

УДК 004.932

Рябоконь Д.И.

**Постобработка результатов стереореконструкции поверхностей как
восстановление скрытых параметров марковского поля**

Введение

Достаточно долгое время для распознавания стереопар использовались алгоритмы, где изображения обрабатывались построчно, и реконструированный рельеф состоял из фрагментов, полученных независимо друг от друга на разных строчках изображения с помощью одномерных алгоритмов динамического программирования [1-5]. Недостатки этих алгоритмов в том, что реконструированный рельеф представлял собой непрерывную поверхность только в одном горизонтальном направлении.

Существенный шаг вперед за последние годы был сделан благодаря использованию новых оптимизационных алгоритмов, которые созданы в рамках теории структурного распознавания образов [6,7]. В частности, удалось обобщить условие непрерывности рельефа с одномерного на двумерный случай. Это дало возможность искать поверхность не в виде набора непрерывных участков, которые восстанавливаются на каждой строке независимо, а в виде цельной непрерывной поверхности.

Особенностями же «двумерных» алгоритмов является то, что они могут решать точно поставленную оптимизационную задачу только для некоторого класса математических моделей. Иначе, оптимизационная задача является NP-полной, и результат работы алгоритма – отказ от распознавания или приближенное решение задачи [8-10].

Целью данной работы является разработка алгоритма восстановления пространственной конфигурации гладких непрерывных поверхностей. Суть подхода состоит в том, что для реконструкции рельефа используются результаты алгоритма предварительной оценки дальности, которые получены без учета непрерывности поверхности. Далее производится постобработка предварительных результатов, которая позволяет получить более качественную модель объекта.

Особенностью работы является использование предположения о том, что поверхность, которую необходимо реконструировать, является случайной реализацией марковского поля. Среди работ, которые внесли существенный вклад в исследование марковских случайных полей для оптимизационных задач текстурной сегментации и машинного стереозрения, следует отметить работы [11-13].

Формулировка задачи

Пусть T – поле зрения, которое определяется как множество $T = \{t \equiv (i, j) \mid 0 < i < H, 0 < j < W\}$, где H – горизонтальный размер поля зрения, W – вертикальный размер поля зрения.

Пусть на каждом элементе поля зрения наблюдается некий сигнал. Множество всех допустимых сигналов обозначим X . **Наблюдением** \bar{x} будем называть функцию, которая каждому элементу поля зрения ставит в соответствие элемент из множества сигналов:

$$\bar{x} : T \rightarrow X . \quad (1)$$

Каждый элемент поля зрения может находиться в одном из нескольких состояний. Множество всех состояний обозначим K . **Разметкой (скрытым состоянием)** \bar{k} будем

называть функцию, которая каждому элементу поля зрения ставит в соответствие элемент из множества состояний:

$$\bar{k} : T \rightarrow K. \quad (2)$$

В дальнейшем сигналы $x(t)$ будут обозначать предварительные измерения дальности до точек некоторого трехмерного объекта. Эти измерения выполнены неточно, или содержат значительные погрешности. Значения состояний $k(t)$, которые неизвестны, будут истинными значениями дальности. Если t – элемент поля зрения, то $x(t)$ обозначает его сигнал, а $k(t)$ – его состояние.

Пусть $\tau \subset T \times T$ – **множество соседних элементов** поля зрения. Соседями элемента $t = (i, j)$ будем считать элементы $(i-1, j)$, $(i, j-1)$, $(i+1, j)$, $(i, j+1)$. Выражение $(t, t') \in \tau$ означает, что элементы $t \in T$ та $t' \in T$ являются соседними.

Для любого элемента t поля зрения **условная вероятность** $P(x(t) | k(t))$ сигнала $x(t)$ при **условии состояния** $k(t)$ не зависит от t и определяется функцией q :

$$q : X \times K \rightarrow R. \quad (3)$$

Разметка \bar{k} представляет собой реализацию случайного марковского поля со структурой соседства τ , нормирующим множителем C и статистическими параметрами, задаваемыми функцией $g : \tau \times K \times K \rightarrow R$. Вероятность разметки вычисляется по формуле:

$$P(k) = C \cdot \prod_{(t,t') \in \tau} g_{t'}(k(t), k(t')). \quad (4)$$

Постановка задачи нахождения дальностей до поверхности следующая. На поле зрения **заданы** наблюдения \bar{x} и значения функции $g_{t'}(k(k), k(t'))$. Необходимо **найти**

функцию q^* , которая максимизирует вероятность наблюдений $P(\bar{x})$ и сгенерировать случайную разметку \bar{k}^* согласно наблюдениям \bar{x} и вероятностям q^* .

$$\begin{aligned}
 q^* &= \arg \max_q P(\bar{x}) = \arg \max_q \sum_{\bar{k} \in K^T} P(\bar{k}, \bar{x}) = \arg \max_q \sum_{\bar{k} \in K^T} P(\bar{k}) \cdot P(\bar{x} | \bar{k}) = \\
 & \arg \max_q \sum_{\bar{k} \in K^T} \prod_{(t,t') \in \tau} P(\bar{k}) \cdot \prod_{t \in T} P(x(t) | k(t)) = \\
 & \arg \max_q \sum_{\bar{k} \in K^T} \prod_{(t,t') \in \tau} g_{u'}(k(t), k(t')) \cdot \prod_{t \in T} q(x(t), k(t)).
 \end{aligned} \tag{5}$$

Существуют другие постановки задачи стереорекострукции поверхности. В работах [6,7] задача ставится как нахождение наиболее вероятной разметки \bar{k}^* для заданных значений функции $g_{u'}(k(k), k(t'))$. В этих же работах показано, что такая задача может быть решена точно в случае, когда множество состояний K является упорядоченным, и функция g удовлетворяет условию монотонности $g_{u'}(k_1, k_2) \cdot g_{u'}(k_2, k_1) \leq g_{u'}(k_1, k_1) \cdot g_{u'}(k_2, k_2)$ для любых $tt' \in \tau$ и $k_1 < k_2, k_1' < k_2'$. Если же это условие не выполняется, то приходится отказываться от точного решения задачи и искать решение приближенно с помощью других алгоритмов, которые еще не исследованы должным образом.

Предложенный метод нахождения дальностей до объекта предназначен именно для случая, когда невозможно описать распознаваемый объект моделью монотонными свойствами функции g .

Алгоритм решения задачи

Алгоритм решения задачи (5) базируется на наиболее правдоподобной оценке параметров системы в режиме самообучения [14, 15] и оценке на этой основе параметров скрытого марковского поля [16]. На каждой итерации $r=1..R$ будет сгенерировано множество разметок $\hat{K}^{[r]}$, для которых подсчитано приближенное значение $P^{[r]}(\bar{k} | \bar{x})$, и подсчитано функцию $q^{[r]}$. Значения $q^{[0]}$ выбираются произвольным случайным образом. Во множество $\hat{K}^{[0]}$ включаются произвольные случайные разметки. Каждая итерация состоит из трех шагов.

Шаг 1. Для всех разметок \bar{k} из множества $\hat{K}^{[r-1]}$ подсчитать приближенные значения условных вероятностей $P^{[r]}(\bar{k} | \bar{x})$:

$$P^{[r]}(\bar{k} | \bar{x}) \approx \frac{P(\bar{x}, \bar{k})}{\sum_{\bar{k}' \in \hat{K}} P(\bar{x}, \bar{k}')} = \frac{P(\bar{k}) \cdot P(\bar{x} | \bar{k})}{\sum_{\bar{k}' \in \hat{K}} P(\bar{k}') \cdot P(\bar{x} | \bar{k}')} = \frac{\prod_{(t,t') \in \tau} g_{tt'}(k(t), k(t')) \cdot \prod_{t \in T} q^{[r-1]}(x(t), k(t))}{\sum_{\bar{k}' \in \hat{K}} P(\bar{k}') \cdot P(\bar{x} | \bar{k}')} . \quad (6)$$

Шаг 2. Для всех $x \in X$, $k \in K$ подсчитать значения функции $q^{[r]}(x, k)$:

$$q^{[r]}(x, k) = P(x | k) = \frac{\sum_{t \in T, t: \bar{x}(t)=x} \sum_{\bar{k}' \in \hat{K}, \bar{k}'(t)=k} P^{[r]}(\bar{k}' | \bar{x})}{\sum_{x' \in X} \sum_{t \in T, t: \bar{x}(t)=x'} \sum_{\bar{k}' \in \hat{K}, \bar{k}'(t)=k} P^{[r]}(\bar{k}' | \bar{x})} . \quad (7)$$

Шаг 3. Регенерация разметок из множества $\hat{K}^{[r]}$. Если $\bar{k}^{[r-1]}$ – это некоторая разметка из множества $\hat{K}^{[r-1]}$, то она будет преобразована в разметку $\bar{k}^{[r]}$ множества $\hat{K}^{[r]}$ по такому принципу.

Для каждого элемента t поля зрения подсчитываются вероятности $p(k | x(t), \bar{k}^{[r-1]}(\tau(t)))$ состояний $k \in K$ в элементе t при условии фиксированных состояний $\bar{k}^{[r-1]}(\tau(t))$ в соседних элементах $\tau(t)$ и условии известного сигнала $x(t)$.

$$p(k | x(t), \bar{k}^{[r-1]}(\tau(t))) = \frac{p(x(t), k, \bar{k}^{[r-1]}(\tau(t)))}{p(x(t))} = \frac{p(k, \bar{k}^{[r-1]}(\tau(t))) \cdot p(x(t) | k)}{\sum_{k' \in K} p(x(t), k')} \quad (8)$$

$$C \cdot p(k, \bar{k}^{[r-1]}(\tau(t))) \cdot p(x(t) | k) = C \cdot \prod_{(t', t) \in \tau} g_{t'}(k, k^{[r-1]}(t')) \cdot q^{[r]}(x(t), k).$$

Далее значения состояния $k^{[r]}(t)$ генерируется как случайная величина с распределением $p(k | x(t), \bar{k}^{[r-1]}(\tau(t)))$.

Исходя из известной теоремы М.И. Шлезингера о монотонности самообучения [14], можно доказать, что при таком итерационном процессе вероятность $P^{[r]}(\bar{x})$ будет неубывающей функцией:

$$P^{[r]}(\bar{x}) = \sum_{\bar{k} = K^r} \prod_{(t, t') \in \tau} g_{t'}(k(t), k(t')) \cdot \prod_{t \in T} q^{[r]}(x(t), k(t)) \leq P^{[r+1]}(\bar{x}). \quad (9)$$

После достаточно большого числа итераций r в качестве разметки \bar{k}^* , которая требовалась в задаче (5), следует взять произвольную разметку из множества $\hat{K}^{[r]}$.

Вычисление исходных данных для алгоритма

Для работы представленного алгоритма необходимо иметь значения $x(t)$ и $g_{t'}(k(k), k(t'))$. Эти данные вычисляются на основании стереопары изображений объекта, рельеф которого необходимо реконструировать. Пусть $y_L(t)$ и $y_R(t)$ значение оттенка серого цвета в элементах поля зрения стереопары. Если элемент t имеет состояние $k(t)$, то это

означает, что параллакс точки t левого изображения равен $k(t)$, и точка t левого изображения соответствует точке $t + k(t)$ правого изображения.

Значения $g_{\mu'}(k(k), k(t'))$ с точностью до нормирующего множителя строятся как сумма меры сходства между точками $y_L(t)$ и $y_R(t + k(t))$, меры сходства между точками $y_L(t')$ и $y_R(t' + k(t'))$ а также слагаемого, который обуславливает гладкость реконструированного рельефа и определяется неотрицательным параметром α :

$$g_{\mu'}(k(k), k(t')) = \mu(t, t + k(t)) + \mu(t', t' + k(t')) + \alpha \cdot \delta_{k(t), k(t')}. \quad (10)$$

В этой формуле символ δ означает символ Кронекера. Как меру μ сходства элемента t_L левого изображения y_L и элемента t_R правого изображения y_R можно использовать следующее значение:

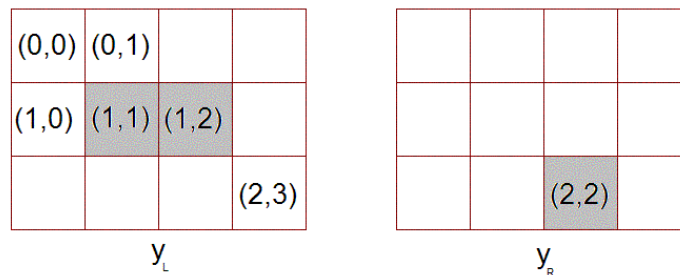
$$\mu(t_L, t_R) = 255 - |y_L(t_L) - y_R(t_R)|. \quad (11)$$

Значения сигналов $x(t)$, которые предварительно оценивают дальности до точек объекта, изображенного на стереопаре, могут быть вычислены по формуле:

$$x(t) = \arg \max_{k \in K} \mu(t, t + k). \quad (12)$$

Не обязательно осуществлять именно такой выбор значений $x(t)$. Данный выбор означает, что каждому элементу левого изображения ставится в соответствие наиболее похожий на него элемент на правом изображении без учета непрерывности рельефа. В качестве сигналов $x(t)$ можно использовать результаты работы какого-нибудь другого алгоритма, который оценивает дальности до точек на изображении. Например, алгоритма, который обрабатывает стереопару построчно и составляет рельеф из участков, которые непрерывны на каждой строке изображения.

Пример построения значений функции g предоставлен на рис. 1. Каждое изображение стереопары определено над полем зрения четыре на три пикселя. При условии известных состояний $k(t) = (1,1)$ и $k(t') = (1,0)$ двух соседних элементов $t = (1,1)$ и $t' = (2,1)$ рассчитаны значения функции $g_{tt'}(k(k), k(t'))$.



$T = \{ t \equiv (i,j) \mid 0 < i < 4, 0 < j < 3 \}$ поле зрения
 $K = \{ (0,0), (0,1), (1,0), (1,1) \}$ состояния элементов
 $t := (1,1)$ соседние элементы
 $t' := (1,2)$
 $k(t) := (1,1)$ состояние элемента t
 $k(t') := (1,0)$ состояние элемента t'

$$g_{tt'}(k(t), k(t')) = \mu(t, t+k(t)) + \mu(t', t'+k(t')) + \alpha \cdot \delta(k(t), k(t'))$$

$$\mu(t, t+k(t)) = 255 - |y_L(t) - y_R(t+k(t))| = 255 - |y_L(1,1) - y_R(2,2)|$$

$$\mu(t', t'+k(t')) = 255 - |y_L(t') - y_R(t'+k(t'))| = 255 - |y_L(1,2) - y_R(2,2)|$$

Рис. 1. Пример подсчета значений функции g .

Результаты работы

На рис. 2 представлена стереопара изображений платка, который создал непрерывную поверхность. На рис. 3 сравниваются трехмерные модели, которые получены по результатам работы двух алгоритмов. Первый алгоритм вычисляет дальностей до объекта без учета его непрерывности. При этом каждой точке на левом снимке ставится в соответствие наиболее похожая на нее точка на правом снимке. Второй алгоритм, который описан в данной работе, на основании предварительных значений дальности строит трехмерную модель с учетом

непрерывности ее рельефа. Как видно, большинство участков поверхности, дальность до которых была предварительно распознана неверно, исчезли.



Рис. 2. Стереопара изображений платка.

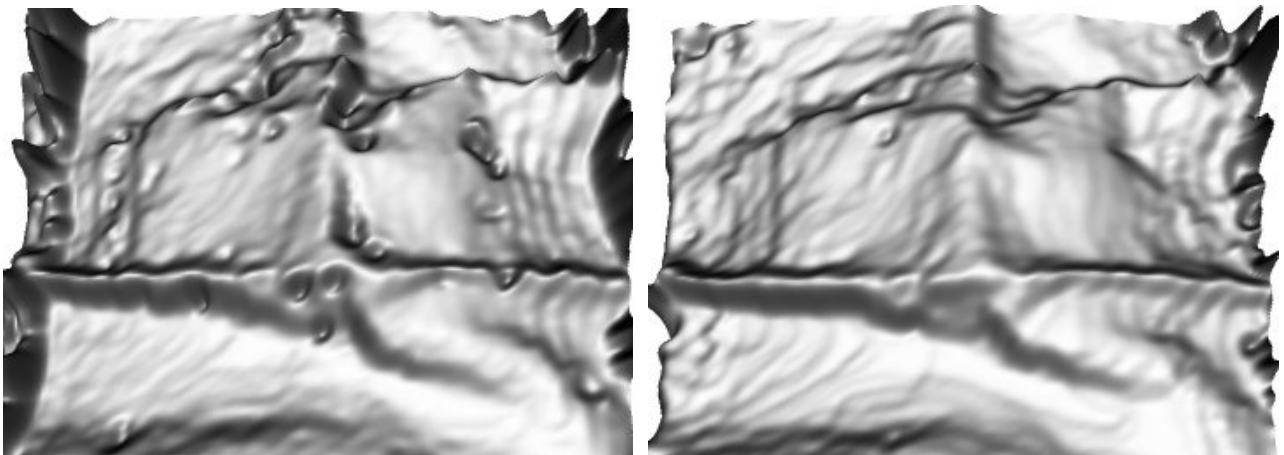


Рис 3. Трехмерные модели, которые построены по результатам работы алгоритма предварительной оценки дальности и алгоритма постобработки.

На рис. 4 изображена стереопара снимков земной поверхности. Результаты ее обработки представлены на рис. 5 и рис. 6. Пространственная модель, полученная без учета непрерывности рельефа местности, содержит много участков, где дальность восстановлена неверно. При использовании постобработки результат стереорекострукции – гладкая непрерывная модель, где высоты точек реконструированы почти всюду верно.

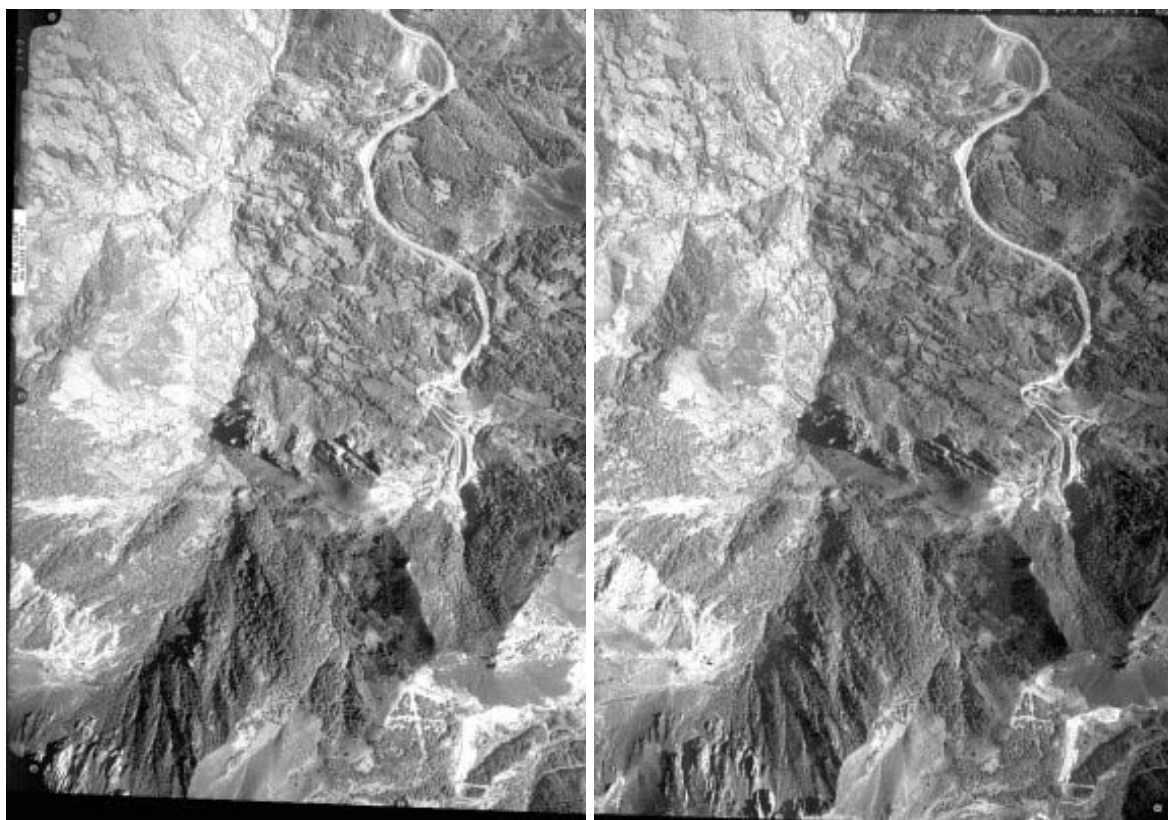


Рис. 4. Стереопара изображений земной поверхности.

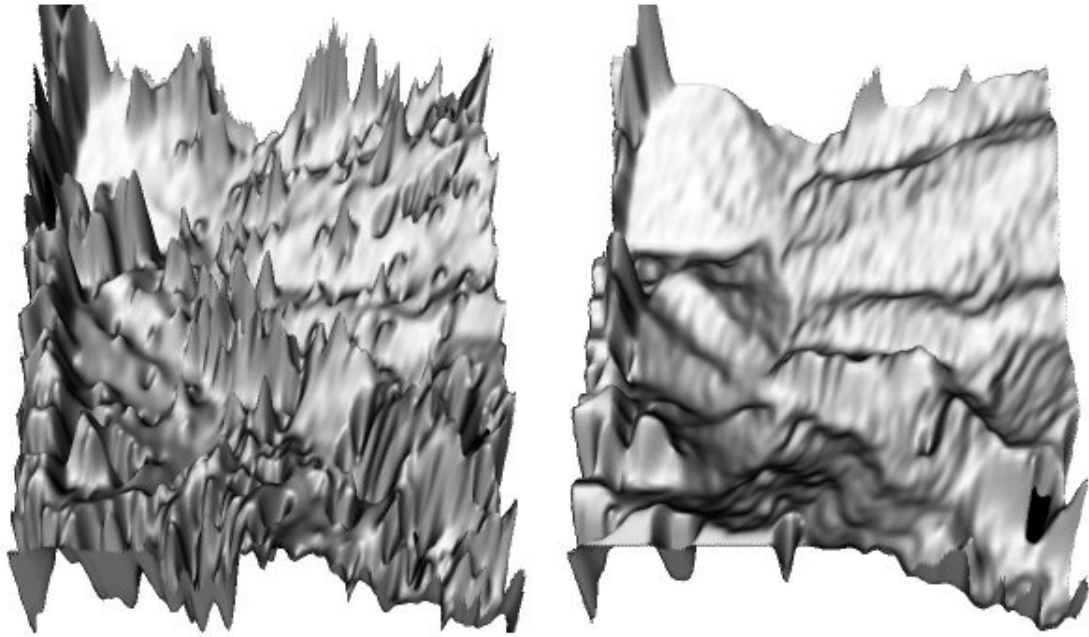


Рис 5. Трехмерные модели, которые построены по результатам работы алгоритма предварительной оценки дальности и алгоритма постобработки(без наложения текстуры).

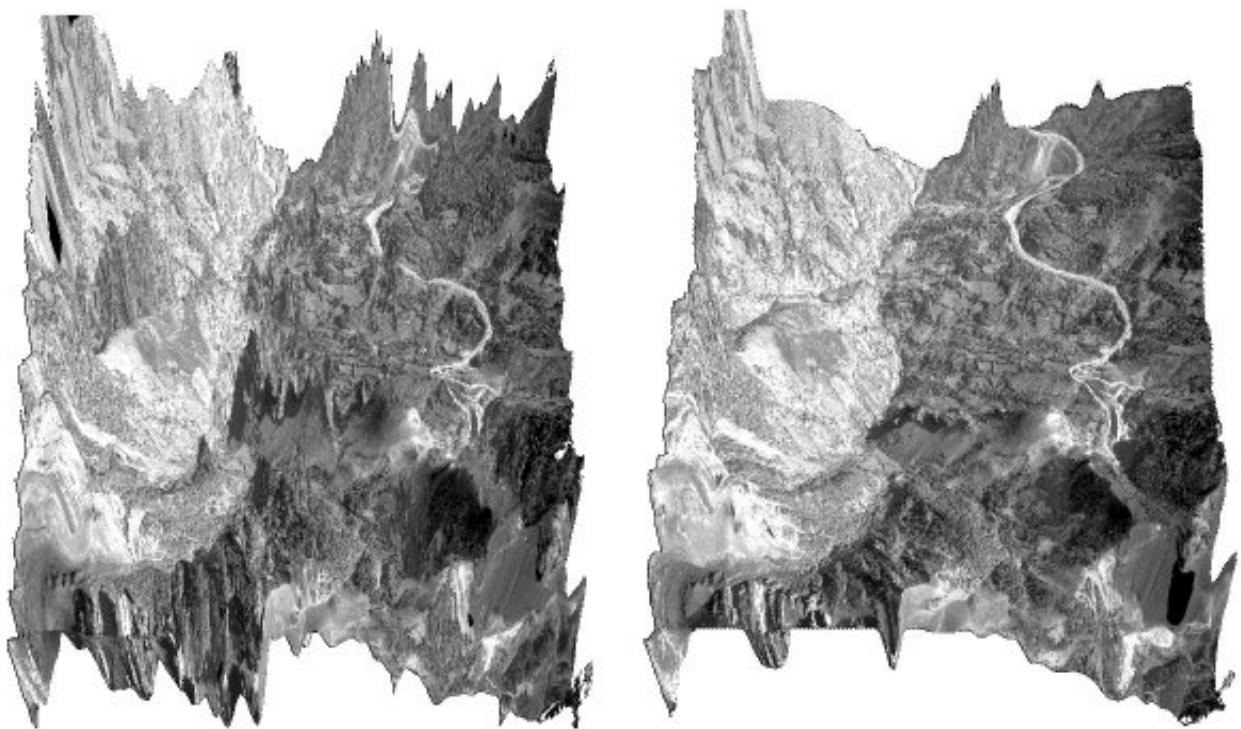


Рис 6. Трехмерные модели, которые построены по результатам работы алгоритма предварительной оценки дальности и алгоритма постобработки (с наложением текстуры).

Выводы

Предложенный алгоритм стереорекострукции позволяет обрабатывать достаточно широкий класс пространственных объектов. Это объекты, состоящие из цельных непрерывных поверхностей. При этом все точки объекта, подлежащие реконструкции, должны содержаться на обоих изображениях стереопары.

Очевидно, что постобработка результатов стереорекострукции для непрерывных поверхностей полезна в случае, когда предварительная оценка дальностей до восстанавливаемого объекта неточна. В результате учета непрерывности объектов, подлежащих восстановлению, и применения постобработки, качество восстанавливаемых моделей улучшается.

Для тестирования алгоритма использован персональный компьютер с тактовой частотой процессора 2GHz. При реконструкции моделей объектов по изображениям размера 1000x1000 элементов время работы алгоритма составило 10-15 мин. Число итераций алгоритма не превышало одной тысячи, скорость сходимости в значительной степени зависела от качества предварительной оценки дальностей до объекта.

Алгоритм также применялся для восстановления пространственной модели задней стенки человеческого глаза и для построения трехмерных моделей человеческих лиц с целью их идентификации.

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. Geman S., Geman D., *Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images*. IEEE Trans. on PAMI, 1984, vol. 6, no. 6, pages 721-741.
12. Gimelfarb G.L. *Texture modeling by multiple pairwise pixel interactions*. IEEE Transaction on PAMI 18(11), pages 1110-1114.
13. Stan Z. Li. *Markov Random Field Modeling in Image Analysis*. Computer Science Workbench, Springer, 2001, 323 pages.
14. Шлезингер М.И. *Взаимосвязь обучения и самообучения в распознавании образов*. Кибернетика – 1968, (2), с. 81-88.
15. 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.
16. Flach B., Kask E., Schlesinger D., Skulysh A. *Unifying Registration and Segmentation for Multi-Sensor Images*. DAGM2002 - Pattern Recognition Symposium, Curriish.