

ИНТЕЛЛЕКТУАЛЬНЫЕ ВИДЕОСИСТЕМЫ СЛЕЖЕНИЯ И ОХРАНЫ ОСОБО ВАЖНЫХ ОБЪЕКТОВ

Л. ПЮШКИ, М.И. ВАСЮХИН, И.М. ВАСЮХИНА

Приведен обзор тенденций, методов и средств построения интеллектуальных видеосистем слежения и охраны особо важных объектов. Рассмотрены подходы к созданию и использованию образного компьютера, методы распознавания алфавитно-цифровой информации, а также методы слежения за объектами на основе интеллектуальной системы видеонаблюдения, разработанной в Институте кибернетики им. В.М. Глушкова НАН Украины.

ТЕНДЕНЦИИ ПОСТРОЕНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ ВИДЕОСИСТЕМ СЛЕЖЕНИЯ И ОХРАНЫ ОСОБО ВАЖНЫХ ОБЪЕКТОВ

При построении интеллектуальных видеосистем слежения и охраны особо важных объектов весьма актуальна проблема создания средств автоматизации интеллектуальных функций человека, связанных с восприятием, обработкой и распознаванием зрительных образов [1–12, 15].

Одним из наиболее мощных направлений создания таких средств можно назвать программу «Образный компьютер» [1]. Образный компьютер (ОК) — это кибернетическая система, где выполняется моделирование интеллектуальной (главное, подсознательной) деятельности человека и всего живого, что связано с восприятием зрительных, слуховых и других образов, анализом сложных динамических сцен, планированием действий и движений, обобщением наблюдений, установлением закономерностей, прогнозированием, принятием решений. ОК оперирует образами и другими сложными понятиями, реализует как образное, так и логическое мышление.

ОК является параллельной мультимодальной системой, которая имеет в своем составе несколько каналов восприятия информации (слуховой, зрительной, текстовой, вкусовой, обонятельной), образную операционную систему, модели внешнего мира (в том числе акустическую, оптическую, геометрическую, лингвистическую, семантическую, канонических форм), развитый интерфейс с человеком, способы взаимодействия с существующими компьютерными и телекоммуникативными сетями.

Образная операционная система «синхронизирует» обработку информации, которая поступает разными каналами восприятия, и, оперируя моделями внешнего мира, выполняет комплексную семантическую интерпретацию всей полученной информации.

Однако особые возможности ОК предполагают и значительные финансовые затраты, часто недопустимые при разработке рассматриваемых авторами периметровых систем [2, 8, 15, 16].

В настоящей работе приведены результаты исследований методов построения периметровых автоматизированных систем слежения и охраны особо важных объектов (АССО ОВО) в реальном времени.

Для периметровых систем охраны в работе [2] предложена структура интегрированной автоматизированной системы, объединяющей достоинства трех подсистем: на основе Leaky feeder, пассивных сигнализаторов и интеллектуальной видеокамеры. Последняя необходима для установки в районе ворот для контроля за въезжающим и выезжающим транспортом.

Характерная задача для периметровых систем — распознавание номерного знака автотранспортных средств. Проведенный анализ условий функционирования таких систем позволяет выделить следующие технические требования:

- инвариантность к изменению в некоторых пределах угла поворота между осью камеры и вектором направления движения автомобиля как в вертикальной, так и в горизонтальной плоскостях;
- инвариантность к изменению масштаба распознаваемого номера, т.е. расстояние между камерой и автомобилем в момент захвата изображения не должно быть строго фиксированным;
- инвариантность к искажению номерного знака, так как при его установке на автомобиле не в строго вертикальной плоскости и наблюдении сбоку знак представляет собой не прямоугольник, а параллелограмм (в этом случае размер символов на изображении несколько искажается);
- сохранение работоспособности в условиях сильных помех (дождь, снег, грязь на номерном знаке).

АНАЛИЗ МЕТОДОВ РАСПОЗНАВАНИЯ ЗРИТЕЛЬНЫХ ОБРАЗОВ В АССО ОВО

Разработка методов распознавания изображений символов посвящено достаточно много работ [2 – 15]. Одной из наиболее удачных, по нашему мнению, можно отметить работу [3], в которой описана оригинальная математическая модель, позволяющая достичь достоверности распознавания одного символа, равной 96%. Структура системы распознавания показана на рис. 1.

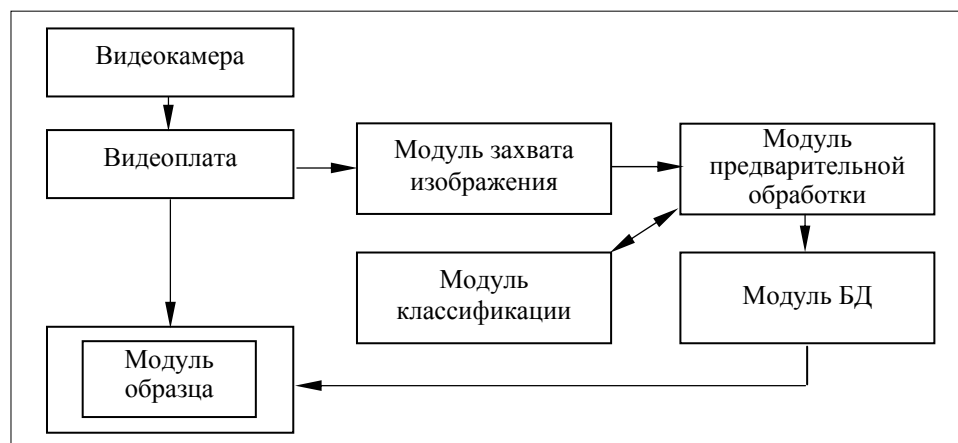


Рис. 1. Структура системы распознавания

Программное обеспечение системы состоит из модулей захвата изображения, предварительной обработки, классификации, а также модуля базы данных.

Модуль захвата изображения представляет собой программу, позволяющую производить покадровую передачу изображения в модуль предварительной обработки области номерного знака (ОНЗ).

Модуль предварительной обработки осуществляет подготовку данных, необходимых для распознавания. В него входят следующие процедуры.

1. *Поиск местоположения ОНЗ* осуществляется при помощи двумерного волнового преобразования Хаара, причем исходное изображение разбивается на клетки 2×2 , для которых выполняется

$$\begin{bmatrix} a_0 & a_1 \\ a_2 & a_3 \end{bmatrix} \rightarrow \begin{bmatrix} \frac{a_0 + a_1 + a_2 + a_3}{2} & \frac{a_0 + a_1 - a_2 - a_3}{2} \\ \frac{a_0 - a_1 + a_2 - a_3}{2} & \frac{a_0 - a_1 - a_2 + a_3}{2} \end{bmatrix} = \begin{bmatrix} b_0 & b_1 \\ b_2 & b_3 \end{bmatrix}.$$

Элементы матрицы b_0, b_1, b_2 и b_3 имеют следующие значения: b_0 — среднее арифметическое в окне 2×2 ; b_1 — горизонтальные контурные перепады яркости; b_2 — вертикальные, а b_3 — диагональные. Размер каждой матрицы в два раза меньше размера исходного изображения.

2. *Бинаризация*. После волнового преобразования получается изображение с подчеркнутыми контурными перепадами, ориентированными вертикально. В ОНЗ количество таких перепадов велико по сравнению со всем изображением, что приводит к появлению области прямоугольной формы, в которой наблюдаются резкие флюктуации интенсивности пикселей. В данном случае можно сформулировать постановку задачи: выделить области, где сигнал, обычно не существенно отличающийся от общего фона, начинает резко колебаться.

Для решения поставленной задачи производится оценка дисперсии $D(x, y)$ сигнала в некоторой области относительно текущего пикселя изображения $I(x, y)$ и сравнивается с некоторым порогом $P_{кр}$. Бинаризация осуществляется согласно

$$I_{бин}(x, y) = \begin{cases} 1, & D(x, y) \geq P_{кр}, \\ 0, & D(x, y) < P_{кр}, \end{cases}$$

где $I_{бин}(x, y)$ — интенсивность пикселя в бинаризованном изображении.

3. *Фильтрация*. В данном случае задачей фильтрации является удаление как можно большего числа объектов на бинаризованном изображении, не являющихся областью номерного знака.

Процедура фильтрации состоит из двух частей:

а) медианная фильтрация в окне 3×3 пикселя

$$B = \begin{cases} 1, & \sum_{i=1}^{N_S} I_i > \frac{N_S}{2}, \\ 0, & \sum_{i=1}^{N_H} I_i < \frac{N_S}{2}, \end{cases}$$

где B — значение бинаризованного изображения в обрабатываемом пикселе после медианной обработки; N_s — количество точек в окрестности обрабатываемого пикселя в направлении сканирования;

б) прослеживание слитных областей, определение их параметров и отсеивание тех областей, которые однозначно не являются ОНЗ.

Для каждой из таких групп областей определяются высота, ширина, площадь и отношение высоты к ширине. Объекты, не проходящие хотя бы по одному критерию, удаляются из бинаризованного изображения.

4. *Выделение номерного знака.* Для выделения ОНЗ используется метод корреляционно-экстремального анализа. Суть метода заключается в плоскопараллельном смещении маски по обрабатываемому изображению и подсчете совпадающих пикселей. Положение маски и ОНЗ, для которых количество таких пикселей максимально, принимается соответствующим реальному положению ОНЗ. Корреляционная функция определяется следующим образом:

$$K(x, y) = \sum_{i=1}^{L_x} \sum_{j=1}^{L_y} B(x+i, y+j) M_{ij},$$

где (x, y) — координаты обрабатываемой точки изображения; L_x — размер маски по горизонтали; L_y — размер маски по вертикали; B — бинаризованное изображение; i, j — координаты пикселей внутри маски; M — маска.

5. *Определение типа изображения.* Считаем позитивом светлые символы на темном фоне. Данная процедура состоит из следующих шагов:

а) вычисление медианы M гистограммы ОНЗ;

б) пороговая бинаризация ОНЗ с использованием медианы M в качестве порога;

в) если число черных пикселей больше числа белых, изображение считается позитивным. Если наоборот, то инвертируется изображение ОНЗ.

6. **Полосовая фильтрация.** Формируем изображение $I'_{ОНЗ}(x, y)$, являющееся результатом обработки изображения ОНЗ $I_{ОНЗ}(x, y)$ двумерным линейным фильтром с окном 31×31 пиксель.

$$I'_{МН}(x, y) = \sum_n \sum_m I_{МН}(n, m) / 31^2,$$

где $n = x - 15, \dots, x + 15$; $m = y - 15, \dots, y + 15$.

Результирующее изображение определяется как

$$I_{ОНЗиф}(x, y) = (I_{ОНЗ}(x, y) - I'_{ОНЗ}(x, y)) * Scale,$$

где $Scale$ — масштабный коэффициент, позволяющий «растянуть» гистограмму изображения ОНЗ на весь диапазон яркости от 0 до 255.

7. *Пороговая бинаризация.* На данном этапе обработки получается хорошо контрастированное изображение ОНЗ. Такое изображение вполне можно бинаризовать обычным пороговым методом. Однако исследования показали, что в случае существенно зашумленного изображения (грязь на номере, дождь, снег) данный метод неэффективен.

Для решения задачи использовался метод пороговой бинаризации.

После бинаризации изображение ОНЗ также подвергается «чистке» путем прослеживания слитных групп, о которой говорилось выше. Предварительный анализ на 50 изображениях автомобилей показал, что после бинаризации изображения отдельных символов не сливаются между собой и с областями, не являющимися символами. Таким образом, в данном случае «чистка» путем прослеживания слитных групп допустима, хотя для полной уверенности требуются более детальные исследования.

8. *Сегментация* является ключевой во всем модуле предварительной обработки. Предшествующая обработка решала задачу получения как можно более качественного изображения символов номерного знака и подавления при этом как можно больше шумов. Поскольку эта задача решена успешно и для сегментации изображения ОНЗ, то вполне успешно можно применить все тот же метод прослеживания слитных групп.

Единственным замеченным недостатком сегментации является невозможность удаления крайней правой/левой области, соответствующей границе номерного знака. При этом после классификации в начале и конце распознанного номера может появиться символ «/». Для решения данной проблемы требуется разработать более интеллектуальные алгоритмы сегментации.

9. *Трансформация сегментов*. Производится трансформация выделенных сегментов к единому масштабу, что обеспечивает инвариантность системы к масштабу символов.

Модуль классификации позволяет непосредственно производить распознавание сегментированного изображения номерного знака на базе нейронной сети.

НЕЙРОСЕТЕВОЙ МЕТОД РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Отметим, что для реализации нейронной сети (НС) достаточно эффективно использовался метод стохастической геометрии [3]. Он предусматривает случайное наложение на изображение символа некоторого количества так называемых «иголочек» некоторой длины. Затем подсчитывается число нулевых, одинарных, двойных и тройных пересечений этих «иголочек» с символом. В построенном пространстве признаков печатный текст хорошо классифицируется без искажений и шумов. Однако при появлении шумов на его изображении классификация затрудняется либо становится вообще невозможной.

Более эффективен ассоциативный метод классификации изображений символов [3] — по неполному их изображению. Для этого нейронную сеть необходимо обучить восстанавливать полную картину, т.е. по части выбранного образа восстановить весь образ. При этом использовалась нейронная сеть Хопфилда.

Ее применение показало три недостатка.

1. Эта сеть оказалась чувствительной к искажениям и смещениям изображений.
2. На «вспоминание» символа, существенно отличающегося от эталонного, требуется много времени.

3. Появление «фантома» при определении указанного символа, т.е. появление такого его изображения, которое не является ни одним из эталонов.

Устранению этих недостатков способствовало применение моделей Кохонена [4], которые напоминают процессы, происходящие в мозгу человека, известные как «самоорганизующиеся карты Кохонена (СКК)».

Для реализации алгоритма необходимо определить меру близости (зону соседства). Зоны соседства уменьшаются с течением времени.

Сеть состоит из M нейронов, а $Nej(t)$ — множество нейронов, которые считаются соседями нейрона j в момент времени t .

Согласно алгоритму Кохонена необходимы такие шаги:

1. Инициализация сети. Весовым коэффициентам сети присваиваются малые случайные значения. Общее число синаптических весов — $M \times N$.

2. Предъявление сети нового входного сигнала.

3. Вычисление расстояния до всех нейронов сети.

Расстояние d_j от входного сигнала до каждого нейрона j определяется соотношением

$$d_j = \sum_{i=0}^{N-1} (x_i(t) - w_{ij}(t))^2,$$

где x_i — i -й элемент входного сигнала в момент времени t ; $w_{ij}(t)$ — вес связи от i -го элемента входного сигнала к нейрону j в момент времени t .

4. Выбор нейрона с наименьшим расстоянием. Выбираем нейрон j^* , для которого расстояние d_j наименьшее.

5. Настройка весов нейрона j^* и его соседей.

Производится подстройка весов для нейрона j^* и всех нейронов из его зоны соседства NE . Новые значения весов

$$w_{ij^*}(t+1) = w_{ij^*}(t) + r(t)(x_i(t) - w_{ij^*}(t)),$$

где $r(t)$ — шаг обучения, уменьшающийся с течением времени (положительное число, меньше единицы).

6. Возвращение к п.2.

Следует отметить, что в общем случае размерность СКК может быть произвольной.

В ходе лабораторных исследований было предложено использовать двумерное стробирование исходного изображения символа. Окно при стробировании представляет собой квадрат размером $L_{стр}$, а величина шага равна $L_{стр} - 1$. Каждый фрагмент изображения, получаемый при стробировании, подается на вход НС. Из получающихся на выходе НС топологических карт отдельных фрагментов исходного изображения формируется карта всего изображения.

Далее производится окончательная классификация, причем полученная топологическая карта сравнивается со всеми эталонами, хранящимися в памяти. В качестве решения выбирается наиболее близкий к карте эталон.

Критерием близости является сумма квадратов отклонений точек реальной карты от эталонной.

Применение метода СКК позволило за несколько десятков секунд достичь вероятности распознавания одного символа автомобильного номера, равной 96%, независимо от положения автомобиля в поле зрения камеры.

Метод обратного распространения ошибки. Существенным развитием описанного выше метода распознавания явился метод обратного распространения ошибки [5]. В отличие от структуры сети, показанной на рис.1, предлагаемая сеть имеет структуру, содержащую две подсистемы: предварительной обработки сегментации изображения и распознавания символов номерного знака (рис. 2).



Рис. 2. Структура нейросетевой системы

Сначала производится процедура улучшения качества изображения не только для улучшения визуального восприятия, но и для упрощения машинного анализа. Для этого применяют методы сглаживания и увеличения контрастности изображений.

Медианный метод сглаживания изображений. Согласно работе [6], наиболее эффективный при наличии импульсного шума — медианный метод. Каждая точка a_{ij} изображения заменяется медианой, вычисленной в заданной окрестности $n \times n$ следующим образом:

1. Все точки окрестности $n \times n$ ранжируются в порядке возрастания своих значений.

2. Медианой такой ранжированной последовательности точек является значение яркости точки, средней по счету в этой последовательности, т.е. значение яркости точки с порядком $\lfloor \frac{n^2}{2} \rfloor$, где $\lfloor \cdot \rfloor$ — операция округления до большего целого числа.

Метод увеличения контрастности осуществляется на основе алгоритма (рис. 3), описываемого следующей формулой:

$$p_{ij}^* = \begin{cases} 0, & p_{i,j} < B, \\ \frac{255}{W-B}(p_{i,j} - B), & p_{i,j} \in [B, W], \\ 255, & p_{i,j} > W, \end{cases}$$

где p_{ij} — исходная яркость точки изображения; p_{ij}^* — результирующая яркость точки изображения; B, W — пороговые значения.

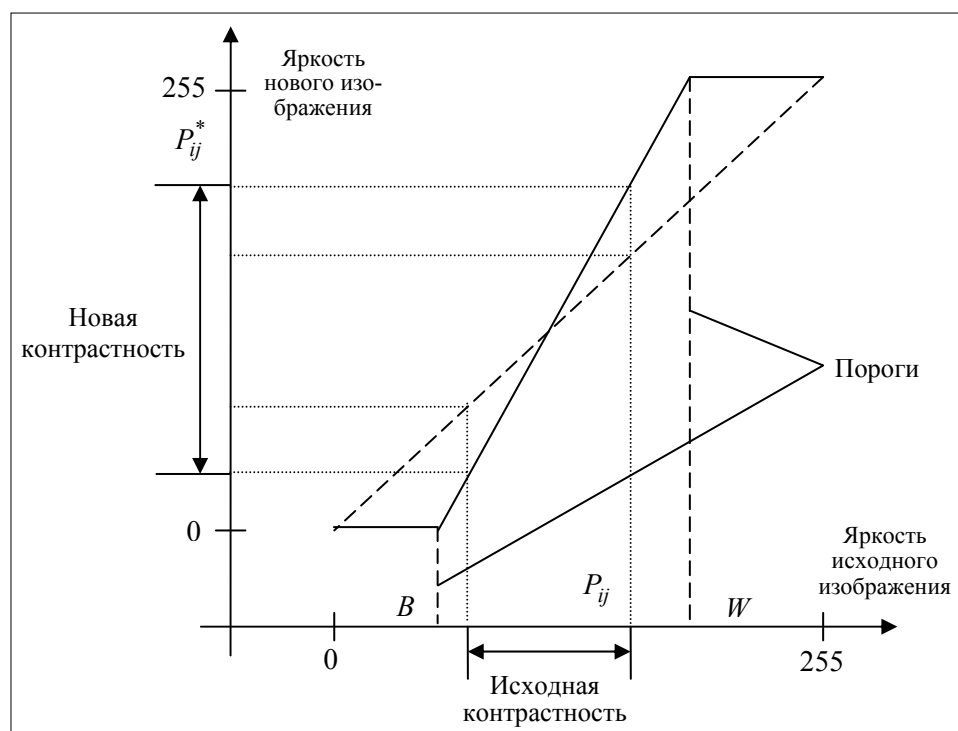


Рис. 3. Алгоритм увеличения контрастности изображения

В результате работы этого алгоритма линейно увеличивается разность яркостей точек, что приводит к тому, что темные области

изображения становятся еще более темными, а наиболее светлые — более светлыми, что обеспечивает решение задачи.

Выделение знака из общего фона изображения осуществляется его построчным сканированием в окне размером $N = 32$ и применением дискретного преобразования Фурье в тех его участках, где содержится символическая информация. Спектр коэффициентов преобразования Фурье имеет пик на низких частотах, для других участков изображения — пик либо на высоких частотах, либо на нулевой частоте (наблюдается хаотический аperiodический сигнал).

Коэффициенты дискретного преобразования Фурье рассчитываются по формуле

$$C_x(k) = \frac{1}{N} \sum_{m=0}^{N-1} X(m)W^{km}, \quad k = 0, 1, \dots, N-1,$$

где $W = e^{-i2\pi/N}$.

Для проведения сегментации предлагается алгоритм, в соответствии с которым изображение сканируется по горизонтали, а затем вычисляется функция как сумма квадратов перепадов яркости точек в каждом вертикальном столбце изображения.

$$f(x) = \sum_i (R_{i,x} - R_{i+1,x})^2.$$

Полученная функция имеет периодический характер, причем вершины соответствуют участкам, где расположены символы номерного знака, а минимумы — промежуткам между ними.

Для распознавания знака предлагается сеть типа многослойного персептрона, обучаемого методом обратного распространения ошибки. Сеть была обучена выборке, состоящей из печатных символов, используемых при написании автомобильных знаков. Особенность метода: в обучающую выборку добавляются искаженные символы, что в целом позволяет повысить надежность и быстродействие системы.

Рассматриваемая нейросетевая система представлена в виде программного модуля на языке Visual C++ 6.0 для ОС WINDOWS. Результаты испытаний показали надежную идентификацию номерных знаков при их смещениях и поворотах в поле зрения, а также при наличии загроможденности и атмосферных помех.

МЕТОДЫ СЛЕЖЕНИЯ ЗА ОБЪЕКТАМИ НА ОСНОВЕ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ ВИДЕОНАБЛЮДЕНИЯ

Первым шагом в построении системы слежения на основе видеокамеры является построение разностного изображения (рис. 4). Для этого строятся матрицы приращений между двумя соседними кадрами. Далее разностное изображение (рис. 4, в) делится на квадраты размером $n \times n$, где $n = 6 \dots 16$. Для каждого из этих квадратов рассчитывается среднее значение яркости по формуле

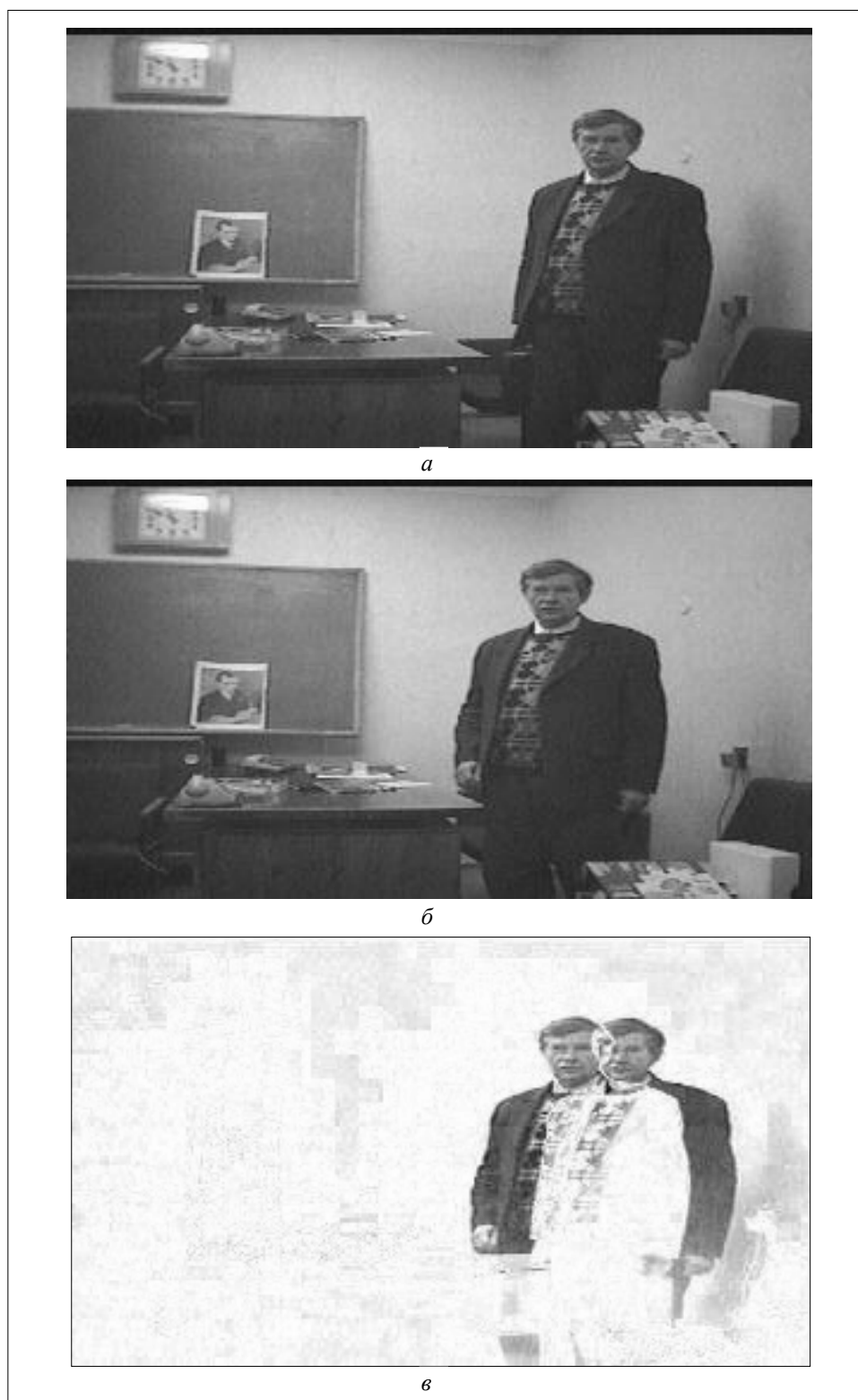


Рис. 4. Соседние кадры видеопоследовательности (*a*, *б*); разностное изображение инвертировано (*в*)

$$b_{\text{cp}} = \frac{\sum_{x=0}^n \sum_{y=0}^n b(x, y)}{n^2},$$

где $b(x, y)$ — значение яркости отдельной точки на разностном изображении.

Затем полученные средние значения яркости сравниваются с некоторым порогом P области разностного изображения. Если подсчитанное среднее значение b_{cp} меньше порога P , то изображения считаются не содержащими межкадровых изменений и в дальнейшем не рассматриваются. Напротив, области, в которых среднее значение b_{cp} больше порога P , указывают на движение или локальное изменение яркости. В результате получаем матрицу M (рис. 5) размерностью в n раз меньше исходного изображения

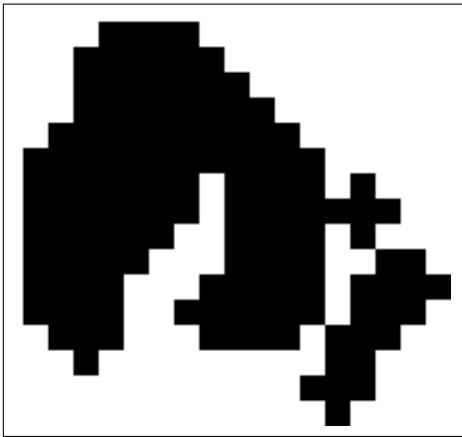


Рис. 5. Пример преобразования разностного изображения с фильтрацией шумовых составляющих (матрица M)

и значениями ячеек, равными 0 при отсутствии межкадровых изменений и 1 при их наличии. По этой матрице легко определить габариты и расположение движущегося объекта — на него указывают ненулевые области. Здесь следует отметить, что на разностном изображении области с максимальными значениями будут приходиться как на исходное положение движущегося объекта в первом кадре, так и на конечное положение во втором. Однако при достаточно высокой частоте кадров перемещение объекта в двух соседних кадрах незначительно, и эти две области будут практически совпадать, указывая на положение движущегося объекта. Разумеется, регистрации подлежат только области, содержащие изменения (рис. 6). При сжатии информации достаточно хранить базовый (опорный) кадр и изменившиеся области для каждого последующего. В условиях редко меняющейся обстановки изменения в кадре происходят редко, и с помощью предлагаемого алгоритма можно достичь степени сжатия в десятки тысяч раз.

Оптимизация алгоритма сжатия может производиться за счет уменьшения количества точек, в которых производится расчет разностного изображения. В таком случае изображение покрывается сеткой с размером ячейки s , и расчет разностного изображения производится только в узлах этой сетки. В результате размеры разностного изображения уменьшаются в s раз, количество обрабатываемых пикселей, а, соответственно, и количество операций по обработке — в s^2 раз. Недостатком такой оптимизации является уменьшение точности определения областей со значительными изменениями, а также возможность пропуска мелких изменений.



Рис. 6. Области кадра 2 (рис. 5), подлежащие сохранению

Для построения траектории движения объекта (объектов) необходимо определить их геометрический центр в каждый момент времени. Координаты $x_{ц}$ и $y_{ц}$ геометрического центра объекта можно определить как центр прямоугольника, ограничивающего объект. Если координаты левого нижнего угла прямоугольника — x_0, y_0 , ширина — b , высота — h , то координаты геометрического центра определяются как $x_0 + b/2 + h/2$.

Возможен также другой вариант определения координат объекта как «центра тяжести» прямоугольника, ограничивающего объект, подсчитанного по разностному изображению. Расчет производится по следующим формулам

$$X_{цт} = \frac{\sum_{x=x_0}^{x_0+b} \sum_{y=y_0}^{y_0+h} (P(x,y)x)}{\sum_{x=x_0}^{x_0+b} \sum_{y=y_0}^{y_0+h} P(x,y)},$$

где $b(x,y)$ — значение яркости разностного изображения в точке с координатами x, y .

Координаты, определенные таким образом, будут тяготеть к части объекта, содержащей максимум изменений от кадра к кадру. Геометрическое место точек таких центров, рассчитанных для кадров видеопоследовательности, и будет определять траекторию движения объекта.

$$Y_{цт} = \frac{\sum_{x=x_0}^{x_0+b} \sum_{y=y_0}^{y_0+h} (P(x,y)y)}{\sum_{x=x_0}^{x_0+b} \sum_{y=y_0}^{y_0+h} P(x,y)}.$$

В случае, если в поле зрения видеокамеры находится два или более движущихся объекта, следует проверить области, содержащие изменения, на принадлежность тому или иному объекту. Для этого можно воспользоваться матрицей M (см. рис. 5). Для разделения нескольких объектов применим алгоритм «наращивания областей». Берем произвольный ненулевой элемент и рекурсивно ищем смежные с ним. Так находим все элементы, принадлежащие одному объекту. Далее берем следующий элемент, не принадлежащий первому объекту, и повторяем операцию для второго объекта и далее для всех объектов.

Траектория движения может дать информацию о том, какие именно события происходят в поле зрения видеокамеры: если траектория движения представляет собой точку или близка к ней, то имеет место локальное изменение яркости или геометрического размера объекта, если траектория — кривая линия, то происходит движение объекта.

В случае движения двух или более объектов в точке пересечения их траекторий (точка *A* на рис. 7) может возникнуть неоднозначность принадлежности последующей после пересечения части траектории тому или иному объекту. Критериями для разрешения неоднозначности могут служить:

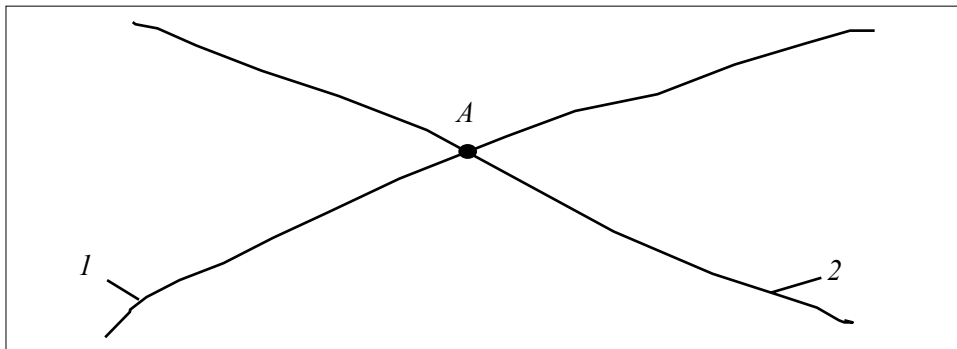


Рис. 7. Траектории движения объектов (1, 2)

а) направление движения объекта: продолжение траектории выбирается без излома (критерием наличия/отсутствия излома могут служить значения производных координат объекта по времени);

б) скорость и ускорение объектов: продолжением считается траектория, в которой объект имеет более близкую скорость (ускорение) по отношению к скорости (ускорению) на предыдущем участке;

в) габаритные размеры движущихся объектов.

Эти критерии, рассмотренные совместно, дают наиболее вероятные траектории движения объектов.

Оптимизация алгоритма слежения может производиться за счет сужения зоны поиска последующего положения объекта, а также уменьшения количества обрабатываемых кадров.

Сужение зоны поиска последующего положения объекта может производиться в двух случаях.

1. Известны координаты движущегося объекта, неизвестны параметры движения (скорость, ускорение). В таком случае зону поиска следует рассчитывать, исходя из максимальной предположительной горизонтальной $V_{x \max}$ и вертикальной $V_{y \max}$ скорости движения объекта в метрах в секунду и минимально возможного расстояния объекта до камеры R_{\min} в метрах. Максимальные перемещения Δx_{\max} и Δy_{\max} между соседними кадрами в пикселях можно выразить в долях от ширины (или высоты) кадра по формулам

$$\Delta x_{\max} = b \frac{V_{x \max}}{R_{\min} \gamma \nu}, \quad \Delta y_{\max} = h \frac{V_{y \max}}{R_{\min} \beta \nu},$$

где b и h — ширина и высота кадра; γ и β — угол захвата камеры по горизонтали и вертикали; ν — частота кадров съемки.

2. Известны координаты движущегося объекта, скорость и направление его движения. На основании этих параметров можно предсказать последующее положение объекта. Двигаясь с той же скоростью, между соседними кадрами объект переместится на расстояние $x = \frac{V}{\nu}$, где V — скорость движения объекта, подсчитанная по двум предыдущим кадрам.

Однако ввиду возможного изменения скорости объекта его положение в следующем кадре может отличаться от предсказанного и, тогда зону поиска следует расширить на величину $\Delta x = \frac{\Delta V_{\max}}{\nu}$, где ΔV_{\max} — максимально возможное изменение скорости объекта.

Во многих случаях можно уменьшить количество вычислений за счет уменьшения количества обрабатываемых кадров. Это достигается равномерным прореживанием потока кадров (для обработки берутся не соседние кадры, а с пропуском некоторого количества).

Побочным эффектом данного способа уменьшения количества вычислений является уменьшение скорости реакции системы видеонаблюдения на перемещение объектов (время реакции равно времени между двумя обрабатываемыми кадрами), возможность пропуска быстро движущихся объектов, а также неравномерность вычислительной загрузки процессора. Для обеспечения равномерности его загрузки вычисление разностного изображения может распределяться по времени и выполняться после приема каждого очередного кадра.

Предлагаемый алгоритм слежения за объектами на основе визуальной информации может применяться в обычных системах видеонаблюдения, а также в тепловизионных системах, основанных на инфракрасном излучении. Приведенные способы оптимизации алгоритма вычислений позволяют значительно уменьшить количество операций, для поиска движущихся объектов.

МЕТОДЫ КРУГОВОГО ОБЗОРА НА ОСНОВЕ РАЗНОСТНОГО ПАНОРАМНОГО ОТОБРАЖЕНИЯ

Интересный способ двухэтапного преобразования видеопоследовательности, полученной с видеокамеры, обращающейся вокруг вертикальной оси. Первый этап — преобразование видеопоследовательности в статичное панорамное изображение. Второй — выделение и регистрация разностного изображения между двумя оборотами видеокамеры.

Главная особенность интеллектуальных систем — умение автоматически выделять перемещение объектов в зоне наблюдения, а также изменение геометрических размеров, яркости и цвета.

Наиболее известные способы компрессии динамических изображений — стандарты MPEG-1 и MPEG-2, которые представляют собой комбинацию наиболее прогрессивных способов компрессии. Их достоинства со-

стоят в том, что они позволяют производить как покадровую компрессию — на основе дискретно-косинусного преобразования и кодирования по способу RLE, так и межкадровую — на основе компенсации перемещения методом оценки вектора перемещения для блока пикселей и двунаправленной интерполяции кадра по двум соседним [9 – 11].

Эти способы универсальны, обеспечивают довольно высокий уровень компрессии динамических изображений, однако требуют значительных аппаратных затрат для обработки больших массивов данных, что зачастую не обеспечивает требований режима реального времени. Выходом из этой ситуации могут быть следующие условия [16]:

- значительное уменьшение избыточных данных при формировании видеопоследовательностей;
- исключение из анализа массивов, представляющих фон.

При круговом обзоре, представляемом серией кадров, следующий кадр содержит значительную часть предыдущего, смещенного в сторону, противоположную направлению движения камеры. При этом некоторая часть изображения выходит из кадра, а такая же часть нового изображения в него входит. Если выделить только новые неповторяющиеся части изображений и «склеить» их, то можно получить широкоформатное панорамное статичное изображение в результате одного поворота камеры.

Основная проблема получения панорамного изображения — это обеспечение стыковки кадров, составляющих видеопоследовательность. Отмечая B как ширину кадра, а b — ширину новой части изображения, для поиска границы предыдущего кадра со следующим предлагается сканирование поля края изображения кадра (область A) шириной в зоне поиска $B - b$ в кадре $n + 1$ (рис.8).

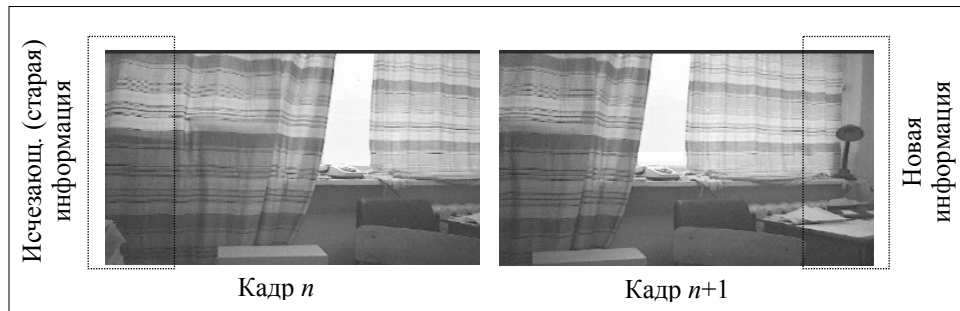


Рис. 8. Кадры видеопоследовательности при вращении видеокамеры вправо

Степень соответствия выбранного поля кадра n в кадре $n + 1$ на каждом шаге сканирования определяется по минимуму функции

$$F = \sum_{(i,j \in R)} |A(i,j) - B(i,j)|,$$

где $A(i, j)$ — значения яркостей элементов поля края изображения кадра n ; $B(i, j)$ — значение яркостей элементов поля, выбираемое в зоне поиска кадра $(n + 1)$.

Минимум функции F будет наблюдаться при определении степени подобия областей A и B (рис. 9).

Сложность процедуры поиска края изображения при неравномерной скорости вращения может быть определена следующим образом.

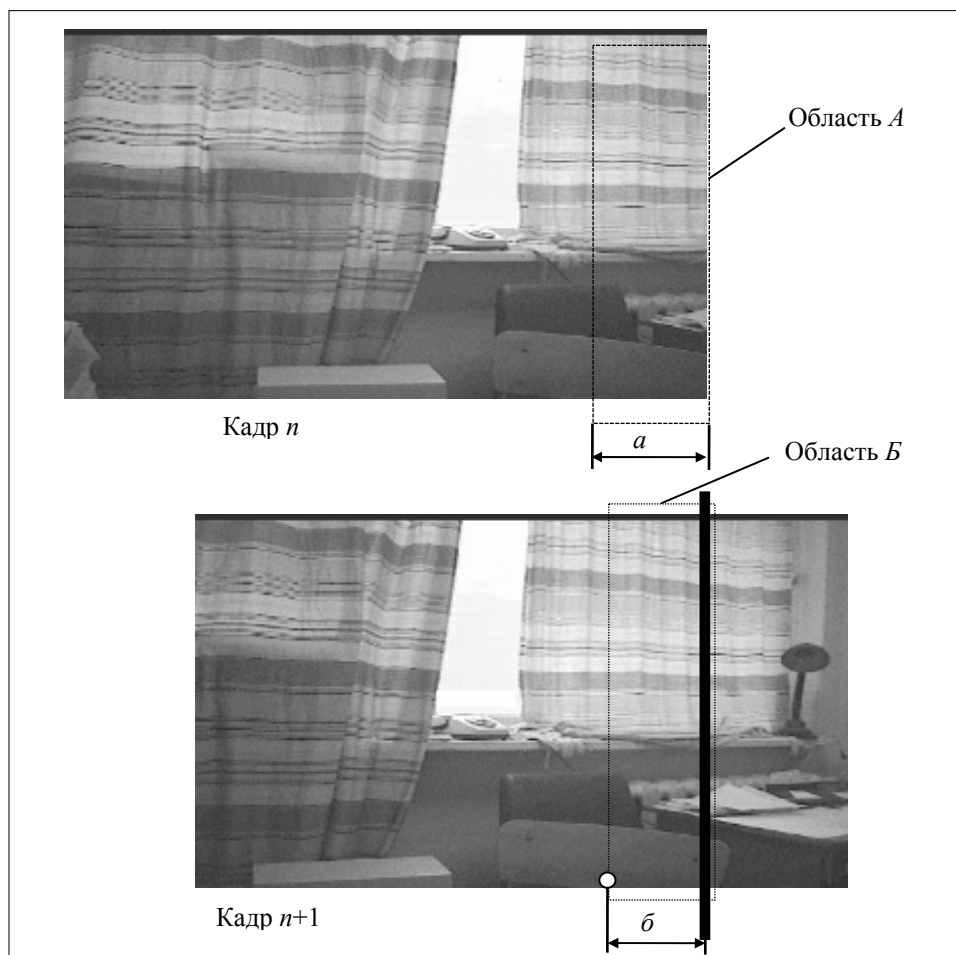


Рис. 9. Поиск границы появления новой информации в следующем кадре. Жирной линией показана граница появления новой части видеоизображения

Для одной операции сравнения с полосой bh пикселей необходимо выполнить $\approx 2bh$ операций типа сложения. Если поиск осуществляется по зоне сканирования шириной $(B - b)$, то общее количество операций определится как $2(B - b)bh$.

Если определить ширину контрольной полосы края изображения в относительных единицах от ширины кадра как $b = kB$, где $k \leq 1$, тогда количество операций для реализации процедуры поиска определится как $2kB^2bh(1 - k)$.

Для улучшения зрительного восприятия панорамного изображения целесообразно выполнить состыковку новой части изображения со сглаживанием границы состыковки [12]. Изображения должны плавно переходить друг в друга (рис. 10).



Рис. 10. Панорамное изображение, представленное суммой кадров, показанных на рис. 9

Описанный способ панорамирования видеопоследовательности может быть использован при проведении съемки с объекта, движущегося прямолинейно (например, с поезда, автомобиля). В результате будет получена широкая панорама и значительная компрессия. Этот подход может эффективно использоваться при построении систем виртуальной реальности. Средний уровень компрессии видеопоследовательности за счет панорамирования достигает 10.

Представление видеопоследовательности в виде панорамного изображения значительно облегчает условия контроля оператором зоны наблюдения.

Полученная последовательность широкоформатных видеоизображений может дополнительно подвергаться сжатию любым алгоритмом компрессии динамических видеоизображений, в том числе и MPEG[10]. При этом обеспечивается степень компрессии значительно выше выходной последовательности кадров (степень компрессии зависит от скорости кругового движения видеокамеры, частоты кадров съемки, а также от динамики смены окружения).

Входное динамическое изображение может быть полностью восстановлено из панорамного последовательным вырезанием из панорамного изображения отдельных областей. Для этого при создании панорамного изображения следует сохранять ширину области, которая дополняется к панорамному изображению от кадра к кадру (это необходимо только в случае сменной угловой скорости видеокамеры). Динамическая последовательность видеокадров может быть восстановлена из панорамного изображения в темпе съемки или в произвольном (замедленном или ускоренном) темпе.

Следующим этапом уменьшения избыточной информации, которая предоставляется оператору, может быть удаление фона с изображения и выделение новой информации между оборотами видеокамеры. Это может быть реализовано путем выделения разностного изображения между панорамными изображениями двух соседних оборотов видеокамеры (рис.11).

На рис.11а, б показана панорама обзора видеокамеры соответственно для первого и второго оборотов камеры. Как видно, изображение фона не подлежит изменениям, разность между этими двумя изображениями заключается в перемещении предмета из положения А в положение Б за время оборота камеры. На основании этих двух изображений построено третье, каждый пиксель которого является разницей значений яркости пикселей этих двух изображений (рис.11, в). Если ограничить это изображение по по-

рогу, то можно автоматически определить не только наличие изменений в обстановке, но и локализовать место, где происходят изменения. В результате автоматического определения получим две области повышенного внимания (на рис. 11 они выделены прямоугольниками). Таким образом осуществляется автоматическая регистрация и сигнализация наличия перемещений или локального изменения яркости (например, при возгорании) в поле зрения видеокамеры кругового обзора.

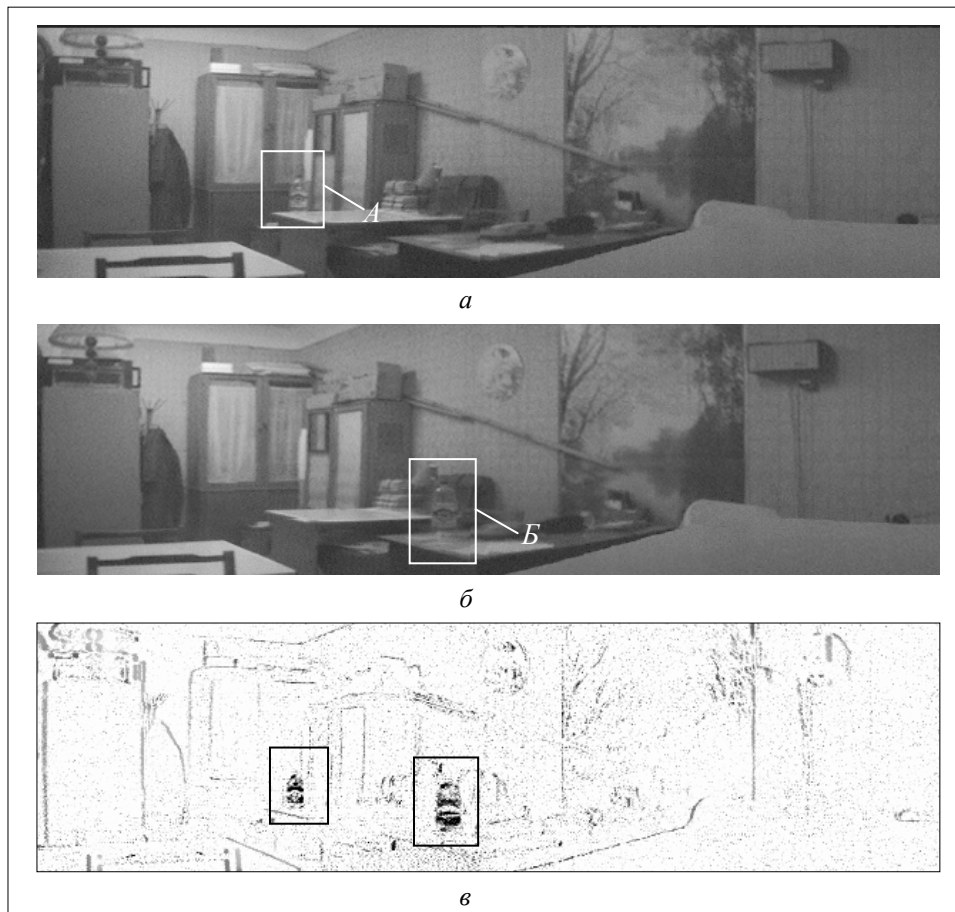


Рис. 11. Пример разностного панорамного изображения, построенного из 160 кадров видеопоследовательности: *а* — панорамное изображение, снятое на первом обороте видеокамеры; *б* — то же на втором; *в* — разница изображений *а* и *б* (значение разницы усилено в четыре раза, изображение инвертировано)

Ошибки построения панорамных изображений, складывающихся из полос, граница между которыми не всегда может быть определена с точностью до пикселя, а также наличие ошибки в механике системы кругового обзора приводят к появлению некоторого «шума контуров» на разностном изображении. Однако такой шум в большинстве случаев устраняется высокочастотной фильтрацией [12–14]. В то же время, исчезнувшие движущиеся предметы или предметы, сменившие геометрическую форму, размеры, яркость или цвет, будут выделены практически полностью (если они отличаются от фона).

Использование разностного панорамного изображения значительно облегчает работу оператора и сокращает объем регистрируемой информации. В зависимости от частоты и масштабности происходящих изменений степень компрессии может достигать сотен тысяч раз. Если выполнялась постоянная регистрация (без фильтрации) разностных изображений, то панорамные изображения, а при необходимости и входная видеопоследовательность, могут быть восстановлены в реальном или произвольном темпе.

Кроме того, полученное разностное изображение создает предпосылки для автоматического раскрытия происходящих изменений, т.е. появляется реальная возможность создания интеллектуальных систем кругового обзора, выдающих сигнал тревоги и регистрирующих возникшие изменения или обращающих внимание оператора на имеющиеся изменения.

ВЫВОДЫ

1. Применение комбинации методов обратного распространения ошибки, медианного сглаживания изображений и увеличения контрастности явилось существенным развитием метода самоорганизующейся карты Кохонена, что позволило сократить время преобразования до нескольких секунд (с помощью известных методов оно длилось в течение десятков минут) с вероятностью распознавания 0,98.

2. Результаты распознавания с помощью нейронной сети, представленной в виде программного модуля на языке Visual C++ 6.0 для ОС Windows показали надежную идентификацию изображения государственных номерных знаков автомобиля при смещениях и поворотах в поле зрения, а также при наличии значительной загрязненности и атмосферных помех.

3. Предложен способ панорамирования видеопоследовательности, который обеспечивает широкую панораму изображений при высокой компрессии данных, в результате чего средний уровень компрессии видеопоследовательности за счет панорамирования достигает 10.

4. Установлено, что использование разностного панорамированного изображения значительно сокращает объем регистрируемой информации. При этом в зависимости от частоты и масштабности происходящих изменений степень компрессии может достигать сотен тысяч.

ЛИТЕРАТУРА

1. *Вінцюк Т.К.* Образный компьютер: концепції, методологія, підходи // Укр ОБРАЗ, 2000. — Київ: ІК ім. В.М. Глушкова НАН України, 2000. — С. 9–16.
2. *Васюхин М.И., Пюшки Л.* Подходы к построению автоматизированной системы охраны особо важных объектов // УСиМ. — 2002. — № 1. — С. 88–91.
3. *Крылов В.Н., Максимов М.В., Давыдов В.О.* Автоматизированная система распознавания // Искусственный интеллект. — 2002. — № 4. — С. 462–469.
4. *Уоссермен Ф.* Нейрокомпьютерная техника: Теория и практика / Пер.с англ. — М.: Мир, 1992. — 240 с.

5. *Галуев Г.А., Тараненко А.С.* Нейросетевая система автоматической идентификации номерных знаков автомобилей // Искусственный интеллект. — 2002. — № 4. — С. 666–674.
6. *O’Gorman L., Nickerson J.V.* An approach to fingerprint filter design // Pattern recognition. — 1989. — **22**, №1. — P. 29–38.
7. *Ту Дж., Гонсалес Р.* Принципы распознавания образов / Пер. с англ. — М.: Мир, 1978. — 411 с.
8. *Боюн В.П., Довгань В.В.* Слежение за объектами на основе системы видеонаблюдения, оптимизация алгоритма слежения // Комп’ютерна математика. Оптимізація обчислень. — Київ: ІК ім. В.М. Глушкова НАН України, 2001. — С. 50–57.
9. *Зелов С.* Цифровое кодирование изображений // Компьютер-Пресс. — 1997. — № 3, 4. — С.13–14.
10. *Сидоренко Ю.* «МРЕГ» — Путь в цифровое завтра // Компьютерное обозрение. — 1997. — №17(90). — С.18–23.
11. *Старовойтов В.В.* Локальные геометрические методы цифровой обработки и анализа изображений. — Минск: Ин-т техн. киберн. НАН Беларуси, 1997. — 284 с.
12. *Прэнтт У.* Цифровая обработка изображений: в 2-х кн. / Пер. с англ. под ред. Д.С.Лебедева. — М.: Мир, 1982. — **2**. — 480 с.
13. *Хироze М.* Создание виртуальных миров на основе реальных изображений // Компьютеры + программы. — 1997. — № 9(42). — С.14–20.
14. *Абламейко С.В., Лагуновский Д.М.* Обработка изображений: технология, методы, применение. — Минск: Ин-т техн. киберн. НАН Беларуси, 1999. — 300 с.
15. *Бородин В.А., Васюхин М.И., Пюшки Л.* Методы повышения надежности автоматизированных систем охраны особо важных объектов, занимающих большие территории // Математичні машини і системи. — 2002. — № 3. — С.90–95.
16. *Боюн В.П., Довгань В.В.* Інтелектуалізація відеосистем колового огляду // Праці Міжнар. конф. УкрОБРАЗ-2000. — Київ: ІК ім. В.М. Глушкова НАН України, 2000. — С.283–286.

Поступила 02.04.2003