

## Раздел V. Техническое зрение

УДК 004.932.72

DOI 10.18522/2311-3103-2020-1-233-245

**А.Е. Щелкунов, В.В. Ковалев, К.И. Морев, И.В. Сидько**

### **МЕТРИКИ ОЦЕНКИ АЛГОРИТМОВ АВТОМАТИЧЕСКОГО СОПРОВОЖДЕНИЯ**

*Работа посвящена обзору существующих метрик для оценки качества выполнения задачи сопровождения объектов на видео различными алгоритмами. При оценке алгоритмов сопровождения для последующего их сравнения недостаточно использовать одну метрику, а следует оценивать алгоритмы по набору различных независимых оценок. С этой целью было проведено исследование существующих метрик оценки алгоритмов, результаты которого приведены в статье. В обзоре участвует множество различных подходов к оценке алгоритмов. Например, подходы основанные на оценке определения центра объекта сопровождения, которые являются одними из первых и популярных до сих пор метрик оценки алгоритмов сопровождения. К основным недостаткам таких подходов можно отнести сложность определения истинного центра объекта, а также интерпретация оценок при различных размерах объекта. Для устранения этих недостатков в статье вводится новая метрика: несмещенная (оконная) ошибка определения центра объекта, которая учитывает постоянную составляющую ошибки определения центра. К другим подходам можно отнести метрики, основанные на анализе коэффициента Жаккара. Так же в статье рассмотрены подходы, основанные на анализе сбоев сопровождения, в которых учитывается длина сопровождения и интенсивность отказов. Предложен новый метод оценки алгоритмов при потере визуального контакта с сопровождаемым объектом, учитывающий количество кадров в которых был потерян визуальный контакт с объектом. В ходе исследования были рассмотрены подходы оценки алгоритмов одновременного сопровождения нескольких объектов. Были предложены интегральные метрики, задача которых получение комплексной оценки алгоритма сопровождения. Для формирования комплексной оценки желательно использование различных некоррелированных метрик. Комплексные оценки предоставляют возможность сравнивать алгоритмы между собой. В качестве комплексной оценки в статье предлагается использование метрики, объединяющей точность и устойчивость алгоритма. Как правило, в качестве метрики точности используется коэффициент Жаккара, однако для задач где точность сопровождения центра объекта является основополагающей, авторами предлагается использовать в качестве метрики точности несмещенную ошибку определения центра.*

*Трекинг; сопровождение объектов; метрики оценки алгоритмов сопровождения.*

**A.E. Shchelkunov, V.V. Kovalev, K.I. Morev, I.V. Sidko**

### **THE METRICS FOR TRACKING ALGORITHMS EVALUATION**

*The work is devoted to a review of existing metrics for assessing the quality of the task of tracking objects on video with various algorithms. When evaluating tracking algorithms for their subsequent comparison, it is not enough to use one metric, and algorithms should be evaluated using a set of different independent estimates. To this end, a study was conducted of existing metrics for evaluating algorithms, the results of which are given in the article. The review involves many different approaches to evaluating algorithms. For example, approaches based on the assessment of the definition of the center of the tracking object, which are one of the first and still popular metrics for evaluating tracking algorithms. The main disadvantages of such approaches include the difficulty of determining the true center of the object, as well as the interpretation of estimates for various sizes of the object. To eliminate these shortcomings, a*

*new metric is introduced in the article: an unbiased (window) error in determining the center of an object, which takes into account the constant component of the error in determining the center. Other approaches include metrics based on the analysis of the intersection over union. Also, the article considers approaches based on the analysis of tracking failures, which take into account the tracking length and failure rate. A new method is proposed for evaluating algorithms in case of loss of visual contact with an tracking object, taking into account the number of frames in which visual contact with the object was lost. During the study, approaches to evaluating algorithms for simultaneous tracking of several objects were considered. Integral metrics were proposed whose task is to obtain a comprehensive assessment of the tracking algorithm. For the formation of a comprehensive assessment, it is desirable to use various uncorrelated metrics. Complex estimates provide the ability to compare algorithms with each other. As a comprehensive assessment, the article proposes the use of a metric combining the accuracy and robustness of the algorithm. As a rule, the intersection over union is used as the accuracy metric, however, for problems where the accuracy of tracking the center of the object is fundamental, the authors propose using an unbiased error in determining the center as the accuracy metric.*

*Tracking; object tracking; metrics for evaluating tracking algorithms.*

**Введение.** Современные темпы развития компьютерного зрения, как области научного знания, а также рост научного сообщества предоставляют исследователям возможность быстро создавать новые подходы и алгоритмы на основе хорошо себя показавших прежних, совмещая их с новыми и прогрессивными. Однако, для целенаправленного движения к повышению качества работы современных алгоритмов трекинга объектов в видеопоследовательности необходимо выработать надежный инструментарий для оценки различных алгоритмов, их сравнения и выбора наилучших по определенным критериям.

**Анализ ошибки определения центра объекта.** Первыми будут рассмотрены оценки, основанные на ошибке определения центра объекта сопровождения, или на разнице между координатами центра объекта предсказанными алгоритмом и истинными.

В общем случае состояние объекта описывается уравнением:

$$A = \{(R_t, x_t)\}_{t=1}^N, \quad (1)$$

где  $x_t \in R$  – центр объекта в момент времени  $t$ ,  $R_t$  – область (регион) объекта в момент времени  $t$ ,  $\{\}_{t=1}^N$  – последовательность (массив) значений во времени от  $t = 1$  до  $N$ ,  $N$  – длина последовательности.  $R_t$  как правило, описывается обрамляющим прямоугольником, но может быть использована более сложная форма для повышения точности.

Центр объекта следует определять отдельно от центра области интереса, так как центр области не всегда соответствует центру объекта (рис. 1).



Рис. 1. Пример объекта

Ошибка определения центра объекта – одна из самых старых метрик для оценки алгоритмов сопровождения. Данная оценка до сих пор имеет широкую популярность. Ошибка рассчитывается на основе предсказанных координат центра объекта алгоритмом сопровождения и истинных (размеченных) координат центра объекта по формуле, указанной ниже:

$$\Delta(L^G, L^T) = \{\delta_t\}_{t=1}^N, \quad (2)$$

где  $L^G$  – положение объекта (размеченное положение),  $L^T$  предсказанное положение объекта алгоритмом сопровождения,  $\delta_t$  – расстояние между центрами объекта предсказанным алгоритмом сопровождения и истинным в момент времени  $t$ ,  $N$  – длина последовательности. Значение  $\delta_t$  определяется по формуле

$$\delta_t = \|x_t^G - x_t^T\|, \quad (3)$$

где  $x_t^G$  – истинные координаты центра объекта в момент времени  $t$ ,  $x_t^T$  – предсказанные координаты центра объекта алгоритмом сопровождения в момент времени  $t$ ,  $\| \cdot \|$  – норма.

Популярность метрик на основе анализа центра объекта обусловлена простотой разметки: необходимо разметить только одну точку на каждом кадре. Результаты оценки алгоритма представляют на графиках как на рисунке 2 или рассчитывают по следующим формулам, как среднюю и среднестатистическую ошибку определений координат центра соответственно

$$\Delta\mu(L^G, L^T) = \frac{1}{N} \sum_{t=1}^N \delta_t, \quad (4)$$

$$RMSE(L^G, L^T) = \sqrt{\frac{1}{N} \sum_{t=1}^N \|x_t^G - x_t^T\|^2}. \quad (5)$$

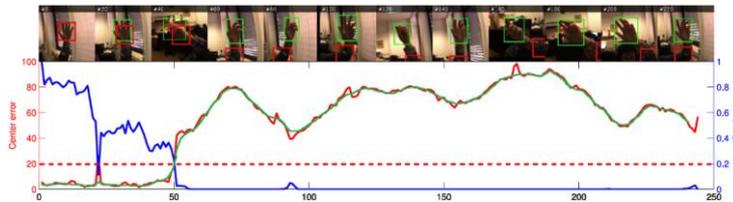


Рис. 2. Ошибка определения центра объекта и площадь перекрытия в зависимости от номера кадра

К недостаткам описанного выше подхода следует отнести чувствительность к необъективной разметке (некорректной работе оператора при разметке видеопоследовательности) и отсутствие учета размеров объекта [1]. Так для объекта, размер которого 100 пикселей ошибка определения координат центра объекта в 5 пикселей является не большой, в то время как для объекта размером 10 пикселей такая ошибка является существенной. В некоторых случаях, на практике сложно определить положение центра объекта. На рис. 1 представлен пример, где затруднительно определить координаты центра объекта.

Для исправления этих недостатков вводится нормализованная ошибка определения центра объекта по отношению к размеру объекта, рассчитываемая по формуле

$$\hat{\Delta}(\Lambda^G, \Lambda^T) = \{\hat{\delta}_t\}_{t=1}^N, \quad (6)$$

где  $\hat{\delta}_t$  – нормализованная ошибка определения центра объекта в момент времени  $t$  и рассчитывается по формуле, представленной ниже:

$$\hat{\delta}_t = \left\| \frac{x_t^G - x_t^T}{\sqrt{S(R_t^G)}} \right\|, \quad (7)$$

где  $S(R_t^G)$  – площадь объекта в момент времени  $t$ .

Однако, даже после проведения нормализации, оценка может дать вводящий в заблуждение результат. Так как ошибка уменьшается пропорционально размеру цели, и для больших объектов, в случае срыва сопровождения, ошибка изменится не существенно.

Существует проблема интерпретации измерений, в которых определение центра объекта сопровождения происходит с некоторым постоянным смещением. Несмотря на то, что объективно алгоритм успешно при этом выполняет задачу слежения за объектом. Для ее решения предлагается перейти к несмещенным (оконным) оценкам ошибки. С учетом гипотезы, о нормальном законе распределения ошибки определения центра, локальная среднестатистическая ошибка определения центра объекта описывается как:

$$\Delta\mu_i(\Lambda^G, \Lambda^T) = \frac{1}{M} \sum_{t=i}^{i+M} \delta_t, \quad (8)$$

где  $i$  – номер окна анализа,  $M$  – ширина окна анализа,  $\delta_t$  – ошибка определения центра в момент времени  $t$ . Пример рассчитанной по формуле (8) локальной ошибки определения центра показан на рисунке 2 зеленым цветом. Ширина окна анализа равна 5 кадрам.

Пользуясь формулой (8) можно рассчитать среднеквадратическую несмещенную ошибку определения центра

$$RMSE_{-i}(\Lambda^G, \Lambda^T) = \sqrt{\frac{1}{N-M} \sum_{t=1}^{N-M} (\delta_t - \Delta\mu_t)^2}, \quad (9)$$

где  $i$  – номер окна,  $M$  – ширина окна анализа,  $\delta_t$  – ошибка определения центра в кадре номер  $t$ ,  $\Delta\mu_i$  – локальное МО ошибки определения центра в окне  $t$ , рассчитанная по формуле (8),  $N$  – длина видеопоследовательности.

Оконный подсчет ошибок может быть полезен, так как в ходе сопровождения объекта интереса его центр может меняться из-за поворотов и изменения дальности до объекта.

Заменив в формулах (8), (9) ошибку определения центра  $\delta_t$ , на нормализованную ошибку определения центра  $\hat{\delta}_t$ , рассчитанную по формуле (7), получим локальную нормализованную среднестатистическую ошибку определения центра

$$\hat{\Delta}\mu_i(\Lambda^G, \Lambda^T) = \frac{1}{M} \sum_{t=i}^{i+M} \hat{\delta}_t, \quad (10)$$

и среднеквадратическую нормализованную ошибку определения центра, рассчитанную по локальному МО

$$\Delta RMSE_{-i}(\Lambda^G, \Lambda^T) = \sqrt{\frac{1}{N-M} \sum_{i=1}^{N-M} (\delta_i^{\wedge} - \Delta \mu_i^{\wedge})^2}, \quad (11)$$

**Анализ области перекрытия.** Проблемы нормализации оценки точности алгоритма могут быть решены при помощи метрик, основанных на анализе области перекрытия. Для реализации данных метрик необходимо иметь информацию об области (регионе), в котором находится объект и предсказанной алгоритмом сопровождения области объекта в кадре. На основе этих данных по формуле, представленной ниже, можно рассчитать оценку области перекрытия

$$\Phi(\Lambda^G, \Lambda^T) = \{\varphi_t\}_{t=1}^N, \quad (12)$$

где  $\varphi_t$  – регион перекрытия (Intersection Over Union или коэффициент Жаккара) в момент времени  $t$ , рассчитанный как

$$\varphi_t = \frac{R_t^G \wedge R_t^T}{R_t^G \vee R_t^T}. \quad (13)$$

В выражении (13) делимое – площадь области пересечения региона, предсказанного алгоритмом сопровождения и регионом объекта, делитель – площадь области объединения регионов. Суть выражения (13) показана на рис. 3



Рис. 3. Формула расчета региона перекрытия (IoU)

Отличительным свойством метрик на основе анализа площади перекрытия является то, что они учитывают, как положение объекта, так и его размер и не дают таких больших ошибок при срыве сопровождения, как при использовании средней и среднестатистической ошибки определений координат центра. Как только алгоритм начинает терять объект и плавно переходит на фон, значение оценки приближается к 0 и становится равным нулю при полном переходе алгоритма на фон, вне зависимости от того, насколько далеко от объекта ушли предсказания алгоритма сопровождения.

Если рассмотреть оценку алгоритма с точки зрения пиксельной классификации то выражение (13) можно записать в виде

$$\frac{R_t^G \wedge R_t^T}{R_t^G \vee R_t^T} = \frac{TP}{TP+FP+FN}, \quad (14)$$

где TP – количество правильно классифицированных пикселей, FP – количество ложно-положительных пикселей, FN – количество ложноотрицательных пикселей. Подобная формулировка похожа на F-меру, которую можно записать в следующем виде

$$F = \frac{2TP}{2TP+FN+FP}. \quad (15)$$

Precision – еще одна метрика, используемая в оценки алгоритмов сопровождения [2], описывается уравнением

$$Precision = \frac{TP}{TP+FP}. \quad (16)$$

Для анализа всей последовательности кадров используют метрику среднего перекрытия [3,4]

$$\bar{\varphi} = \sum_{t=1}^N \frac{\varphi_t}{N}, \quad (17)$$

где  $\varphi_t$  – коэффициент перекрытия в момент времени  $t$ ,  $N$  – длина последовательности. Или метрику длины правильно сопровождаемых кадров, описываемую уравнением

$$N_\tau = \|\{t \mid \varphi_t > \tau\}_{t=1}^N\|, \quad (18)$$

где  $\tau$  – порог перекрытия [5].

Для сравнения оценок работы алгоритмов сопровождения на последовательностях разной длины, количество правильно сопровождаемых кадров делиться на общую длину последовательности в соответствии с формулой

$$P_\tau(\Lambda^G, \Lambda^T) = \frac{\|\{t \mid \varphi_t > \tau\}_{t=1}^N\|}{N}, \quad (19)$$

для случая покadroвого определения  $P_\tau$  – процент правильно сопровождаемых кадров (true positive score). Такая интерпретация стала популярна для оценки алгоритмов сопровождения с появлением новой концепции и использовании алгоритмов сопровождения, основанных на обнаружении (детектировании) объекта в каждом кадре видеопоследовательности.

F-мера – еще одна метрика, которая может быть использована в данном контексте, однако стоит отметить, что метрики, оценивающие качество детекторов, игнорируют последовательный характер проблем сопровождения. Как показано на рис. 4, эти метрики не учитывают всей траектории (не анализируют траекторию). В то время как, траекторный анализ является важным аспектом в алгоритмах сопровождения.

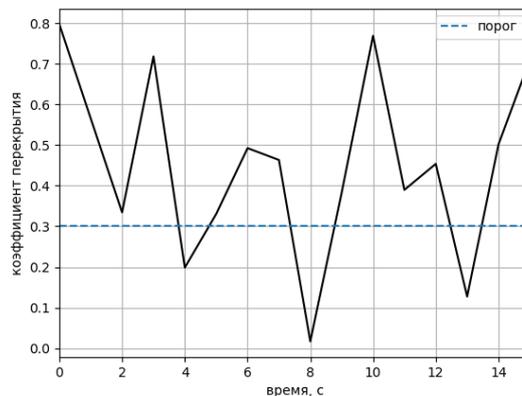


Рис 4. Зависимость перекрытия от времени

**Анализ сбоев сопровождения.** Длина сопровождения – количество верно сопровождаемых кадров с момента инициализации алгоритма сопровождения и до первого срыва сопровождения. В качестве критериев срыва сопровождения может выступать большое количество событий, например, оценка наблюдателя [6]. Однако при этом полученные оценки смещены и не могут быть надежно повторены даже одним и тем же наблюдателем. Лучшим подходом является замена наблюдателя автоматическим детектором срыва сопровождения. Например, путем установки порога оценки ошибки определения координат объекта или значения перекрытия. Выбор критерия может повлиять на результаты сравнения. Так, критерий оценки перекрытия дает более устойчивый результат, чем оценка ошибки определения центра объекта, при изменении размеров объекта.

Однако, замена оператора автоматом приводит к недостатку: если последовательность имеет плохо различимый или плохо видимый объект в ее начале, то это может привести к плохой инициализации алгоритма сопровождения, что повышает вероятность срыва. Из этого следует, что требуется большое количество последовательностей, демонстрирующих различные начальные состояния.

Следующая метрика в значительной степени решает проблему оценки длины сопровождения за счет оператора человека, который реинициализирует алгоритм сопровождения после срыва [6, 7]. Здесь человек–оператор может быть заменен алгоритмом автообнаружения и распознавания. Для формализации метрики вводятся величины:

$$F_{\tau} = |\Phi_{\tau}|, \Phi_{\tau} = \{f_i\}, \quad (20)$$

где  $\Phi_{\tau}$  – массив, содержащий номера кадров в которых наблюдался срыв сопровождения,  $F_{\tau}$  – длина массива  $\Phi_{\tau}$  или количество срывов сопровождения алгоритмом,  $\tau$  – порог срыва сопровождения.

Пользуясь (20), определим интенсивность отказов

$$F_l = \frac{F_{\tau}}{N}, \quad (21)$$

где  $N$  – общее количество кадров в видеопоследовательности.

В отличие от метрики длины правильно сопровождаемых кадров (18), следующая оценка уменьшает влияние начальной части последовательности. Недостатком ее является то, что в оценке не отражается распределение отказов внутри последовательности. Так алгоритмы могут срываться равномерно через определенные промежутки, или могут срываться чаще на определенных участках. Чтобы учесть и минимизировать этот недостаток, следует провести анализ фрагментов траектории объекта между срывами сопровождения. На основе этого вводится понятие коэффициента распределения срывов сопровождения

$$F_r = \frac{1}{\log F_{\tau}} \sum_{f_i \in \Phi_{\tau}} \left( \frac{-\Delta f_i}{N} \log \frac{\Delta f_i}{N} \right), \quad (22)$$

где,  $f_i$  – позиция  $i$ -го сбоя в массиве  $\Phi_{\tau}$ ,  $\Delta f_i$  – определяется выражением

$$\Delta f_i = \begin{cases} f_{i+1} - f_i, & f_i < \max(\Phi_{\tau}) \\ f_1 + N - f_i, & f_i = \max(\Phi_{\tau}) \end{cases} \quad (23)$$

Особый случай для последней ошибки гарантирует, что полученное значение не искажается началом и концом последовательности. Расчет значения коэффициента  $F_r$  имеет смысл только когда есть интервалы между отказами ( $F_{\tau} > 1$ ). Максимальное значение  $F_r = 1$  достигается если срывы сопровождения распределены по равномерному закону распределения.

Данный критерий нельзя использовать как самостоятельную оценку алгоритмов сопровождения, однако его можно использовать как дополнительный критерий устойчивости алгоритмов сопровождения.

**Метрики оценки алгоритмов сопровождения нескольких объектов.** Основными метриками для оценки алгоритмов сопровождения нескольких объектов являются Multiple object detection precision (МОТР) и Multiple object detection accuracy (МОТА) [8] описываемые уравнениями ниже, соответственно

$$МОТР = \frac{\sum_{i=1}^M \sum_{t=1}^N (\varphi_{i,t})}{\sum_{t=1}^N (M_t)}, \quad (24)$$

где  $M$  – количество объектов,  $M_t$  – количество объектов в момент времени  $t$ . В случае сопровождения одного объекта  $M = M_t$ .

$$МОТА = 1 - \frac{\sum_{t=1}^N (c_m MI_t + c_f FP_t + c_s SW_t)}{\sum_{t=1}^N (N_t^G)}, \quad (25)$$

где  $MI_t$  – количество пропусков,  $FP_t$  – количество неправильных обнаружений,  $SW_t$  – количество срывов захватов на другой, ложный объект,  $c_m, c_f, c_s$  – весовые коэффициенты,  $N_t^G$  – количество объектов в момент времени  $t$ .

**Метрики оценки алгоритмов сопровождения при потере визуального контакта с объектом.** В случае кратковременной потери визуального контакта с объектом интереса (скрытие объекта за препятствиями), алгоритмы автоматического сопровождения оцениваются вероятностью срыва сопровождения, оценка вероятности рассчитывается по формуле

$$P^* = \frac{n}{N}, \quad (26)$$

где  $n$  – количество случаев срывов сопровождения,  $N$  – общее количество выхода объекта из-за препятствия. В зависимости от числа кадров, в которых объект находится за препятствием, оценка вероятности срыва сопровождения изменяет свое значение. Для анализа алгоритма и визуализации оценки, предпочтительно строить график зависимости оценки вероятности срыва сопровождения от количества кадров, в которых объект находится за препятствием (рис. 5).

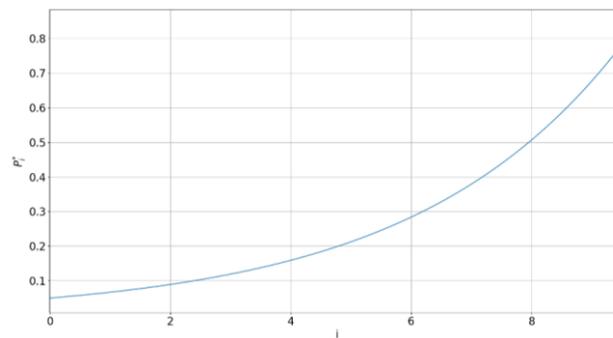


Рис. 5. Зависимость оценки вероятности срыва сопровождения от количества кадров в которых объект находился за препятствием

Так же можно оценивать вероятность для всех, имеющихся в наборе данных, ситуаций скрытия объекта сопровождения за препятствиями, рассчитанную по формуле ниже

$$P_m^* = \frac{1}{N} \sum_{i=1}^N P_i^*, \quad (27)$$

где  $N$  – максимальное количество кадров в которых объект находится за препятствием,  $i$  – количество кадров в которых объект находится за препятствием,  $P_i^*$  – оценка вероятности срыва сопровождения в случае перекрытия объекта на  $i$  кадров. Выражение (31) можно описать как нормализованную площадь под графиком, представленном на рисунке 6. Для анализа точности и надежности оценки вероятности может быть использована формула, описывающая зависимость доверительного интервала и доверительной вероятности от количества экспериментов  $n$ , при условии большого значения  $n$  [9]:

$$\beta = P\{|p^* - p| < \varepsilon\} = 2\Phi\left(\frac{2\sqrt{n}}{\sqrt{p^*q^*}}\right), \quad (28)$$

где  $p^*$  – оценка вероятности,  $p$  – оцениваемая вероятность,  $\varepsilon$  – доверительный интервал,  $n$  – количество экспериментов,  $q^* = 1 - p^*$ ,  $\Phi$  – функция Лапласа,  $\beta$  – доверительная вероятность (вероятность того, что доверительный интервал  $(p^* - \varepsilon, p^* + \varepsilon)$  покрывает оцениваемую вероятность  $p$ ).

Пользуясь графиком из рисунка 6, можно грубо оценить максимальное количество кадров, в которых объект находится за препятствием, и при котором с заданной вероятностью произойдет срыв сопровождения. Для случая, когда  $i$  равно 0 (объект не перекрывается), значение оценки  $P_0^*$  рассчитывается по формуле (26).

Выше дан обзор метрик, используемых для оценки и сравнения алгоритмов сопровождения объектов на видеопоследовательности. Каждая из метрик отражает определенный критерий алгоритма и может быть использована как сама по себе, так и в составе с другими метриками, дополняющими друг друга.

**Комбинированные метрики для оценки алгоритмов сопровождения.** Комбинированная метрика (CoPTS – combined tracking performance score) комбинирует информацию о точности сопровождения и количестве сбоев в одной оценке и описывается выражением, представленным ниже:

$$CoTPS = \beta\Omega + (1 - \beta)\lambda_0, \quad (29)$$

где  $\lambda_0$  – коэффициент не сопровождаемых кадров,  $\beta$  – коэффициент сопровождаемых кадров,  $\Omega$  – коэффициент качества сопровождения,  $\lambda_0, \beta, \Omega$  определяются выражениями, соответственно

$$\lambda_0 = \frac{N_0}{N}, \quad (30)$$

$$\beta = \frac{\hat{N}}{N}, \quad (31)$$

$$\Omega = \sum_{\tau \in (0,1]} \left(\frac{\hat{N}(\tau)}{N}\right), \quad (32)$$

где  $N_0$  – количество кадров где IOU = 0,  $\hat{N}$  – количество кадров где IOU больше 0,  $\hat{N}(\tau)$  – количество кадров, на который перекрытие больше 0, но меньше порога  $\tau$ .  $\hat{N}(\tau)$  определяется выражением

$$N(\tau) = |\{j: \varphi_j \geq 0 \wedge \varphi_j \leq \tau\}|. \quad (33)$$

Выражение (23) на самом деле является аппроксимацией интеграла по порогу и может быть переопределено как

$$\Omega = \int_0^1 \left(\frac{N(\tau)}{N}\right) d\tau = 1 - \int_0^1 \frac{P(\tau)}{N} d\tau, \quad (34)$$

где  $P(\tau) = |\{j: \varphi_j \geq \tau\}|$  количество кадров с перекрытием больше порога.

Также для сравнения алгоритмов сопровождения может использоваться оценка точности (насколько верно алгоритм определяет координаты и область объекта) и устойчивости (сколько раз алгоритм терял объект сопровождения). Далее оценка, основанная на паре точности и устойчивости, называется А–R (Accuracy – Robustness) пара.

Стоит отметить, что есть сходство между метриками CoTPS [1] и А–R. В обеих метриках учитывается два одинаковых аспекта сопровождения. Несмотря на это сходство, А–R пара лучше подходит для сравнения алгоритмов сопровождения по некоторым причинам:

- ◆ Выбранные меры не коррелированы.
- ◆ Протокол контролируемой оценки эффективнее использует последовательность кадров за счет повторного запуска алгоритма сопровождения в случае сбоя.
- ◆ Использование двух значений оценки для описания алгоритма дает больше возможностей для сравнения алгоритмов, за счет увеличения комбинаций результатов.

Согласно данным из [11], для А–R пары в качестве меры точности следует использовать математическое ожидание коэффициента Жаккара, но для алгоритмов основной задачей которых является точное сопровождение центра объекта, желательно использовать нормализованную среднеквадратическую ошибку определения центра, рассчитанную по локальному МО (11), т.к. в [11] указано, что класс оценок, основанных на ошибке определения центра объекта слабо коррелирует с интенсивностью отказов. Очевидно, что для А–R пары следует использовать пары метрик, характеризующихся слабой корреляцией.

В случае использования в качестве меры точности математического ожидания коэффициента Жаккара, А–R пара имеет вид:

$$A - R(L^G, L^T) = (\bar{\Phi}(L^G, L^T), F_0(L^G, L^T)), \quad (33)$$

где  $\bar{\Phi}$  – среднее значение перекрытия  $F_0$  – интенсивность отказов при пороге  $\tau = 0$ . Значение порога  $\tau$  влияет на конечный результат. Если порог установить близким к 1, то алгоритм будет часто реинициализирован из-за возникновения даже слабых отклонений и итоговый результат будет слабо информативен. В формуле (33) используется минимальный порог  $\tau = 0$ , для того, чтобы, повторно запускать алгоритм сопровождения в случае полного несоответствия предсказанного региона с регионом объекта. Теоретически, алгоритм сопровождения может определять большую область вокруг объекта и, тем самым, избегать сбоев сопровождения, но тогда снизится точность определения области объекта. Данный случай является примером того как две оценки дополняют друг друга.

Оценки, полученные метрикой А–R, наиболее наглядно показывают различия между алгоритмами, если их отобразить на графике, как показано на рис. 6.

На рис. 6 каждый из сравниваемых алгоритмов имеет свои координаты в зависимости от значений оценки точности и устойчивости, рассчитанных по формуле (33). Алгоритмы, обладающие большей точностью, находятся выше на графике, а алгоритмы, обладающие большей устойчивостью правее.

Так как количество сбоев алгоритма сопровождения не имеет предела, то для визуализации оценки устойчивости алгоритма используется оценка, рассчитанная по формуле, приведенной ниже

$$R_s = e^{-SM}, \quad (34)$$

где  $S$  – количество кадров после последнего срыва,  $M$  – среднее время между сбоями сопровождения определяется как

$$M = \frac{F_0}{N}, \quad (35)$$

где  $F_0$  – количество сбоев сопровождения,  $N$  – длина последовательности.

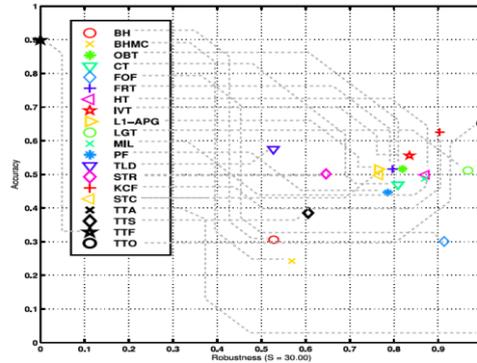


Рис. 6. Оценка точности и устойчивости алгоритмов

Выражение (35) может быть интерпретировано как вероятность того, что алгоритм сопровождения успешно сопроводит объект на протяжении  $S$  кадров с момента последнего срыва сопровождения.

**Выводы.** В связи с тем, что алгоритм автоматического сопровождения должен решать задачу в сложных условиях (смена ракурса объекта интереса, заезд за препятствия, изменения яркости и контрастности объекта интереса и фона), для оценки алгоритма требуется проводить большое количество экспериментов с учетом влияния внешних факторов на алгоритм сопровождения, как по отдельности, так и в совокупности друг с другом. Для оценки степени влияния вышеупомянутых факторов на алгоритм сопровождения рекомендуется пользоваться метриками оценки, описанными выше.

Из обзора следует выделить что, для оценки точности и устойчивости алгоритма, большей описательной способностью обладает метрика пары точности и устойчивости  $A-R$ , где точность, в зависимости от подхода, может оцениваться через метрики, основанные на пересечении областей объекта истинной и предсказанной, так и основанные на ошибке определения центра объекта. Для анализа сбоев сопровождения следует использовать метрику частоты отказов, так как она приближена к использованию алгоритма сопровождения в реальных условиях и позволяет получать оценки на всей длине последовательности кадров. Для анализа алгоритма сопровождения при потере визуального контакта с объектом интереса необходимо использовать оценку вероятности восстановления процесса сопровождения после выхода объекта из-за препятствия или оценку вероятности срыва сопровождения при заходе объекта за препятствие. Стоит отметить что в зависимости от количества кадров в которых объект не виден, меняется оценка вероятностей. Для более детальной оценки алгоритмов сопровождения рекомендуется строить график зависимости оценки вероятности срыва от количества кадров, как на рис. 5 и рассчитывать среднюю оценку вероятностей.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Nawaz T., Cavallaro A. A protocol for evaluating video trackers under real-world conditions // Image Processing IEEE. – 2013. – Vol. 22, No. 4. – P. 1354-1361.
2. Godec M., Roth P., Bischof H. Hough-based tracking of nonrigid objects // ICCV. – Barcelona: IEEE, 2011. – P. 81-88.

3. Wang D., Lu H., Yang M. Online object tracking with sparse prototypes // IEEE Image Processing Conference. – 2013. – Vol. 22, No. 1. – P. 314-325.
4. Wu Y., Lim J., Yang M.-h. Online Object Tracking: A Benchmark // CVPR. – 2013.
5. Everingham M., Gool L., Williams I., Winn J., Zisserman A. The Pascal Visual Object Classes (VOC) Challenge // IJCV. – 2009. – Vol. 88, No. 2. – P. 303-338.
6. Kristan M., Kovačič S., Leonardis A., Pers J. A two-stage dynamic model for visual tracking // Transactions Systems Man Cybernetics Part B. – 2010. – Vol. 40, No. 6. – P. 1505-1520.
7. Khan Z., Balch T., Dellaert F. MCMC-based particle filtering for tracking a variable number of interacting targets // IEEE Transactions Pattern Analysis. – 2005. – Vol. 27. – P. 1805-1819.
8. Kasturi R., Goldgof D., Soundararajan P., Manohar V., Garofolo J., Bowers R., Boonstra M., Korzhova V., Zhang J. Framework for performance evaluation of face, text, and vehicle detection and tracking in video: data, metrics, and protocol // TPAMI. – 2009. – Vol. 31, No. 2. – P. 319-326.
9. Венцель Е.С., Овчаров Л.А. Теория вероятности и ее инженерные приложения: учеб. пособие для вузов. – 2-е изд. – М.: Высш. шк., 2000. – 480 с.
10. Yang M., Wu Y., Hua G. Context-aware visual tracking // IEEE transactions on pattern analysis and machine intelligence. – 2006. – No. 31. – P. 1195-1209.
11. Stauffer C., Grimson W. Learning patterns of activity using real-time tracking // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2000. – P. 747-757.
12. Sundaresan A., Chellappa R. Multi-camera tracking of articulated human motion using shape and motion cues // IEEE Transactions on Image Processing. – 2009. – P. 2114-2126.
13. Smeulders A.M., Chu D., Cucchiara, S. Calderara, A., Shah M. Visual Tracking: an Experimental Survey // TPAMI. – 2013. – Vol. 21, No. 3. – P. 152-168.
14. Henriques F., Caseiro R., Martins P., Batista J. High-Speed Tracking with Kernelized Correlation Filters // TPAMI. – 2014. – Vol. 42, No. 5. – P. 345-362.
15. Frey J. Dueck D. Clustering by Passing Messages Between Data Points // Science Today. – 2007. – Vol. 315. – P. 972-976.
16. Smith K., Gatica-Perez D., Odohez J. Evaluating Multi-Object Tracking // CVPR Work. – 2005. – Vol. 3, IEEE. – P. 32-36.
17. Black J., Ellis T., Rosin P. A novel method for video tracking performance evaluation // VS-PETS. – 2003. – C. 125-132.
18. Kao E., Daggett M., Hurley M. An information theoretic approach for tracker performance evaluation // CVPR. – 2009. – P. 1523-1529.
19. Bashir F. Porikli F. Performance Evaluation of Object Detection and Tracking Systems // PETS. – 2006. – P. 190-203.
20. Checka N., Wilson K., Rangarajan V., Darrell T. A probabilistic framework for multi-modal multi-person tracking // Proceedings of the IEEE Workshop on Multi-Object Tracking (WOMOT '03). – 2003. – P. 203-212.

## REFERENCES

1. Nawaz T., Cavallaro A. A protocol for evaluating video trackers under real-world conditions, *Image Processing IEEE*, 2013, Vol. 22, No. 4, pp. 1354-1361.
2. Godec M., Roth P., Bischof H. Hough-based tracking of nonrigid objects, *ICCV*. Barcelona: IEEE, 2011, pp. 81-88.
3. Wang D., Lu H., Yang M. Online object tracking with sparse prototypes, *IEEE Image Processing Conference*, 2013, Vol. 22, No. 1, pp. 314-325.
4. Wu Y., Lim J., Yang M.-h. Online Object Tracking: A Benchmark, *CVPR*, 2013.
5. Everingham M., Gool L., Williams I., Winn J., Zisserman A. The Pascal Visual Object Classes (VOC) Challenge, *IJCV*, 2009, Vol. 88, No. 2, pp. 303-338.
6. Kristan M., Kovačič S., Leonardis A., Pers J. A two-stage dynamic model for visual tracking, *Transactions Systems Man Cybernetics Part B*, 2010, Vol. 40, No. 6, pp. 1505-1520.
7. Khan Z., Balch T., Dellaert F. MCMC-based particle filtering for tracking a variable number of interacting targets, *IEEE Transactions Pattern Analysis*, 2005, Vol. 27, pp. 1805-1819.
8. Kasturi R., Goldgof D., Soundararajan P., Manohar V., Garofolo J., Bowers R., Boonstra M., Korzhova V., Zhang J. Framework for performance evaluation of face, text, and vehicle detection and tracking in video: data, metrics, and protocol, *TPAMI*, 2009, Vol. 31, No. 2, pp. 319-326.

9. *Ventsel' E.S., Ovcharov L.A.* Teoriya veroyatnosti i ee inzhenernye prilozheniya: ucheb. posobie dlya vuzov [Probability theory and its engineering applications: a textbook for universities]. 2nd ed. Moscow: Vyssh. shk., 2000, 480 p.
10. *Yang M., Wu Y., Hua G.* Context-aware visual tracking, *IEEE transactions on pattern analysis and machine intelligence*, 2006, No. 31, pp. 1195-1209.
11. *Stauffer C., Grimson W.* Learning patterns of activity using real-time tracking // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. – 2000. – P. 747-757.
12. *Sundaresan A., Chellappa R.* Multi-camera tracking of articulated human motion using shape and motion cues, *IEEE Transactions on Image Processing*, 2009, pp. 2114-2126.
13. *Smeulders A. M., Chu D., Cucchiara, S. Calderara, A., Shah M.* Visual Tracking: an Experimental Survey, *TPAMI*, 2013, Vol. 21, No. 3, pp. 152-168.
14. *Henriques F., Caseiro R., Martins P., Batista J.* High-Speed Tracking with Kernelized Correlation Filters, *TPAMI*, 2014, Vol. 42, No. 5, pp. 345-362.
15. *Frey J. Dueck D.* Clustering by Passing Messages Between Data Points, *Science Today*, 2007, Vol. 315, pp. 972-976.
16. *Smith K., Gatica-Perez D., Odobez J.* Evaluating Multi-Object Tracking, *CVPR Work*, 2005, Vol. 3, IEEE, pp. 32-36.
17. *Black J., Ellis T., Rosin P.* A novel method for video tracking performance evaluation, *VS-PETS*, 2003, pp. 125-132.
18. *Kao E., Daggett M., Hurley M.* An information theoretic approach for tracker performance evaluation, *CVPR*, 2009, pp. 1523-1529.
19. *Bashir F. Porikli F.* Performance Evaluation of Object Detection and Tracking Systems, *PETS*, 2006, pp. 190-203.
20. *Checka N., Wilson K., Rangarajan V., Darrell T.* A probabilistic framework for multi-modal multi-person tracking, *Proceedings of the IEEE Workshop on Multi-Object Tracking (WOMOT '03)*, 2003, pp. 203-212.

Статью рекомендовал к опубликованию д.т.н. Н.Е. Сергеев.

**Щелкунов Александр Евгеньевич** – АО «Научно–конструкторское бюро вычислительных систем» (АО НКБ ВС); e-mail: alexandrshcelkunov@mail.ru; г. Таганрог, ул. 1-я Линия; программист.

**Ковалев Владислав Владимирович** – e-mail: vlad.kovalev94@mail.ru; инженер-исследователь.

**Морев Кирилл Иванович** – e-mail: morev-ki@ya.ru; программист.

**Сидько Иван Владимирович** – e-mail: ivan.brokenroll@gmail.com; заведующий отдела; к.т.н.

**Shchelkunov Alexander Evgen'evich** – Joint Stock Company "Scientific Design Bureau of Computing Systems" (JSC SDB CS); e-mail: alexandrshcelkunov@mail.ru; 1st Liniya street, Taganrog, Russia; programmer.

**Kovalev Vladislav Vladimirovich** – e-mail: vlad.kovalev94@mail.ru; research-engineer.

**Morev Kirill Ivanovich** – e-mail: morev-ki@ya.ru; programmer.

**Sidko Ivan Vladimirovich** – e-mail: ivan.brokenroll@gmail.com; department head; cand. of eng. sc.