Обучение на аффинных группах для трекинга изображений объектов

С. Н. Чуканов[1], *С. В. Лейхтер*[2]
ch_sn@mail.ru; leykhter@mail.ru
[1]Омский филиал Института математики им. С. Л. Соболева Сибирского отделения РАН, Россия, г. Омск, ул. Певцова, 13
[2]ФГБОУ ВО «Сибирский государственный автомобильно-дорожный университет», Россия, г. Омск, пр. Мира, 5

Рассмотрены алгоритмы трекинга (отслеживания) объектов и распознавания поведения объектов на основе контроля пространственных и временных изменений параметров с использованием методов обучения. Предложены алгоритмы трекинга, в которых для аффинных преобразований используются лиевы группы. Анализируются параметры движения объекта, которые оптимизируются на многообразии, определяемого с помощью экспоненциального отображения между лиевой группой и ее алгеброй. Представлены алгоритмы совместного обучения и оценивание с помощью наблюдателя Люенбергера для задач трекинга на многообразиях.

Ключевые слова: *трекинг; обучение; аффинные преобразования; лиевы группы; наблюдатель Люенбергера*

DOI: 10.21469/22233792.3.2.01

1 Введение

Распространение компьютеров с производительными видеопроцессорами и необходимость проведения автоматизированного видеоанализа вызвали интерес к алгоритмам трекинга (отслеживания) объекта. Основными этапами видеоанализа кадров изображений являются: обнаружение целевых движущихся объектов, трекинг таких объектов от кадра к кадру и распознавание поведения на основе трекинга. Трекинг объектов представляет собой процесс отслеживания целевой области объекта, движения и положения объекта [1]. Современные алгоритмы трекинга изображений объектов используют методы машинного обучения.

Трекинг объектов широко применяется для визуального управления, наблюдения, роботизированной навигации, взаимодействия человека с компьютером и др. Проблема трекинга быстродвижущихся целей в сложных сценах требует адаптивных алгоритмов для достижения устойчивого и точного трекинга. При традиционном подходе трекинг рассматривается как задача нелинейной оптимизации. Hager и Belhumeur разработали геометрические модели для оценивания движений изображений точек в области цели, где наилучшие параметры оцениваются по минимизации суммы квадратов разностей между шаблоном и областью изображения [2]. В настоящее время некоторые авторы используют лиевы алгебры для систем машинного зрения и робототехники [3, 4]. Одним из преимуществ использования лиевых алгебр является возможность управлять объектом на многообразиях, которые локально (в окрестности каждой точки) могут быть представлены соответствующими евклидовыми пространствами. В работе используется многообразие в качестве инструмента для решения задачи оптимизации при трекинге объекта. Геометрические деформации изображения объекта представляются с использованием аффинной группы. Анализируются параметры движения объекта, которые оптимизируются на многообразии, определяемом с помощью экспоненциального отображения между лиевой группой и ее алгеброй. Методы обучения в алгоритмах трекинга используют описание объекта с помощью вектора значений, соответствующих множеству признаков объекта. Значениями признаков могут быть инварианты многообразий анализируемых данных [5, 6].

Для определения угловой скорости объекта — твердого тела (ТТ) — по информации о векторах направлений необходима информация приборного состава (оптико-электронных приборов, блока магнитометров, системы инерциальной навигации) для определения параметров ориентации [7]. В настоящее время отсутствуют робастные методы определения кинематических характеристик объекта по информации только о векторах заданных направлений (например, на Солнце, звезду или центр Земли) [8]. При значительных интервалах времени квантования поступления информации с приборного состава использование конечно-разностной аппроксимации производных параметров, описывающих ориентацию и движение ТТ, не является корректным. Например, при известной ориентации направления осей объекта — ТТ, т. е. известном кватернионе $\mathbf{q} = \begin{pmatrix} q_1 & q_2 & q_3 & q_4 \end{pmatrix}$ или векторе Гиббса $\mathbf{g} = q_4^{-1} \begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix}$ [7], вектор угловой скорости можно определить из соотношения:

$$\omega = 2 \left(1 + \|\mathbf{g}\|^2\right)^{-1} (\dot{\mathbf{g}} - \mathbf{g} \times \dot{\mathbf{g}}) \, .$$

Однако при больших значениях временно́го интервала Δt конечно-разностная аппроксимация производной $\dot{\mathbf{g}} = [\mathbf{g}(t_k + \Delta t) - \mathbf{g}(t_k)] (\Delta t)^{-1}$ не является корректной. Для определения вектора угловой скорости ω можно использовать обучение методом градиентного спуска с минимизацией целевой функции — нормой рассогласования между фактическим вектором $\mathbf{g}(t_k + \Delta t)$ и вектором $\mathbf{g}_{\mathrm{Aff}}(t_k + \Delta t)$, определенным с помощью аффинного преобразования вектора $\mathbf{g}(t_k)$ при векторе угловой скорости ω.

В настоящей работе предлагается использовать аппарат лиевой алгебры аффинных групп для формирования отображения вектора состояния объекта k-го момента времени на вектор состояния $(k+1)$-го момента времени для определения параметров ориентации объекта — ТТ. В разд. 4 рассмотрен метод совместного обучения на аффинных группах для трекинга изображений объектов и оценивания кинематических характеристик объектов — ТТ — с помощью наблюдателя Люенбергера.

2 Аффинные преобразования объектов

Относительные движения между объектом и точкой наблюдения вызывают деформации изображения объекта в последовательности кадров. Вариации области изображения объекта могут быть описаны с помощью модели аффинного преобразования.

В геометрии аффинное преобразование является функцией между аффинными пространствами, которая сохраняет точки, прямые и плоскости. Примеры аффинных преобразований включают в себя перенос, масштабирование, преобразование подобия, отражение, вращение, сдвиговое отображение. Аффинное преобразование представляет собой комбинацию переноса и линейного преобразования. В конечномерном случае, если линейное преобразование представляется как умножение на матрицу $\mathbf{A}$, а перенос — как сложение с вектором $\mathbf{b}$, то аффинное преобразование f, действующее на вектор $\mathbf{x}$, может быть представлено как $\mathbf{y} = f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$.

Используя расширенную матрицу аффинного преобразования $\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ 0 & 1 \end{bmatrix}$ и расширенный вектор $\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}$, можно представить перенос и линейное преобразование при аффинном преобразовании одним матричным умножением:

$$\begin{bmatrix} \mathbf{y} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix},$$

которое эквивалентно преобразованию: $\mathbf{y} = \mathbf{Ax} + \mathbf{b}$. Множество аффинных преобразований $g_i = \begin{bmatrix} \mathbf{A}_i & \mathbf{b}_i \\ 0 & 1 \end{bmatrix}$ является аффинной группой $\mathrm{Aff}\,(n)$, так как [9]:

– для элементов $g_1 = \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ 0 & 1 \end{bmatrix} \in \mathrm{Aff}\,(n)$ и $g_2 = \begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ 0 & 1 \end{bmatrix} \in \mathrm{Aff}\,(n)$ групповая операция (их комбинация):

$$g_1 \cdot g_2 = \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1\mathbf{A}_2 & \mathbf{b}_1 + \mathbf{A}_1\mathbf{b}_2 \\ 0 & 1 \end{bmatrix} \in \mathrm{Aff}\,(n)\,;$$

– существует единичный элемент группы:

$$\mathbf{I}^{(n+1)\times(n+1)} = \begin{bmatrix} \mathbf{I}^{n\times n} & \mathbf{0}^{1\times n} \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}\,;$$

– для элемента группы $g = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ 0 & 1 \end{bmatrix} \in \mathrm{Aff}\,(n)$ существует обратный элемент:

$$g^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{b} \\ 0 & 1 \end{bmatrix} \in \mathrm{Aff}\,(n)\ ;$$

– операции в группе являются ассоциативными:

$$g_1 \cdot (g_2 \cdot g_3) = (g_1 \cdot g_2) \cdot g_3\,.$$

Рассмотрим алгоритм, который инициализирует трекинг объекта, назначая область, содержащую целевое изображение в первом кадре в качестве исходного шаблона. В частном случае текущая совокупность точек подвергается аффинному преобразованию в последовательных кадрах изображения. Алгоритм трекинга определяет аффинные параметры преобразования, которое совместит текущие точки с эталонными точками — шаблоном. Процесс повторяется для каждых двух последовательных кадров изображений совокупности точек.

Пусть $\mathbf{c}\,(i) = \begin{pmatrix} x\,(i) & y\,(i) & 1 \end{pmatrix}^{\mathrm{T}}, i = 1, \ldots, n,$ — точки текущего положения и $\mathbf{c}_p\,(i) = \begin{pmatrix} x_p\,(i) & y_p\,(i) & 1 \end{pmatrix}^{\mathrm{T}}, i = 1, \ldots, n,$ — соответствующие точки эталонного шаблона. Рассмотрим аффинное преобразование текущего положения в однородных координатах:

$$\mathbf{c}'\,(i) = \begin{pmatrix} x'\,(i) & y'\,(i) & 1 \end{pmatrix}^{\mathrm{T}} = \mathbf{M}\,(\mathbf{p}) \begin{pmatrix} x\,(i) & y\,(i) & 1 \end{pmatrix}^{\mathrm{T}} = \mathbf{M}\,(\mathbf{p})\,\mathbf{c}\,(i)\,,$$

где

$$\mathbf{M}\,(\mathbf{p}) = \exp\left[\begin{pmatrix} p_1 & p_2 & p_5 \\ p_3 & p_4 & p_6 \\ 0 & 0 & 0 \end{pmatrix}\right]\ ;$$

$p_1p_4 - p_2p_3 \neq 0$ — матрица элемента группы аффинного преобразования $\mathrm{Aff}(2)$ плоского изображения [10]; $\mathbf{p} = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 \end{pmatrix}^{\mathrm{T}}$ — вектор параметров аффинного преобразования. Матрица аффинного преобразования $\mathbf{M}(\mathbf{p})$ может быть представлена в виде:

$$\mathbf{M}(\mathbf{p}) = \exp\left(\sum_{j=1}^{6} p_j \mathfrak{g}_j\right),$$

где $\mathfrak{g}_j, j = 1, \ldots, 6,$ — генераторы алгебры $\mathfrak{aff}(2)$:

$$\left.\begin{array}{lll}
\mathfrak{g}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; & \mathfrak{g}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; & \mathfrak{g}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; \\
\mathfrak{g}_4 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}; & \mathfrak{g}_5 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}; & \mathfrak{g}_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.
\end{array}\right\} \quad (1)$$

Используем формулу Г. Цассенхауза (Hans Zassenhaus), полученную из формулы BCH (H. Baker, J. Campbell, F. Hausdorff), которая имеет вид [11]:

$$\exp(t(X+Y)) = \exp(tX) \cdot \exp(tY) \cdot \exp\left(-\frac{t^2}{2}[X,Y]\right) \cdot \exp\left(\frac{t^3}{6}(2[Y,[X,Y]] + [X,[X,Y]])\right) \cdots,$$

где $X, Y \in \mathfrak{g}; t \in \mathbb{R}$.

При малых t, пренебрегая членами $O(t^n), n \geqslant 2$, получим:

$$e^{t(X+Y)} \approx e^{tX} e^{tY} \approx e^{tY} e^{tX},$$

откуда следует, что при малых значениях $|\mathbf{p}|$ можно записать [12, 13]:

$$\mathbf{M}(\mathbf{p}) = \exp\left(\sum_{j=1}^{6} p_j \mathfrak{g}_j\right) \approx \prod_{j=1}^{6} \exp(p_j \mathfrak{g}_j).$$

3 Обучение параметров трекинга методом градиентного спуска

Алгоритм трекинга может быть сведен к минимизации целевой функции [14, 15]:

$$J(\mathbf{p}) = \sum_{i=1}^{n} (\mathbf{c}'(i) - \mathbf{c}_p(i))^{\mathrm{T}} (\mathbf{c}'(i) - \mathbf{c}_p(i)) = \sum_{i=1}^{n} (\mathbf{M}(\mathbf{p})\mathbf{c}(i) - \mathbf{c}_p(i))^{\mathrm{T}} (\mathbf{M}(\mathbf{p})\mathbf{c}(i) - \mathbf{c}_p(i)).$$

В общем случае минимизация $J(\mathbf{p})$ определяет нелинейную задачу оптимизации, которая может быть решена обучением компонент вектора $\mathbf{p}$ с помощью метода градиентного спуска [16, 17]:

$$\left.\begin{array}{l}
\dfrac{\partial J(\mathbf{p})}{\partial(p_j)}\Big|_{\mathbf{p}=\mathbf{p}_{k-1}} = 2\displaystyle\sum_{i=1}^{n} \left[(\mathbf{M}(\mathbf{p}_{k-1})\mathfrak{g}_j \mathbf{c}(i))^{\mathrm{T}} (\mathbf{M}(\mathbf{p}_{k-1})\mathbf{c}(i) - \mathbf{c}_p(i))\right]; \\
(p_j)_k = (p_j)_{k-1} - \dfrac{\gamma}{2}\dfrac{\partial J(\mathbf{p})}{\partial(p_j)}\Big|_{\mathbf{r}=\mathbf{r}_{k-1}} = \\
= (p_j)_{k-1} - \gamma\displaystyle\sum_{i=1}^{n} \left[(\mathbf{M}(\mathbf{p}_{k-1})\mathfrak{g}_j \mathbf{c}(i))^{\mathrm{T}} (\mathbf{M}(\mathbf{p}_{k-1})\mathbf{c}(i) - \mathbf{c}_p(i))\right],
\end{array}\right\} \quad (2)$$

где j — индекс параметра вектора $\mathbf{p}$; $k \geqslant 1$ — номер итерации; γ — параметр обучения; при этом учитывается, что $\partial\mathbf{M}/\partial p_j = \mathbf{M}\mathfrak{g}_j$. Критерием окончания алгоритма градиентного спуска может быть выполнение условия:

$$(J(\mathbf{p}_k) - J(\mathbf{p}_{k-1}))^2 < \varepsilon,$$

где $\varepsilon \in \mathbb{R}^{+}$ — заданное малое число.

Введем ограничение на норму матрицы $\mathbf{M}(\mathbf{p})$ в целевой функции $J(\mathbf{p})$:

$$J(\mathbf{p}) = \sum_{i=1}^{n} (\mathbf{M}(\mathbf{p})\,\mathbf{c}(i) - \mathbf{c}_p(i))^{\mathrm{T}} (\mathbf{M}(\mathbf{p})\,\mathbf{c}(i) - \mathbf{c}_p(i)) + \lambda (\mathbf{M}(\mathbf{p}))^{T} (\mathbf{M}(\mathbf{p})),$$

где λ — коэффициент регуляризации. Тогда решение для обучения компонент вектора $\mathbf{p}$ можно записать в виде:

$$\begin{gathered}(p_j)_k = (p_j)_{k-1} - \\ - \gamma \left\{ \sum_{i=1}^{n} \left[(\mathbf{M}(\mathbf{p}_{k-1})\,\mathfrak{g}_j \mathbf{c}(i))^{\mathrm{T}} (\mathbf{M}(\mathbf{p}_{k-1})\,\mathbf{c}(i) - \mathbf{c}_p(i)) \right] + \lambda (\mathbf{M}(\mathbf{p}_{k-1})\,\mathfrak{g}_j)^{\mathrm{T}} (\mathbf{M}(\mathbf{p}_{k-1})) \right\}.\end{gathered}$$

Пример 1. Аффинные преобразования $\mathrm{Aff}(2)$.

Рассмотрим трекинг изображения объекта — квадрата на плоскости для двух последовательных кадров. Аффинная алгебра $\mathfrak{aff}(2)$ аффинной группы преобразований $\mathrm{Aff}(2)$ содержит генераторы (1) с соответствующими параметрами p_1, p_2, p_3, p_4, p_5 и p_6.

Зададим текущее положение изображения объекта следующими четырьмя точками:

$$\mathbf{c}(1) = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^{\mathrm{T}};\ \mathbf{c}(2) = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}^{\mathrm{T}};\ \mathbf{c}(3) = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}^{\mathrm{T}};\ \mathbf{c}(4) = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}.$$

Положение эталонного шаблона зададим точками:

$$\mathbf{c}_p(1) = \begin{pmatrix} 0{,}10 & 0{,}10 & 1 \end{pmatrix}^{\mathrm{T}}; \quad \mathbf{c}_p(2) = \begin{pmatrix} 0{,}20 & 1{,}21 & 1 \end{pmatrix}^{\mathrm{T}};$$

$$\mathbf{c}_p(3) = \begin{pmatrix} 1{,}20 & 0{,}22 & 1 \end{pmatrix}^{\mathrm{T}}; \quad \mathbf{c}_p(4) = \begin{pmatrix} 1{,}31 & 1{,}33 & 1 \end{pmatrix}^{\mathrm{T}}.$$

Задача заключается в нахождении таких значений параметров преобразования p_1, p_2, p_3, p_4, p_5 и p_6, которые обеспечат совмещение текущего положения объекта с положением эталонного шаблона.

В табл. 1 приведены результаты обучения методом градиентного спуска при выбранном значении параметра обучения $\gamma = 0{,}2$.

Из табл. 1 следует, что требуемые параметры трекинга определены за 16 итераций с точностью $\varepsilon = 0{,}001$. При этом $J(\mathbf{p}) = 3{,}63 \cdot 10^{-6}$; после 32 итераций: $J(\mathbf{p}) = 1{,}19 \cdot 10^{-8}$.

Пример 2. Специальные евклидовы преобразования $\mathbf{SE}$ (3).

Рассмотрим трекинг изображения объекта — куба в трехмерном пространстве для двух последовательных кадров. Алгебра $\mathfrak{se}(3)$ группы специальных евклидовых преобразова-

Таблица 1 Результаты обучения при преобразовании $\mathrm{Aff}(2)$

Итерация	$(p_1)_k$	$(p_2)_k$	$(p_3)_k$	$(p_4)_k$	$(p_5)_k$	$(p_6)_k$
0	0,050	0,050	0,050	0,050	0,050	0,050
1	0,107	0,107	0,114	0,112	0,141	0,149
2	0,080	0,080	0,078	0,077	0,091	0,085
3	0,100	0,099	0,105	0,103	0,120	0,125
...	...	...	...	...	...	...
16	0,099	0,099	0,100	0,099	0,101	0,101

ний $\mathbf{SE}(3)$ (подгруппы аффинной группы $\mathrm{Aff}(3)$) содержит генераторы:

$$\mathfrak{g}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad \mathfrak{g}_2 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad \mathfrak{g}_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix};$$

$$\mathfrak{g}_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad \mathfrak{g}_5 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}; \quad \mathfrak{g}_6 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Матрица преобразования имеет вид:

$$\mathbf{M}(\mathbf{p}) = \exp\left(\sum_{j=1}^{6} p_j \mathfrak{g}_j\right),$$

где p_1, p_2, p_3, p_4, p_5 и p_6 — локальные координаты преобразования.

Зададим текущее положение изображения объекта следующими точками:

$$\mathbf{c}(1) = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}^{\mathrm{T}}; \ \mathbf{c}(2) = \begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix}^{\mathrm{T}}; \ \mathbf{c}(3) = \begin{pmatrix} 0 & 1 & 0 & 1 \end{pmatrix}^{\mathrm{T}};$$

$$\mathbf{c}(4) = \begin{pmatrix} 0 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}; \ \mathbf{c}(5) = \begin{pmatrix} 1 & 0 & 0 & 1 \end{pmatrix}^{\mathrm{T}}; \ \mathbf{c}(6) = \begin{pmatrix} 1 & 0 & 1 & 1 \end{pmatrix}^{\mathrm{T}};$$

$$\mathbf{c}(7) = \begin{pmatrix} 1 & 1 & 0 & 1 \end{pmatrix}^{\mathrm{T}}; \ \mathbf{c}(8) = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}.$$

Положение эталонного шаблона зададим точками $\mathbf{c}_p(i) = \mathbf{M}(\mathbf{p})\,\mathbf{c}(i)$:

$$\mathbf{c}_p(1) = \begin{pmatrix} 0{,}1 & 0{,}1 & 0{,}1 & 1 \end{pmatrix}^{\mathrm{T}}; \quad \mathbf{c}_p(2) = \begin{pmatrix} 0 & 0{,}20 & 1{,}09 & 1 \end{pmatrix}^{\mathrm{T}};$$

$$\mathbf{c}_p(3) = \begin{pmatrix} 0{,}20 & 1{,}09 & 0{,}01 & 1 \end{pmatrix}^{\mathrm{T}}; \ \mathbf{c}_p(4) = \begin{pmatrix} 0{,}10 & 1{,}19 & 1{,}00 & 1 \end{pmatrix}^{\mathrm{T}};$$

$$\mathbf{c}_p(5) = \begin{pmatrix} 1{,}09 & 0{,}01 & 0{,}21 & 1 \end{pmatrix}^{\mathrm{T}}; \ \mathbf{c}_p(6) = \begin{pmatrix} 0{,}99 & 0{,}11 & 1{,}20 & 1 \end{pmatrix}^{\mathrm{T}};$$

$$\mathbf{c}_p(7) = \begin{pmatrix} 1{,}19 & 1{,}00 & 0{,}12 & 1 \end{pmatrix}^{\mathrm{T}}; \ \mathbf{c}_p(8) = \begin{pmatrix} 1{,}09 & 1{,}10 & 1{,}11 & 1 \end{pmatrix}^{\mathrm{T}},$$

при

$$\mathbf{M}(\mathbf{p}) = \begin{pmatrix} 0{,}99 & 0{,}0995 & -0{,}1 & 0{,}1 \\ -0{,}09 & 0{,}99 & 0{,}0995 & 0{,}1 \\ 0{,}11 & -0{,}09 & 0{,}99 & 0{,}1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Задача обучения заключается в нахождении таких значений параметров преобразования p_1, p_2, p_3, p_4, p_5 и p_6, которые обеспечат совмещение текущего положения объекта с положением эталонного шаблона.

В табл. 2 приведены результаты обучения методом градиентного спуска при значении параметра обучения $\gamma = 0{,}1$.

Из табл. 1 следует, что требуемые параметры трекинга определены за 16 итераций с точностью $\varepsilon = 0{,}001$. При этом $J(\mathbf{p}) = 2{,}89 \cdot 10^{-6}$; после 32 итераций — $J(\mathbf{p}) = 9{,}81 \cdot 10^{-10}$. Если измерения проводятся в международной системе единиц СИ, а смена кадров производится через интервал времени $\Delta t = 0{,}1$ с, то $p_1 = p_2 = p_3 = 0{,}1$ рад, $p_4 = p_5 = p_6 = 0{,}1$ м, и можно оценить кинематические параметры движении отслеживаемого объекта: компоненты вектора угловой скорости $\omega_1 = \omega_2 = \omega_3 = 1{,}0$ с$^{-1}$ и компоненты вектора линейной скорости $v_1 = v_2 = v_3 = 0$ м/с.

Таблица 2 Результаты обучения при преобразовании $\mathbf{SE}(3)$

Итерация	$(p_1)_k$	$(p_2)_k$	$(p_3)_k$	$(p_4)_k$	$(p_5)_k$	$(p_6)_k$
0	0,050	0,050	0,050	0,050	0,050	0,050
1	0,067	0,050	0,067	0,087	0,090	0,093
2	0,076	0,074	0,075	0,088	0,108	0,098
3	0,086	0,103	0,088	0,105	0,101	0,103
...	...	...	...	...	...	...
16	0,100	0,100	0,100	0,100	0,100	0,100

4 Совместное обучение и оценивание с помощью наблюдателя Люенбергера

В теории управления наблюдатель состояния системы представляет собой систему, которая обеспечивает оценку внутреннего состояния данной реальной системы по результатам измерений входных и выходных сигналов реальной системы. В большинстве практических случаев состояние системы не может быть определено прямым наблюдением, поэтому используются косвенные методы оценивания состояния системы. Если система является наблюдаемой, то можно полностью восстановить состояние системы измерением входных и выходных сигналов системы с использованием наблюдателя состояния системы.

Рассмотрим линейную систему Σ:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k)\,; \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k)\,, \end{aligned}$$

где $\mathbf{A} \in \mathbb{R}^{n\times n}, \mathbf{B} \in \mathbb{R}^{n\times m}, \mathbf{C} \in \mathbb{R}^{l\times n}$, с дискретным временем k. Здесь $\mathbf{x}(k) \in \mathbb{R}^n$ — вектор состояния системы; $\mathbf{u}(k) \in \mathbb{R}^m$ — вектор сигналов управления; $\mathbf{y}(k) \in \mathbb{R}^l$ — вектор выходных сигналов. Запишем уравнения наблюдателя Люенбергера (David Luenberger [18]) для системы Σ:

$$\begin{aligned} &\hat{\mathbf{x}}(k+1) = (\mathbf{A} - \mathbf{L}\mathbf{C})\,\hat{\mathbf{x}}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{L}\mathbf{y}(k)\,, \quad \mathbf{L} \in \mathbb{R}^{n\times l}; \\ &\hat{\mathbf{y}}(k) = \mathbf{C}\hat{\mathbf{x}}(k)\,, \end{aligned}$$

где $\hat{\mathbf{x}}(k) \in \mathbb{R}^n$ — оцениваемый вектор состояния; $\hat{\mathbf{y}}(k) \in \mathbb{R}^m$ — оцениваемый вектор выходных сигналов. Наблюдатель Люенбергера отличается от оценивания методом фильтрации Калмана тем, что матрица коэффициентов усиления $\mathbf{L}$ выбирается постоянной и обеспечивает устойчивость процессов оценивания.

Если вектором состояния системы является вектор параметров аффинного преобразования: $\mathbf{x}(k) = \mathbf{p}(k) \in \mathbb{R}^n$, вектор управления равен нулю: $\mathbf{u}(k) = 0, \forall k$ и матрица $\mathbf{C} = \mathbf{I} \in \mathbb{R}^{n\times n}$, то можно сформировать наблюдатель Люенбергера:

$$\hat{\mathbf{p}}(k+1) = (\mathbf{A} - \mathbf{L})\,\hat{\mathbf{p}}(k) + \mathbf{L}\mathbf{p}(k)\,, \tag{3}$$

который функционирует совместно с обучением методом градиентного спуска (см. (2)):

$$p_j(k+1) = p_j(k) - \gamma \sum_{i=1}^{n} \Big[(\mathbf{M}\mathfrak{g}_j \mathbf{c}(i))^{\mathrm{T}} (\mathbf{M}\mathbf{c}(i) - \mathbf{c}_p(i)) \Big],$$

где $\mathbf{M} = \mathbf{M}(\mathbf{p}(k))$.

В случае аффинной группы $\mathrm{Aff}(2)$ и изображения — совокупности четырех точек (как в примере 1): $j = 1, \ldots, 6$ — номер компоненты вектора $\mathbf{p}$; $i = 1, \ldots, 4$ — номер точки.

Пример 3. Трекинг изображения при движении свободного твердого тела

Рассмотрим трекинг изображения при движении свободного ТТ с главными моментами инерции $J_x = 2J_y = 2J_z$, уравнения которого подчиняются соотношениям Эйлера:

$$\dot{\omega}_1 = 0\,; \quad \dot{\omega}_2 = -\nu\omega_3\,; \quad \dot{\omega}_3 = \nu\omega_2\,,$$

где $\nu = 0{,}5\omega_1$, $\omega_1 = \omega_x$, $\omega_2 = \omega_y$, и $\omega_3 = \omega_z$ — компоненты вектора угловой скорости ТТ относительно собственной системы координат (ССК) ТТ. Кинематические соотношения вращательного движения ТТ формируются на многообразии лиевой группы $\mathrm{SO}\,(3)$ (подгруппе аффинной группы $\mathrm{Aff}\,(3)$) с алгеброй $\mathfrak{so}\,(3)$.

При интервале квантования Δt можно записать эти уравнения в дискретной форме:

$$\omega\,(k+1) = \exp\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\nu\Delta\mathrm{t} \\ 0 & \nu\Delta\mathrm{t} & 0 \end{pmatrix}\omega\,(k) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\,(\nu\Delta\mathrm{t}) & -\sin\,(\nu\Delta\mathrm{t}) \\ 0 & \sin\,(\nu\Delta\mathrm{t}) & \cos\,(\nu\Delta\mathrm{t}) \end{pmatrix}\omega\,(k)\,.$$

Примем в качестве вектора состояния системы $\mathbf{x}\,(k)$ и вектора параметров аффинного преобразования $\mathbf{p}\,(k)$ компоненты вектора угловой скорости: $\omega\,(k) = \begin{pmatrix} \omega_1\,(k) & \omega_2\,(k) & \omega_3\,(k) \end{pmatrix}^{\mathrm{T}}$, $k \in \mathbb{Z}^+$ — дискретное время. Для получения уравнений наблюдателя Люенбергера (3) примем:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\,(\nu\Delta\mathrm{t}) & -\sin\,(\nu\Delta\mathrm{t}) \\ 0 & \sin\,(\nu\Delta\mathrm{t}) & \cos\,(\nu\Delta\mathrm{t}) \end{pmatrix};\ \mathbf{B} = 0;\ \mathbf{C} = \mathbf{I} \in \mathbb{R}^{3\times 3};\ \mathbf{L} = \lambda \in \mathbb{R}^+.$$

Тогда уравнения наблюдателя Люенбергера будут иметь вид:

$$\hat{\omega}\,(k+1) = \begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & \cos\,(\nu\Delta\mathrm{t})-\lambda & -\sin\,(\nu\Delta\mathrm{t}) \\ 0 & \sin\,(\nu\Delta\mathrm{t}) & \cos\,(\nu\Delta\mathrm{t})-\lambda \end{pmatrix}\hat{\omega}\,(k) + \lambda\omega_m\,(k)\,,$$

где вектор $\omega_m\,(k)$ определяется методом обучения по информации о координатах вершин куба, которые в k-й момент дискретного времени были равны:

$$\mathbf{c}\,(1) = \begin{pmatrix} -1 & -1 & -1 \end{pmatrix}^{\mathrm{T}};\ \mathbf{c}\,(2) = \begin{pmatrix} -1 & -1 & 1 \end{pmatrix}^{\mathrm{T}};\ \mathbf{c}\,(3) = \begin{pmatrix} -1 & 1 & -1 \end{pmatrix}^{\mathrm{T}};$$

$$\mathbf{c}\,(4) = \begin{pmatrix} -1 & 1 & 1 \end{pmatrix}^{\mathrm{T}};\ \mathbf{c}\,(5) = \begin{pmatrix} 1 & -1 & -1 \end{pmatrix}^{\mathrm{T}};\ \mathbf{c}\,(6) = \begin{pmatrix} 1 & -1 & 1 \end{pmatrix}^{\mathrm{T}};$$

$$\mathbf{c}\,(7) = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}^{\mathrm{T}};\ \mathbf{c}\,(8) = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}},$$

а в $(k+1)$-й момент: $\mathbf{c}_p\,(i) = \mathbf{M}\,(\omega)\,\mathbf{c}\,(i)$, где $\mathbf{M}\,(\omega) = \exp\left(\sum_{j=1}^{3}\omega_j\mathfrak{g}_j\right)$; $\mathfrak{g}_j, j = 1, \ldots, 3,$ — генераторы алгебры $\mathfrak{so}\,(3)$:

$$\mathfrak{g}_1 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix};\ \mathfrak{g}_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix};\ \mathfrak{g}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Компоненты вектора $\omega_m\,(k)$ определяются из соотношения (3), в котором $\mathbf{p}\,(k) = \omega_m\,(k)$.

В табл. 3 приведены результаты совместного обучения методом градиентного спуска и на основе наблюдателя Люенбергера при $\omega|_{k=0} = \begin{pmatrix} 0{,}1 & 0{,}01 & 0 \end{pmatrix}^{T}$; $\nu = 0{,}05$; $\gamma = 0{,}05$; $\lambda = 0{,}25$.

5 Заключение

В работе рассмотрены алгоритмы трекинга (отслеживания) объектов и распознавания поведения объектов путем контроля пространственных и временны́х изменений парамет-

Таблица 3 Результаты обучения при преобразовании SO(3)

Шаг	Время	$(\omega(k))^{\mathrm{T}}$	$(\hat{\omega}(k))^{\mathrm{T}}$	$\lvert\omega(k)-\hat{\omega}(k)\rvert$
0	0	$(0{,}1 \quad 1{,}0\cdot10^{-3} \quad 0)$	$(0 \quad 0 \quad 0)$	$1{,}0\cdot10^{-1}$
1	0,1	$(0{,}1 \quad 1{,}0\cdot10^{-3} \quad 0{,}5\cdot10^{-4})$	$(0{,}02 \quad 2{,}0\cdot10^{-3} \quad 1{,}0\cdot10^{-6})$	$8{,}04\cdot10^{-2}$
2	0,2	$(0{,}1 \quad 1{,}0\cdot10^{-3} \quad 1{,}0\cdot10^{-4})$	$(0{,}039 \quad 3{,}90\cdot10^{-3} \quad 3{,}70\cdot10^{-5})$	$6,1\cdot10^{-2}$
3	0,3	$(0{,}1 \quad 1{,}0\cdot10^{-3} \quad 1{,}5\cdot10^{-4})$	$(0{,}054 \quad 5{,}41\cdot10^{-3} \quad 7{,}68\cdot10^{-5})$	$4{,}6\cdot10^{-2}$
…	…	…	…	…
8	0,8	$(0{,}1 \quad 9{,}9\cdot10^{-3} \quad 3{,}99\cdot10^{-4})$	$(0{,}089 \quad 8{,}91\cdot10^{-3} \quad 3{,}46\cdot10^{-4})$	$1{,}1\cdot10^{-2}$
…	…	…	…	…
16	1,6	$(0{,}1 \quad 9{,}97\cdot10^{-3} \quad 7{,}99\cdot10^{-4})$	$(0{,}098 \quad 9{,}86\cdot10^{-3} \quad 7{,}79\cdot10^{-4})$	$1{,}1\cdot10^{-3}$
…	…	…	…	…
32	3,2	$(0{,}1 \quad 9{,}87\cdot10^{-3} \quad 1{,}59\cdot10^{-3})$	$(0{,}1 \quad 9{,}87\cdot10^{-3} \quad 1{,}58\cdot10^{-3})$	$1{,}68\cdot10^{-5}$

Примечания: $\omega(k)$ — вектор угловой скорости ТТ относительно ССК на шаге k; $\hat{\omega}(k)$ — вектор оценки угловой скорости ТТ наблюдателем Люенбергера на шаге k; $\lvert\omega(k)-\hat{\omega}(k)\rvert$ — абсолютное значение рассогласования векторов $\omega(k)$ и $\hat{\omega}(k)$.

ров с использованием методов обучения. Предложены алгоритмы трекинга, в которых для аффинных преобразований используются лиевы группы.

Алгоритмы трекинга могут быть распространены для отслеживания сигналов любой физической природы: электрических, звуковых, радиосигналов и др. Алгоритмы трекинга могут быть использованы при идентификации сбоев и отказов элементов в системах управления объектами.

В настоящей работе многообразие используется в качестве инструмента для решения задачи оптимизации при трекинге объекта. Геометрические деформации изображения объекта представляются с использованием аффинной группы. Анализируются параметры движения объекта, которые оптимизируются на многообразии, определяемом с помощью экспоненциального отображения между лиевой группой и ее алгеброй. Методы обучения в алгоритмах трекинга используют описание объекта с помощью вектора значений, соответствующих множеству признаков объекта, значениями которых могут быть инварианты многообразий анализируемых данных.

Предложен метод определения вектора угловой скорости и параметров ориентации объекта — ТТ — по информации о векторах направлений по информации приборного состава. В настоящее время отсутствуют робастные методы определения кинематических характеристик объекта по информации только о векторах заданных направлений. Указано, что при значительных интервалах времени квантования поступления информации с приборного состава использование конечно-разностной аппроксимации производных параметров, описывающих ориентацию и движение ТТ, не является корректным. Для определения вектора угловой скорости предложено использовать обучение методом градиентного спуска при трекинге изображений объектов — ТТ. В разд. 4 рассмотрен метод совместного обучения на аффинных группах для трекинга изображений объектов и оценивания кинематических характеристик объектов — ТТ — с помощью наблюдателя Люенбергера.

Литература

[1] *Parmar M.* A survey of video object tracking methods // Int. J. Eng. Development Research, 2016. Vol. 14. P. 519–524.

[2] *Hager G.D., Belhumeur P.N.* Efficient region tracking with parametric models of geometry and illumination // IEEE T. Pattern Anal., 1998. Vol. 20. No. 10. P. 1025–1039.

[3] *Drummond T., Cipolla R.* Application of lie algebras to visual servoing // Int. J. Comput. Vision, 2000. Vol. 37. No. 1. P. 21–41.

[4] *Bayro-Corrochano E., Ortegon-Aguilar J.* Lie algebra template tracking // 17th Conference (International) on Pattern Recognition Proceedings. — IEEE, 2004. Vol. 2. P. 56–59.

[5] *Chukanov S.N.* Definitions of invariants for n-dimensional traced vector fields of dynamic systems // Pattern Recognition Image Anal., 2009. Vol. 19. No. 2. P. 303–305.

[6] *Чуканов С. Н.* Формирование инвариантов при визуализации векторных полей, определяемых интегральными кривыми динамических систем // Автометрия, 2011. Т. 47. № 2. С. 58–63.

[7] *Markley F. L., Crassidis J. L.* Fundamentals of spacecraft attitude determination and control. — New York, NY, USA: Springer Science+Business Media, 2014. 486 p.

[8] *Чуканов С. Н.* Определение пространственной ориентации твердого тела по информации приборного состава об одном направлении // Авиакосмическое приборостроение, 2004. № 3. С. 11–14.

[9] *Murray R. M., Li Z., Sastry S. S.* A mathematical introduction to robotic manipulation. — CRC Press, 1994. 456 p.

[10] *Anjyo K., Ochiai H.* Mathematical basics of motion and deformation in computer graphics // Synthesis Lectures on Computer Graphics and Animation, 2014. Vol. 6. No. 3. P. 1–83.

[11] *Hall B. C.* Lie groups, Lie algebras, and representations: An elementary introduction. — New York, NY, USA: Springer, 2003. 353 p.

[12] *Wei J., Norman E.* On global representations of the solutions of linear differential equations as a product of exponentials // P. Am. Math. Soc., 1964. Vol. 15. No. 2. P. 327–334.

[13] *Altafini C.* Explicit Wei–Norman formulae for matrix Lie groups via Putzer's method // Syst. Control Lett., 2005. Vol. 54. No. 11. P. 1121–1130.

[14] *Taylor C.J., Kriegman, D.J.* Minimization on the Lie group SO(3) and related manifolds. — New Haven, CT, USA: Yale University, 1994. Vol. 16. 155 p.

[15] *Absil P.-A., Mahony R., Sepulchre R.* Optimization algorithms on matrix manifolds. — Princeton, NJ, USA: Princeton University Press, 2008. 236 p.

[16] *Liu Y., Shia Z., Lia G.* Deformable target tracking method based on Lie algebra // Proc. SPIE, 2007, Vol. 6786. P. 678617-1.

[17] *Porikli F.* Regression on Lie groups and its application to affine motion tracking // Algorithmic advances in Riemannian geometry and applications / Eds. H. Q. Minh, V. Murino. — Advances in computer vision and pattern recognition ser. — Springer International Publishing, 2016. P. 173–185.

[18] *Kravaris C., Hahn J., Chu Y.* Advances and selected recent developments in state and parameter estimation // Comput. Chem. Eng., 2013. Vol. 51. P. 111–123.

Поступила в редакцию 13.06.2017

Learning on affine groups for tracking images of objects

S. N. Chukanov[1]***, S. V. Leykhter***[2]

ch_sn@mail.ru; leykhter@mail.ru

[1]Omsk Branch of Sobolev Institute of Mathematics of the Siberian Branch of RAS, 13 Pevtsova Str., Omsk, Russia

[2]Siberian State Automobile and Highway University, 5 Mira pr., Omsk, Russia

Algorithms for tracking of objects and recognition of the behavior of objects based on the control of spatial and time changes of parameters using the learning methods are considered in the paper. Tracking algorithms, in which Lie groups are used for affine transformations, are proposed. The parameters of the object's motion determined by means of the exponential mapping between the Lie group and its algebra are analyzed. The parameters are optimized on the manifold. Algorithms for joint learning and estimation with the help of the Luenberger observer for tracking problems on manifolds are presented.

Keywords: *tracking; learning; affine transformations; Lie groups; Lueberberger observer*

DOI: 10.21469/22233792.3.2.01

References

[1] Parmar, M. 2016. A survey of video object tracking methods. *Int. J. Eng. Development Research* 14:519–524.

[2] Hager, G. D., and P. N. Belhumeur. 1998. Efficient region tracking with parametric models of geometry and illumination. *IEEE T. Pattern Anal.* 20(10):1025–1039.

[3] Drummond, T., and R. Cipolla. 2000. Application of lie algebras to visual servoing. *Int. J. Comput. Vision* 37(1):21–41.

[4] Bayro-Corrochano, E., and J. Ortegon-Aguilar. 2004. Lie algebra template tracking. // *17th Conference (International) on Pattern Recognition Proceedings*. IEEE. 2:56–59.

[5] Chukanov, S. N. 2009. Definitions of invariants for n-dimensional traced vector fields of dynamic systems. *Pattern Recognition Image Anal.* 19(2):303–305.

[6] Chukanov, S. N. 2011. Constructing invariants for visualization of vector fields defined by integral curves of dynamic systems. *Optoelectron. Instrumentation Data Processing* 47(2):151–155.

[7] Markley, F. L., and J. L. Crassidis. 2014. *Fundamentals of spacecraft attitude determination and control*. New York, NY: Springer Science+Business Media. 486 p.

[8] Chukanov, S. N. 2004. Opredelenie prostranstvennoy orientatsii tverdogo tela po informatsii pribornogo sostava ob odnom napravlenii [Determination of spatial orientation of a solid body on information about instrumental composition on one direction]. *Aviakosmicheskoe priborostroenie* [Aerospace Instrument Making] 3:11–14.

[9] Murray, R. M., Z. Li,, and S. S. Sastry. 1994. *A mathematical introduction to robotic manipulation*. CRC Press. 456 p.

[10] Anjyo, K., and H. Ochiai. 2014. Mathematical basics of motion and deformation in computer graphics. *Synthesis Lectures on Computer Graphics and Animation* 6(3):1–83.

[11] Hall, B. C. 2003. *Lie groups, Lie algebras, and representations: An elementary introduction*. New York, NY: Springer. 353 p.

[12] Wei, J., and E. Norman. 1964. On global representations of the solutions of linear differential equations as a product of exponentials. *P. Am. Math. Soc.* 15(2):327–334.

[13] Altafini, C. Explicit Wei–Norman formulae for matrix Lie groups via Putzer's method. *Syst. Control Lett.* 54(11):1121–1130.

[14] Taylor, C. J., and D. J. Kriegman. 1994. *Minimization on the Lie group SO(3) and related manifolds*. New Haven, CT: Yale University. Vol. 16. 155 p.

[15] Absil, P. A., R. Mahony, and R. Sepulchre. 2008. *Optimization algorithms on matrix manifolds*. Princeton, NJ: Princeton University Press. 236 p.

[16] Liu, Y., Z. Shia, and G. Lia. 2007. Deformable target tracking method based on Lie algebra. *Proc. SPIE* 6786:678617-1.

[17] Porikli, F. 2016. Regression on Lie groups and its application to affine motion tracking. *Algorithmic advances in Riemannian geometry and applications*. Eds. H. Q. Minh and V. Murino. Advances in computer vision and pattern recognition ser. Springer International Publishing. 173–185.

[18] Kravaris, C., J. Hahn, and Y. Chu. 2013. Advances and selected recent developments in state and parameter estimation. *Comput. Chem. Eng.* 51:111–123.

Received June 13, 2017