

## Нечёткая параметризованная модель классификации блоков небинаризованной маски движения

Мацыпаев Дмитрий Александрович, Таганрогский технологический институт Южного федерального университета, +7 960 4491739, [dmitry.matsypaev@gmail.com](mailto:dmitry.matsypaev@gmail.com)

Броневи́ч Андрей Георгиевич, ОАО «Научно-исследовательский и проектно-конструкторский институт, автоматизации и связи на железнодорожном транспорте», +7 (495) 967 77 06\*329, [brone@mail.ru](mailto:brone@mail.ru)

### Введение

Предметом данной статьи является задача выделения движущихся областей в видеопотоке. Суть предложенного алгоритма заключается в анализе небинаризованной маски движения. Такая маска движения содержит больше информации, чем её бинарный аналог после применения порогового преобразования. Небинаризованная маска движения разбивается на прямоугольные блоки, для каждого из которых считается соответствующее значение статистики, усредняющей меру движения внутри блока. Переход от пиксельного представления маски движения к блочному представлению позволяет сократить размерность области поиска и снизить влияние шумов на итоговый результат. В то же время использование блочного представления маски движения ограничивает геометрическую точность результата линейным размером блока. Для анализа функции распределения значений статистики блоков предлагается параметризованная математическая модель движущихся и покоящихся блоков. В данной модели вводятся функции принадлежности блоков данным множествам, предлагается критерий чёткого разграничения между двумя множествами. Далее вводятся характеристики, которые помогают оценить степень достоверности полученных результатов.

### Краткая постановка задачи

Пусть имеется последовательность кадров видеоряда  $\{F^{(k)}\}_{k=0,1,2,\dots}$ , которая была получена при помощи статической монокулярной видеокамеры с фиксированным фокусным расстоянием. Будем считать каждый кадр полутоновым изображением с разрешением  $w_F \times h_F$  пикселей. Необходимо обнаруживать движущиеся объекты на каждом кадре видеоряда и сопровождать их в видеопотоке. Основными этапами решения этой задачи являются:

1. построение маски движения для текущего кадра видеопоследовательности,
2. сегментация маски движения и выделение на ней областей, соответствующих движущимся объектам или группам объектов,
3. идентификация и сопровождение обнаруженных объектов.

Более подробно постановка задачи рассмотрена в [8]. В рамках данной статьи рассматриваются первый и второй из выше перечисленных этапов решения задачи. В качестве маски движения предлагается использовать абсолютную разность двух соседних кадров видеоряда без применения порогового преобразования. Получаемая в результате небинаризованная маска движения разбивается на прямоугольные блоки, для каждого из которых рассчитывается значение введённой статистики. Анализ функции распределения данной статистики позволяет построить критерий разграничения блоков по признаку наличия движения. Выделение движущихся областей на маске движения осуществляется на основе отношения связности между движущимися блоками.

## Небинаризованная маска движения

Общеизвестным и интуитивно понятным считается факт, что информации, содержащейся непосредственно в самом кадре  $F^{(k)}$  недостаточно для того, чтобы локализовать области движения на нём. Необходимым условием для выделения движения является возможность доступа к информации, содержащейся в двух последних кадрах видеоряда. В настоящее время наиболее популярными являются алгоритмы, в рамках которых используются статистические модели заднего плана для каждого пикселя кадра. Для построения и обновления таких моделей при анализе каждого следующего кадра требуется учитывать информацию, содержащуюся в нескольких десятках или сотнях последних кадров видеоряда [1, 3, 5, 7]. Это может значительно увеличить количество вычислений и объём используемой оперативной памяти.

Независимо от природы используемого подхода, результаты работы таких алгоритмов для кадра  $F^{(k)}$  представляются в виде маски движения  $M^{(k)}$ . Маска движения является полутоновым изображением с разрешением  $w_F \times h_F$ . Интенсивность пикселя в  $M^{(k)}$  отражает степень принадлежности соответствующего пикселя из  $F^{(k)}$  множеству движущихся пикселей. Широко распространённым является стремление получить бинарную маску движения путём применения порогового преобразования, что при удачном результате позволяет сразу же приступить к непосредственной сегментации движущихся объектов. В то же время использование бинарной маски движения имеет ряд недостатков, связанных с потерей существенной доли важной информации, которая содержалась в маске движения непосредственно перед её бинаризацией [8].

В [5] в качестве простейшего способа для построения маски движения рассматривается использование порогового преобразования для абсолютной разности интенсивностей соответствующих пикселей на двух соседних кадрах:

$$M_{ij}^{(k)} = \begin{cases} I_{\max}, & |F_{ij}^{(k)} - F_{ij}^{(k-1)}| \geq h \\ 0, & |F_{ij}^{(k)} - F_{ij}^{(k-1)}| < h \end{cases}.$$

Значение  $I_{\max} > 0$  задаёт диапазон, в котором изменяется величина интенсивности пикселя изображения. Обычно  $I_{\max} = 2^d - 1$ , где  $d$  - глубина представления цвета полутонового изображения в битах. При анализе видео практически всегда приходится работать с кадрами, глубина цвета которых равна  $d = 8$  бит, что соответствует  $I_{\max} = 255$ . Очевидно, что  $0 \leq |F_{ij}^{(k)} - F_{ij}^{(k-1)}| \leq I_{\max}$ . Соответственно,  $0 \leq M_{ij}^{(k)} \leq I_{\max}$ . Величина  $h$  является параметром порогового преобразования.

Очевидным преимуществом данного метода является его высокое быстродействие. В то же время в [5] подчёркивается, что этот метод не применим на практике вследствие слишком высокой чувствительности к значению  $h$ . Отметим также, что для формулы с пороговым преобразованием не предлагается никакой робастной эвристики для выбора значения  $h$ . Таким образом, при решении задачи не используется большое количество априорной информации, которая может существенно ограничить пространство поиска.

В этой статье мы постараемся показать, что абсолютная разность между двумя соседними кадрами может содержать в себе достаточно информации для устойчивой и аккуратной локализации движущихся областей на видео, если вместо применения порогового преобразования работать с небинаризованной маской движения

$$M_{ij}^{(k)} = |F_{ij}^{(k)} - F_{ij}^{(k-1)}|, \quad (1)$$

используя методы обобщённой теории информации[4] и априорную информацию, которая вытекает из постановки задачи.

### **Разбиение маски движения на блоки фиксированного размера и расчёт статистик для блоков**

Использование формулы (1) для расчёта маски движения в идеальном случае отсутствия шумов на кадрах видеопоследовательности  $\{F^{(k)}\}_{k=0,1,2\dots}$  позволяет для всех пикселей маски  $M^{(k)}$  построить отношение порядка, упорядочив их по возрастанию соответствующего значения интенсивности. Введение такого отношения порядка очень важно для разграничения пикселей на множества движущихся и покоящихся. Между тем, в реальных приложениях шум присутствует всегда. Это без применения дополнительных процедур, например, медианной фильтрации, влечёт большое количество ложных срабатываний при обнаружении движущихся пикселей.

Другая проблема попиксельного анализа заключается в том, что даже при самых низких разрешениях кадров видеопоследовательности ( $320 \times 240$ ,  $640 \times 480$ ) количество пикселей исчисляется сотнями тысяч. Поскольку одним из стандартных требований к алгоритмам видеоанализа является способность работать в режиме реального времени при достаточно высоком значении FPS, то это накладывает серьёзные ограничения на сложность и эффективность применяемых алгоритмов. В то же время, априори известно, что движущиеся пиксели являются частями движущихся объектов в кадре и соответственно группируются в кластеры на основе геометрической близости, а шумовые пиксели представляются в виде одиночно располагающихся выбросов. Подобная априорная информация часто используется в компьютерном зрении, например при медианной фильтрации шумов или при уменьшении разрешения исходного изображения или видео. Определение областей движения обычно производится через выделение областей-кандидатов и на основе дискриминативных статистик. В общем случае подобный подход позволяет получить не только более робастные результаты, но и существенно сократить время анализа.

Одним из подходов, учитывающих групповые характеристики пикселей, является разбиение изображения на прямоугольные непересекающиеся области, которые называются блоками. Блочный подход хорошо зарекомендовал себя при сжатии и кодировании видео. В частности в таком современном стандарте сжатия видео как H264 разбиение кадров на блоки размером от  $4 \times 4$  до  $16 \times 16$  позволяет определять векторы движения и достаточно точно выделять области движения на видео [6]. В работе [9] описано применение алгоритма рекурсивного блочного разбиения применительно к произвольной бинарной маске движения, которая получалась после применения одного из нескольких современных методов построения бинарной маски движения, представленных в библиотеке OpenCV [1]. Анализ полученных результатов показал, что данный алгоритм хорошо отделяет блобы на маске движения, если на ней отсутствуют шумы, тени и «призраки». В работе [8] была рассмотрена другая модификация блочного метода (использовались блоки фиксированного размера) применительно к небинаризованной маске движения. Был предложен критерий разграничения неподвижных и движущихся блоков, основанный на применении энтропии Шеннона при разграничении множества блоков, на котором была введена аддитивная мера количества движения. Также в [8] была поставлена задача обнаружения движения на видео и предложены простейшие критерии качества полученного разбиения. Главным недостатком предложенного метода оказался тот факт, что оптимальное значение параметра метода  $k$ , от которого зависит разграничение движущихся и неподвижных блоков, может варьироваться в достаточно широком диапазоне значений при анализе кадров, принадлежащих одной видеопоследовательности. Тем не менее, ана-

лиз результатов показал перспективность обработки небинаризованной маски движения и использования блочного метода.

Недостатком такого подхода является более низкая геометрическая точность получаемых результатов, ограниченная характерным размером анализируемых областей. С другой стороны, при решении далеко не каждой задачи необходима пиксельная точность, а геометрическую точность результатов можно варьировать, изменяя размер блока.

Будем использовать процедуру разбиения на блоки фиксированного размера  $w_B \times h_B$ , аналогичную той, что представлена в [8]. Введём статистику для блока  $B$ , равную среднему значению интенсивности пикселей в блоке  $B$ :

$$I(B) = \frac{1}{w_B h_B} \sum_{i=0}^{w_B-1} \sum_{j=0}^{h_B-1} B_{ij}. \quad (2)$$

Очевидно, что множество значений статистики  $I(B)$  ограничено  $0 \leq I(B) \leq I_{\max}$ .

Относительно статистики  $I$  мы также делаем предположение, что она достаточно хорошо упорядочивает блоки маски движения по степени принадлежности к объектам переднего и заднего планов. Чем больше значение статистики  $I$ , тем с большей достоверностью соответствующий блок относится к переднему плану. Выполнение данного предположения также означает, что для всех блоков  $B \in U$  существует некоторое пороговое значение  $I_{th} \in [0, I_{\max}]$ , которое зависит от параметров алгоритма и для которого верно:

1. если  $I(B) \geq I_{th}$ , то блок  $B$  является движущимся блоком,
2. если  $I(B) < I_{th}$ , то блок  $B$  является покоящимся блоком.

Отметим также, что введение блоков размером  $w_B \times h_B$  позволяет сократить количество анализируемых объектов на изображении с разрешением  $w_F \times h_F$  приблизительно в  $w_B h_B$  раз. Число получающихся при разбиении блоков равно  $n \approx \frac{w_F h_F}{w_B h_B}$ . Это оказывается

весьма существенным вкладом в увеличение производительности в случае использования алгоритмов с суперлинейной арифметической сложностью. Так, ниже мы будем сортировать блоки по значению рассчитанной статистики  $I$ . Арифметическая сложность наилучших алгоритмов сортировки равна  $O(n \log n)$ . Это означает, что при размере блоков  $16 \times 16$  и разрешении изображения в  $640 \times 480$  сортировать блоки примерно в 450 раз эффективнее по быстрдействию, чем сортировать соответствующие пиксели.

## Параметризованная модель заднего и переднего планов

Будем называть множество блоков, соответствующих покоящимся на видео кадре пикселям, множеством блоков заднего плана и обозначать как  $U_{BG}$ . Множество блоков, соответствующих областям движущихся пикселей, будем называть множеством блоков переднего плана и обозначать как  $U_{FG}$ . Всё множество имеющихся блоков будем обозначать символом  $U$ .

Выше уже было изложено предположение о том, что статистика  $I$  в (2) хорошо разделяет движущиеся и покоящиеся блоки маски движения, полученной согласно выражению (1). Чем больше статистика  $I$  соответствующего блока, тем с большей степенью достоверности его можно отнести к множеству  $U_{FG}$ . Это предположение было подтверждено на практике во время проведённых экспериментальных исследований. Также в ходе эксперимента были отмечены ещё две закономерности:

- 1) блоки заднего плана в большинстве случаев количественно доминируют над блоками переднего плана;

2) распределение статистики  $I$  для блоков переднего плана доминирует над соответствующим распределением блоков заднего плана, занимая существенно более значительную часть диапазона значений  $[0, I_{\max}]$ .

Данные закономерности хорошо могут быть проиллюстрированы (см. рис. 1) типичным графиком функции распределения статистики  $F_I(x) = \frac{1}{|U|} \sum_{B \in U | I(B) \leq x} 1$ .

Формализация приведённых выше закономерностей привела к созданию параметризованной модели заднего и переднего планов. Упорядочим блоки по возрастанию соответствующей им статистики  $I$ . Формализуем первую из закономерностей, изложенных выше. Для этого предположим, что к множеству  $U_{BG}$  относится не меньше, чем  $\alpha n$  блоков, где  $\alpha \in (0,1)$  - параметр модели. Поскольку мы имеем упорядоченный по статистике  $I$  массив блоков и предполагаем, что данная статистика хорошо разделяет блоки переднего и заднего плана, введём величину

$$I_\alpha = \min_{x \in [0, I_{\max}]} \{x \mid F_I(x) \geq \alpha\}.$$

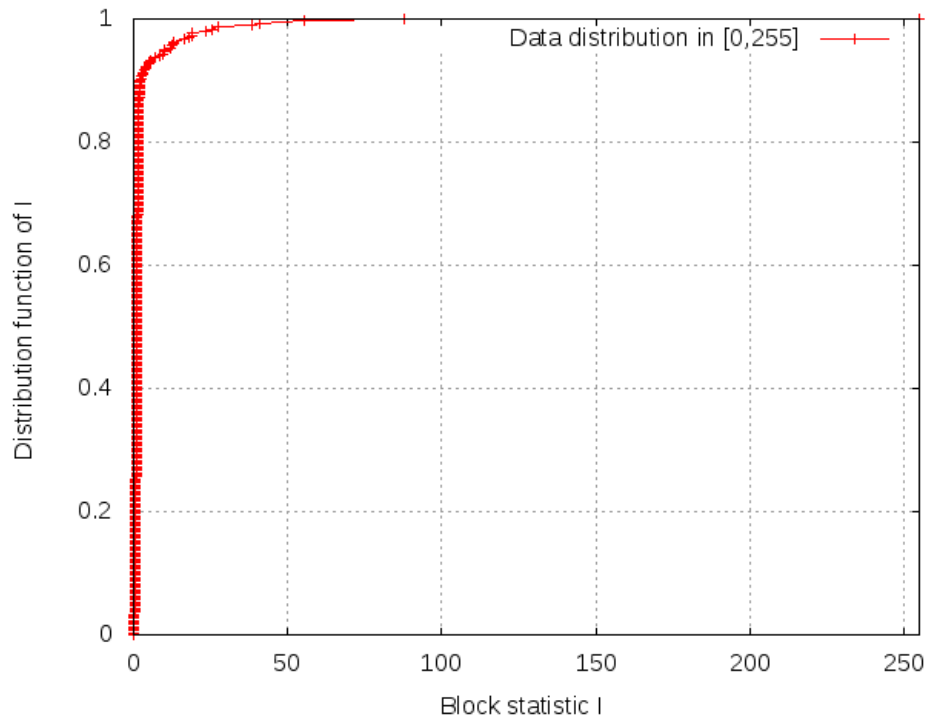


Рис. 1 Функция распределения статистики  $I$

Степень принадлежности блока  $B$  множеству  $U_{BG}$  будем измерять с помощью функции  $\eta_{BG}$  от статистики  $I(B)$ . Будем считать, что функция  $\eta_{BG}$  должна удовлетворять следующим условиям:

1.  $\eta_{BG}(x)$  определена при  $x \in [0, I_{\max}]$ ,
2.  $\eta_{BG}(x) = 1$ ,  $x \in [0, I_\alpha]$ ,
3.  $\eta_{BG}(I_{\max}) = 0$ ,
4.  $\eta_{BG}(x)$  является монотонно невозрастающей функцией на  $[0, I_{\max}]$   
 $\forall x, y \in [0, I_{\max}] \quad x \geq y \rightarrow \eta_{BG}(x) \leq \eta_{BG}(y)$ .

Для формализации второй закономерности предположим, что распределение статистики  $I$  блоков переднего плана занимает долю  $\beta \in (0,1)$  диапазона значений  $[0, I_{\max}]$ ,

где  $\beta$ -параметр модели. Учитывая, что статистика  $I$  хорошо разделяет блоки переднего и заднего плана, введём величину

$$I_\beta = (1 - \beta)I_{\max}.$$

Степенью принадлежности блока  $B$  множеству  $U_{FG}$  будем считать значение функции  $\eta_{FG}$  от соответствующей статистики  $I(B)$ . Будем считать, что функция  $\eta_{FG}$  должна удовлетворять следующим условиям:

1.  $\eta_{FG}(x)$  определена при  $x \in [0, I_{\max}]$ ,
2.  $\eta_{FG}(x) = 1$ ,  $x \in [I_\beta, I_{\max}]$ ,
3.  $\eta_{FG}(0) = 0$ ,
4.  $\eta_{FG}(x)$  является монотонно неубывающей функцией на  $[0, I_{\max}] \forall x, y \in [0, I_{\max}]$   
 $x \geq y \rightarrow \eta_{FG}(x) \geq \eta_{FG}(y)$ .

### Функционал Тверского для оценки степени схожести двух множеств и его применение в модели заднего и переднего планов

Как показал эксперимент, среди всего множества функций  $\eta_{BG}$  и  $\eta_{FG}$ , удовлетворяющих наложенным на них условиям, наиболее удачно согласуются с интуитивным представлением о множествах  $U_{BG}$  и  $U_{FG}$  функции, выведенные из параметризованного функционала Тверского. Параметризованный функционал Тверского используется в теории нечётких множеств [2, 4] для оценки степени схожести или сравнимости двух множеств [2]:

$$T_{a,b}(X, Y) = \frac{f(X \cap Y)}{f(X \cap Y) + af(X \setminus Y) + bf(Y \setminus X)},$$

где

$X, Y$  - множества, схожесть которых необходимо оценить,

$\cup, \cap, \setminus$  - согласованные операции объединения, пересечения и разности двух в общем случае нечётких множеств,

$f$  - функционал, для которого верно  $f(X \cup Y) = f(X) + f(Y)$   
 $\forall X, Y : X \cap Y = \emptyset$ ,

$a, b \in R$  - параметры функционала Тверского.

Введём в рассмотрение множество  $g$  ограниченных чётких подмножеств действительной оси. Очевидно, что  $\forall G_1, G_2 \in g$  справедливы следующие соотношения:

1.  $G_1 \cup G_2 \in g$ ,
2.  $G_1 \cap G_2 \in g$ ,
3.  $G_1 \setminus G_2 \in g$ .

Известно, что  $\forall G_1, G_2 \in g \mid G_1 \cap G_2 = \emptyset \rightarrow L(G_1) + L(G_2) = L(G_1 + G_2)$ , где  $L$ -мера Лебега. Поэтому можно определить функционал  $f$  как  $f(G) = L(G)$  для всех  $G \in g$ . Очевидно, что  $\forall \gamma, x \geq 0$  сегмент  $[-\gamma/2, x] \in g$ , а  $L([-\gamma/2, x]) = x + \gamma/2$ . Не сложно также показать верность следующих соотношений:

1.  $L([-\gamma/2, x] \cap [-\gamma/2, y]) = \min\{x, y\} + \gamma/2 \quad \forall x, y, \gamma \geq 0$ ,
2.  $L([-\gamma/2, x] \setminus [-\gamma/2, y]) = \max\{x - y, 0\} \quad \forall x, y, \gamma \geq 0$ .

Введём функцию  $\varphi(x, y)$ , численно равную значению функционала Тверского при  $f(X) = L(X)$ ,  $a = 1/2$ ,  $b = 0$  от множеств  $[-\gamma/2, x]$  и  $[-\gamma/2, y]$   $\forall x, y, \gamma \geq 0$ :

$$\varphi(x, y) = T_{1/2, 0}([-\gamma/2, x], [-\gamma/2, y]) = \frac{2 \min\{x, y\} + \gamma}{2 \min\{x, y\} + \gamma + \max\{x - y, 0\}}.$$

Нетрудно видеть, что функция  $\varphi(x, I_\alpha)$  как функция от аргумента  $x$  удовлетворяет первому, второму и четвёртому условиям, накладываемым на функцию  $\eta_{BG}(x)$ , но при этом не выполняется третье условие. Чтобы оно выполнялось, определим функцию  $\eta_{BG}(x)$  как:

$$\eta_{BG}(x) = \frac{\varphi(x, I_\alpha) - \varphi(I_{\max}, I_\alpha)}{1 - \varphi(I_{\max}, I_\alpha)}.$$

Рассуждая аналогичным образом, введём функцию  $\psi(x, y)$ , численно равную значению функционала Тверского при  $f(X) = L(X)$ ,  $a = 0$ ,  $b = 1/2$  от множеств  $[-\gamma/2, x]$  и  $[-\gamma/2, y]$   $\forall x, y, \gamma \geq 0$ :

$$\psi(x, y) = T_{0, 1/2}([-\gamma/2, x], [-\gamma/2, y]) = \frac{2 \min\{x, y\} + \gamma}{2 \min\{x, y\} + \gamma + \max\{y - x, 0\}}.$$

Функция  $\psi(x, I_\beta)$  удовлетворяет первому, второму и четвёртому условиям, накладываемым на функцию  $\eta_{FG}(x)$ , но не удовлетворяет третьему условию. Пронормируем значения функции  $\psi(x, I_\beta)$  на сегменте  $[0, I_{\max}]$  и определим функцию  $\eta_{FG}(x)$  как:

$$\eta_{FG}(x) = \frac{\psi(x, I_\beta) - \psi(0, I_\beta)}{1 - \psi(0, I_\beta)}.$$

Продолжая аналогию, убеждаемся в том, что приведённое выше выражение удовлетворяет условиям, накладываемым на функцию  $\eta_{FG}(x)$ .

Параметр  $\gamma > 0$  имеет калибровочный смысл и подбирается таким образом, чтобы выполнялись следующие условия:

1. выражение для функции  $\varphi(x, y)$  корректно при  $x = 0$ ,  $y = 0$ ,
2.  $\varphi(0, 0) = 1$ ,
3. для двух фиксированных значений  $x_0 \in (0, I_{\max}]$  и  $\varphi_0 \in (0, 1)$   $\varphi(x_0, 0) = \varphi_0$ .

Согласно третьему из выше приведённых условий, величины  $x_0$  и  $\varphi_0$  однозначно определяют значение калибровочного параметра  $\gamma$  модели переднего и заднего планов как

$$\gamma = \frac{x_0 \varphi_0}{1 - \varphi_0}.$$

Несложно показать, что выбранное таким образом значение калибровочного параметра  $\gamma$  удовлетворяет также первому и второму условиям. В проведённых экспериментах для расчёта параметра  $\gamma$  использовались значения  $x_0 = 1$  и  $\varphi_0 = 2/3$  при  $I_{\max} = 255$ , что соответствует  $\gamma = 2$ .

## Вероятность принадлежности блока множеству блоков переднего плана

Очевидный теперь способ принятия решения относительно принадлежности блока  $B$  множеству блоков переднего плана  $U_{FG}$  состоит в определении статистики  $I = I(B)$  и в проверке истинности неравенства:

$$\eta_{FG}(I) \geq \eta_{BG}(I). \quad (3)$$

Отметим, что в общем случае  $\eta_{BG}(I) + \eta_{FG}(I) \neq 1$ , что не позволяет трактовать степени принадлежности множествам  $U_{BG}$  и  $U_{FG}$  как соответствующие вероятности. Между тем, вероятностная парадигма предоставляет большое количество инструментов для оценки качества результатов и сложности исходных данных и активно используется в области компьютерного зрения для оценивания степени уверенности в результатах работы алгоритмов.

Заметим, что согласно условиям, накладываемым на функции  $\eta_{BG}(x)$  и  $\eta_{FG}(x)$ ,  $0 \leq \eta_{BG}(I) + \eta_{FG}(I) \leq 2 \quad \forall I \in [0, I_{\max}]$ . Введём характеристическую величину

$$\Delta(B) = 1 - \eta_{BG}(I(B)) - \eta_{FG}(I(B)).$$

Относительно этой величины примем следующие предположения. Если  $\Delta(B) > 0$ , то она указывает на количество недостающей информации в рамках модели переднего и заднего планов для полного описания блока  $B$  и соответствующей ему статистике  $I$ . Если  $\Delta(B) < 0$ , то она указывает на то, что информация о блоке  $B$  избыточна и противоречива. В случае же  $\Delta(B) = 0$  описание блока  $B$  полно и непротиворечиво.

Для перехода к вероятностной парадигме воспользуемся таким инструментом обобщённой теории информации как неточные вероятности [4]. Поскольку нас прежде всего интересует вероятность принадлежности блока  $B$  множеству блоков переднего плана  $U_{FG}$ , введём соответствующую нижнюю вероятность как

$$\underline{p}_{FG}(B) = \underline{p}(B \in U_{FG}) = \min\{\eta_{FG}(I(B)), 1 - \eta_{BG}(I(B))\},$$

а соответствующую верхнюю вероятность как

$$\overline{p}_{FG}(B) = \overline{p}(B \in U_{FG}) = \max\{\eta_{FG}(I(B)), 1 - \eta_{BG}(I(B))\}.$$

Такой переход не противоречит определению неточных вероятностей, поскольку сегмент

$\left[ \underline{p}(B \in U_{FG}), \overline{p}(B \in U_{FG}) \right]$  является выпуклым множеством [4]. Отметим также, что

$$\overline{p}(B \in U_{FG}) - \underline{p}(B \in U_{FG}) = |\Delta(B)|.$$

Пусть  $D$  - множество всех распределений вероятности  $p$ , удовлетворяющих введённым значениям нижней и верхней вероятностей принадлежности блока  $B$  множеству блоков переднего плана  $U_{FG}$ . Вероятностью принадлежности блока  $B$  множеству блоков переднего плана  $U_{FG}$  будем считать ожидаемое значение вероятности на множестве  $D$ :

$$p_{FG}(B) = p(B \in U_{FG}) = \int_{\underline{p}_{FG}(B)}^{\overline{p}_{FG}(B)} p \frac{dp}{\overline{p}_{FG}(B) - \underline{p}_{FG}(B)}.$$

Данное выражение путём элементарных преобразований приведём к виду

$$p_{FG}(B) = \frac{\eta_{FG}(I(B)) + 1 - \eta_{BG}(I(B))}{2}. \quad (4)$$



Обратим внимание, что при  $I(B) = 0$   $p_{FG}(B) = 0$ , а при  $I(B) = I_{\max}$   $p_{FG}(B) = 1$ . Заметим, что неравенство  $p_{FG}(B) \geq 1/2$  равносильно неравенству (3). Вместе с тем, выражение (4) позволяет заполнять множество блоков переднего плана в соответствии с более общим критерием

$$p_{FG}(B) \geq p_{th}, \quad (5)$$

где  $p_{th} \in (0;1]$ .

Отметим также, что критерии (3) и (5) соответствуют выдвинутому выше предположению о том, что достоверность принадлежности блока  $B$  множеству блоков переднего плана  $U_{FG}$  и статистика  $I$ , определённая в (2), связаны друг с другом монотонно неубывающей функцией. Данное соответствие легко показать, если вспомнить, что функция  $\eta_{BG}(x)$  монотонно не возрастает на  $[0, I_{\max}]$ , а функция  $\eta_{FG}(x)$  монотонно не убывает на том же промежутке. Отсюда следует, что функция  $(\eta_{FG} - \eta_{BG})(x)$  монотонно не убывает на  $[0, I_{\max}]$ , поскольку является разностью монотонно неубывающей и монотонно невозрастающей функций. Так как неравенство (3) можно переписать в виде  $(\eta_{FG} - \eta_{BG})(x) \geq 0$ , а неравенство (5) в виде  $(\eta_{FG} - \eta_{BG})(x) \geq 2p_{th} - 1$ , то необходимое соответствие можно считать доказанным.

### Оценка надёжности результатов параметризованной модели переднего и заднего планов, оптимальные параметры модели, оценка качества маски движения

Введение аппарата неточных вероятностей позволяет оценить в битах общий уровень недоверия к результатам классификации блока  $B$  согласно следующей формуле [4]:

$$Mef(B) = S^*(D) = \sup_{p \in D} S(p),$$

где  $S(p)$  - функционал, определяющий численное значение энтропии Шеннона для заданного распределения  $p$  из множества допустимых распределений  $D$ . Применительно к задаче классификации блока  $B$  с заданными классами  $U_{BG}$  и  $U_{FG}$  данное выражение имеет вид:

$$Mef(B) = \begin{cases} -p_{FG} \log_2 p_{FG} - (1 - p_{FG}) \log_2 (1 - p_{FG}), & p_{FG} \leq p_{FG} < 0.5 \\ 1, & p_{FG} \leq 0.5 \leq p_{FG} \\ -p_{FG} \log_2 p_{FG} - (1 - p_{FG}) \log_2 (1 - p_{FG}), & 0.5 < p_{FG} \leq p_{FG} \end{cases}.$$

Зависимость величины  $Mef$  от статистики текущего блока можно проследить на рис. 2. Как и следовало ожидать, распределение данной величины имеет ярко выраженный максимум в той части диапазона  $[0, I_{\max}]$ , которая согласно используемой параметризованной модели блоков переднего и заднего плана соответствует блокам с переходными значениями статистики. Именно для таких блоков вероятность ошибки классификации наиболее велика. Обратим также внимание на то, что в крайних точках диапазона  $[0, I_{\max}]$  неуверенность в результатах классификации равна 0. Чтобы оценить уровень неопределённости, возникающий при обработке маски  $M^{(k)}$ , необходимо сложить неопределённости всех блоков, на которые была разбита  $M^{(k)}$ :

$$Mef(M^{(k)}) = \sum_{B \in M^{(k)}} Mef(B).$$

Эта величина является адекватным отражением надёжности полученных при выделении блоков переднего плана результатом. Максимально возможная неопределённость при анализе маски движения равна количеству блоков  $Mef_{\max} = n$  в битах.

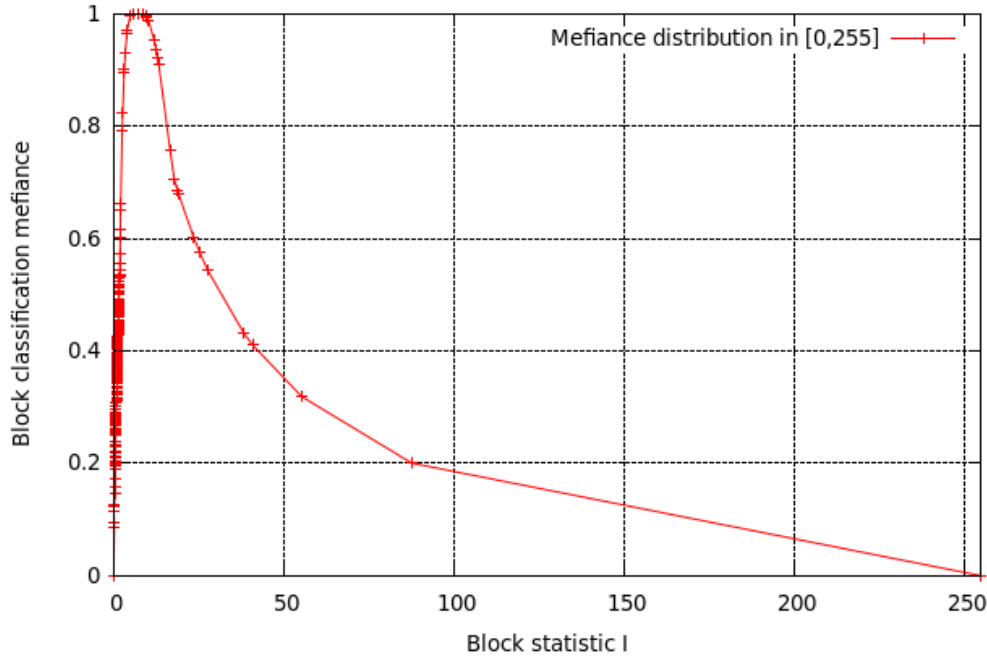


Рис. 2 Зависимость уровня недоверия к результатам классификации от статистики блока

Характеристической величиной маски движения  $M^{(k)}$  будем считать

$$\Delta(M^{(k)}) = \frac{\sum_{B \in M^{(k)}} \Delta(B)Mef(B)}{Mef(M^{(k)})} = \frac{\sum_{B \in M^{(k)}} \Delta(B)Mef(B)}{\sum_{B \in M^{(k)}} Mef(B)}.$$

Характеристическая величина маски движения является оценкой неполноты и противоречивости используемой параметризованной модели блоков переднего и заднего плана в масштабах одного кадра и может помочь при оптимизации параметров модели. Этот вопрос может стать предметом дальнейших исследований.

### Конструирование блобов

Постановка задачи, подробно описанная в [8], подразумевает построение и анализ маски движения, а также выделение на ней блобов, которые соответствуют движущимся объектам или группам объектов. В [8] и [9] блобы строились на основе отношений 4- или 8-связности между движущимися блоками и представлялись в виде описанных вокруг найденных связных компонент прямоугольников. В данной статье блобы представляются в виде выпуклых оболочек произвольной формы, описанных вокруг связных подмножеств движущихся блоков. Множество движущихся блоков выбиралось согласно критерию (5) при различных значениях параметра  $p_{th}$  в диапазоне [0.35, 0.8].

### Результаты работы алгоритма

На рис. 3-5 представлен пример выделения движущихся объектов на одном из кадров видеопотока по маске движения. Небинаризованная маска движения на рис. 3 получена по формуле (1).



Рис. 3 Небинаризованная маска движения

На рис. 4 представлены результаты анализа небинаризованной маски движения. Она разбита на блоки. В центре каждого блока синим цветом написано целое число, которое приближённо равно выраженной в процентах вероятности того, что соответствующий блок принадлежит множеству движущихся блоков (см. формулу (4)). Границы блоков показаны цветом, который плавно меняется от чёрного к красному в зависимости от соответствующей вероятности принадлежности блока множеству движущихся блоков. Согласно критерию (5) происходит отделение движущихся блоков от покоящихся. Из движущихся блоков конструируются блобы, выпуклые оболочки вокруг которых обозначены зелёным цветом.

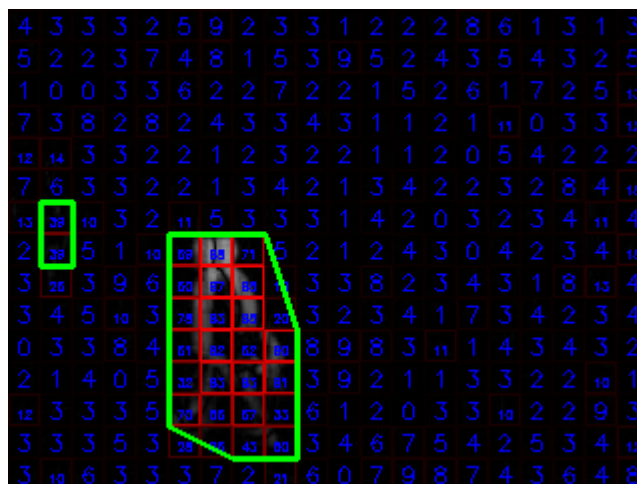


Рис. 4 Результаты анализа небинаризованной маски движения

На рис. 5 выпуклые оболочки блоков обозначены на исходном кадре видеопоследовательности, которому соответствует маска движения, представленная на рис. 3.



**Рис. 5** Выделение движущихся объектов на кадре видеоряда

Ещё несколько примеров выделения областей движения представлено в приложении к данной статье.

### **Заключение**

В данной статье предложен алгоритм анализа небинаризованной маски движения для выделения движущихся объектов и групп объектов в видеопотоке. Главное преимущество небинаризованной маски движения состоит в том, что она может содержать в себе больше информации, нежели её бинарный аналог. Другой особенностью алгоритма является разбиение маски движения на блоки прямоугольной формы и введение статистики, усредняющей интенсивность движения внутри блока. Далее на основе анализа функции распределения статистики блоков строится параметризованная модель движущихся и неподвижных блоков, которая позволяет оценить достоверность наличия движения в каждом из блоков. Также в данной статье предлагаются критерии разграничения движущихся и неподвижных блоков и вводятся величины, помогающие оценить надёжность полученных результатов.

### **Список литературы**

- [1] Bradski G., Kaehler A. Learning OpenCV. Computer vision with the OpenCV library. – Sebastopol: O'Reilly Media, Inc., 2008. – 555 p.
- [2] Cross V., Sudkamp T.A. Similarity and compatibility in fuzzy set theory: assessment and applications. Springer, 2002, 209 p..
- [3] P. KaewTraKuPong and R. Bowden, An improved adaptive background mixture model for real-time tracking with shadow detection, Proc. 2nd European Workshop on Advanced Video-Based Surveillance Systems, 2001
- [4] Klir, G. J. Uncertainty and information. Foundations of generalized information theory. – New York: John Wiley & Sons, Inc., 2006. – 499 p.
- [5] Piccardi M. Background subtraction techniques: a review. The ARC Centre of Excellence for Autonomous Systems (CAS) Faculty of Engineering, UTS, April 15, 2004
- [6] Richardson I.E.G. H.264 and MPEG-4 video compression: video coding for next generation multimedia, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, 2003

[7] Z.Zivkovic, Improved adaptive Gaussian mixture model for background subtraction, International Conference Pattern Recognition, UK, August, 2004

[8] Мацыпаев Д.А., Броневи́ч А.Г. Анализ маски движения в задаче обнаружения и сопровождения движущихся объектов на основе энтропии Шеннона. Труды ОАО "НИИАС". Сборник научных трудов. Выпуск 9. М.: Изд-во ООО "Группа ИДТ", 2011, с. 255-273.

[9] Мацыпаев Д.А. Алгоритм рекурсивного разбиения маски движения для выделения движущихся объектов в видеопотоке // Доклады 12-й Международной конференции «Цифровая обработка сигналов и её применение» ДСПА-2010. -2010. –Т.2. – Стр. 114-116.

#### Приложение.



Рис. 6 Примеры работы алгоритма