

УДК 004.89:531.395

DOI 10.52928/2070-1624-2024-42-1-26-33

## МЕТОДЫ ОБРАБОТКИ ВИДЕОПОСЛЕДОВАТЕЛЬНОСТЕЙ СО СКОПЛЕНИЯМИ ЛЮДЕЙ ДЛЯ ОПРЕДЕЛЕНИЯ ЗАКОНОМЕРНОСТЕЙ ИХ ДВИЖЕНИЯ

С. В. ШОЛТАНЮК<sup>1</sup>, Ц. БУ<sup>2</sup>, д-р техн. наук, доц. А. М. НЕДЗЬВЕДЬ<sup>1,3</sup>

<sup>1</sup>(Белорусский государственный университет, Минск)

<sup>2</sup>(Нанкинский научно-исследовательский институт электроники, Китай)

<sup>3</sup>(Объединенный институт проблем информатики  
Национальной академии наук Беларуси, Минск)

*Движение скоплений однородных объектов на сегодняшний день является одним из наиболее важных и быстро развивающихся приложений компьютерного зрения и машинного обучения. В данной статье рассматривается вопрос определения закономерностей движения людей и их скоплений посредством карт движений, вычисляемых при помощи нейронной сети FlowNet, которая рассматривает движение объектов на видеопоследовательности. Этот подход позволяет получить информацию о направлении и скорости движения толпы относительно других объектов сцены, играя ключевую роль в анализе поведения и обеспечении безопасности. Кроме того, рассматриваются методы предварительной обработки видеопоследовательностей, включая совмещение кадров, для более точного определения карты движений и повышения эффективности анализа динамических сцен.*

**Ключевые слова:** динамический объект, скопление людей, оптический поток, движение скопления объектов, карты движения.

**Введение.** Анализ движения людей на видеопоследовательностях является важной задачей в области компьютерного зрения и обработки изображений. Он находит применение в различных областях, включая видеонаблюдение, управление транспортом, безопасность и анализ поведения толпы. Определение закономерностей движения людей в скоплениях на видеопоследовательностях представляет собой сложную задачу из-за изменчивости условий наблюдения (в т. ч. освещения, положения камеры), разнообразия действий и поведения объектов, а также ограниченной видимости. В последние годы было предложено множество методов и технологий, решающих разные задачи, связанные с наблюдением за скоплениями людей, анализом их движения и поведения, которые опираются в первую очередь на модели компьютерного зрения и машинного обучения. Особенно широко в этой области стали применяться нейросетевые методы благодаря своей способности к извлечению сложных пространственно-временных признаков из видеоданных.

**Движение множества динамических объектов на последовательности кадров.** Большинство методов анализа видеопоследовательностей с движущимися скоплениями объектов рассматривают вопрос движения отдельных компонент этих скоплений, что не позволяет выделить характер и свойства движения самих скоплений как единого целого [1]. Элементарный динамический объект – это небольшой подвижный локализуемый объект с такими физическими параметрами, как объём, плотность и масса. Движение таких объектов может быть вращательным, прямолинейным, ускоренным или даже едва заметным. Сложность обнаружения и определения характера движения определяется размером, формой и самой природой движения объектов, а также изменения этих характеристик.

Динамические объекты могут быть разделены на следующие категории:

- отдельные небольшие объекты, которые являются подвижными компонентами фона и подлежат удалению;
- большие компоненты фона, движение которых обусловлено движением камеры, с которой получена видеопоследовательность. При их наличии необходима предварительная обработка видеопоследовательности (стабилизация, совмещение кадров);
- фрагменты окружающей среды вокруг движущихся объектов, у которых изменяются оптические характеристики (цвет, яркость, насыщенность и др.).

Сами скопления отдельных объектов также можно рассматривать как движущиеся объекты. Их движение может быть отнесено к какой-либо из следующих категорий:

- собственно движение;
- внутреннее смещение компонент скопления;
- изменение формы скопления.

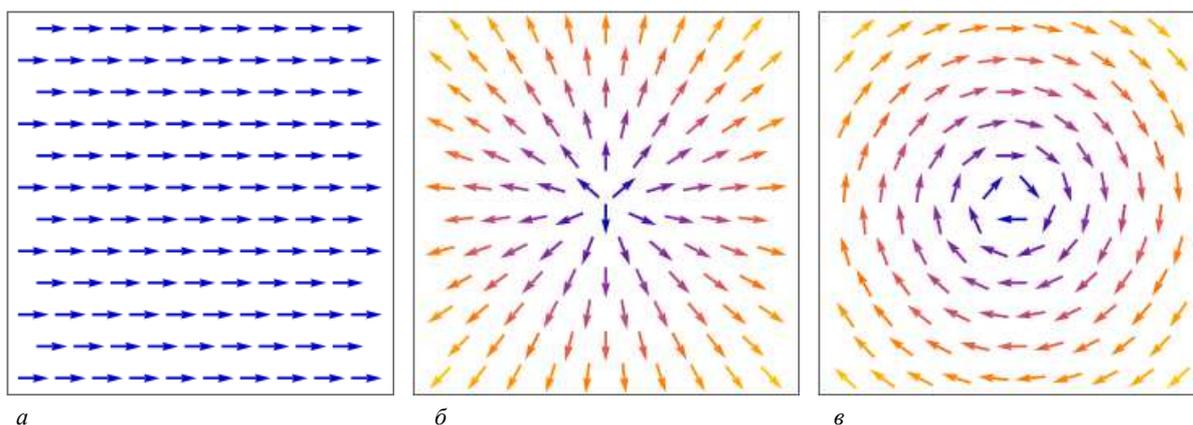
Движение в трёхмерном пространстве можно представить прямым и обратным оптическим потоком, представляющим собой трёхмерное векторное поле.

Непрерывная сцена может быть представлена векторной функцией от 5 аргументов: координат рассматриваемой точки  $\vec{x} = (x, y, z)$  и направления движения, характеризуемого двумя направляющими углами  $(\theta, \varphi)$ . Результатом этой функции является цвет пикселя, который может быть представлен в цифровой модели RGB трёхмерным вектором  $\vec{c} = (r, g, b)$ . На практике движение обычно представляется трёхмерным вектором  $\vec{d}$  с координатами, записанными в некоторой декартовой системе координат. Таким образом, необходимо задать отображение  $\Theta : (\vec{x}, \vec{d}) \mapsto \vec{c}$ .

Одной из основных задач при наблюдении за динамической сценой является отделение фона от рассматриваемых объектов, характеристики которых могут не изменяться или изменяться со временем (форма, положение относительно сцены и т. д.). Учитывая это, задача вычисления направления движения объектов может быть представлена следующими этапами: движение фона, вызванное движением камеры; движение фона вследствие изменений окружающего пространства; движение отдельных объектов; движение внутри скоплений; движение скоплений.

Таким образом, необходимо разделить движение, обнаруженное на видеопоследовательности и представленное в виде векторного поля, по разным категориям по типу движения объектов и/или фона на том или ином участке сцены. Для решения этой задачи часто используются карты движения, которые строятся на основе оптического потока [2; 3].

При определённом движении оптической системы, запечатлевающей сцену, фон может быть представлен в форме постоянного потока, которому соответствует уникальное изображение. В зависимости от характера движения камеры это изображение обладает своими собственными характеристиками (рисунок 1).



*a* – движение камеры влево параллельно сцене, т. е. смещение сцены вправо относительно камеры;  
*б* – приближение камеры к сцене; *в* – поворот камеры

Рисунок 1. – Компоненты оптического потока при движении камеры

**Движение людей и их скопления.** Скопления людей в целом удовлетворяют вышеуказанным закономерностям. Основной особенностью их движения является их перемещение по некоторой (как правило, плоской) поверхности: улица, помещения и т. д. Одним из наиболее используемых инструментов для исследования движения толпы являются искусственные нейронные сети [4–6]. На сегодняшний день существует множество наборов данных из изображений и/или видеопоследовательностей с запечатлёнными на них скоплениями людей, на основе которых можно обучить такие нейронные сети с тем, чтобы получить необходимые характеристики толпы. Так, набор данных HAJJ [7] содержит 18 видео длительностью 20–25 секунд с движущимися скоплениями людей. Эти скопления имеют разную плотность и демонстрируют разные закономерности движения.

Одним из инструментов получения карт движения сцены является FlowNet – искусственная нейронная сеть для оценивания оптического потока по соседним кадрам видеопоследовательности [8]. Она основана на работе двух параллельных ветвей, определяющих разные уровни перемещения, что позволило учитывать все виды движения.

С помощью FlowNet можно оценить оптический поток сцены, отдельных объектов, их скоплений и фона, представленного цветами, насыщенность которых зависит от скорости, оттенки – от направления движения. Другими словами, вместо отображения  $\Theta : (\vec{x}, \vec{d}) \mapsto \vec{c}$  на самом деле оценивается отображение  $\Theta' : (\vec{x}, \vec{d}) \mapsto \vec{c}'$ , где  $\vec{c}'(h, s, v)$  – вектор, представляющий цвет в модели HSV. Для карты движения,

получаемой из FlowNet, яркость  $v$  произвольного пикселя всегда равна 1,  $s = s(\vec{x}, |\vec{d}|)$ ,  $h = h(\vec{x}, \vec{d}/|\vec{d}|)$ .

Таким образом, можно представить искомое отображение ещё одним способом:

$$\Theta' : (\vec{x}, m, \vec{d}_N) \mapsto (h, s),$$

где  $m = |\vec{d}|$  – длина вектора перемещения объекта, которая обычно интерпретируется как его скорость;

$\vec{d}_N = \vec{d}/|\vec{d}|$  – направление движения объекта.

В результате применения этой нейронной сети получились кадры, примеры которых приведены на рисунке 2.

На этих рисунках представлены следующие закономерности движения отдельных людей и их скоплений:

1. Движение толпы в определённом направлении. Так, на рисунке 2, *а* можно выделить два скопления людей, представленных жёлто-зелёной (слева) и фиолетовой (справа) областями. Это два скопления людей, которые движутся соответственно влево-вниз и вверх с точки зрения наблюдателя. Этот тип движения можно описать отображением вида

$$\Theta_{\text{dir}} : (\vec{x}, m, \vec{d}_0) \mapsto (h_0, s),$$

где  $\vec{d}_0 = \text{const}$ ;  $h_0 = \text{const}$  – соответственно направление движения скопления людей и соответствующий этому направлению оттенок этого скопления на карте движения.

2. Движение людей в плотной толпе. На рисунке 2, *б* карта движения состоит преимущественно из ярких цветов, что свидетельствует о примерно одинаковой скорости объектов по всей сцене. Кроме того, встречаются различные цвета, между которыми, впрочем, имеют место плавные переходы, что указывает на отсутствие аномальных закономерностей, таких как столкновение, рассеяние и т. д. Значит, подобное движение можно описать следующим отображением:

$$\Theta_{\text{dense}} : (\vec{x}, m_0, \vec{d}_N) \mapsto (h, 1),$$

где  $m_0 = \text{const}$  – скорость передвижения людей внутри толпы, которой соответствуют максимально насыщенные цвета на карте движения. Кроме того, должно выполняться

$$\lim_{\Delta\vec{x} \rightarrow 0} \Delta\vec{d}_N = \vec{0}. \quad (1)$$



Рисунок 2. – Карты движений для некоторых кадров видеопоследовательностей набора данных НАЖЖ

3. Движение плотной толпы в разные стороны. На рисунке 2, в в силу плавных цветовых переходов нельзя однозначно разметить отдельные скопления людей, однако очевидно, что они движутся в разных направлениях (преимущественно вправо-вниз, вниз и влево-вниз относительно наблюдателя). Значит, для динамической системы должно выполняться равенство (1). Кроме того, подобное движение может быть вызвано перспективными искажениями при наблюдении движения толпы на небольшом расстоянии. В этой ситуации на изображении можно задать точку схода с координатами  $\vec{x}_{cx}$ . Тогда направление движения объекта зависит от его местоположения относительно точки схода:

$$\vec{d}_N = \vec{d}_N(\vec{x}) \parallel (\vec{x} - \vec{x}_{cx}).$$

Все эти закономерности движения можно отнести к нормальному поведению толпы. На рисунке 2, а можно также заметить отдельные области, значительно отличающиеся по цвету от сопредельных областей. Это может свидетельствовать о движении отдельных людей в направлении, отличном от движения основного скопления людей (движение «против течения», пересечение движущегося потока людей). В некоторых ситуациях подобное движение может рассматриваться как аномальное.

Кроме того, в некоторых случаях по карте движения можно обнаружить необходимость дополнительных действий для обработки видеопоследовательности. Например, на рисунке 2, с с одной стороны, отчётливо видно ярко-фиолетовое скопление внизу, соответствующее движению людей вверх и вправо-вверх относительно наблюдателя. С другой стороны, верхняя часть этой карты движения окрашена в тусклые сине-фиолетовые тона. При сверке с исходными кадрами оказывается, что имело место параллельное смещение камеры и вектор этого смещения (см. рисунок 1, а) оказался добавленным к истинному оптическому потоку:

$$\vec{d}(\vec{x}) = \vec{d}_{mot}(\vec{x}) - \vec{d}_{disp},$$

где  $\vec{d}_{mot}$  – вектор перемещения объекта с начальным положением в точке с координатами  $\vec{x}$ ;  
 $\vec{d}_{disp}$  – вектор смещения камеры относительно сцены.

**Постановка задачи совмещения кадров видеопоследовательности.** Одним из способов предварительной обработки видео для предотвращения воздействия внешних факторов, в первую очередь движения камеры (её дрожание, перемещение, приближение, отдаление и т. д.), является совмещение кадров видеопоследовательности. Формально задача совмещения двух изображений может быть поставлена следующим образом. Для двух изображений  $I_1$  (исходный кадр) и  $I_2$  (искажённый) необходимо найти преобразование  $\Phi$ , максимизирующее меру схожести изображений  $R(\Phi(I_2), I_1)$ . В качестве меры схожести в данной статье использована функция  $R = R^{SSD}$  на основе суммы квадратов разностей [9]:

$$R^{SSD}(I_1, I_2) = 1 - \frac{1}{Lwh} \sum_{\substack{i=1,h \\ j=1,w}} (I_{1ij} - I_{2ij})^2, \quad (2)$$

где  $w$  и  $h$  – соответственно ширина и высота исходного кадра;

$I_{1ij}$  и  $I_{2ij}$  – пиксели в  $i$ -й строке,  $j$ -м столбце двух соответствующих изображений;

$L$  – диапазон допустимых значений яркости.

Для сравнения разных методов использовалось значение ошибки, равное  $1 - R^{SSD}(I_1, I_2)$ .

Обычно искомое преобразование выражается матрицей  $\Phi \in \mathbb{R}_{3,3}$  и действует на пиксель с координатами  $(x, y)$  следующим образом:

$$\Phi \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} \omega x' \\ \omega y' \\ \omega \end{pmatrix}, \quad (3)$$

что даёт искомое положение этого пикселя, выраженное координатным столбцом  $(\omega x', \omega y', \omega)^T$  в однородных координатах. Из него легко получить декартовы координаты  $(x', y')$  путём домножения на  $\omega^{-1}$  и взятия первых двух элементов полученного столбца. Несложно видеть, что матрица  $\Phi$  определяется с точностью

до множителя. Поэтому обычно её нижний правый элемент полагают равным 1. Применяя формулу (3) к некоторым четырём пикселям искажённого изображения  $(x_i, y_i)$ ,  $i = \overline{1,4}$ , имеем следующее уравнение:

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

Оно эквивалентно системе из двенадцати линейных алгебраических уравнений относительно двенадцати неизвестных (восемь элементов матрицы  $\Phi$  и четыре коэффициента  $\omega$ ), которая, вообще говоря, имеет не более одного решения. Таким образом, точного решения задача совмещения двух кадров не имеет, и необходимо использование приближённых методов.

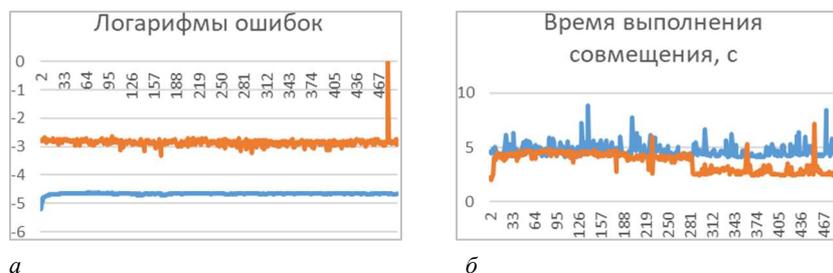
**Методы совмещения кадров видеопоследовательности с динамическими скоплениями людей.** Одним из наиболее распространённых методов совмещения кадров является RANSAC – статистический метод, оценивающий адекватность модели путём проверки статистических гипотез, согласно которым все данные определяются как нормальные либо аномальные [10], а также совмещение по ключевым точкам [11]. Из всех методов определения ключевых точек изображений в данной статье использован метод ORB, работающий быстрее за счёт отсутствия инвариантности относительно поворота изображения [12; 13]. На практике кадры видеопоследовательностей, полученных в результате съёмки сцены при помощи зафиксированных камер (в частности, из набора данных HAJJ), можно считать не претерпевающими подобных искажений.

В наборе данных HAJJ присутствуют четыре видео с заметным движением камеры. Для каждого из них проведено совмещение каждого кадра с первым (исходным) для получения стабилизированной последовательности кадров (рисунок 3).



Рисунок 3. – Пример совмещения некоторого кадра видеопоследовательности с исходным

Полученные значения для меры схожести (2), а также время выполнения методов для каждого кадра приведены на рисунках 4. Исходя из этих данных, можно сделать вывод о том, что использование ключевых точек по методу ORB даёт лучшую точность совмещения кадров. Это подтверждается и самими результатами совмещения (на рисунке 5 в средней строке представлено совмещение одного кадра с исходным, для которого метод RANSAC демонстрирует неадекватный результат). Вместе с тем, исходя из графика на рисунке 4, б, метод RANSAC работает быстрее совмещения кадров по ключевым точкам (в среднем на 27 %). Кроме того, как было отмечено выше, метод ORB работает без учёта ситуаций, когда камера поворачивается вокруг своей фокальной оси. Соответственно, метод RANSAC может применяться, например, при наблюдении сцены при помощи спутников.



*a* – динамика десятичных логарифмов ошибок  $1 - R^{SSD}(I_1, I_2)$  в зависимости от кадра, совмещаемого с исходным; *б* – время выполнения совмещения каждого кадра с исходным

Рисунок 4. – Результаты работы методов совмещения кадров (оранжевая линия – для метода RANSAC, синяя – для метода ORB)



*a* – совмещение кадров методом RANSAC; *б* – совмещение по ключевым точкам ORB

Рисунок 5. – Сравнение результатов наложения некоторых кадров на исходный при использовании разных методов

**Заключение.** В данной статье были представлены основные аспекты классификации закономерностей движения людей и их скоплений на видеопоследовательностях. Использование нейронной сети FlowNet для вычисления карт движений позволяет получать информацию о различных характеристиках движения толпы, в первую очередь направления и скорости относительно других объектов сцены, что является важным шагом в понимании и анализе динамических объектов в сценах с большим количеством людей. Тем не менее такой подход требует предварительной обработки видеопоследовательности. Среди процедур прежде всего можно выделить совмещение кадров, которое для видеопоследовательностей со скоплениями людей наиболее эффективно проводить путём обнаружения и совмещения ключевых точек, которое даёт быстрые и точные результаты.

Исследование выполнено при финансовой поддержке проекта “Technology Development Agreement of developing of algorithms of remote sensing image processing”, Agreement number: 22CETC19-ICN1785.

#### ЛИТЕРАТУРА

1. Ascione F., Caserta S., Perris R. et al. Investigation of Cell Dynamics in vitro by Time Lapse Microscopy and Image Analysis // *Chemical Engineering*. – 2014. – Vol. 38. – P. 517–522. – DOI: [10.3303/CET1438087](https://doi.org/10.3303/CET1438087).
2. Chen, C., Ye, S., Chen, H. et al. Integral optical flow and its application for monitoring dynamic objects from a video sequence // *J. of Applied Spectroscopy*. – 2017. – Vol. 84. – P. 120–128. – DOI: [10.1007/s10812-017-0437-z](https://doi.org/10.1007/s10812-017-0437-z).
3. Sholtanyuk S. Crowd Abnormal Behaviour Patterns: Survey and Detection // *Central European Researchers Journal*. – 2023. – Vol. 9, iss. 1. – P. 48–58, URL: [https://ceres-journal.eu/download.php?file=2023\\_01\\_7.pdf](https://ceres-journal.eu/download.php?file=2023_01_7.pdf) (дата обращения 01.04.2024).
4. Altowairqi S., Luo S., Greer, P. A Review of the Recent Progress on Crowd Anomaly Detection // *Intern. J. of Advanced Computer Science and Applications*. – 2023. – Vol. 14, iss. 4. – P. 659–669. – DOI: [10.14569/IJACSA.2023.0140472](https://doi.org/10.14569/IJACSA.2023.0140472).
5. Choudhry N., Abawajy J., Huda S. et al. Comprehensive Survey of Machine Learning Methods for Surveillance Videos Anomaly Detection // *IEEE Access*. – 2023. – Vol. 11. – P. 114680–114713. – DOI: [10.1109/ACCESS.2023.3321800](https://doi.org/10.1109/ACCESS.2023.3321800).
6. Miao Y., Yang J., Alzahrani B. et al. Abnormal Behavior Learning Based on Edge Computing toward a Crowd Monitoring System // *IEEE Network*. – 2022. – Vol. 36, iss. 3. – P. 90–96. – DOI: [10.1109/MNET.014.2000523](https://doi.org/10.1109/MNET.014.2000523).
7. Alafif, T., Alzahrani, B., Cao, Y. et al. Generative adversarial network based abnormal behavior detection in massive crowd videos: a Hajj case study // *J. of Ambient Intelligence and Humanized Computing*. – 2022. – Vol. 13, iss. 8. – P. 4077–4088. – DOI: [10.1007/s12652-021-03323-5](https://doi.org/10.1007/s12652-021-03323-5).
8. Ilg, E., Mayer, N., Saikia, T. et al. (2017). FlowNet 2.0: Evolution of optical flow estimation with deep networks // *Proc. of the IEEE conf. on computer vision and pattern recognition / Honolulu, HI, USA (21–26 July 2017)*. – 2017. – P. 2462–2470. – DOI: [10.1109/CVPR.2017.179](https://doi.org/10.1109/CVPR.2017.179).
9. Чертков В. М., Железняк, В. К. Алгоритм определения меры схожести идентификационных образов закладных устройств // *Вестн. Полоц. гос. ун-та. Сер. С, Фундам. науки*. – 2018. – № 4. – С. 20–27.
10. Fischler M. A., Bolles, R. C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography // *Communications of the ACM*. – 1981. – Vol. 24, iss. 6. – P. 381–395, DOI: [10.1145/358669.358692](https://doi.org/10.1145/358669.358692).
11. Liu, C., Xu, J., & Wang, F. A Review of Keypoints' Detection and Feature Description in Image Registration // *Scientific programming*. – 2021. – Vol. 2021. – P. 8509164:1–8509164:25. – DOI: [10.1155/2021/8509164](https://doi.org/10.1155/2021/8509164).
12. Rublee E., Rabaud V., Konolige K. et al. G. ORB: An efficient alternative to SIFT or SURF // *2011 Intern. Conf. on Computer Vision / Barcelona, Spain (06–13 Nov. 2011)*. – 2011. – P. 2564–2571. – DOI: [10.1109/ICCV.2011.6126544](https://doi.org/10.1109/ICCV.2011.6126544).
13. Andersson O., Reyna Marquez S. A comparison of object detection algorithms using unmanipulated testing images: Comparing SIFT, KAZE, AKAZE and ORB. – 2016. – 31 p. – URL: <https://www.diva-portal.org/smash/rec-ord.jsf?pid=diva2%3A927480&dswid=9533>. (дата обращения 01.04.2024).

#### REFERENCES

1. Ascione, F., Caserta, S., Perris, R., & Guidoa, S. (2014). Investigation of Cell Dynamics in vitro by Time Lapse Microscopy and Image Analysis. *Chemical Engineering*, (38), 517–522. DOI: [10.3303/CET1438087](https://doi.org/10.3303/CET1438087).
2. Chen C., Ye S., Chen H., Nedzvedz, O. V., & Ablameyko, S. V. (2017). Integral Optical Flow and its Application for Monitoring Dynamic Objects from a Video Sequence. *J. of Applied Spectroscopy*, (84), 120–128. DOI: [10.1007/s10812-017-0437-z](https://doi.org/10.1007/s10812-017-0437-z).
3. Sholtanyuk, S. (2023). Crowd Abnormal Behaviour Patterns: Survey and Detection. *Central European Researchers Journal*, 9(1), 48–58. [https://ceres-journal.eu/download.php?file=2023\\_01\\_7.pdf](https://ceres-journal.eu/download.php?file=2023_01_7.pdf).
4. Altowairqi, S., Luo, S., & Greer, P. (2023). A Review of the Recent Progress on Crowd Anomaly Detection. *Intern. J. of Advanced Computer Science and Applications*, 14(4), 659–669. DOI: [10.14569/IJACSA.2023.0140472](https://doi.org/10.14569/IJACSA.2023.0140472).
5. Choudhry, N., Abawajy, J., Huda, S., & Rao, I. (2023). A Comprehensive Survey of Machine Learning Methods for Surveillance Videos Anomaly Detection. *IEEE Access*, (11), 114680–114713. DOI: [10.1109/ACCESS.2023.3321800](https://doi.org/10.1109/ACCESS.2023.3321800).
6. Miao, Y., Yang, J., Alzahrani, B., Lv, G., Alafif, T., Barnawi, A., & Chen, M. (2022). Abnormal Behavior Learning Based on Edge Computing toward a Crowd Monitoring System. *IEEE Network*, 36(3), 90–96. DOI: [10.1109/MNET.014.2000523](https://doi.org/10.1109/MNET.014.2000523).
7. Alafif, T., Alzahrani, B., Cao, Y., Alotaibi, R., Barnawi, A., & Chen, M. (2022). Generative adversarial network based abnormal behavior detection in massive crowd videos: a Hajj case study. *J. of Ambient Intelligence and Humanized Computing*, 13(8), 4077–4088. DOI: [10.1007/s12652-021-03323-5](https://doi.org/10.1007/s12652-021-03323-5).

8. Ilg, E., Mayer, N., Saikia, T., Keuper, M., Dosovitskiy, A., & Brox, T. (2017). FlowNet 2.0: Evolution of optical flow estimation with deep networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2462–2470). IEEE. DOI: [10.1109/CVPR.2017.179](https://doi.org/10.1109/CVPR.2017.179).
9. Chertkov, V. M., & Zheleznyak, V. K. (2018). Algoritm opredeleniya mery shozhesti identifikatsionnykh obrazov zakladnykh ustrojstv [Algorithm for Determining the Degree of Similarity of Identification Images from Secret Intelligence Device]. *Vestnik Polotskogo gosudarstvennogo universiteta. Seriya C, Fundamental'nye nauki [Herald of Polotsk State University. Series C. Fundamental sciences]*, (4), 20–27. (In Russ., abstr. in Engl.).
10. Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395. DOI: [10.1145/358669.358692](https://doi.org/10.1145/358669.358692).
11. Liu, C., Xu, J., & Wang, F. (2021). A review of keypoints' detection and feature description in image registration. *Scientific programming*, (2021), 8509164:1–8509164:25. DOI: [10.1155/2021/8509164](https://doi.org/10.1155/2021/8509164).
12. Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. In *2011 Intern. Conf. on Computer Vision* (2564–2571). IEEE. DOI: [10.1109/ICCV.2011.6126544](https://doi.org/10.1109/ICCV.2011.6126544).
13. Andersson, O., & Reyna Marquez, S. (2016). A comparison of object detection algorithms using unmanipulated testing images: Comparing SIFT, KAZE, AKAZE and ORB. <https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A927480&dswid=9533>.

Поступила 02.04.2024

## CROWD VIDEO SEQUENCES PROCESSING METHODS FOR DETERMINING THE CROWD MOTION PATTERNS

S. SHOLTANYUK<sup>1</sup>, Q. BU<sup>2</sup>, A. NEDZVED<sup>1,3</sup>

<sup>1</sup>(Belarusian State University, Minsk)

<sup>2</sup>(Nanjing Research Institute of Electronics Engineering, China)

<sup>3</sup>(United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk)

Nowadays, homogeneous objects clusters motion is one of the most important and rapidly developing computer vision and machine learning application. In this paper, we consider the crowd motion patterns determination by using motion maps that we calculate with FlowNet, a neural network examining motion of objects in a video sequence. This approach allows us to get information on the crowd direction and velocity with relation to other objects of scene, which plays the key role in behavior analysis and security establishment. Besides, we consider methods for preliminary video sequence processing, including frame combination, to estimate motion maps more precisely and improve the effectiveness of the dynamic scenes analysis.

**Keywords:** dynamic object, crowd, optical flow, objects cluster motion, motion maps.