

## Определение параметров движения объекта по данным стереокамеры нейростатистическими методами

М.Ю. Артюхов<sup>1</sup>, Н.П. Деменков<sup>1</sup>

<sup>1</sup> МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

*Предложена методика оценки параметров движения подводного аппарата на основе изображений, поступающих со стереопары, с применением вычислительной аппаратуры низкой производительности и с упрощенной процедурой калибровки видеокамер на основе нейронных сетей и генетических алгоритмов.*

**E-mail:** demenkov@iu1.bmstu.ru

**Ключевые слова:** видеонавигация, движение, оценка, параметры, стереопара, видеокамера, изображение, калибровка, нейронная сеть, генетический алгоритм.

Рассматриваемая система видеонавигации состоит из двух видеокамер, закрепленных на подводном аппарате на сферических шарнирах, расстояние между которыми жестко фиксировано. Камеры находятся в герметичных боксах. Система решает задачу определения параметров движения (линейных смещений, скоростей, угловых перемещений) на основании изображений, поступающих с камер.

Достигнутые результаты в области обработки изображений значительно упростили жизнь человека: из средства ретуширования фотографий обработка изображений превратилась в огромную индустрию. Однако, несмотря на заметные успехи в этой области, существует острая потребность в оптимизации имеющихся и реализации новых алгоритмов для применения их в условиях, при которых использование других средств либо слишком дорого, либо не обеспечивает необходимой точности.

Одной из таких областей является навигация под водой. Значительное количество средств, применяемых в подводной среде в целях навигации, характеризуется низкой точностью и большими уходами с течением времени при их довольно высокой стоимости. В качестве примера широко применяемых под водой систем можно назвать системы, основанные на работе ультразвуковых устройств. У теленавигационных систем отсутствует уход с течением времени, они являются относительно дешевым средством измерения, поэтому их разработку и дальнейшее развитие можно рассматривать в качестве важной задачи. Следует также отметить, что место на подводном аппарате ограничено, а тепловыделение должно быть очень низким. Поэтому приходится рассчитывать лишь на маломощные вычислите-

ли и максимально оптимизировать алгоритмы обработки для соответствия данным требованиям. Основной проблемой систем, основанных на анализе изображений, является большая вероятность их малой робастности, в связи с чем требуется обеспечение максимальной сходимости алгоритма оценки параметров движения.

**Специфика определения параметров движения объектов на основании изображений, поступающих с видеоканалов.** Для определения параметров движения подводного аппарата необходимо оценить смещения кадров относительно друг друга. Все современные алгоритмы слежения, за некоторым исключением, опираются на работу [1].

Оригинальный алгоритм Лукаса — Канаде основан на функционале

$$E = \sum [F(x+h) - G(x)]^2, \quad (1)$$

где  $F(x)$ ,  $G(x)$  — два кадра;  $h$  — смещение.

Функцию  $F(x)$  с помощью разложения в ряд Тейлора можно приближенно представить в виде

$$F(x+h) \approx F(x) + h \frac{\partial}{\partial x} F(x), \quad (2)$$

где  $\frac{\partial}{\partial x} = \left[ \frac{\partial}{\partial x_1} \quad \frac{\partial}{\partial x_2} \quad \dots \quad \frac{\partial}{\partial x_n} \right]^T$  — градиент.

Используя приближение (2), минимум функционала  $E$  ищут путем дифференцирования и приравнивания производной к нулю, а смещение  $h$  можно получить из выражения

$$h = \left[ \sum_x \left( \frac{\partial F}{\partial x} \right)^T [G(x) - F(x)] \right] \left[ \sum_x \left( \frac{\partial F}{\partial x} \right)^T \left( \frac{\partial F}{\partial x} \right) \right]^{-1}. \quad (3)$$

В дальнейшем математическая формулировка алгоритма (1), (3) была изменена и стала основой для всех последующих обобщений с учетом аффинных искажений окрестности и освещенности. Путем замены соответствующих переменных на константы любой из используемых алгоритмов превращается в обычный алгоритм Лукаса — Канаде.

Перспективными являются также алгоритмы, оценивающие смещение изображений на основе информации о выделенных сегментах в кадрах. Изображение разбивают на сегменты, которые некоторым образом описывают, после чего похожие сегменты ищут на втором изображении. Описание сегментов проводят с помощью преобразования Фурье на основании таких характеристик, как площадь, центр тяжести, периметр, компактность сегмента, центральные моменты. Тем не менее для работы этих алгоритмов необходима обработка все-

го изображения, в то время как для снижения вычислительной нагрузки желательнее получать оценки смещений лишь некоторых областей.

Таким образом, разрабатываемый алгоритм должен работать быстро и только с некоторыми частями изображения, а кроме того, он должен быть робастным. При работе с качественными изображениями робастность алгоритма обеспечивается калибровкой камеры. Все используемые в настоящее время методы калибровки оценивают некоторые виды нелинейности. Однако не исключено, что в каждой конкретной камере будет нелинейность, присущая только ей одной. Для решения данной проблемы можно использовать настроенную нейронную сеть с радиально-базисными функциями (РБФ-сеть) [2].

**Алгоритм оценки смещений изображений и оценки параметров перемещения объектов на основе анализа данных смещений.** Для оценки смещений изображений алгоритм использует похожий на предложенный Лукасом — Канаде функционал. Он работает с ограниченными областями изображения (довольно малыми), выбираемыми на основании функционала схожести, в качестве которого используется функционал вида

$$F(a, b) = \sum_{x-n}^{x+n} \sum_{y-m}^{y+m} f(x, y, a, b); \quad (4)$$
$$f(x, y, a, b) = | \text{Video}(x + a, y + b) - \text{Sample}(x, y) |,$$

где  $\text{Video}(x, y)$  — яркость изображения в точке  $(x, y)$ ;  $\text{Sample}(x, y)$  — яркость эталонного изображения в точке  $(x, y)$ ;  $m, n = \text{csize}/2$  — половина размера ячейки по соответствующей координате, пиксели.

Функционал схожести (4) имеет вид конуса, является модификацией функционала Лукаса — Канаде и вычисляется вдоль ячеек размером  $\text{csize} \times \text{csize}$  пикселей. Чтобы выяснить, насколько смещены изображения относительно друг друга, находят минимум функционала (4), т. е. вершину конуса.

В целях уменьшения количества расчетов необходимо минимизировать обращение к функционалу (4), поэтому для поиска минимума был предложен следующий алгоритм [3]:

- 1) определяют точку начального приближения координаты вершины конуса  $x(i) = x00$ ,  $i = 1$  — номер итерации;
- 2) вычисляют функционал схожести  $Z(x)$  в точках  $x(i)$ ,  $x(i) + rd$ ,  $x(i) - rd$ , где  $rd$  — положительное целое число, параметр алгоритма, зависящий от области определения  $[m; n]$  функционала схожести;
- 3) определяют вспомогательную величину

$$SL(i) = [Z(x(i) + rd) - Z(x(i) - rd)];$$

- 4) по знаку  $SL(i)$  находят направление на вершину по отношению к  $x(i)$  и определяют наклон образующей конуса

$$\text{Den}(i) = \begin{cases} Z(x(i) - rd) - Z(x(i)), & SL(i) > 0; \\ Z(x(i) + rd) - Z(x(i)), & SL(i) \leq 0; \end{cases}$$

5) подсчитывают коррекцию координаты вершины

$$dx(i) = \frac{SL(i)rd}{2\text{Den}(i)},$$

модуль которой подвергается ограничению

$$|dx(i)| = \min(|dx(i)|, rd);$$

6) вычисляют новое приближение координаты вершины конуса

$$x(i+1) = x(i) + dx(i);$$

7) проверяют условие нахождения крайних точек на разных склонах конуса (условие завершения итераций)

$$SL(i) < \text{Den}(i);$$

8) если условие п. 7 не выполняется, то переходят к п. 2;

9) конец итераций

$$x(i+1) \cong x_0.$$

**Выделение ячеек на сопровождение.** Для выбора ячеек на сопровождение можно использовать автокорреляционную функцию или функционал самоподобия

$$Fs(a, b) = \sum_{x=n}^{x+n} \sum_{y=m}^{y+m} |\text{Sample}(x+a, y+b) - \text{Sample}(x, y)|. \quad (5)$$

Использование автокорреляционной функции не представляется возможным вследствие предъявляемых требований к объему памяти и быстродействию машины. Функционал самоподобия при отсутствии шумов совпадает с функционалом схожести для случая нулевого смещения. В точке нулевого смещения функционал самоподобия всегда равен нулю. Таким образом, наклон функционала самоподобия вблизи главного максимума может быть определен как его значение в точке с единичным смещением. Наклон будет увеличиваться с ростом полезного сигнала. Однако при наличии шумов даже в случае отсутствия полезного сигнала наклон функционала самоподобия будет увеличиваться с ростом амплитуды шума, достигая максимума для белого шума. Это противоречие не позволяет по измерению наклона функционала самоподобия определить, какое изображение содержит ячейка — полезный сигнал или шум.

На практике интенсивность шума обычно слабо изменяется по полю изображения, что позволяет выделить ячейки с самым сильным

полезным сигналом как ячейки с максимальным наклоном функционала (5). После выделения ячеек осуществляется поиск вершины конуса функционала (4), и если в ячейке содержится не шум, то вершина будет найдена, в противном случае ячейка отбрасывается. Дальнейшее отсеивание шумовой информации выполняется в процедуре поиска параметров движения.

**Оценка параметров движения на основе найденных смещений.** Для быстрой оценки параметров движения используется линейная процедура. Лишь в случае отсутствия возможности ее использовать применяется более ресурсоемкая нелинейная оценка.

Для повышения точности оценки параметров движения рассматривают три гипотезы движения аппарата в плоскости: 1) перемещение и одновременный поворот; 2) только поворот; 3) только смещение камеры.

Для выявления наиболее «сложного» случая поворота и одновременного смещения для каждой из характеристических точек можно записать систему вида

$$\begin{cases} x'_i = (\Delta x + x_i \cos \vartheta + y_i \sin \vartheta) \frac{Z + \Delta Z}{Z}; \\ y'_i = (\Delta y + x_i \sin \vartheta - y_i \cos \vartheta) \frac{Z + \Delta Z}{Z}, \end{cases}$$

где  $\Delta x, \Delta y$  — смещение центра кадра по осям  $Ox, Oy$  в системе координат текущего кадра;  $x_i, y_i$  — координаты контрольной точки на предыдущем кадре;  $\frac{Z + \Delta Z}{Z}$  — изменение масштабного коэффициента,

связанное с изменением глубины;  $x'_i, y'_i$  — измененные координаты характеристической точки. Данную систему линеаризуют с учетом того, что угол поворота от кадра к кадру остается малым:

$$\begin{cases} x'_i - Kx_i = (\Delta x + 0\Delta y + y_i\vartheta)K; \\ y'_i + Ky_i = (0\Delta x + \Delta y + x_i\vartheta)K, \quad K = \frac{Z + \Delta Z}{Z}. \end{cases} \quad (6)$$

В результате получим переопределенную систему линейных алгебраических уравнений с неизвестными  $X = [\Delta x, \Delta y, \vartheta]^T$ , для решения которой можно воспользоваться методом наименьших квадратов:

$$X = (A^T A)^{-1} (A^T B); \quad A = K \begin{pmatrix} \dots & & \\ 1 & 0 & y_i \\ 0 & 1 & x_i \\ \dots & & \end{pmatrix}; \quad B = \begin{bmatrix} \dots & & \\ x'_i - Kx_i \\ y'_i + Ky_i \\ \dots & & \end{bmatrix}.$$

Далее каждую из характеристических точек проверяют на соответствие гипотезе одновременного поворота и смещения: характеристические точки  $(x_i, y_i)$  поворачивают на найденный угол  $\vartheta$  относительно центра экрана и смещают на расстояние  $(\Delta x, \Delta y)$ .

Затем вычисляют математическое ожидание и дисперсию остатка посчитанных точек до конечных точек  $(x'_i, y'_i)$ . Фильтрацию точек проводят по критерию непревышения заданного порога от средней величины и дисперсии. Причем точки разбивают на три группы: хорошие, средние, плохие. Каждой из этих групп ставят в соответствие определенное значение порога по среднему отклонению и дисперсии. После такой простой классификации точек на основании наличия заданного числа «хороших» точек по ним решают уточненную систему уравнений (6).

Если же «хороших» точек оказалось мало, то из «хороших» и «средних» снова решают систему уравнений (6) и используют результат как начальное приближение для относительно сложной процедуры при нелинейной оценке параметров. Если «средних» точек также не набралось достаточное количество, то начальное смещение и поворот дополнительно оценивают на основе фундаментальной матрицы, а затем используют нелинейную процедуру.

Если «хороших» точек оказалось много (больше трех), то линейной процедуре оценки можно доверять и остается проверить две оставшиеся гипотезы. Разумно проверять гипотезу на чистый поворот, когда угол, найденный из системы уравнений (6), превышает заданный малый порог, и гипотезу на чистое смещение — в обратном случае. Чистый поворот получается путем обнуления членов  $\Delta x, \Delta y$  из системы (6), которая теперь становится переопределенной с одним неизвестным — углом поворота. Затем каждую точку особенности на предыдущем кадре поворачивают на найденный угол и находят первые два статистических момента. В качестве меры статистического теста используется функция

$$R_{\text{test}} = \bar{X}\sigma_x + \bar{Y}\sigma_y,$$

где  $\bar{X}, \bar{Y}$  — математическое ожидание по  $x$  и  $y$  соответственно,  $\sigma_x, \sigma_y$  — средние квадратичные отклонения по  $x, y$  соответственно.

Гипотеза с меньшим значением функции  $R_{\text{test}}$  считается верной.

Подобную процедуру выполняют и в случае чистого смещения. Для его определения достаточно взять среднее отклонение точек одной от другой на опорном и текущем кадрах.

**Оценка поворота и смещений по фундаментальной матрице.** На основе посчитанных смещений кадров можно определить начальное приближение параметров движения подводного аппарата с помощью фундаментальной матрицы. Пусть камера располагается в

двух разных положениях пространства. Есть некоторая точка  $M$  в пространстве: нам известны экранные координаты данной точки в проекционной плоскости каждого из положений камер (пусть это будут  $m, m'$ ). Тогда фундаментальная матрица  $F$  содержит в себе информацию о преобразовании двух изображений камер и удовлетворяет следующему уравнению:

$$m^T F m' = 0.$$

Таким образом, можно оценкой фундаментальной матрицы получить параметры объектов в пространстве и/или смещение и поворот кадров относительно друг друга.

Для оценки фундаментальной матрицы  $F$  необходимо выделить некоторые точки на исходном изображении, затем по смещению изображений — те же точки на измененном изображении. Точек должно быть как минимум восемь. Все точки нужно нормализовать, иначе их значения будут слишком большими и сильно отличаться друг от друга, вследствие чего оценка фундаментальной матрицы будет весьма неточной. Для нормализации группы точек необходимо найти такую матрицу преобразования, при подстановке которой среднее значение выборки станет равным нулю, а среднеквадратическое отклонение равным  $\sqrt{2}$ . Фундаментальная матрица должна быть вырождена, т. е. ее ранг должен быть равен 2. На практике последнее собственное число несколько отличается от нуля, поэтому его обнуляют, т. е.

$$\begin{aligned} F = USV^T &\Rightarrow S = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}) \Rightarrow \\ &\Rightarrow \hat{S} = \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, 0) \Rightarrow \hat{F} = U\hat{S}V^T. \end{aligned}$$

После этого матрицу денормализуют и используют для нахождения матриц поворота  $R$  и перемещений  $T$ :

$$\begin{cases} R = UR_Z \left( \pm \frac{\pi}{2} \right)^T V^T; \\ T = VR_Z \left( \pm \frac{\pi}{2} \right)^T V^T. \end{cases}$$

**Нелинейная процедура оценки параметров.** В целях улучшения точности полученные матрицы вращения и перемещения используют как начальные приближения. Для более точных оценок можно воспользоваться эпиполярным уравнением

$$|h, \mathbf{m}_\alpha, R\mathbf{m}'_\alpha| = 0.$$

Здесь  $\{\mathbf{m}_\alpha\}$  и  $\{\mathbf{m}'_\alpha\}, \alpha = 1, \dots, N$ , — единичные векторы-направления ( $N$ -векторы) на точки проекции исходного и измененного изображений.

Матрицы поворота и перемещения находят из соотношения

$$\sum_{\alpha=1}^N W_\alpha |h, \mathbf{m}_\alpha, R\mathbf{m}'_\alpha|^2 \rightarrow \min. \quad (7)$$

Выражение (7) может быть модифицировано для определения только поворота. По двум наборам  $N$ -векторов находят матрицу поворота  $R$ , минимизирующую наименьшее собственное значение матрицы:

$$A(R) = \sum_{\alpha=1}^N W_\alpha (\mathbf{m}_\alpha \times R\mathbf{m}'_\alpha)(\mathbf{m}_\alpha \times R\mathbf{m}'_\alpha)^T. \quad (8)$$

Далее определяют вектор  $h$  согласно выражению (7).

Решение нелинейной задачи (8) осуществляют численными методами. Для выполнения итераций необходим градиент целевой функции — наименьшего значения собственного числа. Для его вычисления используется теорема об изменении матрицы. Прежде всего, поворот представляем с помощью кватерниона — четырех параметров. Начальное приближение матрицы поворота находим посредством решения фундаментальной матрицы либо путем линейной процедуры оценки. Из матрицы поворота определяем значения кватерниона. Матрица  $A(R)$  есть функция данного кватерниона (как и матрица  $R$ ). Наименьшее значение собственного числа  $\lambda_m$  есть также функция кватерниона. Из теоремы об изменении матрицы следует

$$\frac{\partial \lambda_m}{\partial q_k} = (l_m, T_k l_m), \quad k = 0, 1, 2, 3, \quad (9)$$

где  $l_m$  — собственный вектор матрицы  $A(R)$ , соответствующий  $\lambda_m$ ;

$$T_k = \sum_{\alpha=1}^N W_\alpha [(\mathbf{m}_\alpha \times D_k \mathbf{m}'_\alpha)(\mathbf{m}_\alpha \times R\mathbf{m}'_\alpha)^T + (\mathbf{m}_\alpha \times R\mathbf{m}'_\alpha)(\mathbf{m}_\alpha \times D_k \mathbf{m}'_\alpha)^T];$$

$$D_k = \frac{\partial R}{\partial q_k}, \quad k = 0, 1, 2, 3.$$

Здесь очень важно получить хорошее начальное приближение, иначе функция (9) может попасть в локальный минимум. Именно поэтому следует использовать предварительно оцененный кватернион из фундаментальной матрицы.



В качестве весовой матрицы используют выражение

$$W_{\alpha} = \frac{\text{const}}{\text{tr}V[m_{\alpha}]\text{tr}V[m'_{\alpha}]},$$

где  $E[\|\Delta m_{\alpha}\|] = \text{tr}V[m_{\alpha}]$ ;  $E[\|\Delta m'_{\alpha}\|] = \text{tr}V[m'_{\alpha}]$ .

Константу выбирают из условия  $\sum_{\alpha=1}^N W_{\alpha} = 1$ .

**Алгоритм для проведения процедуры упрощенной калибровки камер на основе аппарата РБФ-сетей и генетических алгоритмов.** В реальных камерах присутствуют искажения, вызванные несовершенством линз. Устранение неточностей камер осуществляли с помощью РБФ-сети, на вход которой подавали искаженные локальные координаты точки, а с выхода РБФ-сети снимали неискаженные координаты.

В качестве активизационной функции использована функция Гаусса вида

$$f(\bar{X}, \bar{c}) = \exp\left(-\sum_{j=1}^m \frac{[X_j - c_j]^2}{\sigma_j^2}\right), \quad (10)$$

где  $c_j$  —  $j$ -я координата центра активизационной функции нейрона;  $\sigma_j$  — ширина окна активизационной функции  $j$ -го нейрона;  $X_j$  —  $j$ -й вход нейрона;  $m$  — количество нейронов скрытого слоя.

Для точного определения структуры РБФ-сети необходимо вычислить весовые коэффициенты выхода нейронной сети  $\{w_j\}_{j=1}^m$  и параметры активизационной функции Гаусса  $\{(\sigma_j, c_j)\}_{j=1}^m$ .

Весовые коэффициенты находили из решения уравнения

$$\phi w^T = D \Leftrightarrow w^T = \phi^{-1} D,$$

в котором  $\phi = \{f_{ij}\}$ ,  $i = 1, \dots, H$ ;  $j = 1, \dots, H$  — интерполяционная матрица;  $D = \{D_{ij}\}$ ,  $i = 1, \dots, H$ ;  $j = 1, \dots, H$  — матрица выходных шаблонов.

Для нахождения параметров активизационной функции (10) использован генетический алгоритм (ГА).

Согласно стратегии ГА, первый набор (популяция, поколение) решений — кандидатов системы оптимизации создается случайно. Решения представлены в форме строк (хромосом), состоящих из последовательности двоичных разрядов. Каждая битовая строка разделена на сегменты, число которых равно числу оптимизируемых параметров (ширина и центр активизационной функции). Сконструированная таким образом двоичная строка существенно расширяет

возможные решения в многомерном пространстве поиска. Качество каждого решения в популяции зависит от используемого предопределенного критерия годности. После оценки годности всех хромосом в популяции создается новое поколение особей (индивидуумов) с использованием трех генетических операторов: воспроизводства, перехода и мутации.

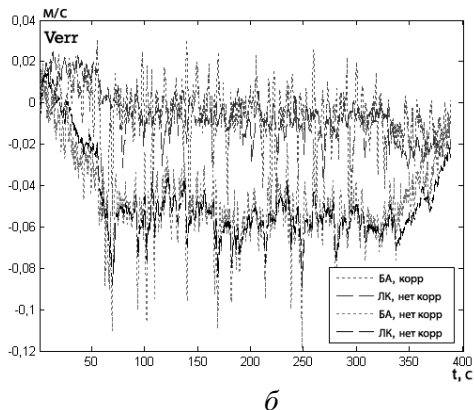
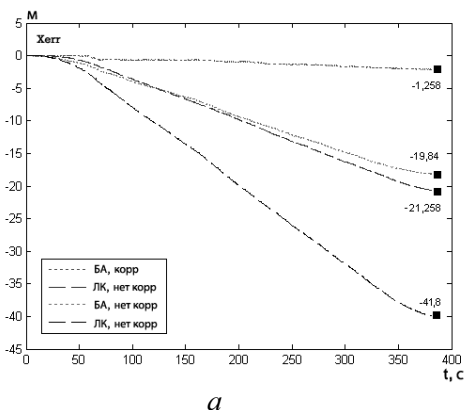
Новая популяция, создаваемая таким образом, подвергается воздействию трех операторов циклично на протяжении многих поколений до тех пор, пока предопределенный критерий остановки — ГА используют в качестве него фиксированное число поколений — не будет выполнен.

Неудобное свойство ГА состоит в том, что требуется точно определить функциональную форму, параметры/переменные которой должны оптимизироваться. Во многих ситуациях на практике, особенно при рассмотрении нелинейных зависимостей, оценка подходящей модели функции годности становится затруднительной. В этом контексте остро необходима методология, приводящая к приемлемой форме функции соответствия посредством самой себя.

В целях модификации исходного алгоритма для уменьшения влияния эффекта переобучения сети множество изменяемых элементов  $\{1,0\}$  дополняют до множества  $\{rRdD\}$ , используя те же операторы, но вводя таблицу разрешения конфликтов.

Результаты тестирования предложенного метода построения алгоритма при сравнении с устоявшимися методами показывают, что данный метод приближается по точности к методу Венга.

**Результаты экспериментов.** Проверка предложенных алгоритмов проведена в три этапа: в лабораторных условиях, в условиях, максимально приближенных к морским (бассейн ЦАГИ г. Москвы), при приемке готовой системы спасательным флотом РФ [4].



### Погрешности измерений перемещений (а) и скоростей (б):

БА — быстрый алгоритм; ЛК — Лукаса — Канаде; корр. — использована коррекция с помощью РБФ-сети; нет корр. — коррекцию не применяли

Обобщенно результаты экспериментов можно представить в виде графика, (см. рисунок), на котором они показаны с применением процедуры коррекции камер и без нее. Как видно, предложенный быстрый алгоритм на реальных кадрах проводит измерения перемещений и скоростей несколько лучше, чем даже хорошо зарекомендовавшие себя методы, учитывающие большое количество параметров.

Созданы и апробированы методика оценки параметров движения объектов на основе данных стереокамеры и методика упрощенной калибровки камер на основе нейрогенетического алгоритма. Проверка работоспособности предложенных методик проведена в различных условиях (как на суше, так и под водой). Созданное навигационное программное обеспечение способно работать с использованием маломощных вычислителей.

#### СПИСОК ЛИТЕРАТУРЫ

1. Lucas B.D., Kanade T. An iterative image registration technique with an application to stereo vision // Proc. of Imaging Understanding Workshop. 1981.
2. Artuerhof M., Demenkov N. The use of genetic algorithms in building's heating systems // Proc. of 13th IFAC Symp. on Information Control Problems in Manufacturing. Moscow, Russia. June, 2009.
3. Артюхов М.Ю. Быстрый алгоритм оценки смещения кадров для решения задачи локальной навигации подводных аппаратов // Автоматизация и современные технологии. 2011. № 7.
4. Опыт создания системы локальной видеонавигации для подводных аппаратов / М.Ю. Артюхов, А.Н. Кропотов, А.А. Макашов, Е.И. Сахарова // Техническое зрение в системах управления мобильными объектами – 2010: Тр. научн.-техн. конф.-семинара. Вып. 4 / под ред. Р.Р. Назирова. М.: КДУ, 2011.

Статья поступила в редакцию 25.10.2012