

# Correction of inertial navigation system's errors by the help of video-based navigator based on Digital Terrarium Map.

**Voronov V.V.**

*R.E.T.Kronshtadt, ZAO, Russia,*  
107076, Moscow,  
Stromynka str., 18  
Tel. +7(495)748 3584  
[voronov@kronshtadt.ru](mailto:voronov@kronshtadt.ru)

**Kupervasser O. Yu.**

*R.E.T.Kronshtadt, ZAO, Russia,*  
107076, Moscow,  
Stromynka str., 18  
Tel. +7(495)748 3584  
[olegkup@yahoo.com](mailto:olegkup@yahoo.com)

## Abstract

*This paper deals with the error analysis of a novel navigation algorithm that uses as input the sequence of images acquired from a moving camera and a Digital Terrain (or Elevation) Map (DTM/DEM). More specifically, it has been shown that the optical flow derived from two consecutive camera frames can be used in combination with a DTM to estimate the position, orientation and ego-motion parameters of the moving camera. As opposed to previous works, the proposed approach does not require an intermediate explicit reconstruction of the 3D world. In the present work the sensitivity of the algorithm outlined above is studied. The main sources for errors are identified to be the optical-flow evaluation and computation, the quality of the information about the terrain, the structure of the observed terrain and the trajectory of the camera. By assuming appropriate characterization of these error sources, a closed form expression for the uncertainty of the pose and motion of the camera is first developed and then the influence of these factors is confirmed using extensive numerical simulations. The main conclusion of this paper is to establish that the proposed navigation algorithm generates accurate estimates for reasonable scenarios and error sources, and thus can be effectively used as part of a navigation system of autonomous vehicles.*

## 1 Introduction

Vision-based algorithms has been a major research issue during the past decades. Two common approaches for the navigation problem are: *landmarks* and *ego-motion integration*. In the landmarks approach several features are located on the image-plane and matched to their known 3D location. Using the 2D and 3D data the camera's pose can be derived. Once the landmarks were found, the pose derivation is simple and can achieve quite accurate estimates. The main difficulty is the detection of the features and their correct matching to the landmarks set.

In ego-motion integration approach the motion of the camera with respect to itself is estimated. The ego-motion can be derived from the optical-flow field, or from instruments such as accelerometers and gyroscopes. Once the ego-motion was obtained, one can integrate this motion to derive the camera's path. One of the factors that make this approach attractive is that no specific features need to be detected, unlike the previous approach. The weakness of ego-motion integration comes from the fact that small errors are accumulated during the integration process. Hence, the estimated camera's path is drifted and the pose estimation accuracy decrease along time. If such approach is used it would be desirable to reduce the drift by activating, once in a while, an additional algorithm that estimates the pose directly. Such navigation-system is being suggested. In that work, like in this work, the drift is being corrected using a Digital Terrain Map (DTM). The DTM is a discrete representation of the observed ground's topography. It contains the altitude over the sea level of the terrain for each geographical location. A patch from the ground was reconstructed using 'structure-from-motion' (SFM) algorithm and was matched to the DTM in order to derive the camera's pose. Using SFM algorithm which does not make any use of the information obtained from the DTM but rather bases its estimate on the flow-field alone, positions their technique under the same critique that applies for SFM algorithms.

The algorithm presented in this work does not require an intermediate explicit reconstruction of the 3D world. By combining the DTM information directly with the images information it is claimed that the algorithm is well-conditioned and generates accurate estimates for reasonable scenarios and error sources. In the present work this claim is explored by performing an error analysis on the algorithm outlined above. By assuming appropriate characterization of these error sources, a closed form expression for the uncertainty of the pose and motion of the camera is first developed and then the influence of different factors is studied using extensive numerical simulations.

## 2 The Navigation Algorithm

The problem can be briefly described as follows: At any given time instance  $t$ , a coordinates system  $C(t)$  is fixed to a camera in such a way that the  $Z$ -axis coincides with the optical-axis and the origin coincides with the camera's projection center. At that time instance the camera is located at some geographical location  $p(t)$  and has a given orientation  $R(t)$  with respect to a global coordinates system  $W$  ( $p(t)$  is a 3D vector,  $R(t)$  is an orthonormal rotation matrix).  $p(t)$  and  $R(t)$  define the transformation from the camera's frame  $C(t)$  to the world's frame  $W$ ,

where if  ${}^c v$  and  ${}^w v$  are vectors in  $C(t)$  and  $W$  respectively, then  ${}^w v = R(t) {}^c v + p(t)$ .

Consider now two sequential time instances  $t_1$  and  $t_2$ : the transformation from  $C(t_1)$  to  $C(t_2)$  is given by the translation vector  $\Delta p(t_1, t_2)$  and the rotation matrix  $\Delta R(t_1, t_2)$ , such that  ${}^{c(t_2)} v = \Delta R(t_1, t_2) {}^{c(t_1)} v + \Delta p(t_1, t_2)$ . A rough estimate of the camera's pose at  $t_1$  and of the ego-motion between the two time instances -  $p_E(t_1)$ ,  $R_E(t_1)$ ,  $\Delta p_E(t_1, t_2)$  and  $\Delta R_E(t_1, t_2)$  - are supplied (the subscript letter ``E'' denotes that this is an estimated quantity).

Also supplied is the optical-flow field:  $\{u_i(t_k)\}$  ( $i=1..n, k=1,2$ ). For the  $i$ 'th feature,  $u_i(t_1) \in \mathbb{R}^2$  and  $u_i(t_2) \in \mathbb{R}^2$  represent its locations at the first and second frame respectively.

Using the above notations, the objective of the proposed algorithm is to estimate the true camera's pose and ego-motion:  $p(t_1)$ ,  $R(t_1)$ ,  $\Delta p(t_1, t_2)$  and  $\Delta R(t_1, t_2)$ , using the optical-flow field  $\{u_i(t_k)\}$ , the DTM and the initial-guess:  $p_E(t_1)$ ,  $R_E(t_1)$ ,  $\Delta p_E(t_1, t_2)$  and  $\Delta R_E(t_1, t_2)$ . (Fig 1.)

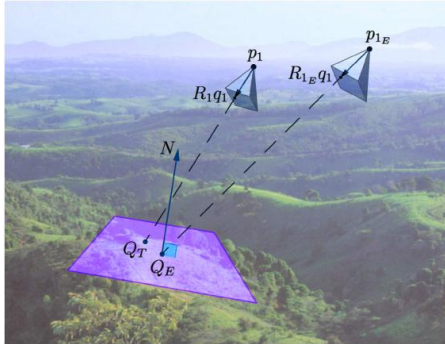


Figure 1: Geometrical description

### 3. Results of numerical simulation for real parameters of flight and camera.

The typical results of numerical simulations can be seen on (Fig.2) for different cases of flight, camera and simulation parameters.

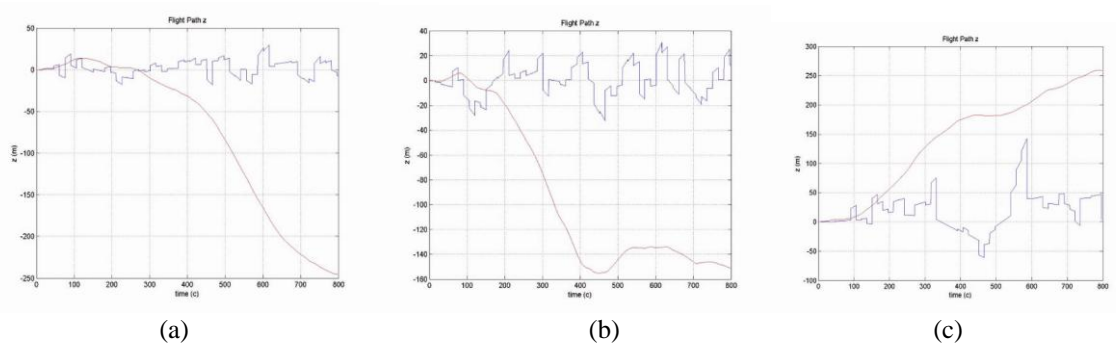


Figure 4: Position errors for x, y, z coordinate of the drift path are marked with a red line, and errors of the corrected path are marked with a blue line. Parameters : FOV 60 degree, Features number 120, Resolution 1000x1000, Baseline=200m,  $\Delta time = 15$  s, Height a) 700m b) 1000m c) 3000m

### 4. Conclusions

An algorithm for pose and motion estimation using corresponding features in images and a DTM was presented with using Kalman filter. The DTM served as a global reference and its data was used for recovering the absolute position and orientation of the camera. In numerical simulations position and velocity estimates were found to be sufficiently accurate in order to bound the accumulated errors and to prevent trajectory drifts.

An error analysis has been performed for a novel algorithm that uses as input the optical flow derived from two

consecutive frames and a DTM. The position, orientation and ego-motion parameters of the camera can be estimated by the proposed algorithm. The main source for errors were identified to be the optical-flow computation, the quality of the information about the terrain, the structure of the observed terrain and the trajectory of the camera. A closed form expression for the uncertainty of the pose and motion was developed. Extensive numerical simulations were performed to study the influence of the above factors.

Tested under reasonable and common scenarios, the algorithm behaved robustly even when confronted with relatively noisy and challenging environment. Following the analysis, it is concluded that the proposed algorithm can be effectively used as part of a navigation system of autonomous vehicles.

On basis results of numerical simulation for real parameters of flight and camera we also can conclude follow:

1) The most important parameter of simulations is FOV: for the small FOV the method diverges. For FOV 60 degree the results are very good. The reason for this is that for small FOV (12 or 6 degree) the situation is close to degenerated state, also we must choose small baseline and observed ground patch is too small and almost flat.

2) Resolution of camera is also very important parameter: for better resolution we have much more better results, because of much more better precision of features detection.

3) The precision of new method depends on flight height. Initially precision increases with height increasing because we can use bigger baseline and can see bigger patch of ground. But for bigger heights precision begin to decrease because of small parallax effect.

### **Acknowledgment**

We would like to thank Ronen Lerner, Ehud Rivlin and Hector Rotstein for very useful consultations.

# КОРРЕКЦИЯ ОШИБОК ИНЕРЦИАЛЬНОЙ НАВИГАЦИОННОЙ СИСТЕМЫ С ПОМОЩЬЮ МЕТОДОВ ВИЗУАЛЬНОЙ НАВИГАЦИИ

**В.В.Воронов**

*ЗАО «Р.Е.Т.Кронштадт», Россия*

107076, Россия, Москва,

Ул. Стромынка, д.18

Тел. +7(495)748 3584

[voronov@kronshtadt.ru](mailto:voronov@kronshtadt.ru)

**О.Ю.Купервассер, к.ф.-м.н.**

*ЗАО «Р.Е.Т.Кронштадт», Россия*

107076, Россия, Москва,

Ул. Стромынка, д.18

Тел. +7(495)748 3584

[olegkup@yahoo.com](mailto:olegkup@yahoo.com)

Предложен алгоритм для нахождения позиции, ориентации и оценки движения, использующий соответствующие точки в изображениях и цифровую карту ландшафта. Использование Цифровой Карты Ландшафта (ЦКЛ) как глобальной справочной информации позволяет восстановление абсолютной позиции и ориентации камеры. Чтобы сделать это, ЦКЛ используется, чтобы сформулировать дополнительные ограничения между соответствующими точками в двух последовательных кадрах. Использование этих данных позволяет улучшить надежность и точность инерциального навигационного алгоритма. Расширенный фильтр Калмана использовался, чтобы объединить результаты инерциального навигационного алгоритма и навигационного алгоритма, основанного на компьютерном зрении. Выполнимость этого алгоритма продемонстрирована путем численного моделирования.

## 1. Введение

Алгоритмы визуальной навигации являются одним из основных объектов исследований в течении последнего десятилетия. Бурное развитие роботизированных систем потребовало выработки специфических подходов для определения в реальном времени координат местоположения и параметров ориентации подвижных объектов, как в закрытых помещениях, так и на улице. Одним из важных требований к современным беспилотным авиационным комплексам является возможность осуществления точной навигации в условиях отсутствия спутниковой информации. Интеграция визуальных методов навигации с традиционными инерциальными системами может обеспечить устойчивое решение навигационной задачи для наземных и летающих автоматических систем.

Существуют три основных подхода к решению задачи определения параметров движения объекта по визуальной информации.

1) Определение параметров движения и ориентации путем отслеживания перемещения в кадре изображений неких характерных (особых) точек местности. После идентификации особых точек на каждом снимке точное определение параметров движения представляет собой достаточно тривиальную задачу. Наиболее сложной проблемой остаётся детектирование особых точек и их соотнесение с набором характерных признаков данной местности, определенным заранее.

2) При подходе “ego motion” оценивается относительное движение камеры по отношению к предыдущему ее положению. Источником данных являются видеокдры. Оценив параметры движения и зная начальные условия, можно путём интегрирования вычислить траекторию движения камеры. Важным достоинством этого метода (в отличие от предыдущего) является отсутствие необходимости выявления особых точек на местности. Проблемной стороной является нарастание ошибки в процессе интегрирования. Таким образом, имеется аналог «дрейфа» оценки местоположения, и точность оценивания деградирует с течением времени (или с ростом пройденного пути).

Два изложенных подхода объединяет необходимость анализа последовательных изображений.

3) Третий метод позволяет компенсировать упомянутый выше эффект «дрейфа», применяя дополнительный алгоритм прямого оценивания местоположения. В представляемой работе используются данные цифрового рельефа местности (ЦРМ). ЦРМ представляет собой цифровое представление топографии местности через матрицу высот над уровнем моря для дискретных координат местоположения. Поставленная задача известна под названием извлечение структуры из движения (structure from motion - SfM). Восстановленный методами SfM рельеф может быть соотнесен с данными ЦРМ и, таким образом, определяется местоположение объекта относительно ЦРМ. В представляемой работе предлагается непосредственно комплексировать данные ЦРМ и видовые данные, полученные с камеры объекта для оценки положения и ориентации объекта, минуя стадию восстановления рельефа по видеоданным. В результате сокращается объём вычислений, и повышается устойчивость алгоритма.

Таким образом, имеются несколько решений навигационной задачи через методы видовой обработки данных. Каждый из предложенных методов базируется на уникальной априорной информации и даёт независимое решение с той или иной степенью точности. Любое из этих решений, как и все вместе, может использоваться для коррекции инерциальной навигационной системы с применением стандартной техники интегрирования данных посредством Калмановской фильтрации.

В представляемой работе изложены алгоритмы получения навигационных данных из видовой информации с использованием ЦРМ, рассмотрены вопросы комплексирования видовых данных с данными инерциальной системы, приводятся результаты компьютерного моделирования.

## 2. Навигационный алгоритм.

Проблема может быть кратко описана следующим образом: В любой момент времени  $t$ , система координат  $C(t)$  установлена на камеру таким способом, что  $Z$ -ось совпадает с оптической осью, и начало координат совпадает с центром проектирования камеры. В этот момент времени камера расположена в некотором географическом местоположении  $p(t)$  и имеет данную ориентацию  $R(t)$  относительно глобальной координатной системы  $W$  ( $p(t)$  - трехмерный вектор,  $R(t)$  - ортонормальная матрица вращения).  $p(t)$  и  $R(t)$  определяют преобразование из координатной системы камеры  $C(t)$  в глобальную координатную систему  $W$ , где, если  ${}^c v$  и  ${}^w v$  являются векторами в  $C(t)$  и  $W$  соответственно, то  ${}^w v = R(t) {}^c v + p(t)$ .

Рассмотрим теперь два последовательных момента времени  $t_1$  и  $t_2$ : преобразование из  $C(t_1)$  в  $C(t_2)$  дано вектором сдвига  $\Delta p(t_1, t_2)$  и матрицей вращения  $\Delta R(t_1, t_2)$  таким образом, что  ${}^{c(t_2)} v = \Delta R(t_1, t_2) {}^{c(t_1)} v + \Delta p(t_1, t_2)$ . Для грубой оценки позы камеры в  $t_1$  и собственное движение камеры между двумя моментами времени -  $p_E(t_1)$ ,  $R_E(t_1)$ ,  $\Delta p_E(t_1, t_2)$  и  $\Delta R_E(t_1, t_2)$  используются. (Символ "E" для нижнего индекса обозначает, что это оценочная (estimated) величина.)

Также используется поле оптического потока:  $\{u_i(t_k)\}$  ( $i=1..n, k=1,2$ ). Для  $i$ 'th характерная точка местности,  $u_i(t_1) \in \mathbb{R}^2$  и  $u_i(t_2) \in \mathbb{R}^2$  представляют её местоположение на первом и втором кадре соответственно.

Используя вышеупомянутые обозначения, цель предложенного алгоритма состоит в том, чтобы оценить истинную позу камеры и её собственное движение:  $p(t_1)$ ,  $R(t_1)$ ,  $\Delta p(t_1, t_2)$  и  $\Delta R(t_1, t_2)$  используя поле оптического потока  $\{u_i(t_k)\}$ , ЦКЛ и приблизительные начальные условия:  $p_E(t_1)$ ,  $R_E(t_1)$ ,  $\Delta p_E(t_1, t_2)$  и  $\Delta R_E(t_1, t_2)$  (Рис. 1)

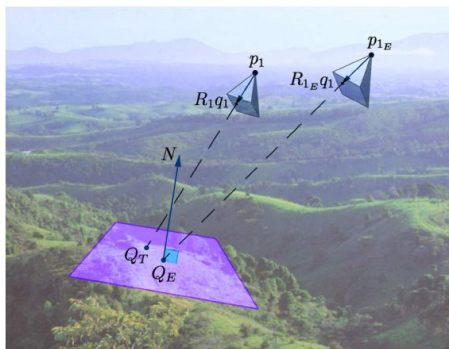


Рис. 1: Геометрическое описание

Типичные результаты числовых моделирований приведены на (Рис. 2) для различных случаев полета, параметров камеры и параметров моделирования.

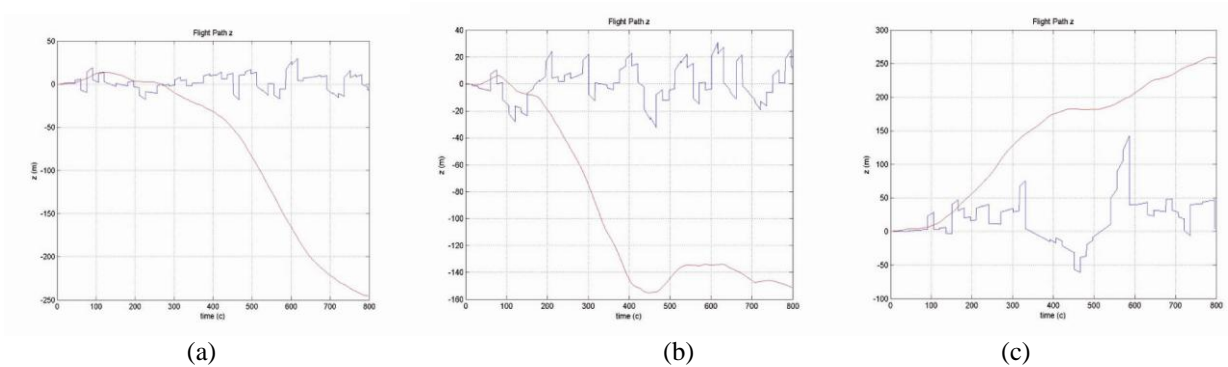


Рис.2: Ошибки позиции  $x$ ,  $y$ ,  $z$  для координат инерциального дрейфа отмечены красной линией, и ошибки исправленного пути отмечены с синей линией. Параметры: FOV  $60^\circ$ , Число характерных точек 120, Разрешение  $1000 \times 1000$ , Расстояние между двумя положениями камер для прведения видеонаблюдения  $=200\text{m}$ , время меду наблюдениями 15 сек, высота а) 700 м. б) 1000 м. с) 3000 м.

### 3. Заключение

Был представлен алгоритм для определения положения и ориентации камеры и оценки движения, использующий характерные точки в изображениях и ЦКЛ, использующий фильтра Калмана. ЦКЛ служит глобальным ориентиром, и ее данные использовались для того, чтобы получить абсолютное положение и ориентацию камеры. Численное моделирование достаточно точно, чтобы предотвратить накопление ошибки и предотвратить дрейф траектории.

Анализ погрешностей был выполнен для нового алгоритма, который использует как оптический поток, полученный из двух последовательных изображений, так и ЦКЛ. Положение, ориентация и параметры эго-движения камера могут быть оценены предложенным алгоритмом. Главные источники ошибок был идентифицированы при вычислении оптического потока - качество информации о ландшафте, структура наблюдаемого ландшафта и траектория камеры. Было разработан анализ погрешностей. Обширные числовые моделирования были выполнены, чтобы изучить влияние вышеупомянутых коэффициентов.

Проверенный согласно разумным и общим сценариям, алгоритм вел себя сильно даже когда сталкивается с относительно шумной и стимулирующей окружающей средой. После анализа приходим к заключению, что предложенный алгоритм может эффективно использоваться как часть навигационной системы автономных транспортных средств.

На результатах числового моделирования для дреальных параметров полета и камеры мы также можем заключить, следующее:

1) Самый важный параметр моделирований – угол зрения камеры (FOV): для небольшого FOV метод расходится. Для FOV  $60^\circ$  результаты очень хороши. Причина для этого - то, что для небольшого FOV ( $12^\circ$  или  $6^\circ$ ) ситуация близка к вырожденной : мы должны выбрать небольшое смещение между камерами, наблюдаемый участок основания является слишком небольшим и почти плоским.

2) Разрешение камеры - также очень важный параметр: для лучшего разрешения у нас есть намного больше лучших результатов, из-за намного большей точности обнаружения особенностей.

3) Точность нового метода зависит от высоты полета. Первоначально точность увеличивается с увеличением высоты, потому что мы можем использовать большие данные и можем видеть больший участок ландшафта. Но на больших высотах точность начинает уменьшаться из-за эффекта параллакса.

### Благодарность

Мы бы хотели поблагодарить Ронена Лернера, Эхуда Ривлина и Хектора Ротштейна за очень полезные консультации.