

УДК 621.391.2

ВОССТАНОВЛЕНИЕ ГЛУБИННЫХ КАРТ НА ОСНОВЕ АТРИБУТОВ ДЕСКРИПТОРА СТЕРЕОПАР ИЗОБРАЖЕНИЙ

В. Ф. Кравченко^{1,2,3}, В. И. Пономарев^{4,*},
академик РАН В. И. Пустовойт², Д. Розас-Миранда⁴

Поступило 25.03.2019 г.

Впервые предложен и обоснован подход, позволяющий восстанавливать глубинные карты по стереопарам цветных изображений при визуализации трёхмерных (3D) объектов. Он использует новый локальный дескриптор изображений, основанный на примитивах и отношениях между ними, характеризующая различия в цветах, положении плоскостей, дистанциях и углах между примитивами. Новый подход сравнивается с общеизвестными системами дескрипторов, в частности с DAISY и SID. Численные эксперименты, анализ и физическая интерпретация результатов в условиях реально существующих радиометрических различий в экспозиции или освещении стереопар изображений показали преимущество нового подхода перед известными.

Ключевые слова: дескриптор изображения, копланарность, соцветность, визуализация, глубинная карта.

DOI: <https://doi.org/10.31857/S0869-56524872130-134>

Стереовизуализация является одним из наиболее активно развивающихся научных направлений в трёхмерной (3D) компьютерной визуализации, где основная проблема — оценка положения объектов в трёхмерном пространстве на основе сформированных глубинных карт (DM).

Различные задачи в компьютерной визуализации включают в себя сегментацию, детектирование объектов в 3D путём извлечения информации из многомерных данных в реальных системах и преобразования этих данных на основе эффективных дескрипторов, которые существенно сокращают размерность проблемы. Контуры изображений и соотношения между ними являются основными элементами в компьютерной визуализации и роботике, представляя обобщённую информацию об анализируемых сценах в виде ряда атрибутов, важных при распознавании объектов. Среди таких дескрипторов Daisy и SID являются наиболее эффективными и широко используемыми в практике.

Используя идеи, изложенные в [1–12], предложен и обоснован оригинальный метод, позволяющий восстанавливать глубинные карты по стереопарам цветных изображений в 3D-визуализации. Новый дескриптор (VPR) изображений основан на примитивах и отношениях между ними, характеризует различия в цветах, положении плоскостей, дистанциях и углах между примитивами. Принципиальное отличие дескриптора VPR от других известных в литературе состоит в том, что он использует структурную и семантическую информацию. Это улучшает его робастные свойства, в частности в неидеальных условиях при регистрации, освещении и отражениях в стереопарах цветных изображений. Численные эксперименты, анализ и физическая интерпретация результатов в условиях реально существующих радиометрических различий в экспозиции и освещении стереопар изображений показали преимущество нового подхода перед известными.

ПОСТАНОВКА ЗАДАЧИ И МЕТОД РЕШЕНИЯ

Теоретическое описание предложенного дескриптора изображения основано на преобразовании Рисса [9]. Здесь для заданного двумерного сигнала $f(x, y)$ формируются две нечётные компоненты сигнала

$$F_{01,02}(w) = \begin{cases} i \frac{w_{x,y}}{\|w\|} F(w), & w \neq 0, \\ 0, & w = 0. \end{cases}$$

Используется Log-Gabor фильтр

¹ Институт радиотехники и электроники
им. В.А. Котельникова

Российской Академии наук, Москва

² Научно-технологический центр
уникального приборостроения

Российской Академии наук, Москва

³ Московский государственный технический
университет им. Н.Э. Баумана

⁴ Национальный политехнический институт Мексики,
Мехико

*E-mail: vponomar@ipn.mx; vponomar@mail.ru;
volodymyr.ponomaryov@gmail.com

$$G_e(w) = \exp\left(-\frac{\log\left(\frac{\|w\|}{w_0}\right)}{2(\log\sigma_0)^2}\right),$$

для которого находятся две нечётные компоненты, $G_{01}(w)$ и $G_{02}(w)$, и изотропный фильтр $G(w) = \sqrt{G_{01}^2(w) + G_{02}^2(w)}$.

После фильтрации формируется вектор из трёх компонент моногенного сигнала (две нечётные и одна чётная) $f_m(x) = [f_e(x), f_{01}(x), f_{02}(x)]$. Этот вектор характеризуется в сферических координатах длиной $A(x)$ и двумя угловыми координатами $\phi(x) = \arctan\left(\frac{f_{02}(x)}{f_e(x)}\right)$, $\theta(x) = \arctan\left(\frac{f_{02}(x)}{f_{01}(x)}\right)$.

Вектор визуальной 2D-примитивы $\Pi_{2D}(x)$ определяется этими тремя параметрами $A(x)$, $\phi(x)$, $\theta(x)$, а 3D-визуальная примитива характеризуется вектором $\Pi_{3D}(x) = (A(x), \theta(x), \phi(x), d(x))$, где d определяет положение примитива в трёхмерном пространстве.

Зависимости между двумя примитивами определяются рядом атрибутов, в частности, соцветность (colority) вычисляется в пространстве цветов CIELab следующим образом:

$$R^C(\Pi_i, \Pi_j) = \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}.$$

Другие атрибуты характеризуются значением цвета, положением и ориентацией, а также соотношениями между ними, а именно:

углом между примитивами

$$R^A(\Pi_i, \Pi_j) = \arccos\left(\frac{\Pi_i \Pi_j}{|\Pi_i| |\Pi_j|}\right);$$

нормализованной дистанцией между примитивами

$$R^{ND}(\Pi_i, \Pi_j) = \frac{1}{2} \left(\frac{|W_i \times \Pi_j|}{|\Pi_j|} + \frac{|W_j \times \Pi_i|}{|\Pi_i|} \right);$$

атрибутом копланарность (coplanarity)

$$R^P(\Pi_i, \Pi_j) = \frac{1}{2} \left[\pi - \arccos\left(\frac{nA^O(\Pi_i)}{|n| |A^O(\Pi_i)|}\right) - \arccos\left(\frac{nA^O(\Pi_j)}{|n| |A^O(\Pi_j)|}\right) \right]$$

[13]. Новый дескриптор (VPR), основанный на отношениях между примитивами, реализуется согласно приведённой на рис. 1 блок-схеме метода, т.е. для данного цветного изображения I_{RGB} проводится преобразование его цветов из RGB- в CIELab-пространство. Соотношение соцветности определяется моногенным сигналом, который формируется

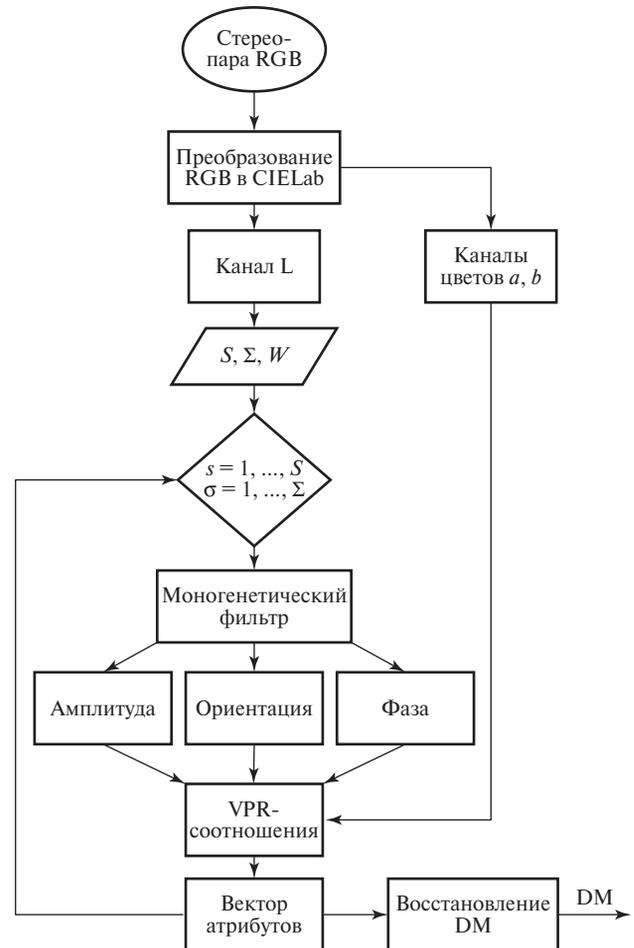


Рис. 1. Блок-схема предлагаемого метода.

для S масштабов и Σ ядер ($i = 1, 2, \dots, S; j = 1, 2, \dots, \Sigma$) с использованием $S \times \Sigma$ гауссовских фильтров. Каждый фильтр реализует конволюцию изображения I_L канала L в CIELab-пространстве и формирует $3 \times S \times \Sigma$ компонент $f_m(x) = [f_e(x), f_{01}(x), f_{02}(x)]$ моногенного сигнала для различных $S \times \Sigma$ фильтров с параметрами s_i и σ_j в гауссовых фильтрах.

Для каждого пикселя новый дескриптор состоит из вектора, определяемого отношениями между примитивами (x, y) и (u, v) в квадратном окне W с центром в (x, y) :

$$H_A(x) = [R_A\{s_i, \sigma_j\}(1, 1), \dots, R_A\{s_i, \sigma_j\}(u, v)],$$

где R_A — угол, определяемый соотношениями между примитивами пикселей (x, y) и (u, v) . Вектор атрибутов для углов между примитивами для $S \times \Sigma$ гауссовых фильтров определяется следующим матричным выражением:

$$D_A(x, y) = \begin{bmatrix} H_A(x, y)_{s_1}^{\sigma_1} & \dots & H_A(x, y)_{s_S}^{\sigma_1} \\ \dots & \dots & \dots \\ H_A(x, y)_{s_1}^{\sigma_\Sigma} & \dots & H_A(x, y)_{s_S}^{\sigma_\Sigma} \end{bmatrix}. \quad (1)$$

Аналогичным образом находятся выражения для атрибутов, характеризующихся нормализованными дистанциями $D_{ND}(x, y)$ и параметрами копланарности $D_{CP}(x, y)$:

$$D_{ND}(x, y) = \begin{bmatrix} H_{ND}(x, y)_{s_1}^{\sigma_1} & \dots & H_{ND}(x, y)_{s_S}^{\sigma_1} \\ \dots & \dots & \dots \\ H_{ND}(x, y)_{s_1}^{\sigma_\Sigma} & \dots & H_{ND}(x, y)_{s_S}^{\sigma_\Sigma} \end{bmatrix}, \quad (2)$$

$$D_{CP}(x, y) = \begin{bmatrix} H_{CP}(x, y)_{s_1}^{\sigma_1} & \dots & H_{CP}(x, y)_{s_S}^{\sigma_1} \\ \dots & \dots & \dots \\ H_{CP}(x, y)_{s_1}^{\sigma_\Sigma} & \dots & H_{CP}(x, y)_{s_S}^{\sigma_\Sigma} \end{bmatrix}.$$

Финальный дескриптор в (x, y) определяется приведёнными выше атрибутами $D_{VPR}(x, y) = [D_A(x, y), D_C(x, y), D_{CP}(x, y), D_{ND}(x, y)]$. В дальнейшем осуществляется восстановление глубинных карт DM на основе процедуры block matching в сравнении атрибутов двух стереопар.

РЕЗУЛЬТАТЫ ЧИСЛЕННОГО МОДЕЛИРОВАНИЯ

Были исследованы тестовые синтетические изображения из Middlebury Stereo Vision сайта [14, 15]. База данных 2005 г. содержит серию различных стереопар совместно с изображениями GT (истинные глубинные карты) в полном формате (1390×1110 пикселей), а также в $1/2$ и $1/3$ форматах. База данных 2014 г., которая использовалась здесь дополнительно, позволяет исследовать робастность нового дескриптора, так как содержит 33 стереопары, разделённые на три группы (10 для обучения, 10 для тестирования и 13 дополнительных пар, которые не иллюстрируются GT). Эта база данных, кроме того, предоставляет две визуализации для каждой

стереопары, снятые в условиях различной иллюминации (L) или различной экспозиции (E).

Все данные в экспериментах обрабатывались совместно с целью подтвердить эффективность нового дескриптора VPR и его робастность при восстановлении глубинных карт в условиях различий регистрации стереопар.

При анализе качества восстановления глубинных карт использовался критерий QBP (количество ошибочных пикселей — Quantity Bad Matching Pixels). Значения критерия QBP вычисляются по формуле для каждого из исследуемых изображений DM:

$$QBP = \frac{1}{N} \sum_{(x,y)} (|DM_I(x, y) - DM_{GT}(x, y)| > \delta_d), \quad (3)$$

где N — число пикселей в изображении или кадре, DM_I и DM_{GT} — оценённые и истинные (GT) карты соответственно и $\delta_d = 1$.

Результаты визуального сравнения восстановленных глубинных карт с использованием дескриптора VPR с лучшими известными DAISY и SID приведены на рис. 2, где все три метода используют традиционную технику block matching при вычислении DM.

В условиях различных иллюминаций (L) и экспозиций (E) приведённые изображения демонстрируют, что SID-дескриптор сильно сглаживает объекты и не позволяет оценивать глубинные карты правильно. В то же время восстановленные глубинные карты на основе VPR и Daisy демонстрируют близкие результаты как в количественной метрике, так и при субъективном анализе сформированных карт. Анализ различных изображений (рис. 2, табл. 1) демонстрирует лучшее качество робастности для

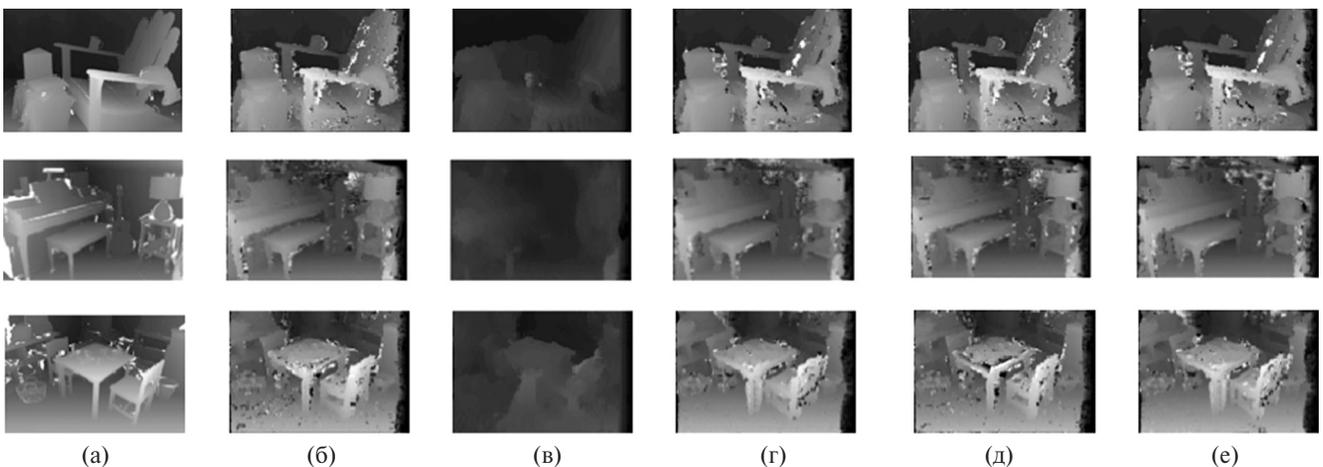


Рис. 2. Результаты восстановления глубинных карт: а — истинная (GT) глубинная карта и восстановленные: б — алгоритм Daisy, в — алгоритм SID, г — предложенный алгоритм VPR, д — алгоритм Daisy E (неидеальное освещение), е — алгоритм VPR E (неидеальное освещение) в случаях изображений (сверху вниз): Adirondack, Piano и Playtable.

Таблица 1. Значения критерия QBP для нового VPR и других дескрипторов в восстановлении глубинных карт в условиях неидеальностей в экспозиции (E) и в освещении (L) стереопар

Стереопара/Дескриптор	Daisy	VPR	Daisy E	VPR E	Daisy L	VPR L
Adirondack	25,76	24,35	25,32	24,69	40,27	34,69
Jadeplant	47,58	45,36	24,81	23,20	28,26	20,59
Piano	22,57	23,37	22,68	24,02	33,05	28,76
Pipes	27,10	25,26	26,48	25,45	41,96	39,28
Playroom	24,31	22,00	23,78	21,96	32,75	29,74
Flowers	43,24	43,95	43,42	44,23	50,87	54,16
Classroom	18,67	20,83	22,22	24,86	44,23	46,98
Sticks	24,64	23,38	24,20	23,49	43,49	46,74
Backpack	10,89	10,28	10,92	10,55	19,44	19,62
Motorcycle	17,06	16,32	16,16	16,34	25,44	21,41
Усреднённые по парам	26,18	25,51	24,00	23,87	39,57	37,15

VPR-дескриптора в обработке неточностей регистрации стереопар изображений.

Так как при восстановлении глубинной карты была использована традиционная техника block matching, проблема окклюзий не может быть разрешена прямо. Новый VPR-дескриптор может быть использован совместно с другими алгоритмами, такими как семиглобальный алгоритм или основанный на укороченных графах.

ВЫВОДЫ

Таким образом, анализ физических результатов численного моделирования нового метода для восстановления глубинных карт позволяет сделать следующие важные выводы:

1. Предложенный и обоснованный метод, основанный на примитивах и отношениях между ними, характеризует различия в цветах, положении плоскостей, дистанциях и углах между примитивами, а также позволяет улучшить качество восстанавливаемых глубинных карт.

2. Новый дескриптор VPR демонстрирует лучшие показатели при восстановлении глубинных карт среди всех известных, в частности при сравнении с DAISY и SID.

3. Экспериментально подтверждено, что VPR-дескриптор демонстрирует лучшую робастность при формировании глубинных карт в условиях радиометрических различий в экспозиции и освещении стереопар изображений.

СПИСОК ЛИТЕРАТУРЫ

1. Ramos-Diaz E., Kravchenko V., Ponomaryov V. // EURASIP J. Adv. Signal Process. 2011. P. 1–10.
2. Кравченко В.Ф., Пономарев В.И., Пустовойт В.И. // ДАН. 2014. Т. 459. № 1. С. 32–36.
3. Кравченко В.Ф., Пономарев В.И., Пустовойт В.И. // ДАН. 2015. Т. 465. № 3. С. 293–297.
4. Кравченко В.Ф., Пономарев В.И., Пустовойт В.И., Садовничий С.Н. // ДАН. 2017. Т. 475. № 5. С. 514–518.
5. Huitron V., Ponomaryov V. // IEEE Lat. Amer. Trans. 2016. V. 14. № 6. P. 2968–2973.
6. Gonzalez-Huilton V., Ponomaryov V., Ramos-Diaz E., Sadovnychiy S. // Signal, Image, Video Proc. 2018. V. 12. № 2. P. 231–238.
7. Rosas D., Ponomaryov V., Reyes R. // Int. J. Comput. 2018. V. 17. № 3. P. 171–179.
8. Kravchenko V.F., Perez-Meana H.M., Ponomaryov V.I. Adaptive Digital Processing of Multidimensional Signals with Applications. Moscow: Fizmatlit, 2009.
9. Felsberg M., Sommer G. // IEEE Trans. Sig. Proc. 2001. V. 49. № 12. P. 3136–3144.
10. Xuanzi Y. // Machine Vision and Appl. 2015. V. 26. № 7/8. P. 975–990.
11. Wan Y., Miao Z., Tang Z., Wan L., Wang Z. // IEICE Trans. Inf. Sys. 2012. V. 95. № 7. P. 2021–2024.
12. Tola E., Lepetit V., Fua P. // IEEE Trans. Pat. Anal. Mach. Intell. 2010. V. 32. № 5. P. 815–830.
13. Pugeault N., Wörgötter F., Krüger N. // Int. J. Human. Robot. 2010. V. 7. P. 379–405.
14. <http://vision.middlebury.edu/stereo/data> (November 2013).
15. <http://vision.middlebury.edu/stereo/>. 2016.

DEPTH MAP RECONSTRUCTION BASED ON FEATURES FORMED BY DESCRIPTOR OF STEREO COLOR PAIRS

V. F. Kravchenko^{1,2,3}, V. I. Ponomaryov⁴,
Academician of the RAS V. I. Pustovoi², D. Rosas-Miranda⁴

¹*Kotelnikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Moscow, Russian Federation*

²*Scientific and Technological Center of Unique Instrumentation, Russian Academy of Sciences, Moscow, Russian Federation*

³*Bauman Moscow State Technical University, Moscow, Russian Federation*

⁴*Instituto Politecnico Nacional de Mexico, Mexico-city, Mexico*

Received March 25, 2019

Novel local image descriptor that is tested in the computer vision problem is proposed. The designed descriptor is based on visual primitives and relations between them, namely, coplanarity, cocolority, distance and angle. The designed feature descriptor covers both geometric and appearance information. Proposed descriptor has demonstrated its ability to compute depth maps from image pairs where the performance evaluation via criterion *Bad Matching Pixels* has shown it superior quality in comparison with other descriptors from state-of-the-art methods.

Keywords: image descriptor, coplanarity, cocolority, visualization, depth map.