

**КЛАССИФИКАЦИЯ ЗАКОНОМЕРНОСТЕЙ ДВИЖЕНИЯ ЛЮДЕЙ И ИХ СКОПЛЕНИЙ
НА ВИДЕОПОСЛЕДОВАТЕЛЬНОСТИ**

С. В. ШОЛТАНЮК

(Белорусский государственный университет, г. Минск);

Цин БУ

(Нанкинский научно-исследовательский институт электроники, Китай);

д-р техн. наук, доц. А. М. НЕДЗЬВЕДЬ

*(Белорусский государственный университет, Объединённый институт
проблем информатики НАН Беларуси, г. Минск)*

Аннотация. *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 движения людей и их скоплений посредством карт движений, вычисляемых при помощи нейронной сети FlowNet, и на основе которых можно судить о различных характеристиках движения и поведения толпы (направление, скорость, изменение структуры и формы и т.д.).*

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

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.).

Скопления отдельных объектов сами 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, представляющий собой трёхмерное векторное поле.

Непрерывная сцена может быть представлена векторной функцией от 5 аргументов: координаты рассматриваемой точки $\vec{x} = (x, y, z)$, и направление движения, характеризуемое двумя направляющими углами (θ, φ) . Результатом этой функции является цвет пикселя, который может быть представлен в цифровой модели RGB 3-мерным 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 $\Theta: (\vec{x}, \vec{d}) \mapsto \vec{c}$.

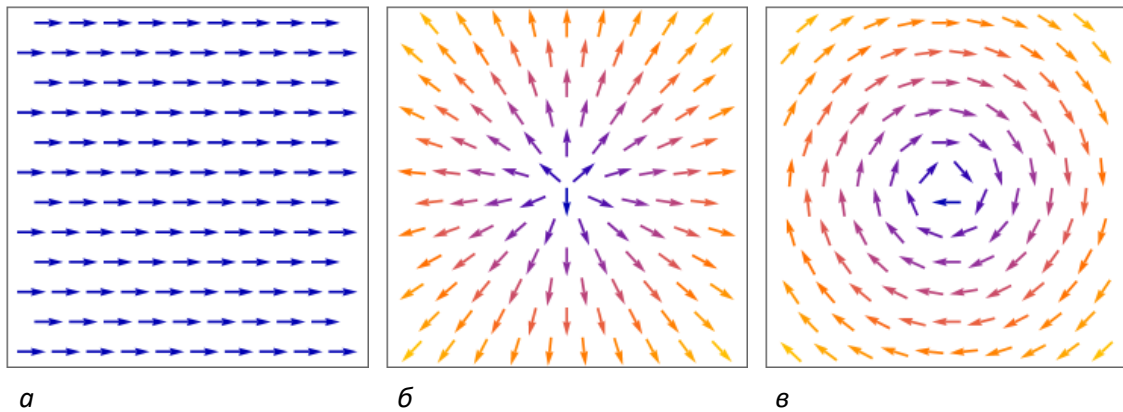
Одной из основных задач при наблюдении за динамической сценой является отделение фона от рассматриваемых объектов, характеристики которых могут не изменяться или изменяться со временем (форма, положение относительно сцены и т.д.). Учитывая это, задача вычисления направления движения объектов может быть 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).

Движение людей и их скоплений. 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 НАJJ [7] содержит 18 видео длительностью 20–25 секунд с движущимися скоплениями людей. Эти скопления имеют разную плотность и демонстрируют разные закономерности движения.



а – 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

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, который представлен цветами, насыщенность которых зависит от скорости, оттенок – 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, $s = s(\vec{x}, |\vec{d}|)$, $h = h\left(\vec{x}, \frac{\vec{d}}{|\vec{d}|}\right)$. Таким образом, можно представить искомое отображение ещё одним образом:

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

где $m = |\vec{d}|$ – длина вектора перемещения объекта, которая обычно интерпретируется как его скорость, $\vec{d}_N = \frac{\vec{d}}{|\vec{d}|}$ – направление движения объекта.

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



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

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

– 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.

– 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

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

– 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 $\vec{x}_{\text{сх}}$. Тогда направление движения объекта зависит от его местоположения относительно точки схода:

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

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, г, 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:

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

где \vec{d}_{mot} – вектор перемещения объекта с начальным положением в точке с координатами \vec{x} ;

\vec{d}_{disp} – вектор смещения камеры относительно сцены.

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.

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. Investigation of cell dynamics *in vitro* by time lapse microscopy and image analysis / F. Ascione [et al.] // Chemical Engineering. – 2014. – Vol. 38. – P. 517–522. – DOI: 10.3303/CET1438087.
2. Chen, Ch. Integral Optical Flow and its Application for Monitoring Dynamic Objects from a Video Sequence / Ch. Chen [et al.] // Journal of Applied Spectroscopy. – 2018. – Vol. 84. – P. 120–128. – DOI: 10.1007/s10812-017-0437-z.
3. Sholtanyuk, S. Crowd Abnormal Behaviour Patterns: Survey and Detection / S. Sholtanyuk // Central European Researchers Journal. – 2023. – Vol. 9, Iss. 1. – P. 48–58.
4. Altowairqi, S. A Review of the Recent Progress on Crowd Anomaly Detection / S. Altowairqi, S. Luo, P. Greer // International Journal of Advanced Computer Science and Applications. – 2023. – Vol. 14, No. 4. – P. 659–669. – DOI: 10.14569/IJACSA.2023.0140472.
5. Choudhry, N. A Comprehensive Survey of Machine Learning Methods for Surveillance Videos Anomaly Detection / N. Choudhry [et al.] // IEEE Access. – 2023. – Vol. 11. – P. 114680–114713. – DOI: 10.1109/ACCESS.2023.3321800.
6. Miao, Y. Abnormal Behavior Learning Based on Edge Computing toward a Crowd Monitoring System / Y. Miao [et al.] // IEEE Network. – 2022. – Vol. 36, Iss. 3. – P. 90–96. – DOI: 10.1109/MNET.014.2000523.
7. Alafif, T. Generative adversarial network based abnormal behavior detection in massive crowd videos: a Hajj case study / T. Alafif [et al.] // Journal of Ambient Intelligence and Humanized Computing. – 2022. – Vol 13. – P. 4077–4088. – DOI: 10.1007/s12652-021-03323-5.
8. Ilg, E. FlowNet 2.0: Evolution of Optical Flow Estimation With Deep Networks / E. Ilg [et al.] // Computer Vision and Pattern Recognition (CVPR) : Proceedings of the IEEE Conference, Honolulu, 22–25 July 2017 / Hawaii Convention Center. – P. 2462–2470.