

Рябокoнь Д. И.

Прoстранственная реконструкция поверхностей по стереопаре изображений с помощью алгоритма поиска минимального сечения на графе

Введение

Большинство известных алгоритмов обработки стереопар обладают той особенностью, что распознают изображения построчно, и составляют рельеф из фрагментов, полученных на различных строках изображения независимо [1-5]. Особенности этих алгоритмов в том, что реконструируемая модель не является непрерывной, а состоит из рваных кусков.

За последние несколько лет в рамках теории структурного распознавания образов были разработаны новые алгоритмы восстановления пространственной конфигурации объектов по стереопарам снимков [6-7]. Эти алгоритмы позволяли восстанавливать поверхность не в виде набора непрерывных участков, полученных в результате независимой обработки строк на стереопаре, а в виде цельной непрерывной поверхности как результата одновременной обработки всех строк на изображениях [13]. Восстановление рельефа в каждом пикселе изображения накладывает ограничения как на рельеф в соседних по горизонтали пикселей, так и на рельеф в соседних по вертикали пикселей.

В представленной работе задача стереорекострукции ставится как задача поиска оптимальной разметки, широко используемая в методах структурного распознавания образов [12]. Далее приводится метод сведения поставленной задачи к задаче поиска минимального сечения на графе. Среди работ, которые внесли существенный вклад в исследование проблемы сведения некоторого разрешимого подкласса задач разметки к известным разрешимым задачам на графе, следует отметить [8-10]. В конце работы представлены результаты восстановления пространственной конфигурации участков земной поверхности по стереопарам изображений.

Формулировка задачи

Имеются две камеры. Элементы поля зрения камер будем называть пикселями, и обозначать их парой переменных (i, j) , или же одной переменной t . Множество соседей пикселя $t \equiv (i, j)$ будем обозначать $\tau(t) = \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\}$.

На снимках, поступающих из камер, содержится изображение одного и того же объекта, являющегося непрерывной поверхностью. Значения сигналов в элементах поля зрения камер обозначим $F_L(i, j)$ и $F_R(i, j)$ соответственно.

Будем считать, что стереопара (F_L, F_R) является идеальной: всякому пикселю (i, j) в j -ой строке левого изображения соответствует пиксель (i', j) j -ой строке правого изображения.

Параллаксом (состоянием, меткой) пикселя (i, j) левого изображения будем называть разность $k \equiv i - i'$, где i' – это горизонтальная координата соответствующего ему пикселя на правом изображении. Будем считать, что значения параллакса пикселей левого изображения принимают конечное число значений из множества $\{0, 1, \dots, K-1\}$. Функцию, которая ставит в соответствие каждому пикселю его параллакс, обозначим \bar{k} , а значение параллакса в пикселе $t \equiv (i, j)$ левого изображения обозначим $k(t)$.

Для каждого пикселя $t \equiv (i, j)$ левого изображения заданы значения функции потерь $q_t(k)$, определяющие штраф за принятие решения о том, что пиксель t имеет параллакс k .

Для каждой пары соседних пикселей t и t' заданы значения функции потерь $g_{tt'}(k, k')$, определяющие штраф за принятие решения о том, что пиксель t имеет параллакс k и вместе с тем его сосед t' имеет параллакс k' .

Задача стереореконструкции (нахождения параллаксов) ставится как задача оптимальной разметки.

Необходимо для каждого пикселя t левого изображения указать такое значение параллакса $k^*(t)$, чтобы суммарные потери были минимальны:

$$\bar{k}^* = \arg \min_k \sum_{t, t' \in \tau(t)} g_{tt'}(k(t), k(t')) + \sum_t q_t(k(t)). \quad (1)$$

В общем случае задача (1) является NP-полной. В работах [6, 7] показано, что задача (1) может быть решена за полиномиальное время в случае, когда для любых $t, t' \in \tau(t)$ и $k_1 < k_2, k'_1 < k'_2$ выполняется условие монотонности:

$$g_{tt'}(k_1, k_2) + g_{tt'}(k_2, k'_1) \geq g_{tt'}(k_1, k'_1) + g_{tt'}(k_2, k'_2). \quad (2)$$

Выберем значения функции $g_{tt'}(k, k')$ следующим образом:

$$g_{tt'}(k, k') = \begin{cases} 0 & \text{при } |k - k'| \leq 1, \\ \infty & \text{иначе} \end{cases}.$$

Такой выбор означает, что в параллакс соседних пикселей должен отличаться не более чем на единицу. Можно убедиться, что функция g удовлетворяет условию монотонности (2) и задача (1) решается за полиномиальное время.

Иллюстрация формулировки задачи

Формулировку задачи (1) проиллюстрируем на примере графа (рис. 1, 2).

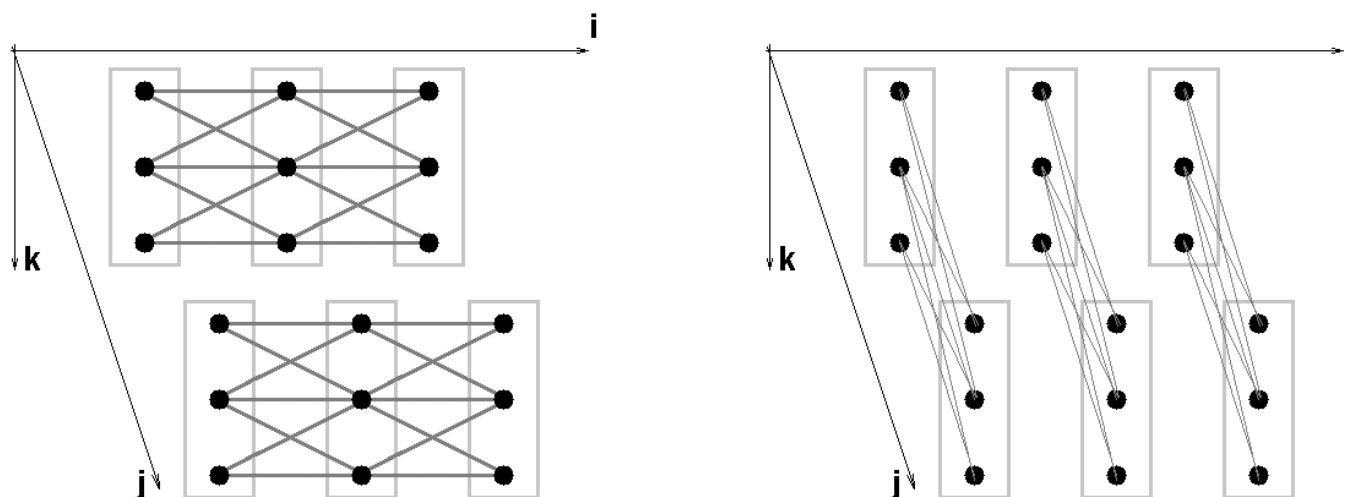


Рис. 1. Горизонтальные ребра графа.

Рис. 2. Вертикальные ребра графа.

На рис. 1 содержится массив пикселей размером 3×2 . Пиксели обозначены клетками. Каждый пиксель может находиться в одном из трех состояний. Состояния обозначены круглыми метками. Множество вершин графа $v_{i,j}^k$ состоит из 18 вершин: оно включает в себя все возможные состояния всех пикселей. Ограничение на состояния соседних пикселей осуществляется с помощью ребер. Две вершины графа $v_{i,j}^k$ и $v_{i',j'}^{k'}$ соединены ребром, если эти вершины являются состояниями k и k' соседних пикселей (i, j) и (i', j') , и разность параллаксов $|k - k'|$ не превышает единицы.

Каждому состоянию k в каждом пикселе (i, j) приписывается число $q_{i,j}(k)$ – штраф за принятие решения о том, что пиксель (i, j) имеет параллакс k .

В терминах графов задача (1) формулируется следующим образом. Необходимо в каждой клетке выбрать одно состояние так, чтобы сумма чисел, приписанных выбранным состояниям, была минимальна, и при этом в соседних клетках выбранные состояния были бы соединены ребром.

Сведение задачи разметки к задаче поиска минимального сечения на графе

Один из методов решения задачи (1) заключается в ее сведении к задаче поиска минимального сечения на графе.

Строится ориентированный граф $\langle V, E, C \rangle$ с вершинами V , ребрами E и весами на ребрах C . Множество вершин V состоит из вершины-источника S , вершины-приемника R , а также трехмерного массива вершин $v_{i,j}^k$.

Множество ребер E и их веса C строятся по следующим правилам (рис. 3):

1. Из вершины-источника S выходят вниз ребра к вершинам $v_{i,j}^0$. Вес этих ребер определяется по формуле $c(S \rightarrow v_{i,j}^0) = q_{i,j}(0)$.
2. Из каждой вершины $v_{i,j}^k$ выходит направленное вниз ребро в вершину $v_{i,j}^{k+1}$. Вес этого ребра определяется по формуле $c(v_{i,j}^k \rightarrow v_{i,j}^{k+1}) = q_{i,j}(k+1)$.
3. К вершине-приемнику R приходят сверху ребра из вершин $v_{i,j}^{K-2}$. Вес этих ребер определяется по формуле $c(v_{i,j}^{K-2} \rightarrow R) = q_{i,j}(K-1)$.
4. Две вершины $v_{i,j}^k$ и $v_{i',j'}^{k-1}$ соседних пикселей (i, j) и (i', j') соединены ребром с весом $c(v_{i,j}^k \rightarrow v_{i',j'}^{k-1}) = \infty$.

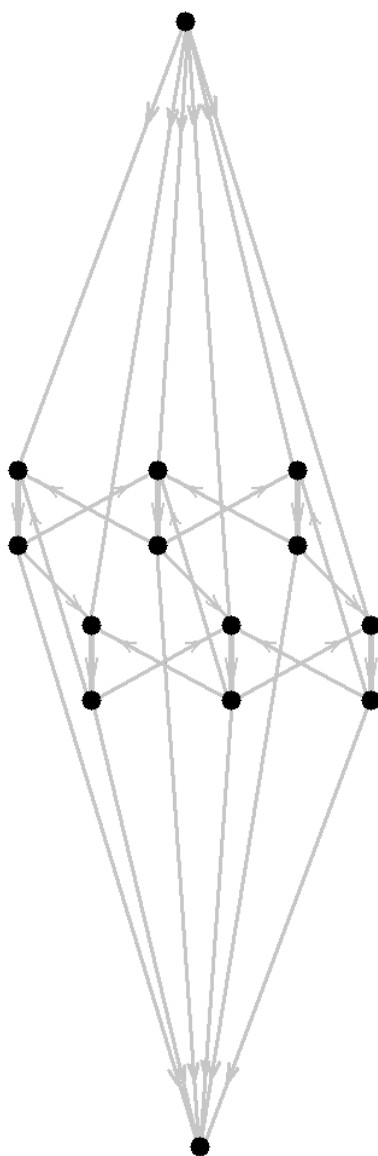


Рис. 3. Пример графа.

Сечением графа $\langle V, E, C \rangle$ будем называть разбиение его множества вершин V на два непересекающихся подмножества V_S и V_R , включающих в себя источник S и приемник R соответственно. Пропускная способность сечения будет определяться суммой весов ребер, направленных из вершин множества V_S к вершинам множества V_R .

В работах [8-10] показано, что поиск оптимальной разметки k^* в задаче (1) эквивалентен поиску на графе $\langle V, E, C \rangle$ сечения (V_S^*, V_R^*) с минимальной пропускной способностью (поиску минимального сечения).

Емкость сечения с минимальной пропускной способностью совпадает с качеством наилучшей разметки. Если $k^*(i, j)$ – это разметка, являющаяся решением задачи (1), то минимальное сечение определяется как $V_S^* = S \cup \left(\bigcup_{(i,j)} \bigcup_{k < k^*(i,j)} v_{j,j}^k \right)$, $V_R^* = V \setminus V_S^*$. И наоборот, если (V_S^*, V_R^*) – это минимальное сечение, то решение задачи (1) определяется как $k^*(i, j) = k \mid (v_{i,j}^{k-1} \in V_S^*) \& (v_{i,j}^k \in V_R^*)$.

Алгоритмы поиска минимального сечения приведены в источнике [11]. Наиболее известные алгоритмы решения задачи о минимальном сечении – это алгоритм Эдмондса–Карпа, алгоритм «проталкивания предпотока» и алгоритм «поднять-и-в-начало»(lift-to-front). Время работы перечисленных алгоритмов на графе, что содержит V вершин и E ребер, составляет $O(V \cdot E^2)$, $O(V^2 \cdot E)$ и $O(V^3)$ соответственно. Таким образом, задачу (1) можно решить за время $O(I^3 \cdot J^3 \cdot K^3)$, где I – горизонтальный размер поля зрения камер, J – вертикальный размер поля зрения камер, а K – количество возможных состояний в каждом пикселе. Для реализации алгоритмов требуется память объемом $O(E)$

Результаты стереорекострукции

На рис. 4-6 представлены результаты пространственной реконструкции непрерывных поверхностей по стереопарам снимков.

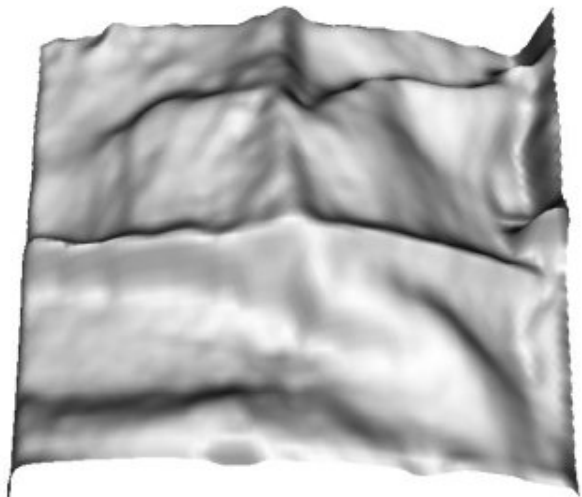


Рис. 4. Стереореконструкция скатерти.

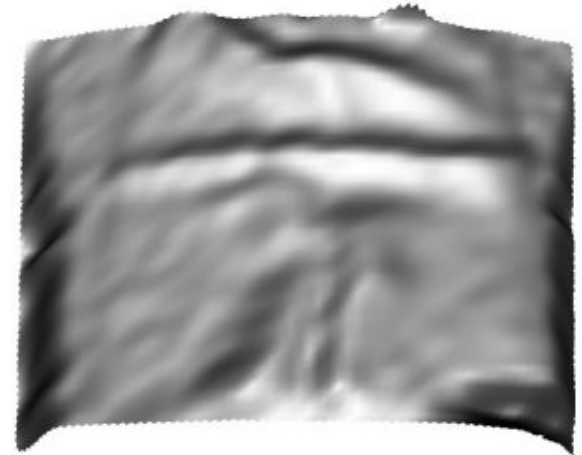
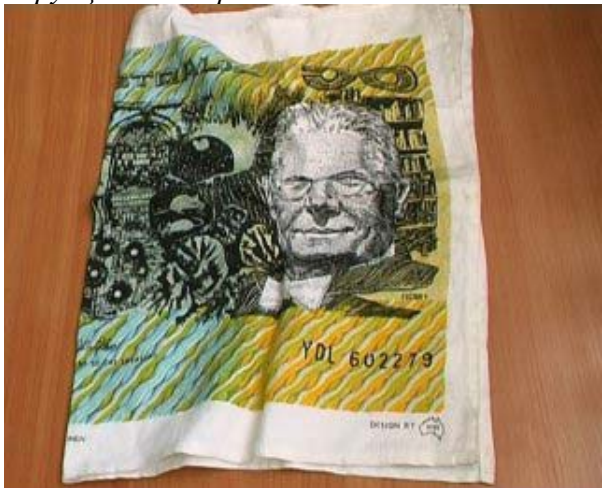


Рис. 5. Стереореконструкция скатерти

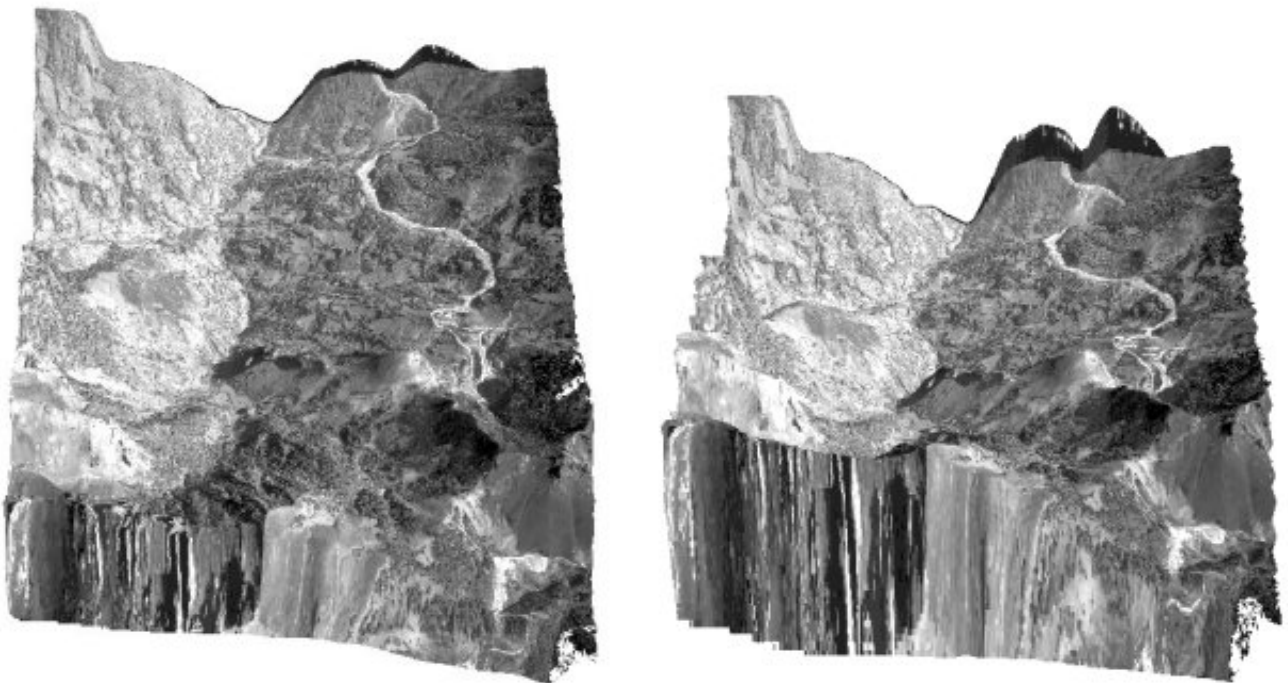
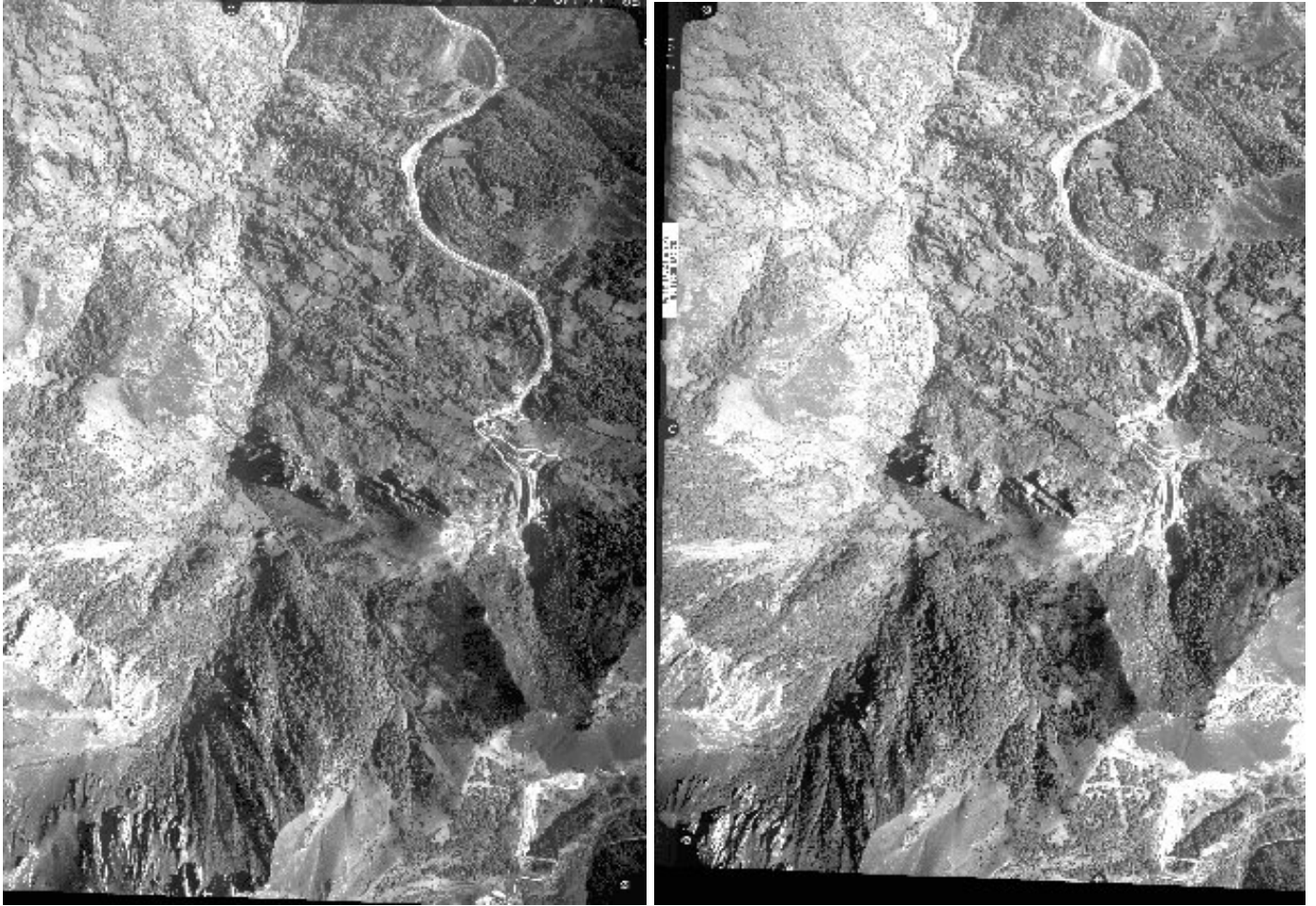


Рис. 6. Стереореконструкция участка земной поверхности.

В верхней части рис. 4 и рис. 5 содержатся стереопары скатерти, полученные с помощью двух видеокамер. Скатерть образовала на столе непрерывную поверхность, моделируя тем самым поверхность земли. В нижней части рисунков содержатся пространственные модели – раскрашенная модель и не раскрашенная «гипсовая» модель, – восстановленные по стереопаре снимков. Достоинство данного эксперимента в том, что результаты стереорекострукции можно легко проверить на реальном рельефе, образованном скатертью.

В верхней части рис. 6 содержатся стереопары земной поверхности, полученные в результате аэрофотосъемки. В нижней части рис. 6 содержатся две пространственные модели, сгенерированные с разных точек зрения.

Выводы

Предложен алгоритм пространственной реконструкции участков земной поверхности по стереопаре снимков. Отличительная особенность алгоритма заключается в том, что при восстановлении рельефа учитывается его непрерывность как в горизонтальном, так и в вертикальном направлении. Поиск трехмерной модели местности выполняется среди таких пространственных моделей, которые являются непрерывными поверхностями. Все точки стереопары должны быть бинокулярно наблюдаемы: если некоторая точка поверхности видима на одном снимке, то она должна быть видима и на другом снимке.

По сравнению с алгоритмами построчной стереорекострукции, предложенный алгоритм позволяет улучшить визуальное качество восстанавливаемых моделей объектов – модели выглядят более гладкими и не содержат скачкообразных участков. Для объективного же сравнения результатов восстановления рельефа разными алгоритмами следует сопоставлять такие характеристики как доверительные интервалы восстановленного рельефа или его отклонение от реального рельефа в модельном эксперименте. Разработка критериев, оценивающих качество восстановления рельефа, будет проводиться в дальнейших исследованиях. Следует отметить, что в приведенных примерах восстановления рельефа результат работы предложенного алгоритма отличается от результата

алгоритма построчной обработки изображений в небольшом количестве пикселей (порядка 20-30%). Трудоемкость алгоритма – требуемая память $O(I \cdot J \cdot K)$ и время работы $O(I^3 \cdot J^3 \cdot K^3)$ – значительно превышают трудоемкость построчных алгоритмов, которые требуют память объемом $O(I \cdot K)$ и решают задачу за время $O(I \cdot J \cdot K)$.

Результаты работы могут представлять интерес специалистам в области геодезии.

Литература

1. Гимельфарб Г.Л. *Симметричный подход к задаче автоматических стереоскопических измерений в фотометрии*. Кибернетика 1979, (2), с.73-82.
2. Гимельфарб Г.Л., Марченко В.Б., Рыбак В.И. *Автоматическое отождествление идентичных точек стереоснимков с учетом неравномерных фотометрических искажений*. Кибернетика 1976, (4), с.107-112.
3. Шлезингер М.І., Флах Б., Онищенко І.В. *Оптимізація кусково-лінійної деформації знімків у задачах машинного стереозору*. П'ята Всеукраїнська міжнародна конференція оброблення сигналів та розпізнавання образів (УкрОБРАЗ 2000), с. 177-178.
4. Савчинський Б.Д. *Порівняльний аналіз алгоритмів стереозору в рамках байєсівської теорії прийняття рішень*. П'ята Всеукраїнська міжнародна конференція оброблення сигналів та розпізнавання образів (УкрОБРАЗ 2000), с. 25-28.
5. Рябоконт Д.І. *Створення тривимірних моделей об'єктів за стереопарами зображень для електронних музеїв*. Електронні зображення та візуальні мистецтва EVA 2002, праці конференції. 22-24 травня 2002 р., Київ, с. 61-68.
6. Schlesinger M.I., Flach B. *Some solvable subclass of structural recognition problems*. Czech Pattern Recognition Workshop 2000, pages 55-62.
7. Schlesinger M.I., Flach B. *Analysis of optimal labeling problems and their application to image segmentation and binocular stereovision*. Шоста Всеукраїнська міжнародна конференція оброблення сигналів та розпізнавання образів (УкрОБРАЗ 2002), с. 15-18.

8. Ishikawa H., Geiger D. *Segmentation by Grouping Junctions*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA, June 1998.
9. Boykov, Yu., Veksler, O., and Zabih, R. *Fast Approximate Energy Minimization via Graph Cuts*. In International Conference of Computer Vision, pages 377-384, 1999.
10. Kolmogorov, V., Zabih, R. *Computing visual correspondence with occlusions via graph cuts*. In International Conference of Computer Vision pages 508-515, 2001.
11. Кормен Т., Лейзерсон Ч., Ривест Р. Алгоритмы: построение и анализ. М.: МЦНМО, 2000. С. 535-578.
12. Schlesinger M.I and Hlavac V. *Ten Lectures on Statistical and Structural Recognition*. Computational Imaging and Vision, Kluwer Academic Publishers, Dordrecht / Boston / London, 522 pages, 2002.
13. Рябоконтъ Д.И. *Постобработка результатов стереорекострукции поверхностей как восстановление скрытых параметров марковского поля*. Управляющие системы и машины – 2003, (4), с. 56-61.