

МЕТОДЫ СОВМЕЩЕНИЯ КАДРОВ ПРИ НАБЛЮДЕНИИ СЦЕНЫ С ПОМОЩЬЮ ДВИЖУЩЕГОСЯ ЛЕТАТЕЛЬНОГО АППАРАТА

С. В. Шолтанюк

*Белорусский государственный университет,
пр. Независимости, 4, 220030, г. Минск, Беларусь, ssholtanyuk@bsu.by*

В данной статье сравниваются методы совмещения кадров при наблюдении сцены с помощью движущегося летательного аппарата. Для этого сформулирована формальная постановка задачи совмещения двух кадров, рассмотрены и применены некоторые численные и статистические методы её решения. Проведено сравнение этих методов по точности и времени их работы.

Ключевые слова: обработка видеопоследовательностей; компьютерное зрение; совмещение кадров; спутниковые снимки; геонаблюдение.

Введение. Постановка задачи

В современном мире использование беспилотных летательных аппаратов (дронов) для наблюдения за окружающей средой становится все более распространенным [1, 2]. Однако одним из основных вызовов при таком наблюдении является объединение видеоданных с различных камер, установленных на летательном аппарате, для создания цельной и информативной картины сцены. В этом контексте исследователи активно разрабатывают методы и алгоритмы для эффективного совмещения кадров, полученных с помощью дронов. Кроме того, зачастую требуется обработка видеопоследовательностей с целью анализа сцены, за которой невозможно наблюдать при помощи неподвижной камеры или при различных эффектах, вызывающих нежелательное движение камеры [3].

Основными ситуациями, требующими совмещения кадров видеопоследовательности, являются различные движения камеры, например, смещение в сторону (рис. 1), приближение и отдаление (рис. 2), поворот (рис. 3). Геометрические преобразования сцены, продемонстрированные на рис. 1 и 2, обычно имеют место при наблюдении сцены летательными аппаратами, на рис. 3 – при наблюдении со спутников, «висящих» над фиксированной точкой Земли.

Формально задача совмещения двух кадров может быть поставлена следующим образом. Для заданных изображений I_1 и I_2 необходимо найти преобразование φ , минимизирующее ошибку



Рис. 1. Наблюдение сцены при смещении камеры

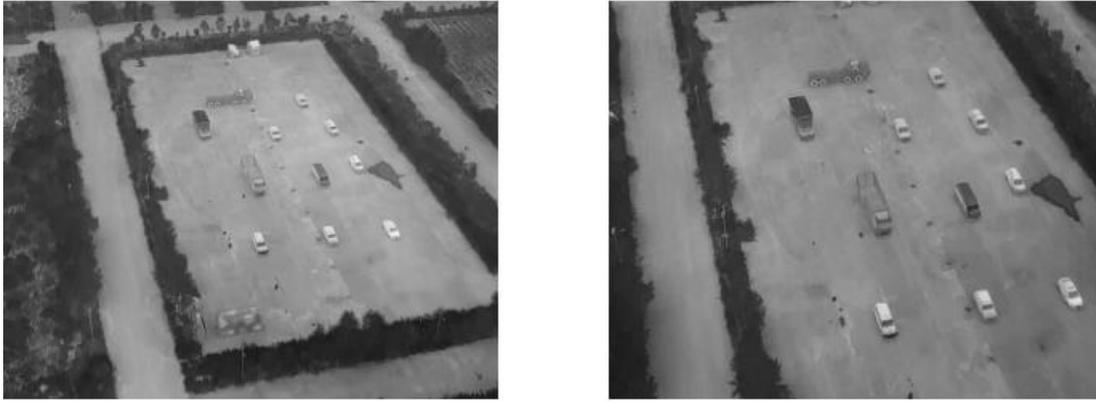


Рис. 2. Наблюдение сцены при приближении камеры

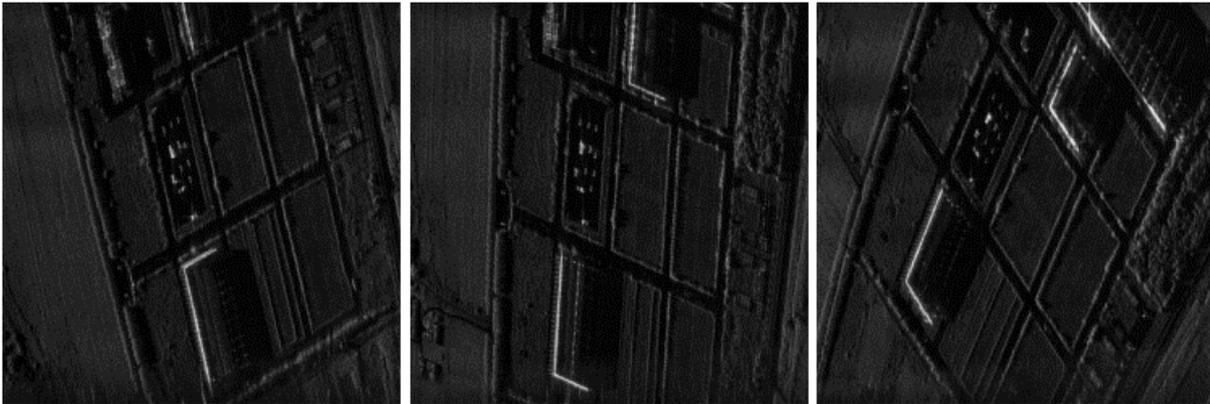


Рис. 3. Наблюдение сцены при повороте камеры

$$\varepsilon = E(\varphi(I_2), I_1),$$

где E – некоторая функция ошибки, которая вычисляется путём сравнения положений одних и тех же объектов на изображениях $\varphi(I_2)$ и I_1 . Обычно искомое преобразование φ выражается некоторой матрицей $\Phi \in \mathbb{R}_{3,3}$ и действует на точку с координатами $(x, y) \in \mathbb{R}^2$ по следующей формуле:

$$\Phi \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \omega x' \\ \omega y' \\ \omega \end{pmatrix} \sim \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix},$$

$$\varphi(x, y) = (x', y').$$

Матрица Φ восстанавливается с точностью до коэффициента по четырём точкам (x_i, y_i) , не лежащим на одной прямой, и их образам (x'_i, y'_i) при преобразовании φ :

$$\begin{pmatrix} \varphi_{11} & \varphi_{12} & \varphi_{13} \\ \varphi_{21} & \varphi_{22} & \varphi_{23} \\ \varphi_{31} & \varphi_{32} & \varphi_{33} \end{pmatrix} \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} \omega_1 x'_1 & \omega_2 x'_2 & \omega_3 x'_3 & \omega_4 x'_4 \\ \omega_1 y'_1 & \omega_2 y'_2 & \omega_3 y'_3 & \omega_4 y'_4 \\ \omega_1 & \omega_2 & \omega_3 & \omega_4 \end{pmatrix} \quad (1)$$

Полагая в уравнении (1) $\varphi_{33} = 1$, получаем, что его можно переписать в виде системы двенадцати линейных алгебраических уравнений относительно двенадцати неизвестных: φ_{ij} , $(i, j) \in \{1, 2, 3\}^2 \setminus \{(3, 3)\}$, ω_i , $i = \overline{1, 4}$. В общем случае она имеет не более одного решения. Следовательно, задача совмещения двух кадров не имеет точного решения и обычно решается численными и статистическими методами.

Вычислительный эксперимент и результаты

В рамках эксперимента рассмотрены четыре монохромные видеопоследовательности, полученные при съёмке местности с летательного аппарата (рис. 1 и 2). Для каждой из них в результате совмещения первого кадра со всеми остальными получены видеопоследовательности с выделенными на них зонами видимости камеры (рис. 4).



Рис. 4. Визуализация изменения зоны видимости камеры при наблюдении сцены с помощью подвижного летательного аппарата

Для сравнения использовались следующие методы:

- RANSAC [4],
- решение матричного уравнения (1) на основе сингулярного разложения матрицы (SVD),
- решение матричного уравнения (1) на основе метода градиентного спуска (GD),
- совмещение кадров по ключевым точкам SURF [5].

При помощи этих методов минимизировалась *средняя абсолютная ошибка выравнивания*:

$$E(\varphi(I_2), I_1) = \frac{1}{wh} \sum_{\substack{i=1,h \\ j=1,w}} \|\varphi(i, j) - \hat{I}_1(i, j)\|_2, \quad (2)$$

где w – ширина кадра, h – его высота, $\varphi(i, j)$ – позиция пикселя изображения I_2 после применения к нему преобразования φ , $\hat{I}_1(i, j)$ – истинная позиция этого же пикселя на изображении I_1 .

На рис. 5 представлена зависимость ошибок (2) (вертикальная ось, на которой отмечены логарифмы ошибок) от номера кадра, совмещаемого с первым (горизонтальная ось) при использовании каждого из четырёх методов для одной из видеопоследовательностей. Результаты совмещения некоторых кадров с первым для каждого из методов приведены на рис. 6 (каждая строка соответствует конкретному кадру исходной видеопоследовательности).

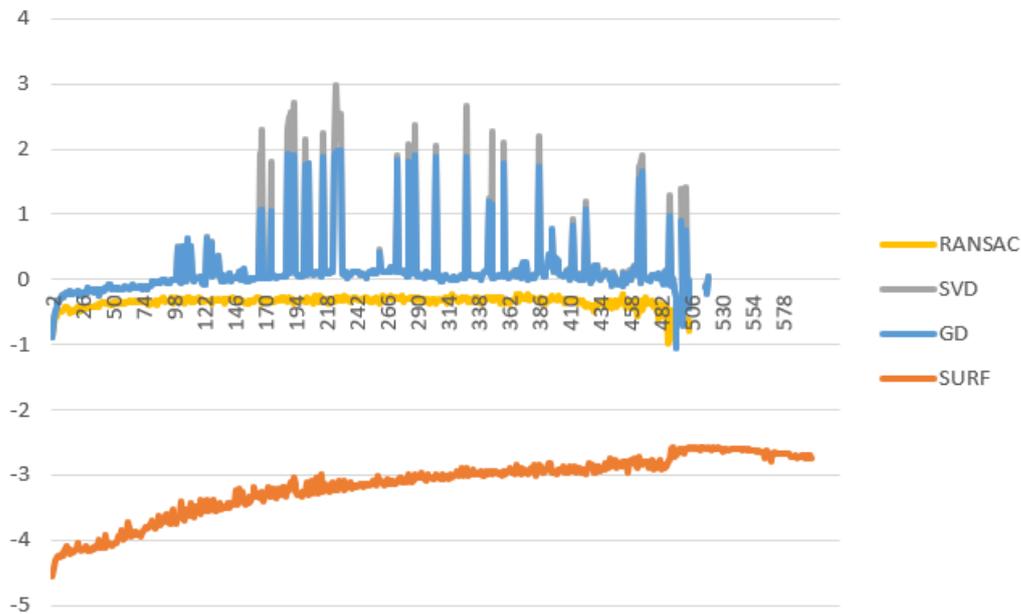


Рис. 5. Сравнение ошибок совмещения кадров при использовании разных методов *

Из графика на рис. 5 видно, что совмещение кадров по ключевым точкам даёт наиболее точные результаты. Большие невязки, полученные при использовании методов SVD и GD, приводят к неадекватному наложению кадров (рис. 6б и 6в, 3-я строка). Кроме того, для большинства кадров, начиная с кадра 495, все методы, кроме SURF, оказались не способны совместить их с исходным кадром, что выражено на графике с рис. 5 отсутствием данных. Это объясняется тем, что зона видимости камеры для последующих кадров слишком мала по сравнению с первым кадром вследствие приближения летательного аппарата к данному участку сцены. Примечательно то, что для этих же кадров совмещение по ключевым точкам также показало неадекватные результаты (рис. 6г, последняя строка).

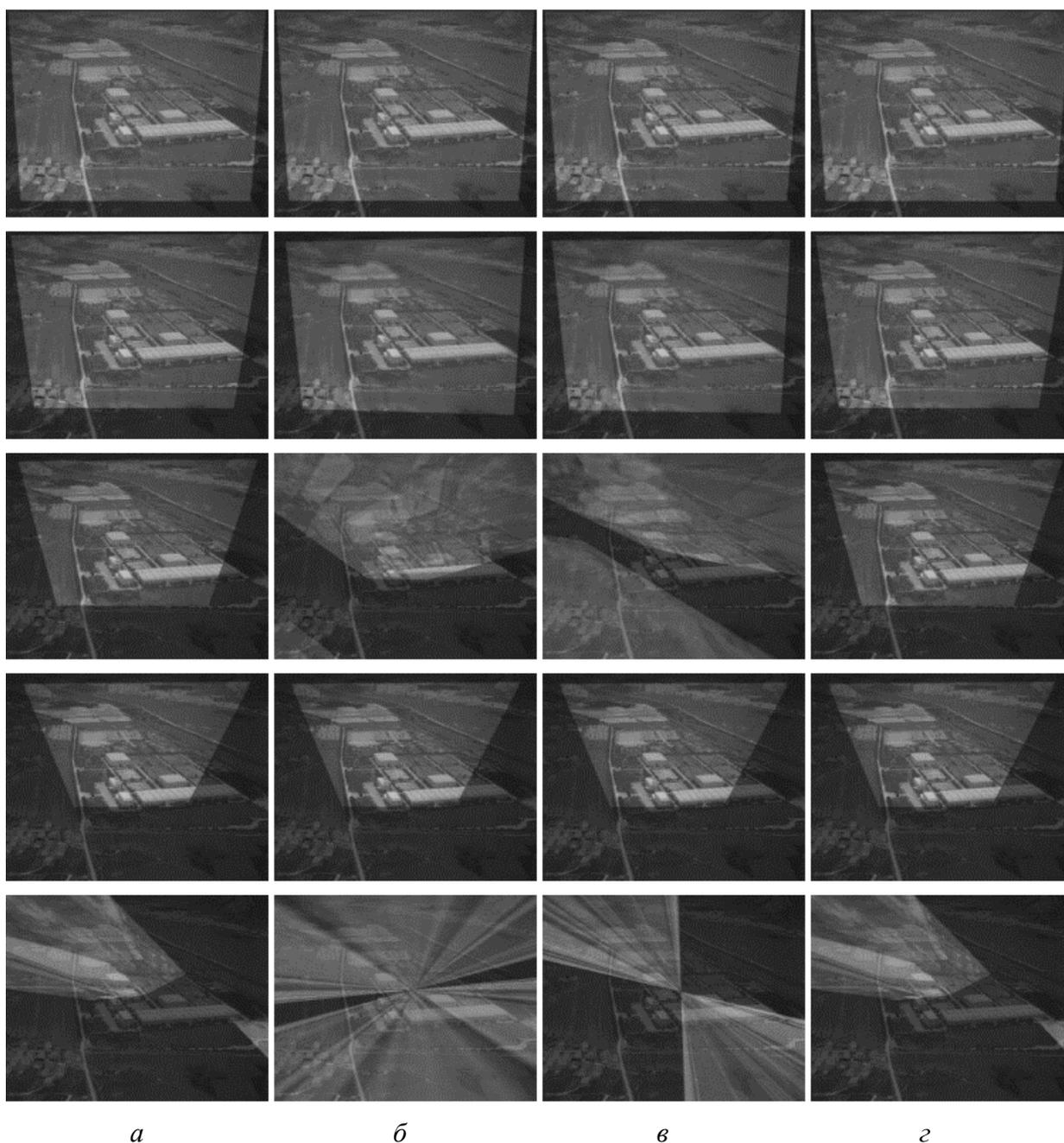


Рис. 6. Результаты совмещения разных кадров видеопоследовательностей с исходным (а – методом RANSAC, б – при помощи сингулярного разложения матрицы, в – методом градиентного спуска, г – при помощи ключевых точек)

Из рис. 6 видно, что в целом методы SVD и GD демонстрируют неадекватное наложение кадров чаще, чем RANSAC и SURF. Кроме того, можно отметить, что результаты при использовании методов RANSAC и SURF несмотря на значимую разницу между ошибками практически не отличаются друг от друга.

Рассмотренные методы показали следующее среднее время совмещения одной пары кадров (в секундах): SURF – 1,45, RANSAC – 0,80, GD – 0,55, SVD – 0,51. С учётом отмеченного выше, таким образом, можно сделать вывод, что для решения задачи совмещения кадров видеопоследовательности наиболее подходящим является RANSAC, который, будучи не самым времязатратным, демонстрирует адекватные результаты. Вместе с тем более медленный метод

с использованием ключевых точек может быть использован для уточнения совмещения отдельных кадров.

Библиографические ссылки

1. Коротаяев А.А, Новопашин Л.А. Применение беспилотных летательных аппаратов для мониторинга сельскохозяйственных угодий и посевных площадей в аграрном секторе // Аграрный вестник Урала. 2015. №12 (142). С. 38-42.

2. Аскеров Э.С., Абдулаева А.А., Ухумаалиева А.М. Перспективы использования беспилотных летательных аппаратов при обследовании земель и земельном надзоре // Аграрное и земельное право. 2022. №2 (206). С. 108-111. DOI: 10.47643/1815-1329_2022_2_108.

3. Буряченко, В.В. Пространственно-ориентированная стабилизация изображений // DSPA: вопросы применения цифровой обработки сигналов. 2016. Т. 6, № 4. С. 763-767.

4. Fischer M. A., Bolles R. C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography // Comm. Of the ACM. 1981. Vol. 24. P. 381-395. DOI: 10.1145/358669.358692.

5. Bay H., Tuytelaars T., Van Gool L. SURF: Speeded Up Robust Features // Computer Vision ECCV 2006 : proceedings of the 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006 / eds.: A. Leonardis, H. Bischof, A. Pinz. Springer, Berlin, Heidelberg, 2006. P. 404-417. DOI: 10.1007/11744023_32.