения.

Реконструкция 3D-объекта по данным одного ракурса изображения. Пусть $V \subset \mathbb{R}^3$ — объёмный домен, окружающий входное изображение $I: P \rightarrow \mathbb{R}^3$ с представлением плоскости $P \subset V$. Необходимо найти замкнутую поверхность $S \subset V$, которая соответствует объекту на изображении I и его контуру C . Ортогональная проекция определяется с помощью $\pi: V \rightarrow P$. Для обработки произвольных топологий, поверхность S неявно представляется в виде функции $f: V \rightarrow \{0, 1\}$, задающей внутреннюю часть объекта $f = 1$ и внешнюю $f = 0$. Поверхность получается за счёт минимизации энергетического функционала:

$$E(f) = \int_V f(x)(\delta_v(x) + \delta_c(x)) d^3x. \quad (1)$$

Контур вычисляется с помощью выражения

$$\delta_c(x) = \begin{cases} -\infty, & \text{если } \eta_c(\pi(x)) = 1, \text{ и } x \in P; \\ +\infty, & \text{если } \eta_c(\pi(x)) = 0; \\ 0 & \text{иначе,} \end{cases} \quad (2)$$

где характеристическая функция $\eta_c: P \rightarrow \{0, 1\}$ указывает на внешнюю или внутреннюю часть контура.

Функция $\delta_c(x)$ обеспечивает согласованность контура. Точки, выступающие за пределы контура, будут отнесены к фону $f = 0$, а точкам, находящимся на плоскости изображения и внутри объекта, будет присвоен объект $f = 1$.

Функция инфляции объёма δ_v позволяет сделать некоторое предположение о форме объекта. Эта функция может быть принята для достижения любой желаемой формы объекта и также может быть изменена в результате взаимодействия с пользователем. Делаем простое предположение, что толщина наблюдаемого объекта увеличивается по мере того, как продвигаемся внутрь от его контура. Для любой точки $p \in P$ пусть выражение

$$\{d_f(p, \partial C)\} = \min_{c \in \partial C} \|p - c\| \quad (3)$$

определяет расстояние до контура $\partial C \subset P$. Затем

$$\delta_v(x) = \begin{cases} -1, & \text{если } d_f(x, P) \leq h(\pi(x)); \\ +1 & \text{иначе.} \end{cases} \quad (4)$$

Карта высот $h: P \rightarrow \mathbb{R}$ определяется с помощью функции расстояния

$$h(p) = \min \{\nu_{co}, \nu_{os} + \nu_f d_f(p, \partial C)^n\}. \quad (5)$$

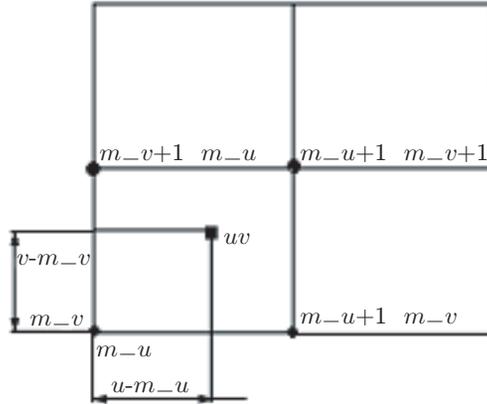


Рис. 1. Таблица чисел карты высот

Параметры $\{\nu_{co}, \nu_{os}, \nu_f, n\}$ модифицируют форму h .

Например, с помощью параметра $\{\nu_{co}\}$ можно клиппировать плоскостью карту высот, с помощью параметра $\{n\}$ (5) вытягивать или сплющивать карту высот, с помощью параметра $\{\nu_{os}\}$ расширять или сужать карту высот, с помощью параметра $\{\nu_f\}$ определять пологость и крутизну карты высот относительно базовой плоскости по границе контура.

Карта высот (рис. 1) представляет собой двумерный прямоугольник, который далее будет именоваться областью возмущения D_P базовой плоскости P , внутри которого задана функция возмущения $h(u, v)$. Карта высот, в свою очередь, определяет само возмущение. Область определения функции $h(u, v)_{D_{h(u,v)}} = \{U, V\}$, где U и V — размеры прямоугольника. Карта высот связана с базовой плоскостью следующим образом: существует преобразование $G (\mathbb{R}^3 \Rightarrow \mathbb{R}^2)$ из системы координат, в которой находятся S и P , в систему координат карты высот. Как правило, таким преобразованием является параллельная проекция.

Значение функции $h(G(d_S))$ характеризует отклонение точки d_S , лежащей на поверхности S , от точки d_P , являющейся проекцией этой точки на плоскость P . Другими словами, значение функции $h(G(d_S))$ равно модулю вектора

$$\mathbf{v} = \mathbf{d}_S - \mathbf{d}_P. \quad (6)$$

Следовательно, область сложной поверхности можно задать множеством точек в \mathbb{R}^3 , определяемых векторным уравнением

$$\mathbf{S} = G(\mathbf{v}) + \mathbf{n}h(G(\mathbf{v})); \quad \forall \mathbf{v} \in \mathbb{R}^3, \quad (7)$$

где \mathbf{n} — нормаль базовой плоскости.

Если \mathbf{v} находится за пределами области возмущения, то вектор $\mathbf{n}h(G(\mathbf{v})) = 0$, а вектор \mathbf{S} является вектором на базовой плоскости.

Таким образом, для задания формы возмущающей поверхности можно использовать таблицу чисел, а в качестве функции h можно взять функцию интерполяции по узловым значениям, которые берутся из таблицы. Можно полагать, что в области возмущения D_P будет задано скалярное поле.

Задание сложной поверхности S на основе скалярного поля есть: базовая поверхность, преобразование из системы координат базовой поверхности в систему координат карты высот, функция возмущения и таблица чисел, характеризующих отклонения поверхности S относительно базовой плоскости, в контрольных точках (рис. 2).

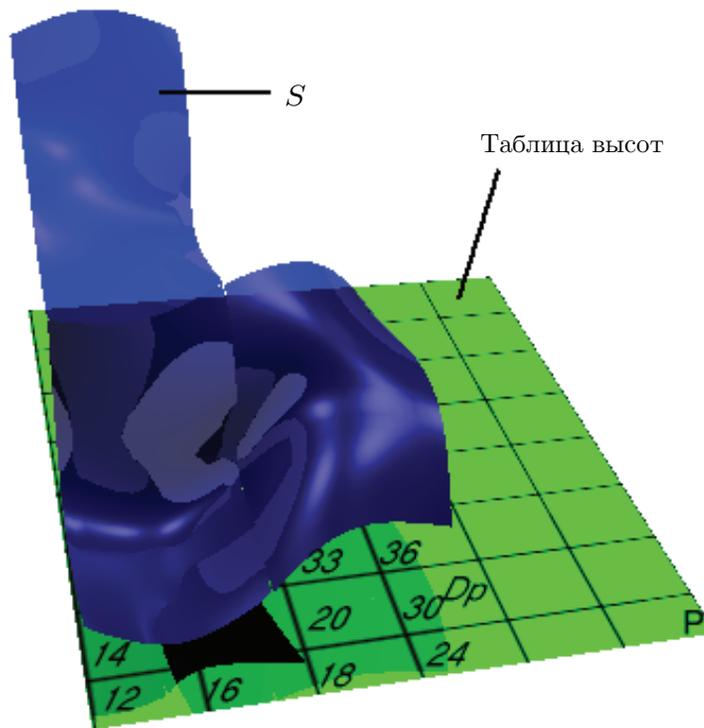


Рис. 2. Поверхность и базовая плоскость



Рис. 3. Базовая плоскость и возмущение

Однако объекты, смоделированные этим методом, имеют прямоугольную границу (рис. 3).

Поэтому метод построения поверхностей был модифицирован так, чтобы стало возможным отображать объекты с «нелинейной» границей. Для этого введено понятие порогового значения (рис. 4).

Для эффективности вычислений предварительно составляются карты уровней детальности (рис. 5). Исходные данные задают уровень n , если размерность сетки равна $2^n \times 2^n$. Данные для уровня $n - 1$ получаются выбором максимума из четырёх смежных значений уровня n , остальные три далее не учитываются, т. е. получаем размерность сетки $2^{n-1} \times 2^{n-1}$. Уровень 0 состоит из единственного значения — максимума по всей карте высот. Таким образом, можно рассматривать карту высот как текстуру формы.

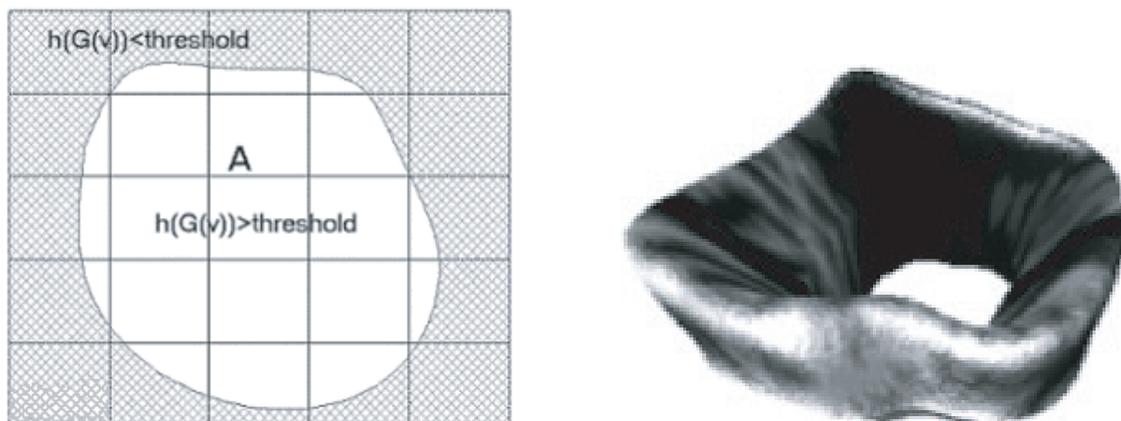


Рис. 4. Область возмущения (слева), поверхность (справа)

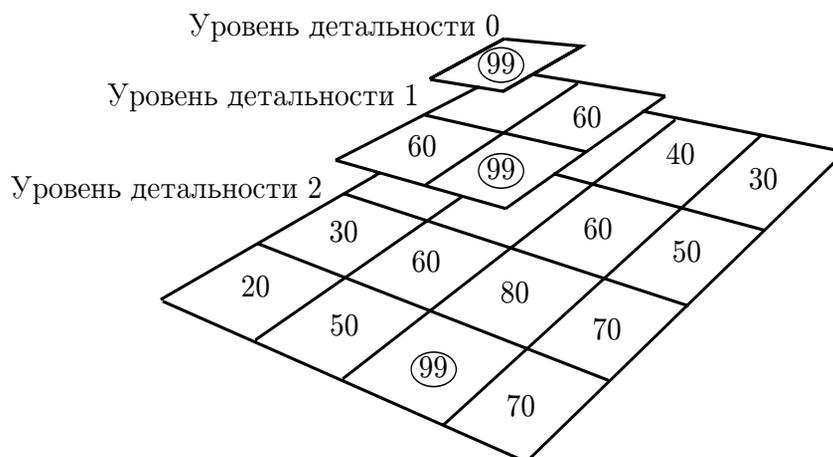


Рис. 5. Карты уровней детальности

Визуализация. Применяется метод отслеживания лучей, в котором через каждый пиксель плоскости изображения пускается луч от взгляда наблюдателя сквозь пирамиду видимости (объектное пространство). Эти лучи проецируются на базовую плоскость.

Выделение формы поверхности происходит на этапе деления луча (воксела, который ограничен передней и дальней гранями пирамиды видимости) по z -координате. Одномерный брусок — воксел V_0 — задаётся парой векторов $P_0 = (x_0, y_0, z_0)$ и $P_1 = (x_1, y_1, z_1)$, $V_0 = \{P_0, P_1\}$. Координаты воксела V_0 с помощью матрицы G преобразуются в систему координат карты высот:

$$\{(x_0, y_0, z_0), (x_1, y_1, z_1)\} \Rightarrow \{(u_0, v_0, h_0), (u_1, v_1, h_1)\}. \quad (8)$$

Для этого вводится матрица T преобразования в систему координат карты высот, которая, будучи умноженной на матрицу геометрических преобразований M , даёт результирующую матрицу преобразования G , $G = TM$:

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (9)$$

Преобразованные координаты воксела (u, v, h, a) в системе координат карты высот вычисляются из координат воксела (x, y, z) в объектном пространстве умножением вектора точки на матрицу G :

$$G \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} u \\ v \\ h \\ a \end{bmatrix}. \quad (10)$$

Далее осуществляется деление воксела по координате z или бинарное деление воксела. На этом этапе для текущего уровня рекурсии вектор конца воксела, являющегося ближайшим по отношению к наблюдателю, полагается равным вектору ближайшего конца воксела предыдущего уровня деления. Вектор дальнего конца воксела вычисляется как полусумма векторов ближнего и дальнего концов воксела предыдущего уровня деления:

$$P_{ni} = P_{ni-1}, \quad P_{fi} = (P_{ni-1} + P_{fi-1})/2, \quad V_i = \{P_{ni}, P_{fi}\}, \quad (11)$$

где V_i — воксел i -го уровня рекурсии, а P_{ni}, P_{fi} — координаты ближнего по отношению к наблюдателю и дальнего концов воксела на i -м уровне рекурсии. Таким образом, геометрические преобразования для воксела делаются всего лишь 1 раз на первом уровне бинарного деления, а далее идёт просто процесс деления воксела на две части и выбор ближайшего пересечённого воксела.

По размерам проекции воксела соответствующего уровня рекурсии вычисляется уровень детальности, который, по сути, является тоже картой высот, но меньшего разрешения, и по координатам u и v точек P_{ni} и P_{fi} осуществляется выборка максимального значения из таблицы чисел, представляющих данный уровень детальности. Полученное таким образом число и является значением функции возмущения базовой поверхности.

На каждом этапе бинарного деления воксела по его размерам вычисляется уровень детальности — level. Если level — не последний уровень детальности, то полученная высота h сравнивается со значением высоты данного уровня H_{\max} , и если $h > H_{\max}$, то деление воксела прекращается. Таким образом, вычисления происходят в два этапа.

1. Определяются габариты прямоугольника, являющегося проекцией воксела на карту высот, как максимум расстояния от точки $\{u_0, v_0\}$ до точки $\{u_1, v_1\}$ — уровень детальности L_p .

2. Из неравенства

$$\frac{1}{2^{\text{level}}} < L_p < \frac{1}{2^{\text{level}+1}}$$

определяется уровень детальности level.

На рис. 6 и 7 показаны результаты визуализации реконструируемых объектов.

Результаты работы. Вычислительный процесс для реализации метода был распараллелен на графическом процессоре GPU 470 GTX. Это включало в себя шаги проекции и бинарного деления вокселей. Был протестирован метод на нескольких реальных изображениях, сравнены полученные результаты с другими известными методами, оценён внешний вид, время вычисления, объём данных пользовательского ввода [12]. Так как нельзя получить истинные значения глубины из одного ракурса, то не ставилась задача решения абсолютного сходства с исходными данными, скорее, стремились к правдоподобию реконструкции. Поскольку обратные стороны объектов естественно невидимы во входном изображении, то реконструкции должны быть симметричными. Это может быть достигнуто



Рис. 6. Слева представлен реальный объект, справа — реконструированный и повернутый объект относительно оси Y



Рис. 7. Слева представлен реальный объект, справа — реконструированный и повернутый объект относительно оси X

путём простого зеркального отображения вычисленных значений глубины вдоль плоскости изображения. С помощью этой эвристики можно получить замкнутые представления объектов из карт глубины.

Сравним результаты с полученными с помощью методов [7] и [8]. Метод [8] присутствует в сравнении, потому что он ограничен реконструкцией карты глубины, в то время как другие методы фокусируются на криволинейных объектах. Все подходы являются глобально оптимальными и вычисляют реконструкции на интерактивных частотах кадров. В основном эти методы отличаются необходимым количеством вводимых пользователем данных. В методе [8] у пользователя есть разнообразие выбора для поверхностных манипуляций, таких как положение и нормальные ограничения, разрывные ограничения, плоская область ограничения и ручное разделение сетки. Многие из этих ограничений необходимы для сложных реконструкций, приводящих к временам моделирования от нескольких минут до нескольких часов, даже для опытных пользователей. Для метода [7] необходимы описание входных данных и экспертные знания. Пользователь должен назначить детали от контурных линий к линиям в пространстве параметров, что становится сложным для объектов более высокого уровня. В результате этого топология ограничена вторым уровнем. Поэтому при реконструкции объектов высшего порядка проявляются колебания поверхности. Более того, для объёмной инфляции пользователю необходимо определить

набор ограничений интерполяции. На последующих этапах пользователю могут потребоваться дополнительные ограничения для реконструкции изгибов поверхности. В отличие от предлагаемого подхода метод [7] может справиться лучше с изображениями, в которых плоскость симметрии объекта не параллельна плоскости изображения. Поскольку для предлагаемого метода требуется меньше памяти и вычислительного времени, можно использовать входные изображения со значительно более высоким разрешением. Это приводит к более высоким детализированным силуэтам и реконструкциям. Кроме того, когда вычисляются непрерывные значения глубины, можно добиться хорошей гладкости поверхности.

Заключение. Представлен метод реконструкции по одному ракурсу, основанный на картах высот. В отличие от известного объёмного подхода, предложенное $2D$ -решение имеет три преимущества:

- 1) разрешение в трёх измерениях имеет двузначную точность;
- 2) вычисления являются оптимальными;
- 3) $2D$ -решение резко сокращает объём памяти и время вычислений (примерно на порядок).

Для большого разнообразия объектов с хорошим разрешением изображения правдоподобные реконструкции вычисляются за доли секунды. Этот метод хорошо подходит для интерактивного $3D$ -моделирования. Поскольку многоуровневая карта высот является единственным свободным параметром модели, предложенный метод также превосходит многие известные подходы в отношении необходимой суммы параметров из пользовательского ввода.

СПИСОК ЛИТЕРАТУРЫ

1. **Liebowitz D., Criminisi A., Zisserman A.** Creating architectural models from images // Eurographics. 1999. **18**, N 3. P. 39–50.
2. **Li D., Hu D., Sun Y., Hu Y.** 3D scene reconstruction using a texture probabilistic grammar // Multimedia Tools and Applications. 2018. **77**, N 21. P. 28417–28440. DOI: 10.1007/s11042-018-6052-z.
3. **Zheng Q., Chellappa R.** Estimation of illuminant direction, albedo, and shape from shading // Physics-Based Vision: Principles and Practice: Shape Recovery /Eds. by Lawrence B. Wolff, Steven A. Shafer, Glenn E. Healey. 2002. Vol. 3. 544 p.
4. **Favaro P., Soatto S.** A geometric approach to shape from defocus // IEEE Trans. Pattern Analysis and Machine Intell. 2005. **27**, N 3. P. 406–417.
5. **Вяткин С. И., Долговесов Б. С.** Метод реконструкции функционально заданных поверхностей по стерео изображениям реальных объектов // Автометрия. 2020. **56**, № 6. С. 19–26. DOI: 10.15372/AUT20200603.
6. **Chuang J. H., Ahuja N., Lin C. C. et al.** A potential-based generalized cylinder representation // Computers & Graphics. 2004. **28**, Iss. 6. P. 907–918.
7. **Prasad M., Fitzgibbon A.** Single view reconstruction of curved surfaces // Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2006). New York, USA, 17–22 June, 2006. P. 1345–1354. DOI: 10.1109/CVPR.2006.281.
8. **Zhang L., Dugas-Phocion G., Samson J.-S., Seitz S. M.** Single view modeling of free-form scenes // Proc. of the Int. Conference on Computer Vision and Pattern Recognition. Kauai, USA, 08–14 Dec., 2001. P. 990–997. DOI: 10.1109/CVPR.2001.990638.
9. **Joshi P., Carr N. A.** Repousse: Automatic Inflation of 2D Artwork // Proc. of the EUROGRAPHICS Workshop on Sketch-Based Interfaces and Modeling. Aire-la-Ville,

Switzerland, 11 June, 2008. P. 49–55. URL: http://pushkarjoshi.org/pdf/JC_SBIM2008.pdf (дата обращения: 02.06.2023).

10. **Hassner T., Basri R.** Example based 3D reconstruction from single 2D images // Proc. of the Int. Conference on Computer Vision and Pattern Recognition. Workshop. IEEE Computer Society. New York, USA, 17-22 June, 2006. P. 15. DOI: 10.1109/CVPRW.2006.76.
11. **Топпе Е., Освальд М. Р., Кремерс Д., Ротер С.** Image-based 3D modeling via Cheeger sets // Proc. Asian Conference on Computer Vision (ACCV 2010). Queenstown, New-Zealand, Nov. 2010. P. 53–64. Vol. 2. P. 189–193.
12. **Вяткин С. И., Долговесов Б. С.** Методы интерактивного моделирования и визуализации функционально заданных объектов для 3D-Web приложений // Автометрия. 2022. **58**, № 1. С. 111–118. DOI: 10.15372/AUT20220112.

Поступила в редакцию 16.01.2023

После доработки 31.05.2023

Принята к публикации 02.06.2023
