

УДК 004.93

Э.А. Троценко, Н.А. Миняйло, С.А. Икол, П.Н. Веропотвелян

СИСТЕМА УПРАВЛЕНИЯ БНТС В ДИСТАНЦИОННОМ РЕЖИМЕ ПРИ ПОМОЩИ ЖЕСТОВ

Аннотация. В статье рассматривается подход к построению системы управления беспилотным наземным техническим средством в дистанционном режиме при помощи выдачи жестов. Комплексно решены задачи по обработке видеопотока, сегментации изображений, детектированию движения и выделению объектов в зоне интереса, распознаванию человека-регулирующего. Представлена реализация логических моделей определения управляющих статических и динамических жестов на уровне программируемого логического контроллера.

Ключевые слова: беспилотное транспортное средство, система управления, регулировщик, жест, распознавание, логика, массив, временная булевая функция.

Введение. Одним из способов бесконтактного человеко-машинного взаимодействия (англ. Human-machine interface, HMI) является управление техническим средством в дистанционном режиме при помощи выдачи жестов.

Необходимость реализации данного режима для беспилотных наземных технических средств (БНТС) обоснована управлением машинами в сложных условиях при совершении маневров в ограниченном пространстве, например, на стоянках, в укрытиях (гаражах), при погрузке/разгрузке транспортных средств и т. д. Также при перемещении БНТС по дорогам общего пользования необходимо решать задачу распознавания и определения типов жестов, которые могут выдаваться регулировщиками дорожного движения.

Использование БНТС в аварийно-спасательных подразделениях Министерства чрезвычайных ситуаций и в военной отрасли предполагает движение машин в режиме «тишины», когда выдавать звуковые, световые или другие виды сигналов запрещено или затруднительно.

© Троценко Э.А., Миняйло Н.А., Икол С.А., Веропотвелян П.Н., 2016

Поэтому единственным интерфейсом человеко-машинного взаимодействия остаются жесты.

Состояние вопроса. Развитие направления создания беспилотных технических средств и оснащения их различными видами сенсорных устройств предполагает также использование современных подходов в распознавании объектов и их положения в пространстве, построенных на методах и алгоритмах, использующие разные математические аппараты. При решении задач распознавания жестов рук выделяют три основных подхода [1-3].

В основе скелетных методов обычно используются алгоритмы поиска границ объектов с целью определения замкнутого контура. При анализе формы контуров в данном случае учитываются углы, выступы, впадины и другие точки с высокими значениями кривизны. После определения контура объекта строится скелет и находятся опорные точки, за перемещение которых в дальнейшем следит система [1-3].

Более сложными являются методы построения 3D моделей рук в виде трехмерных поверхностей и их классификации с помощью нейронных сетей.

Метод 2D распознавания жестов рук схож с предыдущим, но вместо объемных моделей использует построение двумерных изображений.

На основе разработанных методов на сегодняшний день созданы практические приложения, наиболее значимыми из которых считаются [4]:

- программные библиотеки Kinect SDK (Microsoft), OpenNI/NITE (PrimeSense), PMD SDK и SoftKinetic IISU для распознавания ключевых точек тела человека посредством применения трехмерных сенсоров;

- сенсоры компаний LeapMotion, SoftKinetic для распознавания ключевых точек руки человека;

- компьютерные приложения Flutter, PointGrab HGRS, XTR3D, EyeSight PC для управления компьютером посредством простых жестов рук.

Несмотря на наличие разносторонних подходов к решению задач распознавания жестов, каждый из методов имеет свои преимущества и недостатки. Применение подхода построения 3D моделей и

обучение нейронной сети предполагает получение информации со стереокамер с возможностью определения глубины изображения, а также требует значительных ресурсов для выполнения расчетных задач, что усложняет его использование в системах реального времени. Чаще всего используют 2D методы, т.к. для их реализации достаточно информации, поступающей с монокамеры и в результате упрощения расчетных задач снижается вычислительная мощность и скорость обработки данных. При этом, данные методы считаются менее точными и требуют поиска путей усовершенствования алгоритмов распознавания как динамических, так и статических жестов рук.

Таким образом, актуальной является задача создания новых комплексных подходов, методов и алгоритмов распознавания жестов, которые бы также могли учитывать высокую чувствительность изображений, получаемых с видеокамер, к изменениям освещённости, сложность в подборе адаптивных алгоритмов сегментации, частичную потерю данных при бинаризации изображений, а также решать поставленные задачи за минимально возможный период времени.

Постановка задачи. Разработать систему управления (СУ) БНТС в дистанционном режиме при помощи выдачи жестов человеком-регулирующим. При этом обработку видеоинформации с целью распознавания объектов выполнять на уровне ЭВМ, а определение самих жестов, формирование и выдачу управляющих воздействий на исполнительные устройства БНТС осуществлять на уровне программируемого логического контроллера (сокр. ПЛК; англ. programmable logic controller, сокр. PLC).

Основная часть. На первом этапе проектирования системы управления БНТС в дистанционном режиме выдачи жестов была разработана функциональная схема (рис.1).

Предложенная структура СУ включает в себя три базовых узла: персональный компьютер (сокр. РС; англ. personal computer) РС-box и два программируемых логических контроллера: PLC1, PLC2. РС-box принимает и обрабатывает информацию от системы ориентации БНТС в пространстве, сенсорных устройств ближнего и дальнего действия, а также через каналы связи обменивается данными с устройствами дистанционного управления. PLC1 выполняет функции принятия решения по управлению движением БНТС. Сформированные управляющие воздействия передаются на приводы исполнитель-

ных механизмов транспортного средства. В качестве обратной связи на контроллер поступают сигналы о положении всех механических устройств, входящих в состав БНТС. Второй PLC2 выполняет обработку информации с радаров, установленных спереди и сзади БНТС.

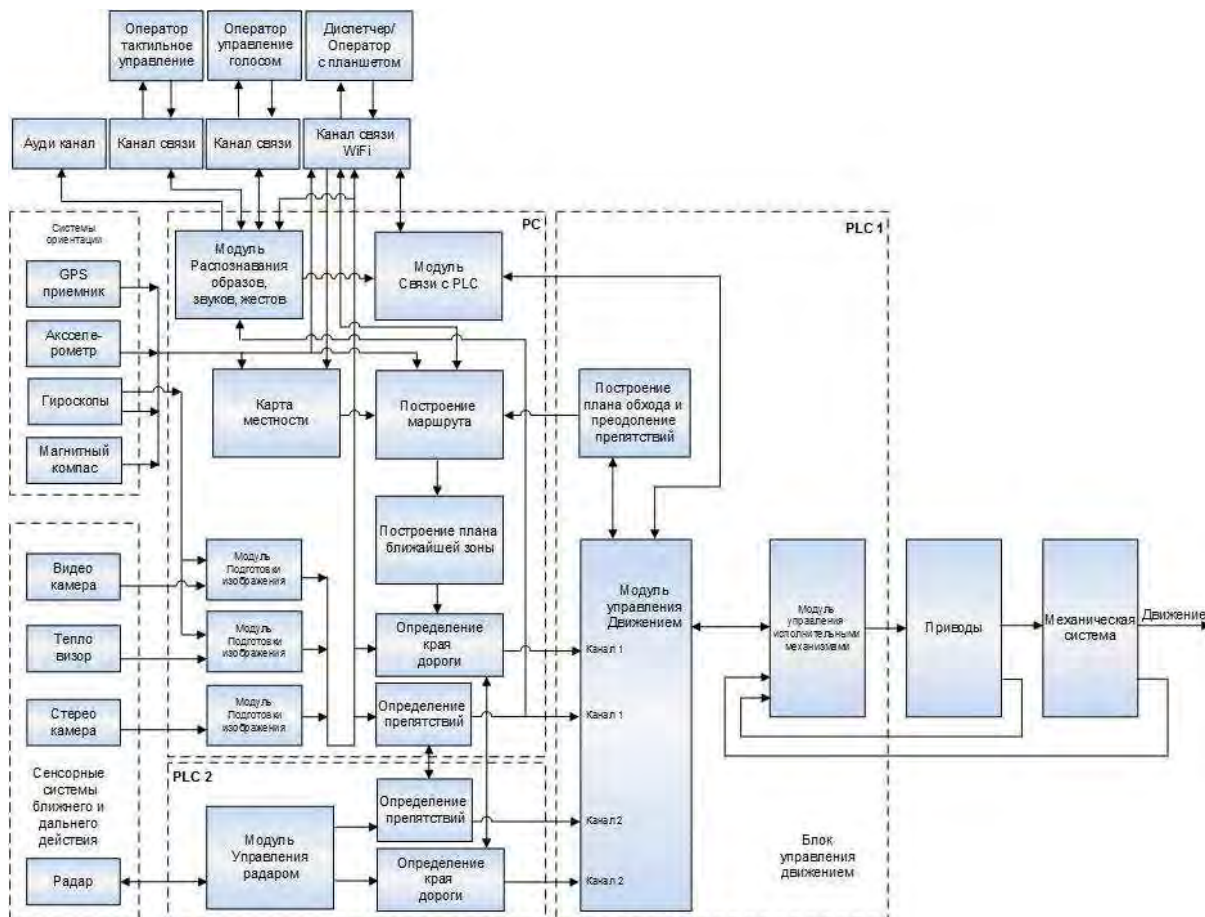


Рисунок 1 – Функциональная схема управления движением БНТС

Как говорилось ранее, одним из необходимых режимов управления движением БНТС является дистанционная выдача команд при помощи жестов человеком-регулирующим. Переключение СУ БНТС в данный режим может осуществляться диспетчером/оператором с РС планшета либо в условиях автономного движения по сигналу наличия человека-регулирующего от программного модуля распознавания препятствий.

Выбор сигналов управления БНТС, подаваемых снаружи руками, основывался на общепринятых правилах управления машинами на стоянках, для постановки их в укрытие, при погрузке и т. д. В качестве основных приняты команды «Внимание», «Заводи», «Глуши двигатель», «Вперед», «Направо», «Налево», «Меньший ход», «Стоять», «Задний ход» (таблица 1).

Типы и коды жестов для управления БНТС

Жест													Произвольное положение рук
Имя	Нет фигуры	Ожидание	Внимание	На лево	На право	Заводи	Меньше ход	Стой	Глуши двигатель	Ко мне вперед	Задний ход	Идентификация человека	Ошибка жест
Код F_	00	01	02	03	04	05	06	07	08	09	0A	0B	FF

Исходя из правил подачи каждого из жестов их можно разделить на простые и комплексные, а также на статические и динамические [2]. Статические жесты характеризуются четкой фиксацией по времени крайних положений рук. Динамические жесты выполняются непрерывно на протяжении всего времени их выдачи, или до момента пока они не будут приняты системой для выполнения.

Согласно принятым правилам, управление движением машины при помощи жестов должно осуществляться человеком-регулирующим, находящимся на расстоянии не ближе, чем 5м от транспортного средства. Для исключения вероятности выдачи команд управления посторонним человеком, а также человеком находящимся на опасном расстоянии от машины, в системе введено понятие «зоны интереса» (рис.2). Эта область пространства, которая ограничивается расстоянием до БНТС от 5 до 10 метров. Также принято, что выдавать команды в виде жестов может человек, находящийся ближе всего к оптической оси видеокамеры, т.е. попадающий в область на изображении $A_1A_2D_2D_1$.

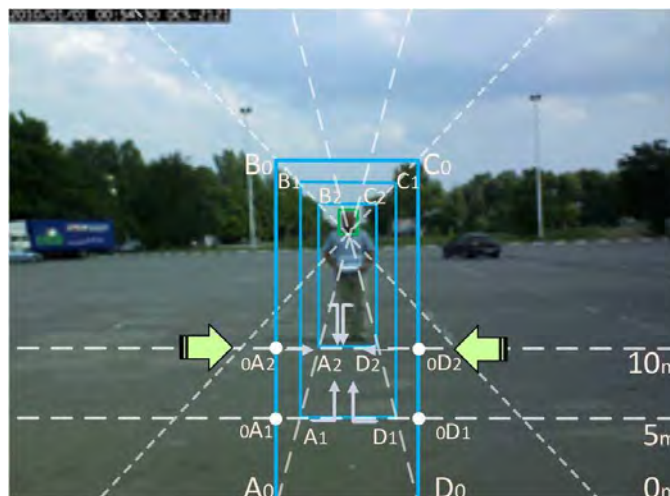


Рисунок 2 – Определение на изображении «зоны интереса»

Для выполнения функций по распознаванию положения человека-регулировщика на уровне PC-box разработано программное обеспечение, которое включает два основных модуля: модуль определения зоны интереса и распознавания движущихся объектов в данной области (рис.3), а также модуль построения и анализа изменения полей динамической сетки с целью формирования битового массива для дальнейшего определения типа жеста (рис.4).

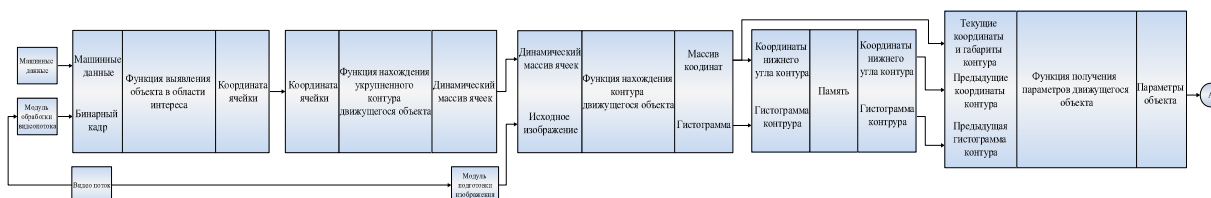


Рисунок 3 – Структурная схема программного модуля определения зоны интереса и распознавания движущихся объектов

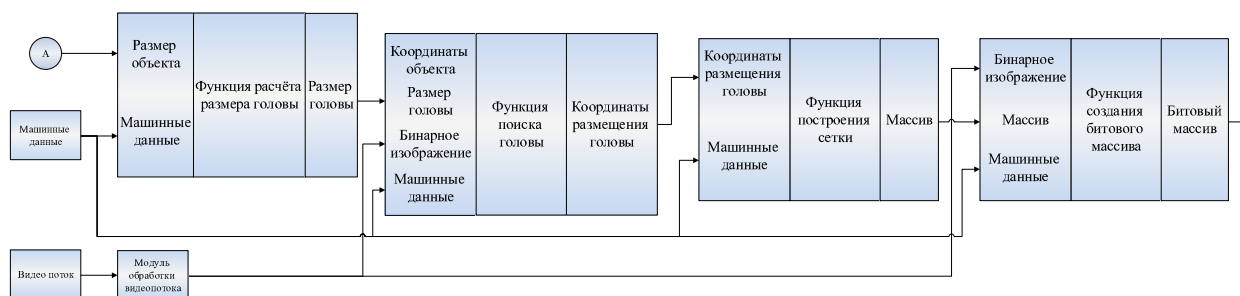


Рисунок 4 – Структурная схема программного модуля построения и анализа изменения полей динамической маски и формирования битового массива

Исходными данными для модуля определения зоны интереса и распознавания движущихся объектов в данной области являются параметры фокальной плоскости и вес пикселей в видеокдрах, которые рассчитываются на основании данных о расположении камеры и ее технических характеристиках.

Для построения зоны интереса вначале производится расчет ее ширины A_0D_0 и с учетом особенностей построения изображений оптической системой, ее проекция A_2D_2 на расстоянии A_0A_2 , D_0D_2 . Определение координат точек A_1, A_2, D_2 и D_1 позволяет получить уравнение двух прямых A_0A_2 , D_0D_2 , которые ограничивают зону интереса справа и слева.

Поскольку координата $У$ меняется дискретно (координаты нижней и верхней границ строки виртуальной сетки), находятся координаты столбцов, которые будут принадлежать прямым A_0A_2, D_0D_2 .

Используя уравнения прямых A_0A_2, D_0D_2 , проходящих через координаты нижних и верхних точек, ограничивающих зону безопасности, и зная координаты, ограничивающие высоту каждой строки, определяются координаты крайних левых и правых ячеек сетки (рис.5а). Далее для всей «зоны интереса» строится статическая сетка (рис.5б), которая для расчетов представляется как массив данных $B(i,j)$.



Рисунок 5 – Определение зоны интереса (а) и разбиение ее на отдельные области (б)

Распознавание движущихся объектов в области интереса происходит на основании бинарного изображения и анализа изменения степени заполнения отдельных ячеек сетки белыми пикселями (рис.6):

$$B(i, j) = \begin{cases} 1, & R_{ij} > 20\%; \\ 0, & R_{ij} < 20\%, \end{cases} \quad (1)$$

где $B(i, j)$ – i, j -ая ячейка сетки; R_{ij} – суммарное количество белых пикселей в i, j ячейке сетки, pix .

Если $B(i, j)=1$, то система выполняет анализ наличия таких же ячеек в соседних областях сетки по 8-ми направлениям относительно первой найденной заполненной ячейки. Поиск объекта продолжается до тех пор, пока по тем же направлениям будут найдены области с $R_{ij} < 20\%$. По координатам крайних ячеек строится прямоугольная об-

ласть, в которой был определен объект и рассчитываются ее параметры: R^x_{object} , R^y_{object} – координаты (x, y) верхней левой точки; w – ширина прямоугольника (объекта), h – высота прямоугольника объекта, pix (рис.6.).

В качестве исходной информации для модуля построения динамической сетки является бинарное изображение и параметры области, в которой определен человек-регулировщик (рис.6).

Принцип построения динамической сетки основан на физиологических особенностях строения тела человека и пропорциях его фигуры [6]. Обычно за единицу измерения человеческого тела принимается «длина головы». Рост человека в среднем равен 7,5 – 9 длины головы, ширина в плечах – 2 длины [6].

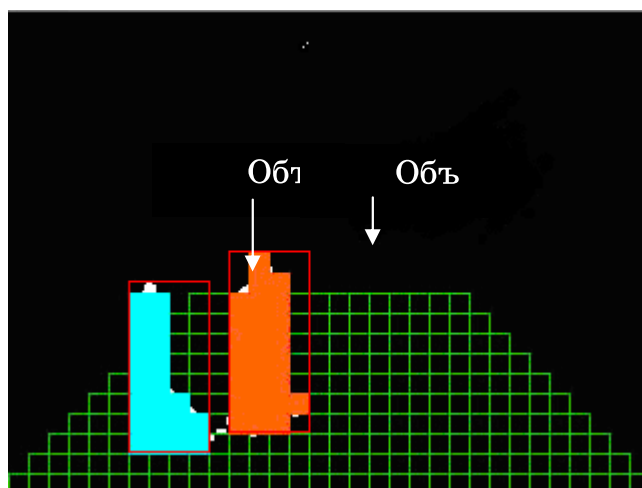


Рисунок 6 – Результат обработки изображения модулем определения зоны интереса и распознавания двух движущихся объектов

При управлении БНТС в режиме выдачи жестов в качестве регулировщика может быть человек любого телосложения, поэтому на начальном этапе в разработанной системе предусмотрена функция инициализации его параметров. Уточнение пропорций человека выполняется на основе информации об изменении ширины (w) и высоты (h) области, в которой определен регулировщик, при исходном его положении (рис.7а), в момент максимального размаха рук в сторону (рис. 7б) и при поднятии рук вверх (рис. 7в):

$$w = \max\{w_1, w_2, \dots, w_n\}; \quad h = \max\{h_1, h_2, \dots, h_n\} \quad (2)$$

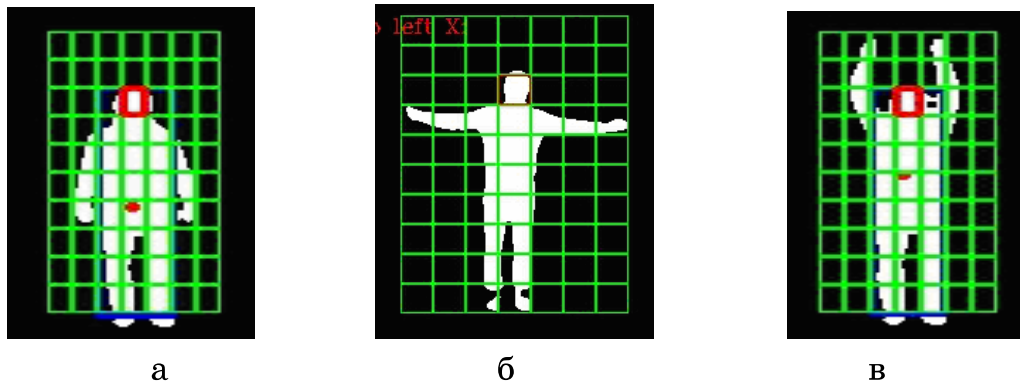


Рисунок 7 – Основные этапы инициализации пропорций регулировщика: а – исходное положение человека; б – положение человека с максимальным размахом рук в сторону; в – положение человека с максимально поднятыми руками вверх

В разрабатываемом подходе за базовую точку расчета построения динамической сетки выбрано расположение головы человека. Данное решение обосновывается следующими положениями:

- голова является неотъемлемой частью человека;
- высота и ширина человека пропорционально соотносится с размерами головы;
- регулировщик выдает жесты в положении стоя-прямо, что обуславливает нахождение головы в верхней части туловища;
- голова имеет наименьшую степень свободы перемещения по сравнению с руками и ногами.

Поиск расположения головы регулировщика начинается с определения точки, которая соответствует середине туловища (рис.7а). Координаты данной точки на первом этапе рассчитываются по формулам:

$$P_{center}^x = \frac{R_{object}^x + w}{2} ; \quad (3)$$

$$P_{center}^y = \frac{R_{object}^y + h}{2} , \quad (4)$$

где P_{center}^x и P_{center}^y – координаты середины объекта; R_{object} , $R_{yobject}$ – координаты вершины прямоугольника описывающего объект; h , w – высота и ширина прямоугольника, px .

В связи с тем, что человек в выделенной области может располагаться не по центру, производится уточнение его ширины и высоты

относительно найденной точки (P_{center}^x и P_{center}^y). Для этого в направлении вправо и влево производится анализ бинарного изображения и находятся координаты крайних точек туловища человека P_{right}^x, P_{right}^y и P_{left}^x, P_{left}^y . Далее уточняются координаты середины объекта:

$$P_{centerObj}^x = \frac{P_{right}^x - P_{left}^x}{2}; \quad (5)$$

$$P_{centerObj}^y = P_{center}^y. \quad (6)$$

Аналогичным путем выполняется уточнение координат точки ($P_{centerObj}^x$ и $P_{centerObj}^y$) по оси ОУ, в результате анализа бинарного изображения вверх и вниз (рис.7в). После уточнения расположения туловища человека находим координаты соответствующего ему прямоугольника:

$$R_{human}^x = P_{right}^x; \quad (7)$$

$$R_{human}^y = P_{top}^y; \quad (8)$$

$$W_{human} = P_{right}^x - P_{left}^x; \quad (9)$$

$$h_{human} = (R_{object}^y + h_{object}) - P_{top}^y, \quad (10)$$

где R_{human}^x, R_{human}^y – координаты верхней левой вершины прямоугольника; W_{human} – ширина прямоугольника, рix; h_{human} – высота прямоугольника, рix; P_{top}^y – координата ОУ верхней точки объекта.

Исходя из указанных ранее пропорций человека, голова соответствует области размером:

$$W_{head} = \frac{w}{7}; \quad (11)$$

$$h_{head} = \frac{h}{9}, \quad (12)$$

где W_{head} – ширина области головы, рix; h_{head} – высота области головы, рix.

Поиск координат головы производится в 1/7 верхней части прямоугольника R_{human} , вдоль координатной оси ОХ :

$$R_{head}^x = \begin{cases} R_{human}^x + \left(\frac{w_{human}}{2} - \frac{w_{head}}{2} \right), & \max = R_{head}^1 \\ R_{head}^x - w_{head}, & \max = R_{head}^2 \\ R_{head}^x + 2w_{head}, & \max = R_{head}^3 \end{cases}, \quad (13)$$

где $R_{head}^1, R_{head}^2, R_{head}^3$ – степень заполнения области головы белыми пикселями при расположении прямоугольника R_{head} в 1, 2 и 3 позиции соответственно; $\max = \max \{R_{head}^1, R_{head}^2, R_{head}^3\}$ – максимальное значение степени заполнения области головы белыми пикселями относительно трех позиций прямоугольника.

Поскольку анализ проводится вдоль оси ОХ, то во всех случаях координата по ОУ равняется $R_{yhead} = R_{yhuman}$.

Далее относительно найденных координат расположения области головы регулировщика строится динамическая сетка (рис.7), которая представляет собой двумерную матрицу, каждый элемент которой является прямоугольником на изображении:

$$A = \begin{pmatrix} R_{1A} & R_{1B} & \dots & R_{1G} \\ R_{2A} & R_{2B} & \dots & R_{2G} \\ \dots & \dots & \dots & \dots \\ R_{9A} & R_{9B} & \dots & R_{9G} \end{pmatrix} \quad (14)$$

Формирование из матрицы (14) битового массива для контроллера осуществляется согласно формуле:

$$C(i, j) = \begin{cases} 1, & R_{ij} > 20\% \\ 0, & R_{ij} < 20\% \end{cases} \rightarrow$$

	A	B	C	D	E	F	G
1	0	0	0	0	1	0	0
2	0	0	0	0	1	1	0
3	0	0	0	1	1	1	0
4	0	0	1	1	1	0	0
5	0	0	1	1	1	0	0
6	0	0	1	1	1	0	0
7	0	0	1	1	1	0	0

Распознавание типа жеста на уровне контроллера осуществляется в результате последовательной обработки бинарных массивов, которые поступают с РС-box. Каждый жест описан логической моделью на основе временных булевых функций и состояний [7]. Например, для жеста «Стой» определено три состояния: S_0 – исходное положение регулировщика, руки опущены вниз; S_1 – второе состояние, левая рука поднята вверх; S_3 – исходное положение, руки опущены вниз, но в отличие от S_0 данному состоянию должно обязательно

предшествовать состоянию S_1 (рис.6б). В модели учитывается также время перехода из одного состояния в другое (t_1, t_2):
 $S_0 \xrightarrow{t_1} S_1 \xrightarrow{t_2} S_2$.

Для учета возможных вариаций положения рук относительно туловища в момент выдачи жестов в каждом состоянии рассматривается несколько комбинаций по схеме «ИЛИ». При описании логики обработки массивов жеста «Стой» в состоянии S_1 за истинное значение принимается любая из следующих комбинаций заполнения «1» ячеек сетки: $(E_4, E_3, E_2) \mid (E_4, F_3, F_2) \mid (E_4, E_3, F_2)$ (рис.8а).

В логических моделях комплексных динамических жестов учитывается также то, что регулировщик может повторять некоторые положения рук (состояния) N -ое количество раз.

Программная реализация логики распознавания статических и динамических жестов выполнена с применением стандартных функциональных блоков (Function blocks – FB) языка STEP 7.

Каждому жесту в соответствие поставлен код (таблица 1), на основании которого формируется набор управляющих команд и выполняется передача их на исполнительные устройства БНТС.

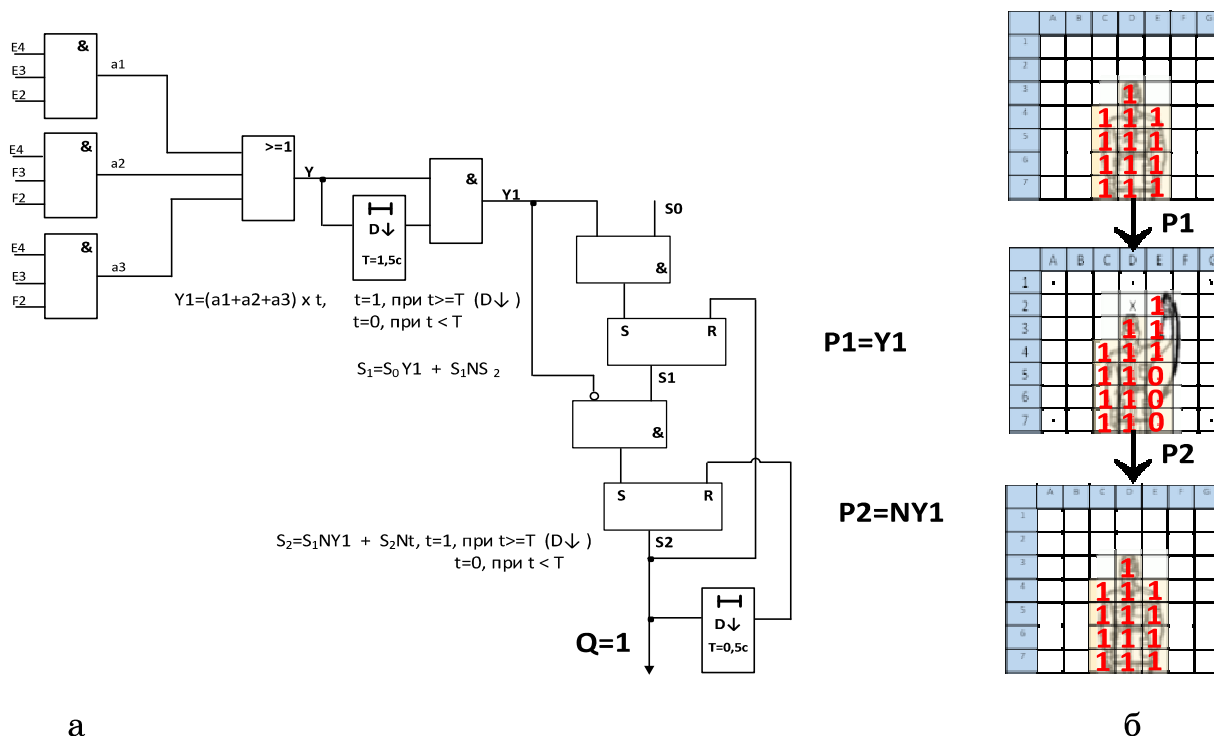


Рисунок 8 – Жест «Стой»: а – логика обработки данных с помощью временных булевых функций; б – соответствие позиций регулировщика состояниям и бинарным массивам

Разработанная система управления БНТС с помощи жестов тестировалась в условиях дневного освещения, когда информация об окружающем пространстве поступает через видеопоток стереокамеры и в ночное время, когда основным источником данных является видео с тепловизора (рис.9а,б). Отличительной особенностью работы с видеоизображениями различных сенсорных систем является процедура предварительной обработки кадров и применение алгоритмов сегментации изображений, а представленное в работе комплексное решение задач построения зоны интереса, выделения человека-регулирующего и распознавания жестов остается для всех условий одинаковым (рис.9в-д).

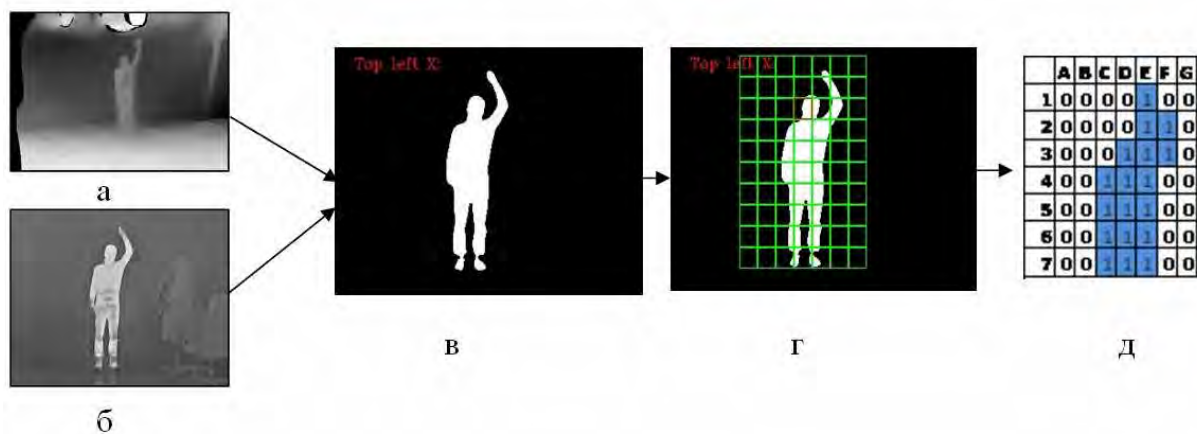


Рисунок 9 – Результаты работы программного обеспечения на уровне РС-box: а – исходное изображение со стереокамеры; б – исходное изображение с тепловизора; в – бинарное изображение; г – построение динамической сетки; д – сформированный массив данных для уровня PLC.

Таким образом, предложенный подход к распознаванию команд человека-регулирующего на уровне логических функций позволил распределить задачи между базовыми узлами системы управления БНТС, повысить скорость обработки данных, формирования и выдачи управляющих команд, что особенно актуально при управлении движущимся транспортным средством.

ЛИТЕРАТУРА

1. Мурлин А.Г. Алгоритм и методы обнаружения и распознавания жестов руки на видео в режиме реального времени / А.Г. Мурлин, Д.Л. Пиотровский, Е.А. Руденко, М.В. Янаева [Электронный ресурс] // Научный журнал КубГАУ – 2014. – № 97(03). – С.284 – 293. // Режим доступа: <http://ej.kubagro.ru/2014/03/pdf/20.pdf>.
2. Антощук С.Г. Людино-машинный интерфейс на основі розпізнавання жестів /С.Г. Антощук, О.С. Яровенко [Електронний ресурс]// Електромашинобудування та електрообладнання. – 2010. – № 76. – С. 15-20. // Режим доступа: http://nbuv.gov.ua/UJRN/etks_2010_76_5.
3. Кладов С.А. Распознавание жестов / С.А. Кладов, О.В. Рогозин [Электронный ресурс] // Молодежный научно-технический вестник. Издатель ФГБОУ ВПО "МГТУ им. Н.Э. Баумана". – 2013. – № 05. // Режим доступа: <http://sntbul.bmstu.ru/doc/568859.html>
4. Нагапетян В.Э. Методы распознавания жестов руки на основе анализа дальностных изображений: дис. канд. физ.–матем. наук: 05.13.17 / Нагапетян Ваагн Эдвардович. Российский университет дружбы народов – Москва, 2013. — 117 с.
5. Fast and Accurate Hand Pose Detection for Human-Robot Interaction / Luis Antyn-Canalns, Elena Sónchez-Nielsen, Modesto Castrillyn Santana // IbPRIA – 2005. – S. 553-560.
6. Хамм Д. Как рисовать голову и фигуру человека [Текст] / Джек Хамм. – Минск, 2007. – 126 с.
7. Поляков В.Е. Теоретические основы построения логической части релейной защиты и автоматики энергосистем [Текст] / Поляков В.Е., Жуков С.Ф., Троценко А.А. и др.; Под ред. В.Е. Полякова. – М.: «Энергия». – 1979. – 240с.