



Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В. Келдыша
Российской академии наук

А.К.Платонов Н.С.Сербенюк

Выявление
с помощью TV-камеры
препятствий движению робота

Препринт №

Москва

Аннотация

Выявление с помощью TV-камеры
препятствий движению робота

А.К.Платонов Н.С.Сербенюк

(platonov@keldysh.ru serb@keldysh.ru)

Работа посвящена описанию разработки монокулярной телевизионной системы для управления движением робота среди препятствий. Описывается устройство дешёвой системы технического зрения на базе черно-белой телевизионной камеры. Приводится оценка погрешностей привязки телевизионных кадров к параметрам навигационной системы мобильного робота. Описывается взаимодействие зрительной системы с системой управления движением робота. Разработанная система машинного зрения показала возможность использования локальных преобразований яркостной картины кадра для детектирования возникновения препятствия движению робота.

Работа поддержана грантами РФФИ № 02-01-00671, НШ-1835.2003.1 и 02-07-90425

The Obstacle Detecting by TV Camera during Mobile Robot Moving

A.K.Platonov, N.S. Serbenyuk

KIAM RAS

(platonov@keldysh.ru serb@keldysh.ru)

The work deals with algorithms of video scene analysis for obstacle avoidance during mobile robot navigation. The robot's visual and motion control systems are described. Some estimations of precision of visual scene interpretation are given. The created visual system proves the possibility of the obstacle detecting by only local transformation of photometric data.

The work was supported by RFFI grants # 02-01-00671, 02-07-90425, НШ-1835.2003.1

Содержание

Введение.....	3
1. СТЗ робота "ТРИКОЛ" с ВСК	5
2. Проблема привязки сцен ВСК к параметрам движения робота	9
3. Оценка погрешности знания параметров движения робота	11
4. Проблема быстрого детектирования факта препятствия	12
Автоматное или комбинационное преобразование сцен?.....	13
Способы анализа фотометрических данных	16
6. Зрительное управление движением.....	21
7. Результаты экспериментов	22
Заключение	27
Литература	28

Введение

На текущем этапе развития робототехники главные её проблемы связаны с повышением уровня адаптации систем управления роботов к состоянию операционной обстановки. Эти проблемы имеют две сильно взаимозависимые составляющие: распознавание операционной ситуации и выбор способа двигательного поведения. Для этого наиболее информативным средством, безусловно, является видеокамера, обеспечивающая принципиальную возможность осмотра пространства в направлении движения робота. При этом, первой актуальной задачей, связанной с выполнением работ в заранее неизвестной среде, является рассматриваемая ниже задача разработки методов и средств информационного обеспечения безопасности движения.

Попытки разработать эффективную робототехническую систему машинного видения (СМВ) на базе телевизионных (ТВ) систем технического зрения (СТЗ)^{*)} с разной степенью успеха делятся уже более 30 лет (в нашей стране первые попытки связи видеокамеры с ЭВМ были выполнены в ИППИ АН СССР в 1970 г., а в США - чуть ранее). Как известно, основные трудности в создании таких систем лежат в сложности анализа получаемого большого объема сигнальной информации, подчас лишней и, как правило, зашумлённой.

Кроме этого, при всём богатстве яркостных сцен высокого пространственного и спектрального разрешения, если речь идёт о *распознавании содержания* визуальных сцен, только одного фотометрического отражения внешней среды явно недостаточно. Например, тень объекта не всегда отделима от самого объекта или его яркость не несет информации о его твёрдости. Здесь требуется использование других датчиков и/или априорных знаний, без которых невозможно найти нужный ответ. Поэтому задача *распознавания пространственных и физических свойств* объектов сцены заключается в поиске алгоритмов сложного *логического* преобразования разнородных сигнальных множеств большой кардинальности.

Однако более простой оказывается, рассматривая ниже, задача *детектирования* присутствия в наблюдаемой части пространства искомого объекта, - без детализации его свойств и пространственного расположения. Эта задача является обязательной частью первичного анализа сцены, обеспечивая быструю реакцию на изменения операционной обстановки робота и выбор необходимых в связи с таким изменением алгоритмов и действий.

Автоматическое выполнение работы, недоступной для человека в экстремальных условиях окружения или из-за других обстоятельств, означает, что робот должен в этом отношении быть "сильнее" человека. Одно из наиболее важных ограничений человеческих возможностей, давно изученное в психологии, которое в этом случае должна преодолеть система машинного виде-

^{*)} Здесь и ниже под СТЗ понимается фотометрическое преобразование яркостной картины окружения робота в цифровую форму двумерных сигналов (обычно – телевизионного стандарта). Под СМВ понимается СТЗ, дополненная алгоритмами распознавания этих двумерных цифровых сигналов в интересах решаемой задачи.

ния, это - предельная *скорость реакции* человека на предвиденные или неожиданные обстоятельства (100...200мсек).

Психологи разделяют время "простой двигательной реакции" человека - наиболее быстрого ответа простым и заранее известным движением на внезапно появляющийся, но известный сигнал и её информационную составляющую – "время простой реакции" – от момента появления сигнала до момента начала двигательного ответа [1].

Для обеспечения наиболее часто требуемой функции автоматического перемещения робота в заранее неизвестном пространстве зрительный канал должен, прежде всего, обеспечивать безопасность такого перемещения. Очевидно, что одной из главных опасностей движения любого подвижного средства является его наезд на какое-либо непреодолимое им препятствие. В случае торможения перед таким препятствием безопасность движения определяется длиной тормозного пути, в которую, как параметр, входит время "простой реакции" системы машинного зрения на неожиданное появление препятствия.

Обычно для распознавания препятствий движению предполагается использование информации стереопары, лазерного или ультразвукового дальномеров. Но крайне желательна разработка и чисто зрительного алгоритма для экстренного торможения - по сигналам *только одной видеокамеры*.

Для решения этой задачи был использован лабораторный робот "Трикол" [2-10], оснащенный "Вперёд смотрящей камерой" (ВСК) с ограниченным полем зрения. СТЗ робота обеспечивала период захвата ТВ-кадров, равный 40 мсек (25 кадров в сек). В этом режиме с разрешением ТВ-сцены (640px)×(480строк) удалось получить время простой двигательной реакции



робота 60мс, соизмеримое с таким временем у человека (рис1.).

Ниже описывается найденный способ использования телевизионной камеры для организации быстрых "простых двигательных реакций" робота при возникновении препятствий его движению.

1. СТЗ робота "ТРИКОЛ" с ВСК

Система технического зрения робота предназначена, как было сказано выше, для формирования цифрового представления визуальной сцены перед роботом. Разработанная с этой целью СТЗ [10] содержала черно-белую аналоговую ТВ-камеру, укрепленную на корпусе робота (см. рис.1 и 2), плату фрейм-граббера, и бортовую ЭВМ на базе PC VIA-C3-866МГц с операционной

Рис.1 Сравнение зрительного канала роботом системой Windows-98 и специализированным программным обеспечением (Про), описанным в [9]. Камера оснащена средствами подсветки наблюдаемых объектов.

Параметры ТВ-камеры ВСК приведены в Таблице 1, параметры фрейм-граббера – в Таблице 2. Заметим, что в аппаратуре СТЗ последовательно проходят преобразования: свет→сигналПЗС→ТВсигнал→цифровой код, влияющие на временную и пространственную привязку сцен. Параметры пятна поля зрения камеры перед роботом показаны на рис.3. Архитектура СТЗ показана на



Рис.2 "Вперед смотрящая камера"

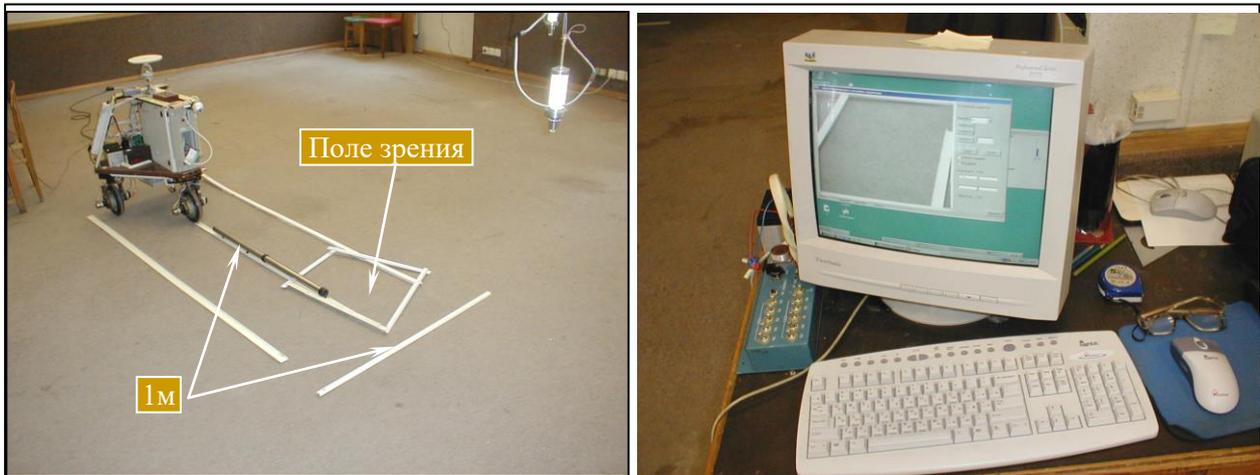


Рис. 3 Колея, её осевая линия и поле зрения ВСК перед роботом и на экране монитора

Рис.4.

Небольшое поле зрения ВСК связано с мотивами выбора объектива в момент создания робота. Этот выбор диктовался соображениями линейности пространственной картины сцен ВСК, с одной стороны, и желанием отработать алгоритмические механизмы двигательного поведения робота для осмотра окружающего пространства. Это было важно, ввиду того, что те или иные ограничения поля зрения всегда существуют, и это требует развития

процедур двигательной активности для осмотра окружения робота в процессе его движения. Таким образом, выбор такого поля зрения был небольшой "методической хитростью" для получения нужных результатов в проблеме организации управления движением колёс робота в интересах осмотра пространства для его передвижения.

Параметры вперёд смотрящей видео-камеры

Таблица 1

ТВ камера:		
Тип		SK-1004 (ПЗС)
Изготовитель		Юж. Корея
ПЗС-матрица		1/3', Samsung
Цветность		Черно-белая
Чувствительность	[лк]	0,15
Количество пикселей (формат ПЗС-матрицы: 4/3)		270040 (628x430)
Количество ТВ-элементов (формат ТВ-кадра: 4/3)		441600 (768x576)
Электронный затвор	[сек]	1/50 - 1/100000
Синхронизация (ТВ-стандарт)		Внутренняя
Частота кадров	[1/сек]	25
Видеосигнал	[В], [Ом]	1; 75
Разрешение	[ТВЛ]	430
Напряжение питания постоянного тока	[В]	12
Температурный режим	[°С]	-10 ...+50
Потребляемый ток	[мА]	125
Габариты и масса (с объективом)	[мм], [г]	32x32x30, 25
Объектив:		
Фокусное расстояние	[мм]	9,3
Светосила		1:1,2
Угол поля зрения	[Град.]	28
Система освещения сцены:		
Тип светодиодов		ИРС-1-870
Длина волны излучения	[мк]	0,87 (ИКдиапазон)
Мощность излучения	[мВт]	22
Число светодиодов		6
Суммарный угол диаграммы направленности излучения		15°

Параметры фрейм-граббера

Таблица 2

Изготовитель		Lifeview inc.
Синхронизация процесса получения строк и полукадров		От камеры
Запуск процесса захвата полукадров или кадров		При старте робота
Привязка момента конца захвата кадра		По таймеру ОС
Дискретность отсчёта момента времени захвата кадра	[нс]	280
Задержка захвата кадра (обратный ход луча)	[мс]	1,6
Длительность процесса захвата полукадра	[мс]	18,4
Задержка захвата полукадра (обратный ход луча)	[мс]	1,6
Длительность процесса захвата строки	[мкс]	52
Задержка захвата строки (обратный ход луча)	[мкс]	12

Задержка обработки прерывания камеры	[мс]	10
Связь с памятью ЭВМ		DMA
Тип шины		PCI
Рабочее разрешение	[рх]	640x480

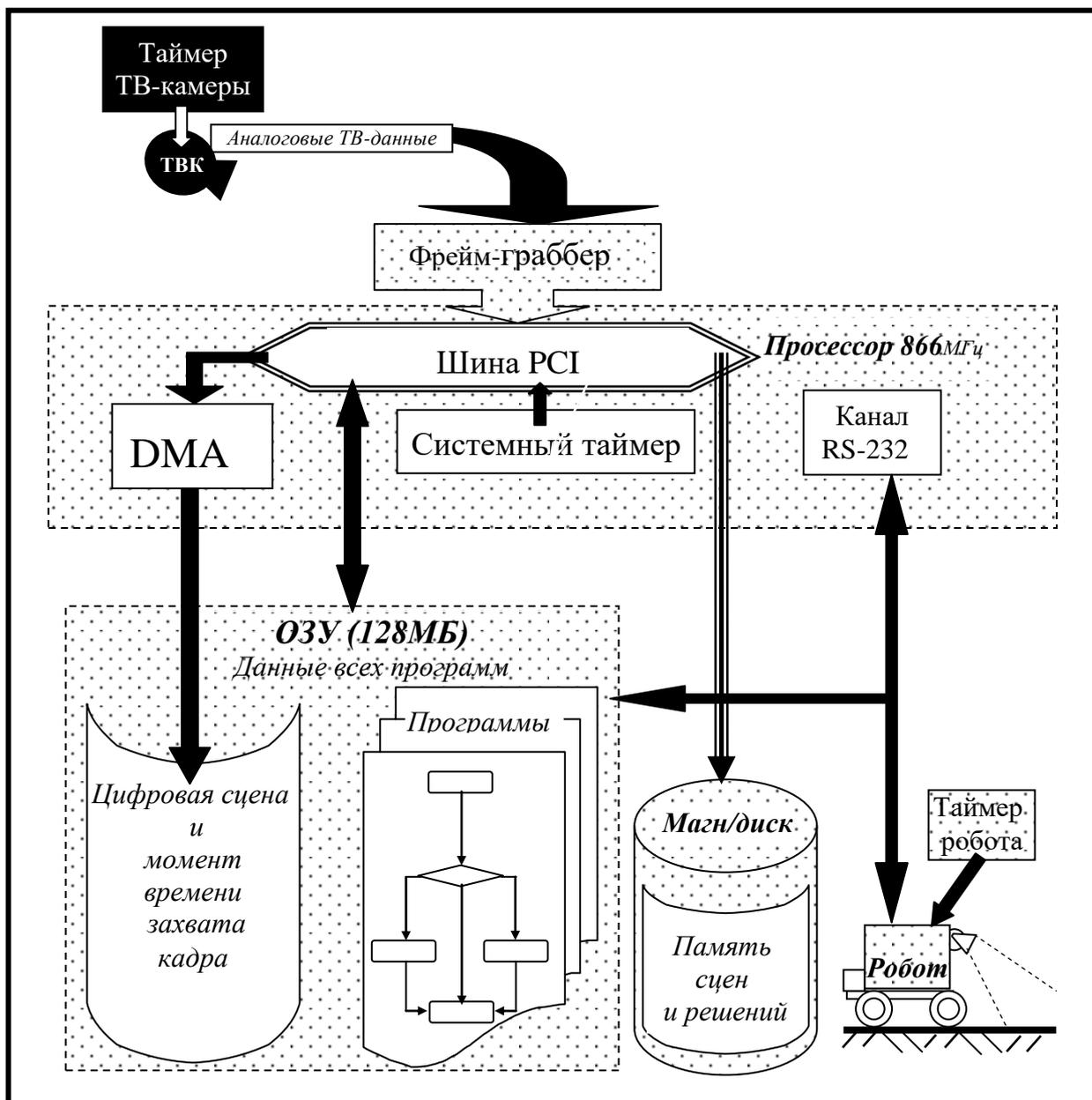


Рис.5 Архитектура аппаратных средств технического зрения робота "Трикол".

Аналоговые сигналы телевизионной камеры (ТВК) оцифровываются фрейм-граббером, и каждые 20мс полукадры размером $640 \times 240 \text{px} \times 1 \text{Байт}$ измеренной яркости по каналу прямого доступа поступают в оперативную память робота. Захват сцены происходит в среднем за 30мс до этого момента (с учётом задержки в 10мс - времени обработки прерываний драйвера фрейм-граббера). Этот процесс оформлен в виде отдельного потока ОС Windows с привязкой к её системному времени. Он протекает квазипараллельно относительно остальных процессов системы управления робота. Моменты захвата кадров изображения операционного пространства робота известны в числе дискрет системного таймера (с шагом 280 нсек) от момента запуска драйвера фрейм-граббера (с учётом задержки до 50мс - времени ожидания синхроимпульсов первого кадра).

2. Проблема привязки сцен ВСК к параметрам движения робота

Принципиальным свойством СТЗ роботов является, как правило, наличие нескольких независимых источников текущего времени – таймеров ("часов") у робота, телевизионной камеры и управляющей ЭВМ. Какой-нибудь один из этих приборных таймеров выбирается в качестве источника "абсолютного времени" робота. Тогда сигналы остальных таймеров должны интерпретироваться через параметры совпадения времени их отсчётов с абсолютным временем робота. Сложности этого процесса связаны с неизвестными ошибками частот тиков таймеров, вызывающими на больших интервалах времени расхождения их привязок к абсолютному времени.

Абсолютное время всех процессов робота "Трикол" формируется по сигналам ("тикам") таймера нижнего уровня системы управления [5], поступающим в управляющую ЭВМ верхнего уровня по последовательному каналу. В этом абсолютном времени номера тика с периодом 32 мсек в систему управления поступают данные о позиции и ориентации робота в стартовой абсолютной системе координат.

Моменты наблюдения телевизионных сцен окружения робота в процессе движения робота должны быть привязаны к его позиции в момент наблюдения. Для этого требуется согласовывать приборные "часы" ТВК и системы счисления пути. Это происходит следующим образом:

Поступление данных от ТВК синхронизируется её таймером, но привязка получаемых данных к абсолютному времени системы выполняется (после некоторой задержки в системе прерывания) в программном драйвере фрейм-граббера в моменты обработки аппаратного прерывания *по сигналу конца поступления очередного кадра* ТВК (но до начала обратного хода луча). Время захвата сцены фиксируется в памяти драйвера по меткам ("дискретам") системного таймера ОС Windows, отсчитываемым от начального момента запуска работы драйвера при старте системы управления. Процесс захвата сцен ВСК по сигналам ТВК организован в отдельном потоке процессов ОС Windows и протекает в режиме реального времени независимо от времени исполнения других программ. При этом, в зависимости от статически заданного требуемого пространственного разрешения сцены фрейм-граббер передает в фиксированную область памяти робота или полные кадры черезстрочной развёртки ТВК с высоким разрешением (но с временным шагом 40 мсек) или отдельные полукадры более низкого разрешения (с временным шагом 20 мсек и привязкой времени только второго полукадра каждого кадра).

Программа анализа сцен ВСК и привязки их к параметрам движения робота выполняется на верхнем уровне системы управления ("ВУСУ") движением робота в её отдельном потоке и вне реального времени поступления видеоданных (см. рис.6). Это происходит в некий момент - либо в момент её очередной готовности, либо раньше - в силу "предаварийной" необходимости обработки зрительных данных в процессе движения среди препятствий. При её вызове инициируется обращение к памяти робота и считывается послед-

ний захваченный полукадр или кадр телевизионных данных с его системным временем (числом дискрет системного таймера, прошедших от включения робота до момента конца предыдущего кадра). В программе выполняется анализ того, что именно лежит в памяти (первый или второй полукадр кадра) и соответственно вычисляется момент времени захвата сцены в дискретах ("дисках") системного времени ОС Windows. Перевод этого момента в единицы ("тики") прошлого абсолютного времени робота выполняется по формуле:

$$T_{\text{ТИК}} = (T_{\text{К}} - T_{\text{К0}}) \times d_{\text{СИСТТИК}} + \Delta T_{\text{ТИК}} + T_{\text{ТИК0}}, \quad \text{где:}$$

$T_{\text{ТИК}}$ - абсолютное время захвата сцены в тиках;

$T_{\text{К}}$ - системное время конца предыдущего кадра драйвера [дис.];

$T_{\text{К0}}$ - системное время запуска драйвера фрейм-граббера [дис.];

$d_{\text{СИСТТИК}}$ - длительность дискрет системного времени в тиках, равная:

$$d_{\text{СИСТ}}/d_{\text{ТИК}} = 280 \text{нсек}/16 \text{мсек} = 17,5 \cdot 10^{-6} [\text{ТИК}] = 2,8 \cdot 10^{-4} [\text{мсек}];$$

$\Delta T_{\text{ТИК}}$ - одна из трёх констант привязки полукадров к тикам часов робота, равная половине промежутка времени до поступления видеоданных на матрицу ПЗС (без учета неизвестного времени накопления и времен обратного хода луча, соответствующего 1,6 мм пути в худшем случае):

- $-40 \text{мсек}/16 \text{мсек} = -2,5$ тика при захвате 1-го полукадра;
- $-20 \text{мсек}/16 \text{мсек} = -1,2$ тика при захвате 2-го полукадра;
- $-30 \text{мсек}/16 \text{мсек} = -1,9$ тика при захвате целого кадра.

$T_{\text{ТИК0}}$ - значение момента запуска фрейм-граббера ВСК [тик] с учётом неопределенности (50 мсек) момента опроса первого кадра.

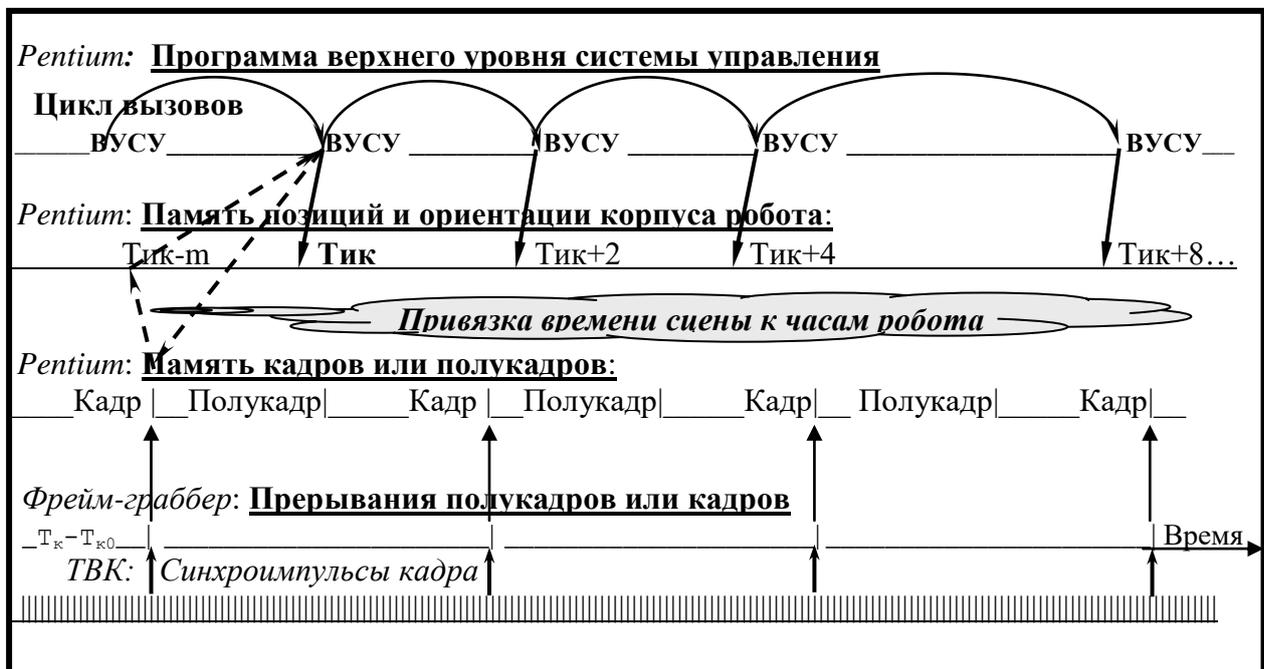


Рис.6 Временная диаграмма процессов получения зрительных данных. Обработка видеоданных ВСК выполняется на верхнем уровне системы управления робота в цикле его вызовов в фоновом режиме остальных процессов. Получаемые видеоданные относятся к предыдущим моментам движения робота. Параметры движения определяются интерполяцией данных о состоянии робота на момент захвата кадра.

3. Оценка погрешности знания параметров движения робота

Оценим максимальную погрешность δT привязки момента захвата сцены к позиции робота в стартовой системе координат. Линейная вариация $T_{\text{тик}}$, выраженная в миллисекундах имеет вид:

$$\delta T_{\text{тик}} = \delta(T_{\text{к}} - T_{\text{к0}}) \times d_{\text{сис}} + (T_{\text{к}} - T_{\text{к0}}) \times \delta d_{\text{систик}} + \delta \Delta T_{\text{тик}} + \delta T_{\text{тик0}} + \delta T_{\text{поз}} \text{ [мсек]}.$$

Здесь:

$$\delta(T_{\text{к}} - T_{\text{к0}}) \times d_{\text{сис}} = 1 \text{ диск} \times d_{\text{сис}} = 2,8 \cdot 10^{-4} \text{ [мсек]} - \text{максимум случайной суммарной погрешности двух дискрет.}$$

$$\delta d_{\text{систик}} = d_{\text{систик}} (\delta d_{\text{сис}} / d_{\text{сис}} + \delta d_{\text{тик}} / d_{\text{тик}}) = 2,8 \cdot 10^{-4} \times 2 \cdot 10^{-5} = 5,6 \cdot 10^{-9} \text{ [мсек]} - \text{погрешность несовпадения частот таймеров нижнего и верхнего уровней системы управления (среднеквадратичная оценка равна } 3,96 \cdot 10^{-9} \text{ [мсек])};$$

$$\delta \Delta T_{\text{тик}} = 20 \text{ мсек} - \text{случайная погрешность привязки полукадров к тикам часов робота};$$

$$\delta T_{\text{тик0}} = 25 \text{ мсек} - \text{случайная величина систематической погрешности нуля отсчета времени фрейм-граббера};$$

$$\delta T_{\text{поз}} = 8 \text{ мсек} - \text{случайная погрешность привязки к абсолютному времени момента знания позиции робота.}$$

Маловероятная верхняя суммарная оценка ошибок времени привязки сцены ВСК для длительности $(T_{\text{к}} - T_{\text{к0}}) = 10^{11}$ диск, соответствующей времени непрерывной езды на запасе хода робота (равного ~ 10 часов и определяемого емкостью его аккумуляторов) равна 610 мсек. Более достоверная среднеквадратичная оценка дает значение ~ 400 мсек (40 см пути робота на максимальной скорости). При этом хорошо видно, что главной погрешностью привязки видеосцен на большом времени непрерывного движения является случайная относительная погрешность кварцевых стабилизаторов частот таймеров (принятая равной 10^{-5}). За 10 часов езды таймеры в этих предположениях разбегаются на 396 мс, в то время как все остальные погрешности вместе взятые квадратично добавляют чуть больше 1 мсек.

С другой стороны, при более частом случае малого времени езды, например порядка 10 мин ($214 \cdot 10^4$ диск), таймеры разбегаются всего на 8 мсек, и остальные погрешности становятся главными. Среднеквадратичная оценка их суммарной погрешности привязки кадра равна 30 мсек (3 см пути робота).

Для полученного момента времени захвата анализируемой сцены необходимо определить, какие были координаты и ориентация робота в прошлом в момент захвата. Следует учесть, что координаты и ориентация робота определяются в независимом от процесса захвата кадров процессе управления движением колёс робота [5]. Последний, как уже говорилось, протекает с шагом в 32 мсек по "тикам" таймера нижнего уровня системы управления, как правило, не совпадающими с моментами захвата сцен (см. рис.6). Это обстоятельство требует запоминания прошлых параметров движения робота на длине достаточной для построения интерполяционного процесса определения параметров движения робота в момент захвата анализируемой сцены.

Здесь возникает другая ошибка, влияющая на точность интерпретации сцены ВСК: погрешность счисления пути на нижнем уровне системы управления. Эта погрешность связана как с погрешностями тарировки и юстировки робота (см. [6]), так и с неизвестным проскальзыванием колес в процессе движения. Эксперименты показали, что при движении робота по паркетному полу ошибка счисления пути на 5 м составляет 1 см по пути и малые доли градуса по углу ориентации корпуса. При этом ошибка ориентации практически объясняется углом смещения пути. Полагая эти погрешности случайными и линейно-зависимыми от пройденного пути, получим, что на запасе хода робота 1км суммарная максимальная погрешность знания позиции робота составит около 2м. Отсюда следует, что для обеспечения требуемой точности движения (10 см) время движения на максимальной скорости без дополнительных навигационных привязок ограничено всего 1мин (50м пути робота).

4. Проблема быстрого детектирования факта препятствия

При движении робота первичной (и одной из важнейших) функций зрительного анализа является обнаружение препятствий движению. К этой функции предъявляются требования наибольшей надежности и наибольшего быстродействия. Надежность обнаружения должна заведомо исключать ошибки первого рода (необнаружение возникшего препятствия движению) и максимально сокращать ошибки второго рода (ложные сигналы наличия препятствия движению).

Быстродействие процесса обнаружения препятствия на пути движения робота, как было упомянуто, определяется соотношением дистанций обзора и торможения. Например, при скорости $V=90$ км/час (25 м/сек) автомобиля его тормозной путь с ускорением $3g$ равен 10 м. и сам процесс торможения длится 0,85сек. При этом время реакции водителя 0,2 сек (время его "простой двигательной реакции") увеличивает тормозной путь на 5 м.

Если перевести эти данные в относительные линейные размеры по параметру длины L корпуса автомобиля ($L=5$ м $\Rightarrow V=5L/\text{сек}$, тормозной путь $S=2L$ и $V^2/(2S)=30L/\text{сек}^2$), то в пересчете на характерный размер робота "Трикол" (70см) его тормозной путь может быть порядка 1,5м (при скорости движения 3,5 м/сек). В действительности скорость робота "Трикол" ниже, - порядка 1 м/сек. На этой скорости измеренный тормозной путь (от момента подачи команды торможения) равен 15-20см.

Сделаем оценку максимального допустимого времени анализа кадра.

Аппаратное время получения фотометрического изображения сцены перед роботом (256 градаций яркости для 640x480 точек изображения) в соответствии с телевизионным стандартом требует 0,04сек (4см пути при скорости движения 1м/сек). Измеренный тормозной путь робота в худшем случае составляет 20см, а дальность осмотра (см. рис.3) – около 1м. Поэтому время распознавания наличия препятствия движению для реализации экстренного торможения с 10см запасом не должно превышать 0,7сек (70 см пути).

Автоматное или комбинационное преобразование сцен?

Функция информационного канала в цепи обратной связи на верхнем уровне процесса управления движением робота заключается больше чем в простой перекодировке первичных сигналов^{*)} датчиков информации в цифровые коды, "понятные" цифровым модулям системы управления. Применительно к зрительной обратной связи главная задача возникает на следующем этапе: по наблюдаемому изменению цифровых значений сенсорных кодов сцены определить характер изменения операционной ситуации. Эта задача может решаться как в простейшей постановке *детектирования* возникновения ожидаемой внешней ситуации, так и в постановке более сложного *определения параметров* возникшей ситуации.

Заметим, что в психологии человека четко выделены два уровня решения мозгом человека этой информационной задачи: *ощущения* (сенсорный уровень) и *восприятие* (уровень сознания) [1,11]. Осознанное восприятие окружающего мира требует сложной обработки различных источников информации с использованием глубин памяти. Процессы ощущений значительно проще, здесь человек воспринимает сигналы внешнего мира только на уровне простейших чувств, без осознания смысла и значения этих сигналов. Иными словами – в терминах теории автоматического управления, если восприятие, как правило, - результат целевого интегрального процесса в *автоматной* постановке, то ощущения могут быть ограничены локальными сигналами единственных сенсоров и их простым *комбинационным преобразованием*^{**)}.

Если в пределах аналогии приложить эти психологические понятия к процессам обработки телевизионных сцен в робототехнике, то аналог ощущения - процесс быстрого детектирования одной из двух ситуаций: *совпадения* и *несовпадения* сенсорной картины с ожидаемой. Этот процесс может быть реализован как в комбинационной, так и в автоматной постановке. В первом случае требуется построить детектор отличия зрительной картины от стандартной *статической* (т.е. – задаваемой до старта) ситуации. Во втором – построить детектор отличия наблюдаемой зрительной картины от *динамической* картины в памяти робота, формируемой *в процессе* движения алгоритмом предсказания её ожидаемого изменения (алгоритмом *акцептора действия* - см. [8]). В робототехнике, где важную роль играют как ограничения времени решения зрительных задач, так и адаптивность к "тонким" изменениям в операционной обстановке, скорее всего, придётся реализовывать оба аналога ощущения - в комбинационном и автоматном подходах.

Аналогом восприятия, в свою очередь, могут быть *контекстные* (т.е. - с использованием прошлых сцен) процессы выделения содержательных параметров сцены и отношений между ними. Попробуем проанализировать такой подход к проблеме более полного распознавания видеоданных.

*) Напомним, что "сигнал" – это физический процесс, обеспечивающий передачу информации

***) Выходной сигнал комбинационного преобразователя зависит только от входного сигнала.

Выходной сигнал автоматного преобразователя зависит от входного сигнала и от состояния автомата.

Искомыми параметрами на сцене в этом случае не обязательно должны быть параметры пространственного расположения пиксельных яркостей на сцене. Ими могут быть другие – внутренние параметры фотометрической картины, например, - параметры фреймовых^{*)} структур, определяющих не только более сложное качественное ("знаковое") описание типизируемых фреймами ситуаций, но и количественную их параметризацию. Пример такого подхода к сцене на рис.7 был описан в [10]. Простейшие случаи такого "семиотико-параметрического" описания ситуации можно квалифицировать

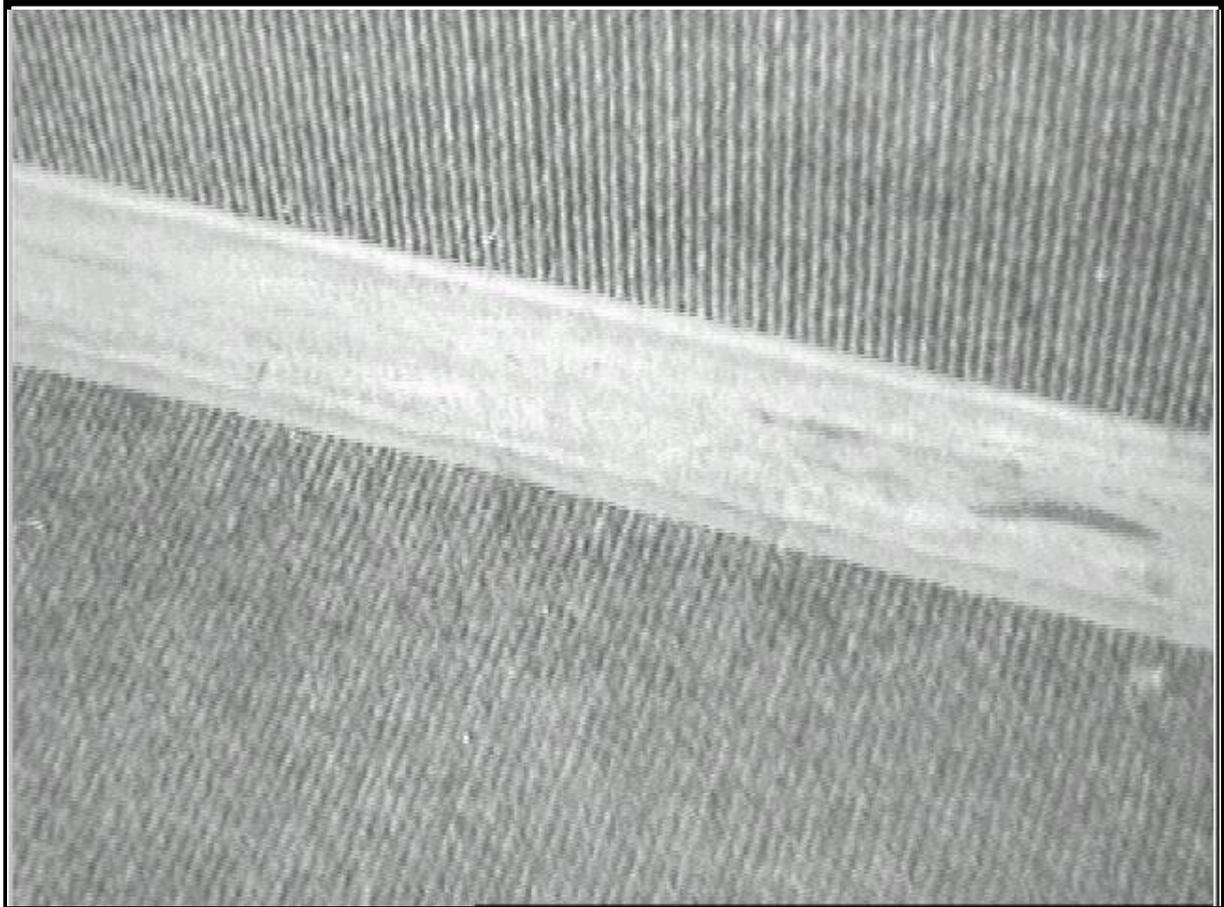


Рис.7 Сцена перед роботом (часть пола, плинтус, часть стены).

Это изображение, при условии распознавания плинтуса, позволяет вычислить расстояние и угол ориентации робота относительно стен[10]. С другой стороны, распознавание ориентации периодической структуры ковра пола позволяет вычислить ориентацию робота относительно стен помещения вне видимости самих стен.

как выделение неких простых ситуационных "образов" в операционной среде. В самом предельном случае увеличения сложности используемых фреймовых структур из памяти робота можно говорить о "понятийной" обработке зрительной сцены. Точную границу такого деления, по-видимому, установить невозможно, да и хороших примеров реализации образного или понятийного зрительного восприятия пока ещё не найдено.

^{*)} Под фреймом в данном случае понимается описание типичной ситуации в виде фиксированного перечня и структуры данных, записываемой по логическим условиям процесса разбора сцены [13].

Но можно посмотреть на эти аналогии и иначе, так сказать, - с технологической стороны. Очевидно, что аналогией ощущения является обработка лишь "чувственной" *фотометрической* информации - без анализа её *геометрического расположения* на плоскости сцены. С другой стороны, процессу восприятия сцены более соответствует именно анализ характеристик *пространственного* расположения фотометрических "пятен" в плоскости сцены, их формы и связности с безусловным привлечением памяти прошлых сцен и прогноза движения робота.

Почему для восприятия важна именно геометрия зрительного сенсорного пространства^{*)} сцены? Потому, что его геометрические характеристики основаны на искомым содержательных свойствах воспринимаемого операционного пространства, и *только они* отображают содержательную информацию о последнем. В самом множестве пиксельных сигналов такой информации нет. Поэтому содержательная обработка сцен по необходимости должна работать с совокупными элементами пиксельного множества с тем, чтобы найти отношения между его элементами, адекватные отношениям элементов операционного пространства.

Но принципиальным моментом "восприятия" сцены (её содержательного анализа) является то, что эти искомые отношения на сигнальном множестве можно получить лишь с привлечением семантического анализа операционной ситуации и, возможно, лишь в её текущем контексте. Действительно, при движении робота многие изменения в операционном пространстве робота, так или иначе, отражаются в пиксельном множестве, но отбираться должны только те наблюдаемые изменения, которые важны с точки зрения решаемой задачи.

Т.е. элементы сигнального множества связаны в отношения именно ***условиями и обстоятельствами решаемой задачи***. Отсюда следует, что в алгоритме анализа текущего содержания сцены по необходимости должно присутствовать преобразование в сигнальное множество содержательного описания текущих обстоятельств решаемой задачи в операционном пространстве робота. Это преобразование является обратным искомому преобразованию фотометрической картины в обстоятельства состояния операционной обстановки.

Можно утверждать, что разработка методов преобразования в сигнальное множество содержания операционной обстановки робота с точки зрения решаемой задачи (допускающего, к тому же, невырожденное искомое обратное преобразование сцены) является главной проблемой развития средств информационного обеспечения в робототехнике. Одним из эффективных методов здесь является *использование моделей и законов механики* для описания и/или прогнозирования яркостной картины операционного окружения робота (удачные примеры использования механики в процедурах обработки телевизионных сцен можно найти в [14]).

^{*)} Напомним, что "пространство" – это множество с отношениями между его элементами. Эти отношения определяют "геометрию" пространства.

Искомая пара взаимно-обратных избирательно-контекстных отображений зависящих от предыстории, текущего состояния и, возможно, прогноза изменений операционной обстановки безусловно является *автоматным преобразованием* операционного и пиксельного пространств. И оно, безусловно, требует как затрат памяти для хранения фазовых состояний робота, так и процессорного времени для прогнозирования будущих фазовых состояний (и для отображения этой информации в пиксельный формат кадра) и для неоднократных просмотров кадра при сравнительном анализе всех этих элементов. Это усложняет его использование при больших скоростях движения робота.

С другой стороны, более бедное *комбинационное преобразование* лишь фотометрических обстоятельств операционной обстановки робота целиком определяется только текущим состоянием параметров операционной обстановки робота, отражаемым телевизионной камерой. Оно заведомо не даёт ответа на тонкие вопросы анализа зрительной картины. Но, как это будет продемонстрировано ниже, оно всё же позволяет *быстро* решить важную задачу грубого первоначального детектирования факта возникновения или отсутствия нужной или опасной ситуации. И тогда следующим шагом является активное замедление движения робота с одновременным включением мощных автоматных процедур содержательного распознавания операционной обстановки и принятия не стандартных, а *аварийных* решений о направлении и/или режиме движения.

Это означает, что не менее важной проблемой развития средств информационного обеспечения в робототехнике является поиск "грубых" комбинационных методов обработки чисто фотометрических данных в интересах быстрого первоначального детектирования отклонения операционной обстановки от стандартной ситуации. Рассмотрим это подробнее

Способы анализа фотометрических данных

Чисто фотометрические данные телевизионного сигнала представляют собой пространственно неструктурированную информацию об окружении робота в пределах поля зрения объектива телевизионного приемника (см. рис.7 и 8). В интересах сокращения времени распознавания внешней ситуации, целесообразно проанализировать, что полезного можно извлечь из этой первичной информации и какие затраты времени процессора для этого нужны?

Функциями только яркости сцены являются величины её максимума и минимума, величина суммарного светового потока, падающего на плоскость изображения (суммарная яркость), распределение отдельных значений яркостей по диапазону её измерения (гистограмма яркостей) и различные параметры этого распределения: её мода (максимальное значение числа одинаковых (или примерно одинаковых) значений яркости, шумовая составляющая.(изрезанность гистограммы) и моменты гистограммы.

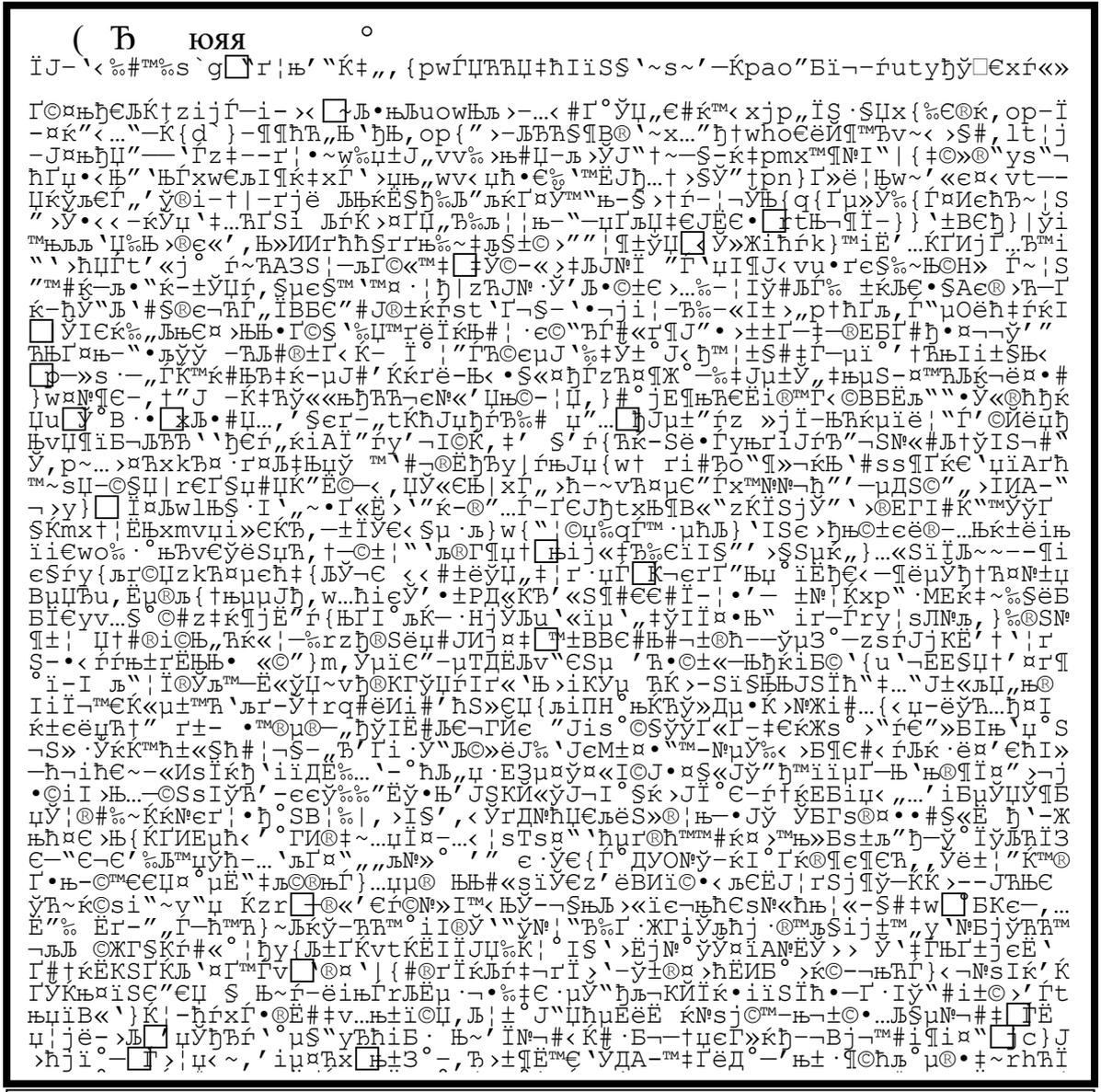


Рис.8 Фрагмент (10%) цифрового изображения поля зрения общим размером 640x480px (левого верхнего угла предыдущего рис.7). Здесь показано только 64 левые точки каждой строки из 48 начальных строк кадра. Каждое пиксельное значение изображено символом, код которого соответствует "измеренной яркости" в диапазоне 0...255 относительных градаций (чернота символа не соответствует яркости пиксела).

Физический диапазон измерений и форма и крутизна относительной шкалы градаций формируется аппаратно путём выбором перед стартом работа параметров кривой коэффициента усиления входного сигнала ("яркость" и "контраст").

Короткая верхняя строка – заголовок файла кадра. Вторая однородная строка из скрытых символов () – данные для настройки программы обработки изображения.

Вычисления первых названных функций, включая гистограмму, выполняются за один перебор всех точек сцены с числом операций не более двух в каждой точке для каждой функции. Для вычисления функций гистограммы (изрезанности и моментов) требуется дополнительный перебор точек гистограммы с не более, чем с пятью командами для каждой яркости..

Из опыта работы с телевизионными изображениями хорошо известно, что абсолютные значения яркости являются трудно распознаваемым параметром, ввиду их зависимости не столько от содержания сцены, сколько от переменных условий её освещения. Поэтому максимальный и/или минимальный код яркости, как и сумма пиксельных кодов сами по себе не достаточны для оценки внешней ситуации. Это тем более очевидно, что в большинстве телевизионных датчиков используется аппаратно встроенная функция "АРУ" - автоматического регулирования усиления входного сигнала, делающая обратное преобразование практически вырожденным. Однако от определения максимальной яркости, как и от вычисления суммарной яркости сцены, отказываться, всё же, не следует ввиду их полезности для нормирования относительной яркости отдельных пикселей и отдельных областей сцены.

Итак, наиболее быстро получаемые параметры сцены - номер пиксела с максимальной яркостью и суммарная яркость всех пикселей сцены непосредственно не целесообразно использовать для определения есть или нет на пути движения робота препятствие его движению.

Рассмотрим возможности использования для этой цели информации, содержащейся в гистограмме яркостей изображения. Очевидно, что изменения внешнего освещения сцены должны, прежде всего, влиять на её смещение вдоль оси яркостей пикселей. Действительно, ввиду достаточно узкого поля зрения ВСК, в большинстве случаев вся сцена равномерно освещена внешними источниками света в достаточно узком диапазоне соответствующих углов падения и отражения света. Поэтому изменение силы внешнего света примерно в равной мере изменяет количество отражаемого света, проходящего через объектив камеры, что и приводит к простому сдвигу гистограммы к более светлым или более тёмным значениям оси яркостей. Если какая-то часть сцены не освещена ни прямым, ни переотражённым светом, то на гистограмме должны присутствовать некие "пики тёмного и светлого", соответствующие числу пикселей теневой и освещенной частей сцены. При изменении силы внешнего освещения (без изменения обстоятельств возникновения тени) можно ожидать смещения вдоль оси яркостей только светлого пика. Заметим, что сам тёмный пик никакой информации не несёт (кроме той, что в поле зрения есть область, где ничего не видно).

Другая ситуация возникает, когда условия освещения сцены меняются из-за возникновения в поле зрения предметов внешнего мира или их теней, а не от общего потемнения или просветления окружающего робота пространства. Каждый новый предмет в поле зрения порождает новые области пикселей с новым уровнем яркости. Поэтому в таком случае меняется также и *соотно-*

шение количества светлых и тёмных точек сцены, что приводит не только к сдвигу, но и к изменению формы гистограммы.

Из этих соображений вытекает, что при анализе гистограммы на предмет наличия в поле зрения препятствий движению, не следует обращать внимание ни на положение гистограммы вдоль оси яркостей, ни на положение границ гистограммы (значения самых ярких и самых темных яркостей в сцене), ни на наличие самого левого пика темного (ниже порога яркости теней), если только он вдруг не возник именно на этой гистограмме. Главным предметом анализа гистограммы тогда должны быть *изменения* её формы и соотношения количеств "светлого" и "серого" в плоскости изображения.

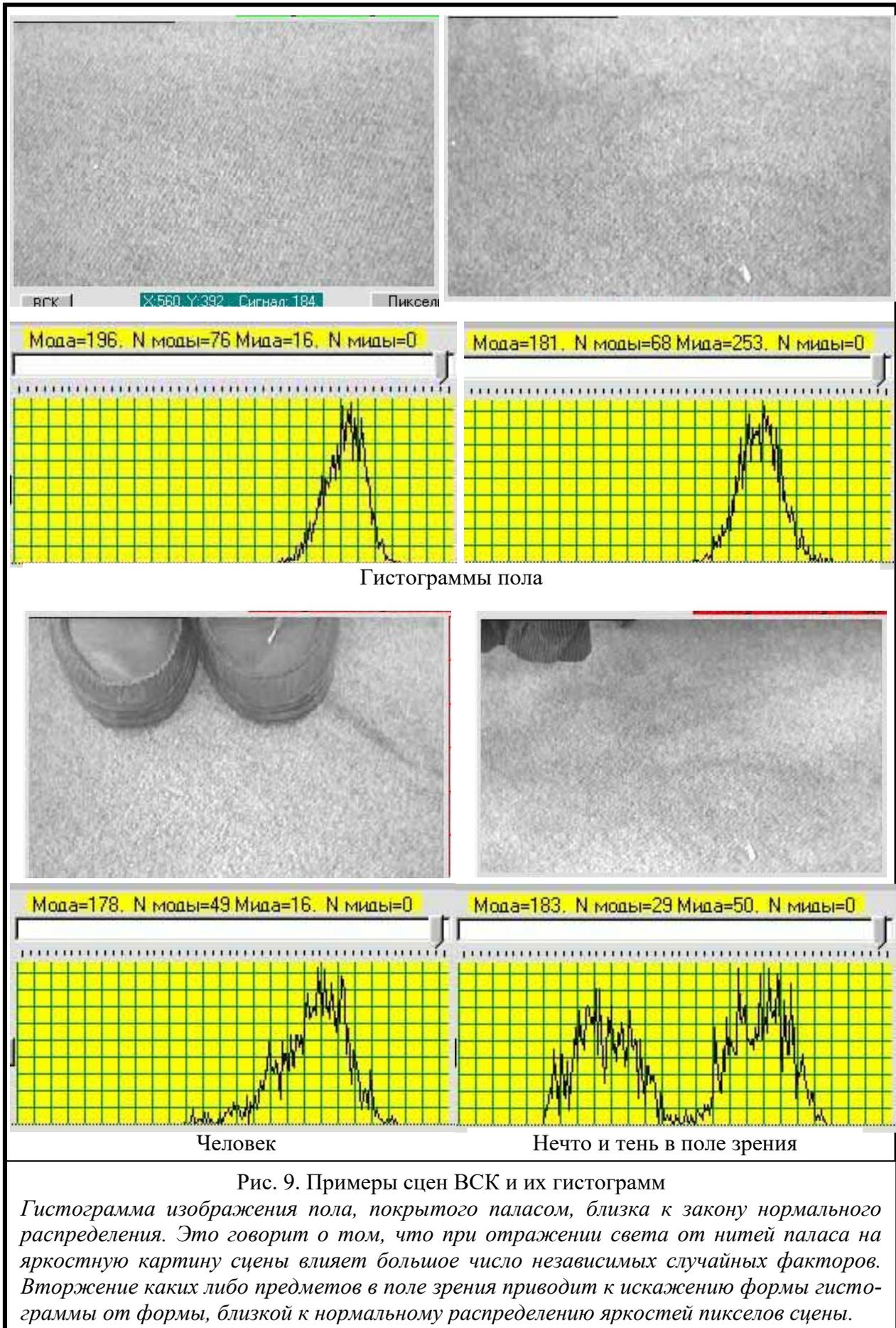
На рис.9 и 10 приводятся примеры сцен без препятствий движению робота и сцен с появляющимися в поле зрения объектами, яркостная картина которых отличается от картины свободного пола перед роботом. Следует обратить внимание на значения моды. Заметим, что словом "мида" (как оппозит моде гистограммы) обозначена максимальная яркость с минимальным количеством пикселей, больше уровня шума. В разработанном алгоритме "быстрой гистограммы" из гистограммы вычитается измеренный шум, поэтому $N_{\text{МИДЫ}}$ близко к 0. Параметр $N_{\text{МОДЫ}}$ в алгоритме играет нормирующую роль.

Хорошо видно, что распределение яркостей пикселей на сценах пола, покрытых ковровым паласом, близко к нормальному распределению вероятностей. Всякое отличие от такой формы гистограммы должно настораживать, но оно еще не является заведомо верным признаком наличия препятствия на пути движения. Изменение яркостной картины может быть вызвано бликом света или пятнами на полу. Примером этому может служить правое изображение свободного пола, в верхней части которого хорошо видно светлое пятно. в результате этого гистограмма правой сцены пола отличается от гистограммы сцены слева. Хотя это отличие невелико, но оно оказалось достаточным, чтобы система "почувствовала" здесь возможное препятствие и притормозила.

Возможна и обратная ситуация, когда препятствие существует, но его контрастность и размеры практически не влияют на форму гистограммы. Пример этого показан на рис.10, где в поле зрения видна почти сливающаяся с полом тонкая ножка стула. Эта сцена при отладке алгоритма не вызвала сигнала тревоги.

Борьба с бликами и тенями составляет одну из главных задач быстрой обработки телевизионных сцен. Исключение влияния бликов и слабых контрастов на первом этапе возможно путём построения "яркостного образа" сцены в виде фреймового описания особенностей её наблюдаемой гистограммной функции.

Для следующего, более надежного распознавания факта появления препятствия необходимо использовать дополнительные признаки, доступные уже только из анализа формы гистограммы *в контексте её изменения* в процессе движения. Именно движение позволяет в ряде случаев по мере приближения к большому препятствию надёжно его почувствовать.



6. Зрительное управление движением

Ограничения поля зрения ВСК требуют активного управления осмотром окружения робота в процессе его движения. Ввиду фиксированного крепления ВСК на корпусе робота осмотр местности по пути движения возможен лишь через изменение управления движением колёс робота. Для этого потребовалось разработать механизмы автоматического управления ориентацией корпуса робота в интересах требуемого направления осмотра - в зависимости от результата распознавания получаемых сцен ВСК.

Как видно на рис.10 кинематика робота позволяет организовать поступательно-вращательное движение корпуса. Однако движение колёс при изменении направления движения или ориентации корпуса робота заметно ограничено условиями согласования их рулевых углов и взаимных скоростей вращения. При управлении движением колёс необходимо учитывать параметры быстродействия и инерционных характеристик приводных механизмов робота. Кроме этого число оборотов рулевых механизмов колёс ограничено длиной силовых и сигнальных проводов их двигателей с датчиками.

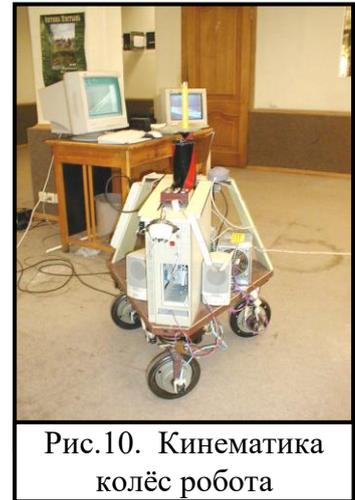


Рис.10. Кинематика колёс робота

Поэтому для формирования требуемого движения колёс потребовалась разработка тонких алгоритмов управления переходными процессами в системе построения движения (на верхнем уровне системы управления). Разработанные алгоритмы построения движения колёс реализованы в виде ряда фиксированных режимов движения (поворот на месте, поворот в движении, движение вбок или назад, торможение или разгон, поступательное или "автомобильное" движение к цели), которые обеспечивают как учёт всех ограничений, так и минимальные погрешности счисления пути в навигационной системе робота. Эти режимы исполняются параллельно с работой ВСК в четырёхпроцессорной системе управления шестью степенями подвижности приводов колёс (системой исполнения требуемого движения) [7].

Дополнительной проблемой, хорошо известной в робототехнике, было необходимость согласования (калибровки) систем координат зрительных осей ВСК, корпуса робота и приводных систем колёс. Решение этой задачи описано в [6].

В результате, - в зависимости от обстоятельств движения и сигналов ВСК – формируется необходимое движение робота, после чего для его реализации выбирается один из возможных режимов движения корпуса робота и его колёс. Выбор и последовательность формирования двигательной активности выполняется каждые 0,12сек на самом верхнем уровне управления

(системой управления поведением робота) и последовательно исполняются упомянутыми системами построения движения и исполнения движения.

7. Результаты экспериментов

Эксперименты выполнялись в помещении (рис.11), размеры которого, казалось бы, были достаточны для организации сложного движения робота с препятствиями на его пути. Тем не менее, в процессе отладки алгоритмов *обработки сцены, фреймовой структуры* выходных данных, подбора её *пороговых значений* и, наконец, - отладки логических условий *смены режимов движения*, робот несколько раз бился об стены помещения и неоднократно наезжал на предметы на его пути. Поэтому одним из выводов работы является необходимость в достаточно большом пространстве для проведения экспериментов с движущимся роботом. Без этого разработка таких систем оказывается практически незавершенной.



Рис.11. Полигон для отладки алгоритмов

Дополнительно к этому, для поиска методов решения задач, отмеченных выше курсивом, и необходимость отладки соответствующих алгоритмов и программ потребовали разработки специального инструментального комплекса (рис.12.). Были созданы программные процедуры для визуализации



Рис.12. Инструментальный комплекс (что есть на пути, что видно, что и почему решено?)

сцен, протоколирования экспериментов и эмуляции протоколов. Эти инструментальные средства обеспечивают возможность как реальной (in-line), так и последующей камеральной (of-line) обработки картины движения с получением объяснений о характере и причинах движения по ветвям сложного алгоритмического дерева. Можно утверждать, что без таких средств (их формирование требует значительных усилий и времени) создание надёжных зрительных систем управления автоматическим движением машин практически невозможно.

В экспериментах в качестве препятствий движению использовались в основном различные стулья.



Рис.13. Типы препятствий

Это диктовалось рядом соображений и прежде всего тем обстоятельством, что стулья являются одним из характерных и наиболее трудных для распознавания видов препятствий, не описываемых картой помещения из памяти робота. Другие перемещаемые, но объёмные, объекты (ящики, люди и т.п.) с резким отличием их фотометрического образа надёжно идентифицировались разработанным алгоритмом распознавания аварийной ситуации. Однако, как показали эксперименты, ножки стульев не всегда замечались, как препятствия движению, ввиду их малых размеров и часто - их малого контраста по отношению к сцене пола, (см. рис.14).

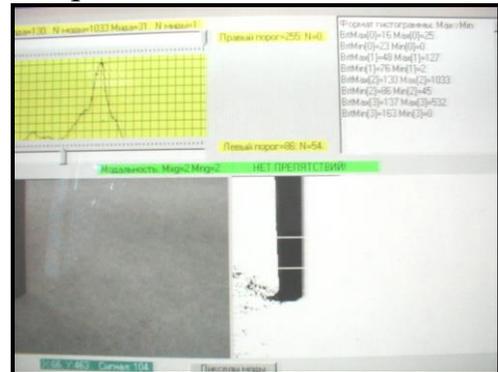


Рис.14. Пример малококонтрастного препятствия движению робота.
Ответ системы: "НЕТ ПРЕПЯТСТВИЙ"

Выше было упомянуто, что именно движение робота позволяет облегчить проблему зрительной обратной связи в контуре управления его движением. В случае наезда на стул малая площадь и контрастность элементов его изображения исчезала по мере приближения к нему. Практически во всех выполненных экспериментах подобранные уровни фреймовых параметров гистограммного образа надёжно обеспечивали его идентификацию, как препятствия движению.

Основные трудности, поэтому были связаны с организацией способа объезда препятствия с учётом размеров поля зрения ВСК. На рис.15 показана типичная ситуация: робот едет на третий стул, который на достаточно большом расстоянии от него не был зафиксирован в процессе поворота на месте для осмотра пути объезда зафиксированного препятствия в виде двух дальних стульев. Эта ситуация потребовала обогащения процедуры осмотра окружения робота маневрами согласованного сближения с препятствием при прямом повороте от него и последующего выполнения обратного поворота для осмотра текущей колеи в процессе движения.



Рис.15. Робот отвернулся от стула!

В процессе проведения экспериментов оказалось, что необходима более глубокая разработка алгоритмов управления двигательным поведением робота при объезде препятствия. Разработанная с этой целью процедура, запускаемая при идентификации наличия препятствия на пути движения, выполняет последовательно несколько стандартных этапов управления движением колёс (*торможение, выбор режима движения, поворот колес, вращение с осмотром, объезд препятствия, движение к цели*). Каждый из них может быть прерван появлением нового препятствия или заведомым исчезновением объезжаемого объекта.

В результате описанной работы был снят видео-ролик, показывающий функционирование робота "Трикол" при объезде стульев, столбов и людей.

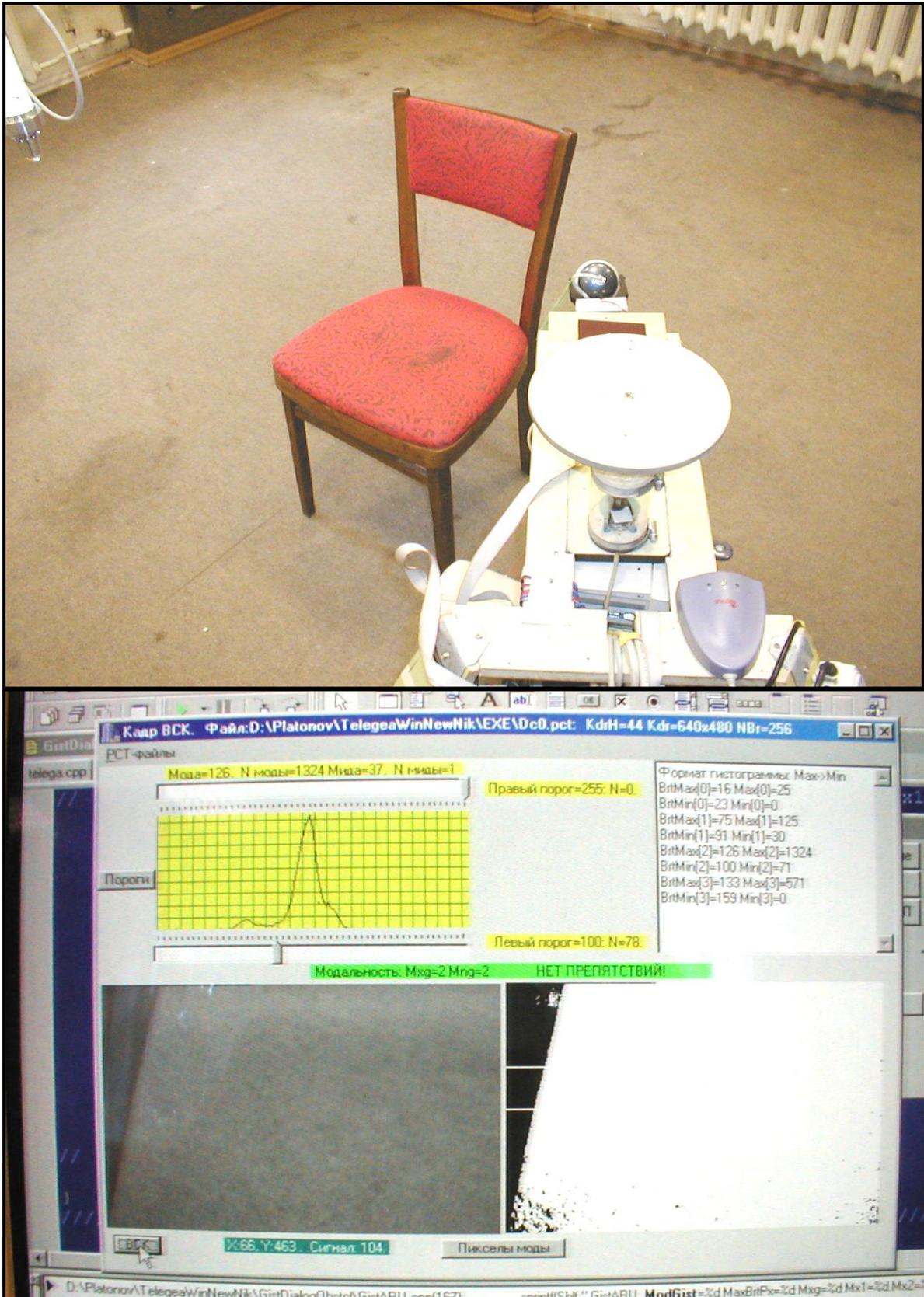


Рис.16. Ситуация наезда на близко стоящий стул.

Гистограммный образ изменился, но недостаточно сильно, т.к. мало точек верха спинки стула попало в поле зрения ВСК, и большая часть этих пикселей лежит вне колеи на полу (ср. рис. 3 и 17). Этот случай потребовал расширения запретной области поля зрения (области построения гистограммного образа сцены).

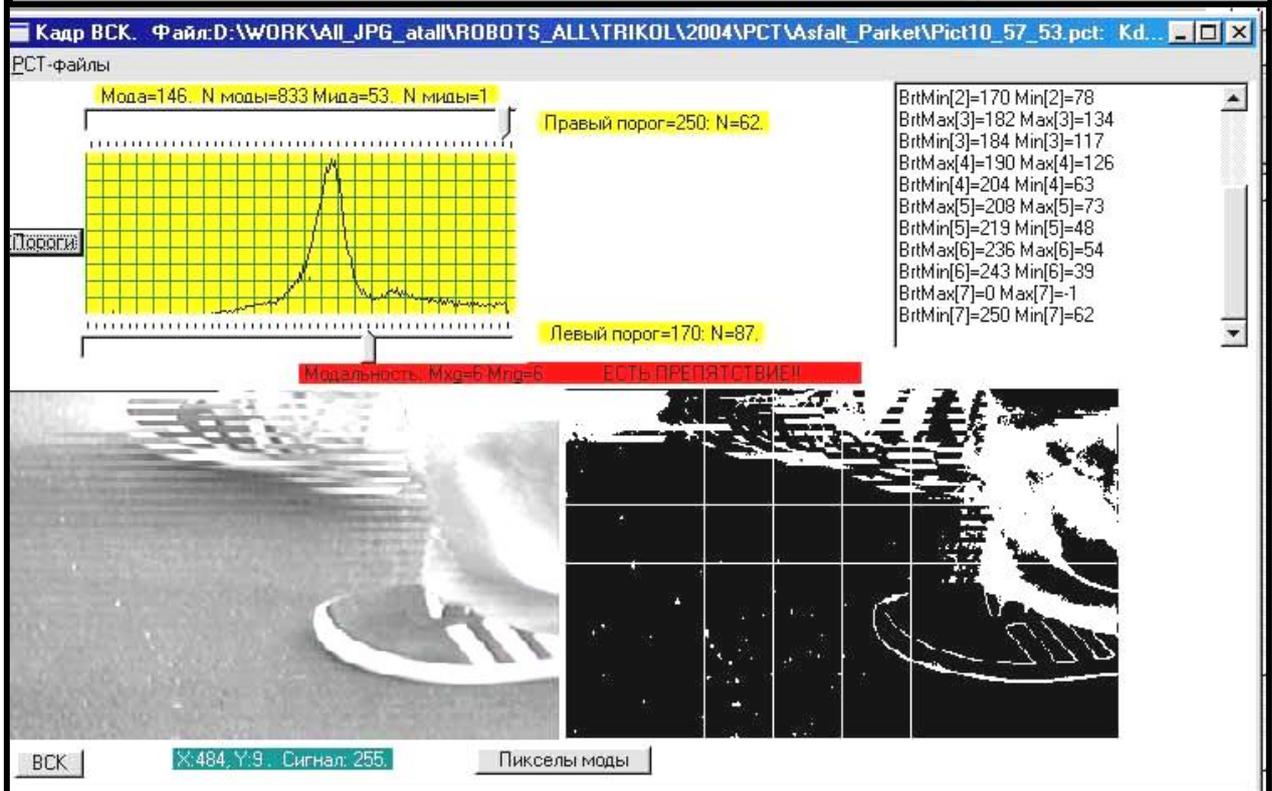
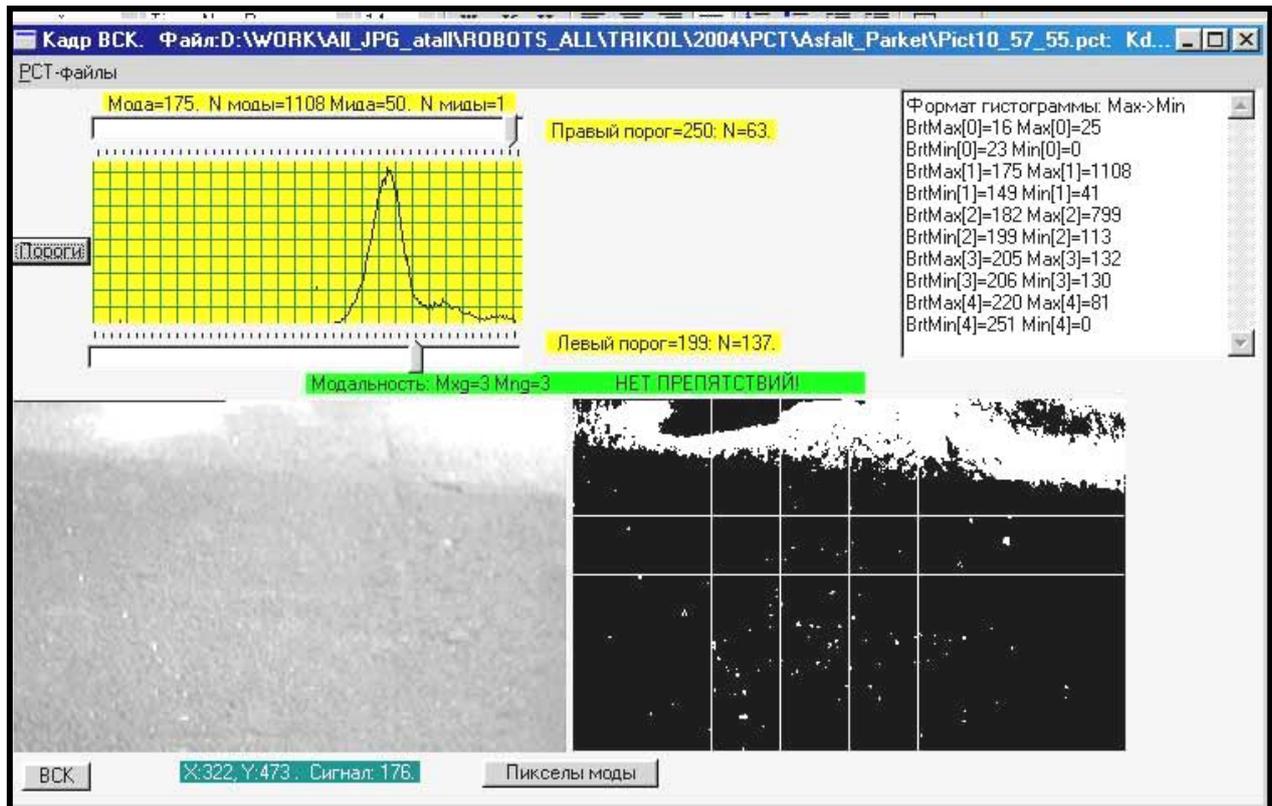


Рис.17. Пример гистограммы реальной сцены асфальта с тенью, бликом и человеком. Рисунки представляют собой копии экранов инструментальной системы ВСК. Правое верхнее окно экранов содержит параметры гистограммной функции, быстрый анализ которых определяет параметры фрейма ситуации "Препятствие!" Как видно в этом случае система правильно справилась с помехами типа "Блик" и "Тень" (это бывает не всегда) Ситуация "Препятствие", как и обычно, определена правильно.

Заключение

* Разработанная система управления мобильного робота со зрительной обратной связью способна идентифицировать наличие препятствий на пути движения и, объезжая их, приезжать к заданной точке области передвижения. Информация о внешней среде поступает от телевизионной камеры, неподвижно установленной на роботе ВСК - "Вперёд смотрящей камеры". Телевизионный сигнал ВСК оцифровывается и обрабатывается в системе управления в режиме реального времени по разработанным алгоритмам.

* Аппаратное время получения фотометрического изображения сцены перед роботом (256 градаций яркости для 640x480 точек изображения) в соответствии с телевизионным стандартом требует 0,04сек, что соответствует 4см пути при скорости движения 1м/сек. Тормозной путь робота в худшем случае составляет 15см, а дальность осмотра – около 1м. Как было показано выше при таких параметрах двигательных возможностей робота время распознавания наличия препятствия движению для реализации экстренного торможения желательно иметь не более 0,7сек.

* Испытания показали, что реализованное время приема изображения, распознавания и выдачи сигнала на торможение (или на поворот) не превышает 0,08сек, при среднем времени порядка 0,02сек, что обеспечивает более чем десятикратный запас в процедурах осмотра препятствий.

* Ограничения поля зрения ВСК требуют активного управления осмотром окружения робота в процессе его движения. Осмотр местности по пути движения реализуется через управление движением колёс робота, которое в зависимости от результата распознавания получаемых сцен ВСК формируется автоматически в алфавите доступных по кинематическим и конструктивным ограничениям режимов движения корпуса робота.

* В результате проделанной работы реализована эффективная система управления многорежимным поведением робота при движении среди препятствий (стулья, люди, стены, столбы) с использованием зрительной обратной связи от монокулярной дешёвой черно-белой ТВ-камеры аналогового типа.

* Опыт работы показал, что разработка систем управления автоматическим движением мобильных роботов требует использования большого помещения для безопасных экспериментов с движущимся роботом на стадии отработки алгоритмов управления и отладки программного обеспечения.

* Для отработки алгоритмов и управляющих программ робота с системой машинного зрения разработаны принципы, алгоритмы и средства инструментального программного обеспечения, поддерживающего процедуры протоколирования работы систем робота в процессе экспериментов и процедуры последующего моделирования работы робота с эмуляцией протоколов в режиме off-line.

Литература

1. Платонов К.К. *"Занимательная психология"*. М.: Молодая гвардия, 1962г., 325с.
2. Емельянов С.Н., Платонов А.К., Ярошевский В.С.. *"Система управления полноприводного трехколесного движителя"* Труды школы-конференции "Мобильные роботы и мехатронные системы", Москва: Изд-во МГУ, 2000г., сс. 89-100.
3. Богуславский А.А., Сербенюк Н.С., Соколов С.М. *"Конический сенсор для навигации подвижного робота по маякам"*. Мобильные роботы и мехатронные системы: Материалы научной школы-конференции Москва: Изд-во МГУ, 2000г.. 42-56с.
4. Сербенюк Н.С. *"Экспериментальное исследование свойств конического сенсора"*. Мобильные роботы и мехатронные системы: Материалы научной школы-конференции (Москва, 5-6 декабря 2000г.). – Москва: Изд-во МГУ, 2000г.. 56-68с.
5. Ярошевский В.С. *"Согласованное управление движением колес и корпуса полноприводного колесного движителя"*. В сб. Мобильные роботы и мехатронные системы. М.:МГУ, 2001г.,70-78с.
6. Емельянов С.Н., Платонов А.К. *"Метод согласования электрических сенсорных сигналов и геометрических характеристик мобильного робота"*. В сб. Мобильные роботы и мехатронные системы. М.:МГУ, 2001г., сс. 78-84.
7. Ярошевский В.С. *"Организация процессов пятипроцессорной системы управления мобильного трёхколёсного робота"*. Труды школы-конференции "Мобильные роботы и мехатронные системы", Москва: ИМех. МГУ, 2002г.
8. Платонов А.К. *"Система управления поведением мобильного робота"*. Труды школы-конференции "Мобильные роботы и мехатронные системы", Москва 2002г., ИМех. МГУ.
9. Богуславский А.А., Соколов С.М. *"Компонентные методы разработки программного обеспечения СТЗ"*. Труды школы-конференции "Мобильные роботы и мехатронные системы", Москва 2003г., ИМех. МГУ.
10. Богуславский А.А., Платонов А.К., Сербенюк Н.С.. *"Система машинного видения робота Трикол"*. Труды школы-конференции "Мобильные роботы и мехатронные системы", Москва 2003г., ИМех. МГУ.
11. *"Восприятие – механизмы и модели"*(сб.). М.: "Мир", 1974г., 368с.
12. Анохин П.К. *"Общая теория функциональных систем организма"* в сб. "Прогресс биолог. и медиц. кибернетики". М.: Медицина, 1974, сс 52-110..
13. Минский М. *"Структура для представления знаний"*. в сб. "Психология машинного зрения". М.: Мир, 1978г., сс 58-71.
14. Белоусов И.Р., Богуславский А.А. и др. *Автоматический захват подвижного объекта роботом-манипулятором* Препринт ИПМ им. М.В.Келдыша РАН, 1996г., № 78, 23с.