

Алгоритм множественного трекинга лабораторных животных*

Ветров Д. П., Кропотов Д. А.

vetrovd@yandex.ru, dkropotov@yandex.ru

Москва, ВМиК МГУ, Вычислительный Центр РАН

В работе рассматривается задача сопровождения (трекинга) нескольких лабораторных мышей, находящихся в клетке со встроенной системой видеонаблюдения. Предложены алгоритмы отделения изображений мышей от фона, выделения отдельных особей и их последующей идентификации. Также в работе предложен специальный способ тестирования алгоритма множественного сопровождения.

Введение

Одним из основных инструментов в современных когнитивных исследованиях являются методы анализа поведения человека или животного. Построение описания поведения вручную является очень трудоемким процессом, поэтому все большую популярность получают системы автоматического анализа поведения. Такие системы, как правило, включают в себя модуль видеонаблюдения, модуль сегментации видеосигнала (выделения в нем элементарных структурных единиц поведения — поведенческих актов) и модуль поиска закономерностей в поведении. Первичным элементом любой системы автоматического анализа поведения является модуль видеонаблюдения, позволяющий осуществлять сопровождение наблюдаемого объекта (трекинг) и, возможно, оценивать ряд его характеристик, например, геометрическую форму, скорость и т.п. Наибольшее распространение такие системы получили для наблюдения за лабораторными мышами.

Серьезным ограничением современных систем видеонаблюдения за животными является их неспособность осуществлять наблюдение и идентификацию одновременно для нескольких животных, находящихся в клетке. Это сильно ограничивает область их применения, т.к. не позволяет анализировать социальное поведение. Кроме того, большую часть времени вне экспериментов животное проводит в т.н. домашней клетке, содержащей обычно 3–5 мышей. Для изучения изменений поведения животного в течение долгих интервалов времени (недели, месяцы) желательно осуществлять видеонаблюдение за ним и в домашней клетке. Современные технологии позволяют это сделать аппаратно, но надежных методов для множественного трекинга и идентификации мышей в клетке до сих пор не создано. Это связано в первую очередь со сложностями при разделении мышей, которые любят сбиваться в кучку и находиться в тесном телесном контакте, а также с персонификацией найденных особей.

Существующие аналоги обычно либо используют вид сбоку [3], либо сильно ограничивают возможные перекрытия тел мышей [6].

В настоящей работе для решения первой проблемы предложено использовать ЕМ-подобную процедуру разделения смеси распределений. При этом Е-шаг отличается от классического подсчета вероятности принадлежности к компоненте смеси в том смысле, что одна и та же точка на изображении может принадлежать двум и более мышам (например, когда одна мышь забралась на спину другой). Для идентификации предложен алгоритм, основанный на определении близости к положению мыши на предыдущем кадре.

Вычитание фона

Входными данными для рассматриваемых алгоритмов являлись видеоролики с разрешением 500×550 пикселей, содержащих запись движения одиночной мыши в клетке в течение 15 минут. На основе этих записей был создан виртуальный видеоролик, позволяющий отлаживать систему множественного видеотрекинга.

Для отслеживания животных внутри сцены используется традиционный метод видеослежения, основанный на «вычитании фона» [4]. Изображение было переведено в серую шкалу, и для каждого пикселя оценивалась медиана интенсивности серого по N кадрам видеосъемки сцены при том же положении камеры и освещении, что и в процессе эксперимента, но без животных. Далее для кадров с животными пиксели, интенсивность которых составляла менее 60% от соответствующего ему медианного значения, помечались как точки объекта. Найденные пиксели объекта фильтровались с помощью методов математической морфологии [7]. Это позволило отсеять одиночные выбросы, помехи и отдельные мелкие объекты, например, насекомых, попавших в кадр. Значение порога было выбрано так, чтобы к объекту не причислялись тени, отбрасываемые мышами, и их отражения от стен клетки. В результате вычитания фона выделялась маска, содержащая изображения мышей.

Разделение особей

В данной работе предполагалось, что общее количество мышей, находящихся в клетке,

*Работа выполнена при финансовой поддержке РФФИ, проект № 08-01-00405.



Рис. 1. Пример работы системы трекинга одной мыши в клетке. Белым показан выделенный контур мыши. Также отмечены найденные ключевые точки: звездочкой — центр масс, плюсиком — нос, крестиком — хвост.

известно. В том случае, если в результате выделения масок животных на видеокadre количество компонент связности равнялось количеству мышей, находящихся в клетке, то задача сводилась к определению контуров и характерных точек (носа, точки крепления хвоста), аналогичному случаю наблюдения за одной мышью [1]. Пример работы алгоритма определения контура и характерных точек показан на рис. 1.

Также для каждого контура находилась эллипс, наилучшим образом его приближающий. Для этого для каждой маски мыши M_k рассчитывалась матрица

$$A_k^{-1} = \frac{\sum_{\mathbf{x}_i \in M_k} (\mathbf{x}_i - \mathbb{E}_k \mathbf{x})(\mathbf{x}_i - \mathbb{E}_k \mathbf{x})^\top}{\sum_{\mathbf{x}_i \in M_k} 1},$$

где

$$\mathbb{E}_k \mathbf{x} = \frac{\sum_{\mathbf{x}_i \in M_k} \mathbf{x}_i}{\sum_{\mathbf{x}_i \in M_k} 1},$$

а \mathbf{x}_i — пиксели кадра. Уравнение эллипса, наилучшим образом приближающего контур мыши, имело вид

$$(\mathbf{x} - \mathbb{E}_k \mathbf{x})^\top A_k (\mathbf{x} - \mathbb{E}_k \mathbf{x}) = \frac{1}{4}. \quad (1)$$

Если количество компонент связности оказывалось меньше количества мышей, то анализировались эллипсы, вписанные в контур

каждой мыши на предыдущем кадре. Для каждой компоненты связности определялись те особи, контуры которых оказывались внутри компоненты связности. Это определялось путем подсчета ближайшей компоненты связности к эллипсам с предыдущего кадра.

Далее для каждой компоненты связности, содержащей более одной мыши, производилось разделение мышей с помощью процедуры, представляющей собой модификацию EM-алгоритма разделения смеси распределений [5]. Обозначим через C_k — внутренность эллипса (1). Алгоритм разделения особей представляет собой двухшаговую итеративную процедуру. На первом шаге (аналог E-шага EM-алгоритма) осуществляется расчет переменных $\gamma(z_{ik})$, определяющих принадлежность данной точки маски \mathbf{x}_i к k -ой особи. Заметим, что в отличие от классического EM-алгоритма, сумма принадлежностей к различным особям не обязана равняться единице, т.к. одна и та же точка может принадлежать двум и более животным одновременно, например, из-за того, что одно животное влезло на спину другому:

$$\gamma(z_{ik}) = \begin{cases} 1, & \mathbf{x}_i \in C_k \text{ или } k = \arg \min_j \rho_M(\mathbf{x}_i, \mathbb{E}_j \mathbf{x}); \\ 0, & \text{иначе;} \end{cases}$$

где $\rho_M(\mathbf{x}_i, \mathbb{E}_j \mathbf{x})$ — расстояние Махаланобиса, задаваемое матрицей A_j .

На втором шаге (аналог M-шага EM-алгоритма) осуществляется уточнение эллипсов. Обозначим через α_k угол наклона большей полуоси эллипса (направления, соответствующего носу¹), взятого с предыдущего кадра. Для уточнения матрицы A_k сначала центрируем систему координат и выполняем преобразование поворота на угол $-\alpha_k$:

$$\mathbf{y}_i = R_{-\alpha_k}(\mathbf{x}_i - \mu_k),$$

где

$$\mu_k = \frac{\sum_i \gamma(z_{ik}) \mathbf{x}_i}{\sum_i \gamma(z_{ik})} = \frac{\sum_i \gamma(z_{ik}) \mathbf{x}_i}{n_k}.$$

Затем осуществляется поиск наиболее вероятной матрицы A_k путем максимизации апостериорной плотности на матрицу A'_k , представляющей собой матрицу A_k , повернутую на угол $-\alpha_k$. Функция правдоподобия на A'_k принимает следующий вид:

$$p(Y|A'_k) = \left(\sqrt{\det A'_k} \right)^{n_k} \exp\left(-\frac{1}{2} \sum_i \mathbf{y}_i^\top A'_k \mathbf{y}_i \right).$$

¹Об идентификации точек носа и крепления хвоста см., например, работу [1]

Для того, чтобы учесть априорные представления о характерной форме эллипса, приближающего контур мыши, введем априорное распределение Уишарта на множестве матриц

$$\begin{aligned} p(A'_k) &= \\ &= B(V, \nu) (\det A'_k)^{\frac{\nu-3}{2}} \exp\left(-\frac{1}{2} \text{tr}(V^{-1} A'_k)\right) \sim \\ &\sim \mathcal{W}(A'_k | V, \nu), \end{aligned} \quad (2)$$

где $B(V, \nu)$ — (несущественная для нас) нормировочная константа. Вопрос выбора параметров распределения Уишарта будет рассмотрен в следующем разделе. В связи с тем, что распределение Уишарта является сопряженным для нормального распределения с матрицей A'_k , то апостериорное распределение может быть вычислено аналитически. Оно также будет представлять собой распределение Уишарта. Выражение для максимума апостериорной плотности для A'_k имеет вид

$$A'_k = (n_k + \nu - 3) \left(\sum_i \gamma(z_{ik}) \mathbf{y}_i^T \mathbf{y}_i + V^{-1} \right)^{-1}.$$

Чтобы получить матрицу A_k , осуществляем обратный поворот системы координат на угол α_k

$$A_k = R_{\alpha_k} A'_k R_{\alpha_k}^T.$$

Описанные два шага повторяются итерационно до сходимости. В качестве начального приближения для каждой особи, чей контур входит в данную компоненту связности, используются матрицы и центры эллипсов с предыдущего кадра.

Настройка параметров априорного распределения

Введение априорного распределения позволяет отсекать заведомо нереалистичные эллипсы, например, слишком длинные или слишком узкие, которые не могут описывать контур реальной мыши. Распределение Уишарта было выбрано в качестве априорного в том числе и потому, что оно позволяет получить выражение для максимума апостериорной плотности в явном виде при использовании гауссовского распределения в качестве функции правдоподобия (т.е. является сопряженным к последнему). Для набора априорной информации мы воспользовались несколькими видеороликами с одиночной мышью в клетке, снятыми в тех же условиях. Для каждого кадра решалась задача вписывания эллипса в контур мыши. Затем полученные эллипсы поворачивались так, чтобы их первый собственный вектор (т.е. собственный вектор, отвечающий большему собственному значению) был параллелен оси абсцисс. Обозначим



Рис. 2. Пример кадра из виртуального ролика, содержащего наложенные изображения трех мышей.

через $\mathbb{E}A$ среднее арифметическое получившихся матриц. Известно, что математическое ожидание распределения Уишарта выражается формулой $\mathbb{E}A = \nu V$, отсюда

$$V = \frac{\mathbb{E}A}{\nu}.$$

Для оценки параметра ν воспользуемся информацией об изменении длины большей полуоси эллипса r_1 . Легко показать, что если неотрицательно определенная матрица имеет распределение Уишарта, то ее собственные значения имеют гамма-распределение. Пусть r_1 имеет гамма-распределение с параметрами a и b , т.е. $r_1 \sim \mathcal{G}(r_1 | a, b)$. Оценим их значения, используя известные соотношения между ними и первыми двумя моментами распределения

$$\mathbb{E}r_1 = \frac{a}{b}, \quad \mathbb{D}r_1 = \frac{a}{b^2}.$$

С другой стороны, известно, что распределение Уишарта является многомерным обобщением гамма-распределения, причем параметру a соответствует значение $\frac{1}{2}\nu$ [2]. Приняв предположение о близости характеристик разброса длины большей полуоси и характеристики изменчивости эллипса, в качестве значения ν примем величину

$$\nu = 2a = 2 \frac{(\mathbb{E}r_1)^2}{\mathbb{D}r_1}.$$

Построение тестового видеоролика и результаты экспериментов

Для оценки качества разработанного алгоритма видеотрекинга был разработан специальный



Рис. 3. Пример кадра из виртуального ролика с разделением особей с помощью предложенной EM-подобной процедуры. Белыми эллипсами показаны спрогнозированные формы мышей, плюсами отмечены точки носа.

виртуальный видеоролик, содержащий изображения многих мышей, информацию об их форме и идентификаторы особей. Данный видеоролик был получен путем обработки видеоклипа, содержащего запись поведения одной мыши в течение сравнительно долгого промежутка времени. Этот клип был разбит на p равных частей. Из каждого кадра был выделен контур мыши, а затем был получен новый ролик путем наложения p контуров, взятых с соответствующих кадров каждой из частей, на статическую модель фона. Кадр из

получившегося ролика показан на рис. 2. Пример работы алгоритма разделения особей показан на рисунке 3. Алгоритм успешно отработал на видеороликах, содержащих 3 мыши в течение 15 минут, не сделав при этом ни одного сбоя сопровождения или идентификации. Примеры его работы будут показаны в ходе доклада.

Литература

- [1] *Конущин А. С., Ветров Д. П., Воронин П. А., Синдеев М. С., Ломакина-Румянцева Е. И., Кропотов Д. А., Зарайская И. Ю., Анохин К. В.* Система видеонаблюдения за поведением лабораторных животных с автоматической сегментацией на поведенческие акты // Труды межд. конф. Графикон-2008, М.: МАКС Пресс, 2008. — С. 199–206.
- [2] *Bishop, C.* Pattern Recognition and Machine Learning, Springer, 2006.
- [3] *Branson K., Belongie S.* Tracking multiple mouse contours (without too many samples) // Proc. of the 2005 IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR'05) 2005, — Vol. 1, Pp. 1039–1046.
- [4] *Cheung S., Kamath C.* Robust techniques for background subtraction in urban traffic video // Visual Communications and Image Processing, 2004, Pp. 881–892.
- [5] *Dempster A., Laird N., Rubin D.* Maximum likelihood from incomplete data via the EM algorithm // Journal of the Royal Statistical Society, 1977, В 39, 1, 1–38.
- [6] *Goncalves W. N., Monteiro J. B. O., de Andrade Silva J., Machado B. B., Pistori H., Odakura V.* Multiple mice tracking using a combination of particle filter and k -means // Proc. of XX Brazilian Symp. on Computer Graphics and Image Processing, 2007, Pp. 173–178.
- [7] *Serra J.* Image Analysis and Mathematical Morphology. New York: Academic, 1982.