

# Компьютерное слежение с масштабированием, основанное на градиентном спуске

Альберт Шапошников, Елена Шапошникова  
ООО «Tomsklabs», Томский государственный университет, Томск, Россия  
albertelena@mail.ru

## АННОТАЦИЯ

В статье построена функция схожести двух изображений одного объекта на разных кадрах видеопоследовательности. Найдены частные производные этой функции по координатам центра и размерам области слежения. На основании этих производных построен алгоритм слежения, функционирующий при независимых изменениях длины и ширины области слежения. Алгоритм основан на градиентном спуске.

**Ключевые слова:** слежение, цветовые гистограммы, сдвиг среднего, функция схожести, масштабирование.

## ANNOTATION

A similarity function between the images of the object in the different video frames is constructed. The partial derivatives of this similarity function with respect to the abscissa and ordinate of the center of the object and the width and height of the object are taken. Using these derivatives the tracking algorithm is constructed. This algorithm takes into account the independent changes of the width and height of the object. This algorithm is based on the gradient descent and not uses the Mean Shift.

**Keywords:** tracking, histograms, mean shift, gradient descent, scaling.

## 1. ВВЕДЕНИЕ

Компьютерная обработка и анализ видеоизображений – многогранная и быстроразвивающаяся область. Использование видеoinформации в современном мире стремительно нарастает. Это и различные системы мониторинга, технического зрения, видеотелефонии, регистрирующие и передающие огромные объемы видеоданных, и различные автономные системы (роботы), принимающие решения на основе анализа видеоизображений, и программы изучения космоса с беспилотными летательными аппаратами, и медицина, и интернет, и многие другие направления.

Примеры успешного применения обработки и анализа видеоизображений можно найти в физике, астрономии, биологии, медицинской радиологии, промышленности, в оборонной и правоохранительной сферах.

При этом наряду со значительным повышением уровня развития техники, весьма существенную роль играют методы обработки видеoinформации.

В задачах анализа видеоизображений одним из ключевых элементов является поиск на изображении некоторого объекта. Такая задача распадается на две части. Во-первых, нужно дать описание образца объекта, пригодное для эффективного сравнения этого образца с другим изображением объекта. Во-вторых, разработать алгоритм поиска области изображения (окна), для которой и производить это сравнение. Вопрос описания образца объекта

(получения характеристического вектора) изучен широко [8,9]. Описание образца объекта может быть выбрано инвариантным к повороту объекта, изменению его цвета или размера. Для выбора области изображения (окна) подходы тоже имеются [8,9].

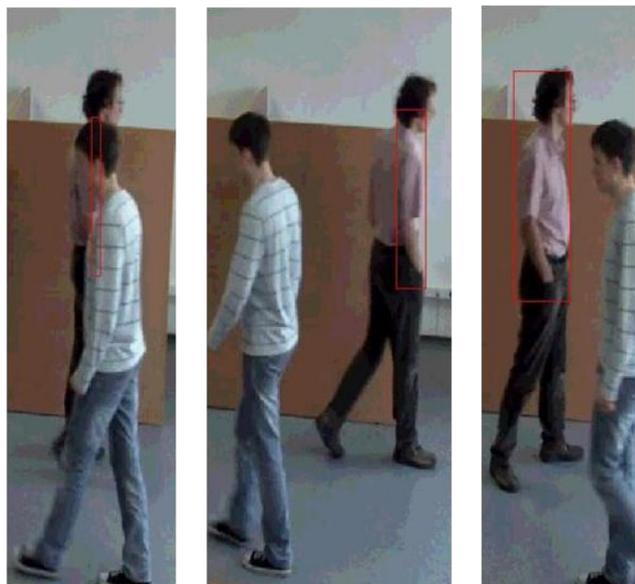
Для задач слежения поиск и подходящего описания объекта, и окна можно сузить, поскольку анализу подвергается последовательность контекстно связанных между собой изображений (кадров). Так, например, в алгоритме сдвига среднего Mean Shift [1,3–6] поиск наиболее подходящей области изображения начинается с области, зафиксированной на предыдущем кадре. В процедуре Mean Shift, введенной и подробно описанной в [4,5], объекты описываются с помощью нормированных цветовых гистограмм, то есть характеристическим вектором в данном случае является нормированная цветовая гистограмма. Для гистограмм на основании коэффициента Бхаттачария строится вещественнозначная функция схожести. На первом кадре последовательности (видеоряда) каким-либо образом отмечается объект, а затем для каждого следующего кадра видеоряда определяется ближайший локальный максимум функции схожести. Процедура Mean Shift оказалась эффективной и получила дальнейшее развитие в работах [1,2,3,6,7]. Однако при изменении пропорций и размеров объекта при его движении длительное слежение производить не удавалось, так как «нечеткое» окно захватывает слишком много фона и «соскальзывает» с объекта. В [3,6] эта проблема решена в частных случаях и найден способ слежения за объектом, у которого изменяются размеры, но не изменяются пропорции (человек приближающийся лицом не должен повернуться боком).

В предлагаемой работе построен алгоритм отыскания области кадра, имеющей наибольшую схожесть (максимум функции схожести в смысле коэффициента Бхаттачария) с заранее выделенным образцом. Данный алгоритм учитывает изменения как положения, так и размеров объекта. В отличие от [1–7] для отыскания локальных максимумов применен градиентный спуск.

Описанный в статье алгоритм определения области с гистограммой, максимально похожей на заданную, был реализован в MatLab'e. Для эксперимента был выбран ролик со свободно прохаживающимися людьми. Целью алгоритма было слежение за произвольно выбранным человеком. На первом кадре молодой человек в зеленой кофте был обведен вручную прямоугольником, и для получившейся области была посчитана гистограмма. Затем был запущен алгоритм, и на каждом следующем кадре по максимуму схожести цветовых гистограмм определялись размеры и положение области слежения.

На рисунке 1 приведены первый, 86-й, 131-й, 165-й и 168 с момента начала слежения кадры. На первом кадре размер прямоугольника 85×296 пикселей. Размеры прямоугольников на других кадрах получились 23×296, 58×271, 24×273, 24×273. Как видно, на 131 кадре молодой человек, за

которым ведется наблюдение, частично заслонен другим человеком – прямоугольник сузился. На 165-ом кадре молодой человек виден полностью, но лицом к нам и дальше от исходного положения – прямоугольник изменил размеры, причем снова не пропорционально первоначальным размерам, а согласно цели – максимуму схожести с первоначальной гистограммой.



Кадр 241.

Кадр 261.

Кадр 231

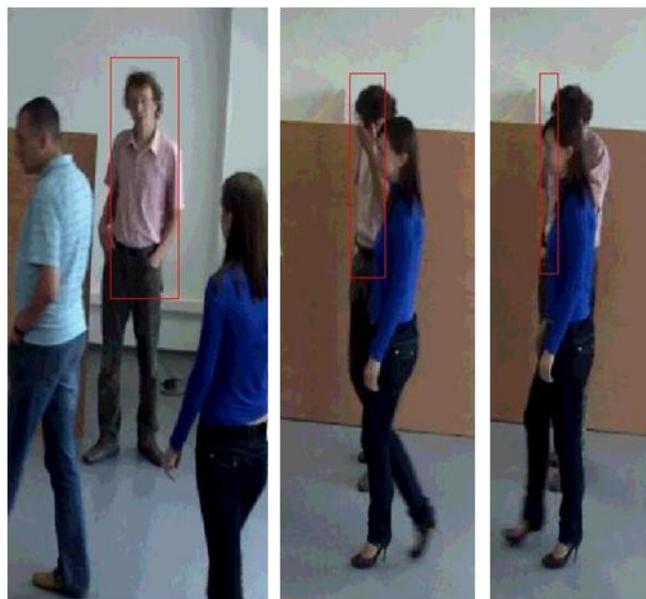
**Рис 1.**



**Рис 2.**

Отметим также, что в прямоугольник на 131 кадре не попадает часть темного фона от первоначальной области на первом кадре – изменился фон, а лица на первоначально обведенной области на первом кадре вообще видно не было,

поэтому коэффициент схожести значительно снизился до 0,88.



Кадр 315.

Кадр 412.

Кадр 420.

**Рис 3.**

Однако, и положение, и размеры области, найденной на текущем кадре, адекватны размерам и положению видимой части области, отмеченной на первом кадре. На 165 и 168 кадрах объект слежения заслонен, но не потерян.



Кадр 1.

Кадр 87.

Кадр 98.

**Рис 4.**

На рисунках 2,3, 4 и 5 другой пример слежения. На первом кадре объект слежения был обведен вручную красным прямоугольником. Затем по цветовой гистограмме обнаруживалась область наибольшей схожести. В нарезке представлены 1, 87, 98, 231, 241, 261, 315, 412, 420, 421, 426 и 432 кадры. Особо отметим, что изображения не были масштабированы, но только обрезаны

На 1 кадре объект стоит близко спиной. На 87 и 98 кадрах он же, но дальше и левым боком. Размеры обводящего прямоугольника заметно изменились, причем они взаимно независимы. Аналогичную картину можно наблюдать на кадрах 231, 261, 315 и 432.



Кадр 421.

Кадр 426.

Кадр 432.

Рис 5.

На кадрах 412–432 объект слежения заслонили. Заметно, как по мере того, как слева остается видна все меньшая часть объекта, прямоугольник постепенно сокращается. Затем обнаруживается другая часть вновь появившегося объекта наблюдения. И по мере того, как объект все больше открывается, прямоугольник разрастается.

## 2. ОПИСАНИЕ ОБЪЕКТА ЦВЕТОВОЙ ГИСТОГРАММОЙ И ПОСТРОЕНИЕ ФУНКЦИИ СХОЖЕСТИ

Пусть задана функция  $f: P \rightarrow B$ , где  $P = [0, d_1] \times [0, d_2]$  – прямоугольник,  $B$  – конечное множество. Обозначим через  $\chi_{f^{-1}(b)}(x)$ , где  $x \in P$ , характеристическую функцию полного прообраза  $f^{-1}(b)$  элемента  $b \in B$ , то есть  $\chi_{f^{-1}(b)}(x) = 1$ , если  $f(x) = b$ , и  $\chi_{f^{-1}(b)}(x) = 0$  в противном случае.

Для каждого  $b \in B$  определим величину  $H_b = \iint_P \chi_{f^{-1}(b)}(x) dx_1 dx_2$ . Можно предполагать, что этот интеграл существует (это выполнено, если цветовые пятна плотные и непрерывные, то есть множество точек разрыва функции  $\chi_{f^{-1}(b)}(x)$ , а все они первого рода, имеет меру нуль).

Множество  $\{H_b | b \in B\}$  называется гистограммой кадра.

Введем в рассмотрение семейство функций  $g(x) = \max\left(0, 1 - \left(\frac{x_1 - y_1}{l_1}\right)^2 - \left(\frac{x_2 - y_2}{l_2}\right)^2\right)$ , где  $x = (x_1, x_2) \in P$ ,

$x$  – переменная;  $y = (y_1, y_2) \in P, l_1, l_2 \in R$ ,  $y_1, y_2, l_1, l_2$  – параметры.

Пусть теперь  $G \subset P$ ,  $G$  – эллиптическая область с центром в точке  $y = (y_1, y_2)$  и полуосями  $l_1, l_2$ . Выбор эллиптической области обусловлен тем, что объект в кадре определяется своим центром  $y = (y_1, y_2)$  и размерами  $l = (l_1, l_2)$ , то есть эллиптическая область хорошо подходит для выделения объекта в кадре.

Для определения гистограммы эллиптической области  $G$  рассмотрим функции  $H_b(y, l) = \iint_P \chi_{f^{-1}(b)}(x) g(x) dx_1 dx_2$ , где

$y = (y_1, y_2)$ ,  $l = (l_1, l_2)$ . Величина  $H_b = H_b(y_0, l_0)$  есть вес той части области  $G$  с центром в точке  $y_0$  и размерами  $l_0$ , которая имеет данный цвет  $b$ , – вес цвета  $b$ . Цвет, который в центре области, вносит больший вклад в гистограмму. Заметим, что  $\sum_{b \in B} H_b$  есть величина, характеризующая область

$G$ . Будем называть ее весом области.

Тогда множество чисел  $H = \{H_b | b \in B\}$  будет гистограммой области. Ясно, что гистограмма  $H$  изменяется при изменении положения центра  $y = (y_1, y_2)$  и размеров  $l = (l_1, l_2)$  области.

Однако для построения функции схожести сами гистограммы  $H = \{H_b | b \in B\}$  непригодны, так как  $H_b$  существенно зависят от размеров  $l$  области. А при пропорциональном увеличении или уменьшении (масштабировании) изображения объекта функция схожести должна идентифицировать получившиеся изображения как изображения одно и того же объекта.

Поэтому рассмотрим  $h_b = \frac{H_b}{\sum_{b \in B} H_b}$  – отношение веса цвета  $b$  к

весу всей области, то есть относительный вес цвета.

Тогда  $h = \{h_b | b \in B\}$  – нормированная гистограмма области  $G$ . Ясно, что нормированная гистограмма  $h$  изменяется при изменении  $y = (y_1, y_2)$  и  $l = (l_1, l_2)$ , то есть является функцией переменных  $y_1, y_2, l_1, l_2$ .

Для двух нормированных гистограмм  $h_0$  и  $h$  определим, следуя [4,5], функцию схожести  $\rho = \sum_{b \in B} \sqrt{h_{0b} \cdot h_b}$ , где

$h_0 = \{h_{0b} | b \in B\}$  – некоторая гистограмма, с которой производим сравнение.

Тогда  $\rho$  есть функция, зависящая от параметров  $y = (y_1, y_2)$  и  $l = (l_1, l_2)$ . Для этой функции  $\rho = \rho(y, l)$  получаем  $\rho: R^3 \rightarrow R$ , если размеры объекта  $l_1$  и  $l_2$  связаны каким-либо соотношением, например, пропорциональны как при приближении одноцветного шара, или  $\rho: R^4 \rightarrow R$ , если  $l_1, l_2$  не зависят друг от друга, или  $\rho: R^5 \rightarrow R$ , если ввести еще один параметр  $\alpha$ , который обеспечивает изменение угла поворота области. При введении в рассмотрение параметра  $\alpha$  в равенстве

$H_b(y, l) = \iint_P \chi_{f^{-1}(b)}(x) g(x) dx_1 dx_2$  в качестве функции  $g(x)$

можно использовать

$$g = \max\left(0, 1 - \left(\frac{(x_1 - y_1) \cos \alpha - (x_2 - y_2) \sin \alpha}{l_1}\right)^2 - \left(\frac{(x_1 - y_1) \sin \alpha + (x_2 - y_2) \cos \alpha}{l_2}\right)^2\right),$$

$$\alpha \in [-90^\circ, 90^\circ].$$

Для простоты запишем  $\rho = \sum_{b \in B} \rho_b$ , где  $\rho_b = \sqrt{h_{0b} \cdot h_b}$ .

### 3. ГРАДИЕНТНЫЙ СПУСК

Найдем градиент функции схожести  $\rho: R^4 \rightarrow R$

$$\text{grad} \rho = \left( \frac{\partial \rho}{\partial y_1}, \frac{\partial \rho}{\partial y_2}, \frac{\partial \rho}{\partial l_1}, \frac{\partial \rho}{\partial l_2} \right)$$

Сначала находим

$$\frac{\partial H_b}{\partial y_i} = \frac{2}{l_i^2} \iint_P \chi_{f^{-1}(b)}(x) \chi_{\text{supp}(g(x))}(x) (x_i - y_i) dx_1 dx_2, \text{ где } \text{supp}(g) -$$

носитель функции  $g$ , то есть множество точек, в которых функция принимает ненулевое значение,

$$\frac{\partial H_b}{\partial l_i} = \frac{2}{l_i^3} \iint_P \chi_{f^{-1}(b)}(x) \chi_{\text{supp}(g(x))}(x) (x_i - y_i)^2 dx_1 dx_2,$$

$$\frac{\partial h_b}{\partial y_i} = \frac{\frac{\partial H_b}{\partial y_i} - h_b \sum_{b \in B} \frac{\partial H_b}{\partial y_i}}{\sum_{b \in B} H_b},$$

$$\frac{\partial h_b}{\partial l_i} = \frac{\frac{\partial H_b}{\partial l_i} - h_b \sum_{b \in B} \frac{\partial H_b}{\partial l_i}}{\sum_{b \in B} H_b}, \text{ где } i = 1, 2.$$

$$\text{Тогда имеем } \frac{\partial \rho}{\partial y_i} = \sum_{b \in B} \frac{1}{2h_b} \frac{\partial h_b}{\partial y_i} \rho_b, \quad \frac{\partial \rho}{\partial l_i} = \sum_{b \in B} \frac{1}{2h_b} \frac{\partial h_b}{\partial l_i} \rho_b, \quad i = 1, 2.$$

Для функции схожести  $\rho$  (в предположении ее дифференцируемости) имеем

$$\rho \approx \rho(y_0, l_0) + \text{grad} \rho|_{(y_0, l_0)} \cdot (\Delta y_1, \Delta y_2, \Delta l_1, \Delta l_2) =$$

$$= \rho(y_0, l_0) + t \left| \text{grad} \rho|_{(y_0, l_0)} \right|^2, \quad \text{где}$$

$$(\Delta y_1, \Delta y_2, \Delta l_1, \Delta l_2) = t \cdot \text{grad} \rho|_{(y_0, l_0)}. \text{ То есть вектор приращений}$$

выбираем в направлении градиента, так как в каждой точке вектор градиента по численному значению и по направлению характеризует наибольшую скорость возрастания величины, в данном случае схожести.

Вообще говоря, знание градиента в точке дает лишь направление шага, а для определения величины шага этого недостаточно. Однако знание максимального значения функции схожести  $\rho = 1$  дает приблизительную оценку для

$$t \approx \frac{1 - \rho(y_0, l_0)}{|\text{grad} \rho|^2}. \text{ Откуда для величины шага имеем}$$

$$(\Delta y_1, \Delta y_2, \Delta l_1, \Delta l_2) \approx \frac{1 - \rho(y_0, l_0)}{|\text{grad} \rho|^2} \text{grad} \rho.$$

### 4. ЗАКЛЮЧЕНИЕ

Отметим, что предложенный выше метод слежения не учитывает угол поворота объекта относительно оси, перпендикулярной плоскости кадра. То есть он эффективен, если объект слежения движется, оставаясь параллельным самому себе (поступательно) и его наибольшее измерение параллельно одной из сторон кадра. Например, такая ситуация складывается при слежении камерой вдоль коридора, по которому движутся люди. При слежении за вращающимся некруглым объектом следует ввести угол поворота. В этой задаче функция схожести  $\rho: R^5 \rightarrow R$  и мы будем иметь, как указано выше, пять параметров. Рассмотрение этой модели позволит расширить класс обрабатываемых видеопоследовательностей. Например, можно будет производить компьютерное слежение сверху за автомобилями на перекрестках.

### 5. ССЫЛКИ

- [1] R. Szeliski, Computer Vision: Algorithms and Applications. Microsoft Research. URL: <http://szeliski.org/Book>.
- [2] D. Comaniciu, V. Ramesh, and P. Meer. Real-Time Tracking of Non-Rigid Objects Using Mean Shift // Proc. Conf. Computer Vision and Pattern Recognition, 2000. Vol. 2. P. 142-149.
- [3] D. Comaniciu, P. Meer. Mean Shift: A Robust Approach Toward Feature Space Analysis // IEEE Transactions on Pattern Analysis and Machine Intelligence, May 2002. Vol. 24. № 5. P. 603-619.
- [4] R. Collins, Mean-shift Blob Tracking through Scale Space // IEEE Computer Vision and Pattern Recognition (CVPR03), Madison, WI, June 16-22, 2003. P. 234-240.
- [5] D. Comaniciu, V. Ramesh, and P. Meer, The Variable Bandwidth Mean Shift and Data-Driven Scale Selection // Proc Eighth Int'l Conf. Computer Vision, July 2001. Vol. I. P. 438-445.
- [6] S. T. Birchfield, S. Rangarajan. Spatiograms Versus Histograms for Region-Based Tracking // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, California, June 2005. Vol. 2. P. 1158-1163.
- [7] C. O'Conaire, N. E. O'Connor, A. F. Smeaton. An improved spatiogram similarity measure for robust object localization // Proceedings of the 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2007. Vol. 1. P. 1069-1072.
- [8] Ning He, Jiaheng Cao, Lin Song, Kernel Based Spatiogram Tracking Using Improved Similarity Measure // Ninth IEEE International Conference on Computer and Information Technology, 2009. Vol. 1. P. 124-127.
- [9] Р. Гонсалес, Р. Вудс, С. Эддинс. Цифровая обработка изображений. М.: Техносфера, 2006.

### ОБ АВТОРАХ

Елена Васильевна Шапошникова - доцент каф. общей математики ММФ ТГУ. e-mail: [albertelena@sibmail.com](mailto:albertelena@sibmail.com).

Альберт Игоревич Шапошников - научный консультант ООО «Tomsklabs». e-mail: [albertelena@mail.ru](mailto:albertelena@mail.ru)