

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ДОБАВЛЕНИЯ ВИРТУАЛЬНОГО ОБЪЕКТА В РЕАЛЬНЫЙ МИР. ЧАСТЬ 2*

Ф. М. Кулаков, Е. Н. Смирнов, А. Е. Липатов, В. Н. Латыпов

Санкт-Петербургский институт информатики и автоматизации РАН
199178, Санкт-Петербург, 14 линия ВО, д. 39
<kul@mail.iias.spb.su>

УДК 528.8

Ф. М. Кулаков, Е. Н. Смирнов, А. Е. Липатов, В. Н. Латыпов. **Информационная технология добавления виртуального объекта в реальный мир. Часть 2** // Труды СПИИРАН, Вып. 2, т. 2. — СПб.: Наука, 2005.

Аннотация. В статье описана информационная технология "погружения" произвольного синтезированного на компьютере виртуального объекта, в том числе робота-манипулятора, в реальную внешнюю среду. Разработанная технология обеспечивает как визуальный, так и тактильно-силовой эффект восприятия. Эта технология реализует так называемую добавленную реальность (Augmented Reality), быстро развивающееся направление виртуальной реальности (Virtual Reality). Виртуальный объект в этом случае является добавлением к реальности. В отличие от известных вариантов, описанная в статье технология обеспечивает более реалистичное восприятие виртуального объекта в реальном мире и является менее затратной. В этом выпуске представлена вторая часть статьи, содержащая экспериментальные результаты. — Библиограф. 2 назв.

UDC 528.8

F. M. Kulakov, E. N. Smirnov, A. E. Lipatov, V. N. Latipov. **Information technology of addition of virtual object in the real world. Part 2** // SPIIRAS Proceedings. Issue 2, vol. 2. — SPb.: Nauka, 2005.

Abstract. The paper presents information technology for immersion of arbitrary computer-synthesized virtual object, in particularly, robot-manipulator, into real environment. Developed technology provides visual and tactile-force effects of perception. The technology realizes so called Augmented Reality which is fast developing branch of Virtual Reality. Virtual object in this case is augmentation to Reality. Unlike from the known variants the technology presented by the paper provides more realistic perception of the virtual object into the real environment and this technology requests more small expenditure. In this is presented description of experimental results. — Bibl. 2 item.

5. Результаты экспериментального исследования

Основными целями экспериментальной проверки разработанной технологии формирования добавленной реальности являлось выяснение степени реализма восприятия человеком как через визуальные, так и через тактильно-силовые ощущения того, что физически несуществующий объект как бы является существующим и находящимся среди объектов реального мира.

5.1. Проверка визуального эффекта "погружения"

Экспериментально проверялся визуальный эффект погружения виртуального объекта в реальный мир для случая, когда человек наблюдает реальный мир с помощью видеокамер.

Реализм визуального ощущения при безусловном обеспечении экранного эффекта (загораживания частей объектов, находящихся дальше от наблюда-

* Часть 1 см.: Труды СПИИРАН. Вып. 2, т. 1. С. 236–256.

теля, частями других объектов, находящихся к нему ближе), а также идентичности освещенности виртуального объекта и объектов реальной внешней среды, зависит от двух факторов: **непрерывности восприятия и точности со-вмещения** изображения видимых фрагментов внешней среды с соответствующими фрагментами изображений их геометрических моделей.

Непрерывность восприятия определяется временем, необходимым для формирования кадра изображения дополненной реальности. Это время не должно превышать 50–60 мсек, что соответствует частоте смены кадров 15–20 Гц.

Формирование кадра изображения дополненной реальности требует реализации трех процессов:

- формирование видео кадра изображения реальной внешней среды;
- формирование текущих координат позиции и ориентации головы;
- захват видеоизображения, сформированного TV-камерами, и формирование по нему цифрового изображения внешней среды, а также формирование цифрового изображения виртуального объекта, дополняющего изображение внешней среды с выводом результирующего изображения дополненной реальности на экран дисплеев для левого и правого глаз.

Известно, что для формирования кадра видеоизображения при сетевой частоте 50 Гц стандартная видеосистема требует 40 мсек, а для формирования координат позиции/ориентации головы с помощью оптико-телевизионной системы отслеживания позиции/ориентации головы (HTS), как показали эксперименты, также требуется не более 40 мсек.

Таким образом, для достижения минимального времени формирования кадра дополненной реальности необходимо, чтобы перечисленные процессы выполнялись одновременно, причем время выполнения третьего процесса не должно превышать 40 мсек.

Для выполнения указанных процессов были использованы нижеследующие устройства, входящие в аппаратно-программный комплекс, реализующий технологию погружения виртуального объекта в реальный мир, блок-схема которого изображена на рис. 1:

- две мобильные видеокамеры, генерирующие изображение реальной внешней среды (блок 1 и блок 2) для левого и правого глаз, реализующие первую операцию;
- персональный компьютер (блок 3) для вычисления текущих пространственных координат позиции/ориентации головы, реализующий вторую операцию;
- графические станции (блок 7 и блок 8), каждая из которых имеет два центральных процессора, работающих параллельно, карту захвата видеоизображения и карту вывода изображения; станции реализуют третью из выше перечисленных операций.

Очевидно, при параллельной работе этих устройств максимальная частота смены кадров изображения дополненной реальности будет равна 25 Гц, если время выполнения третьего процесса не будет превышать 40 мсек. Это достаточно для обеспечения непрерывности восприятия изображения.

При проведении экспериментов в качестве геометрической модели внешней среды была использована модель поверхности орбитальной станции.

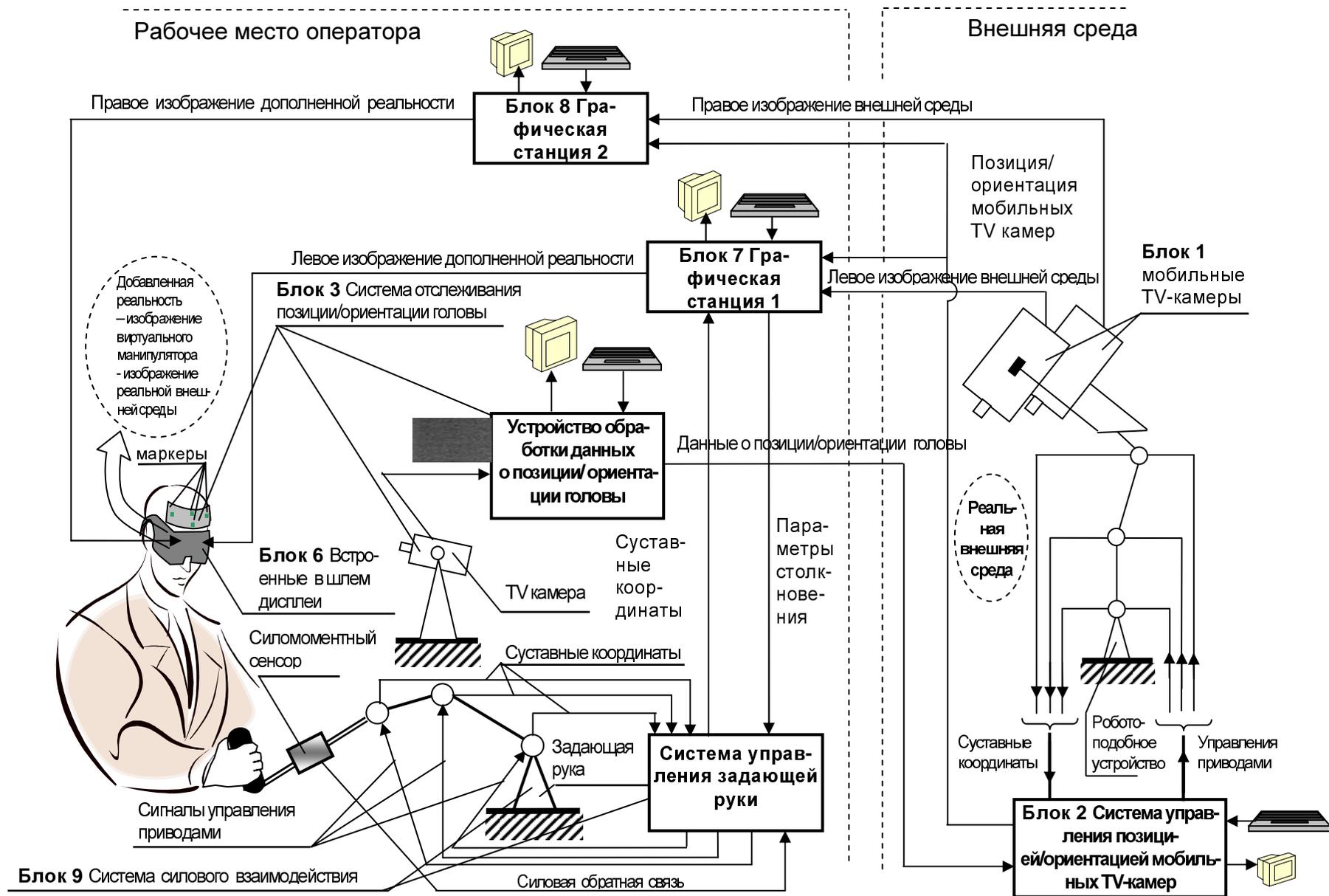


Рис. 1. Структура аппаратно-программного комплекса для реализации эффекта "погружения" виртуального манипулятора в реальный мир.

Рис. 2с иллюстрирует изображение проволочной модели орбитальной станции. Рис. 3а, 3б представляют изображение геометрической модели сборочного участка орбитальной станции с виртуальным манипулятором.

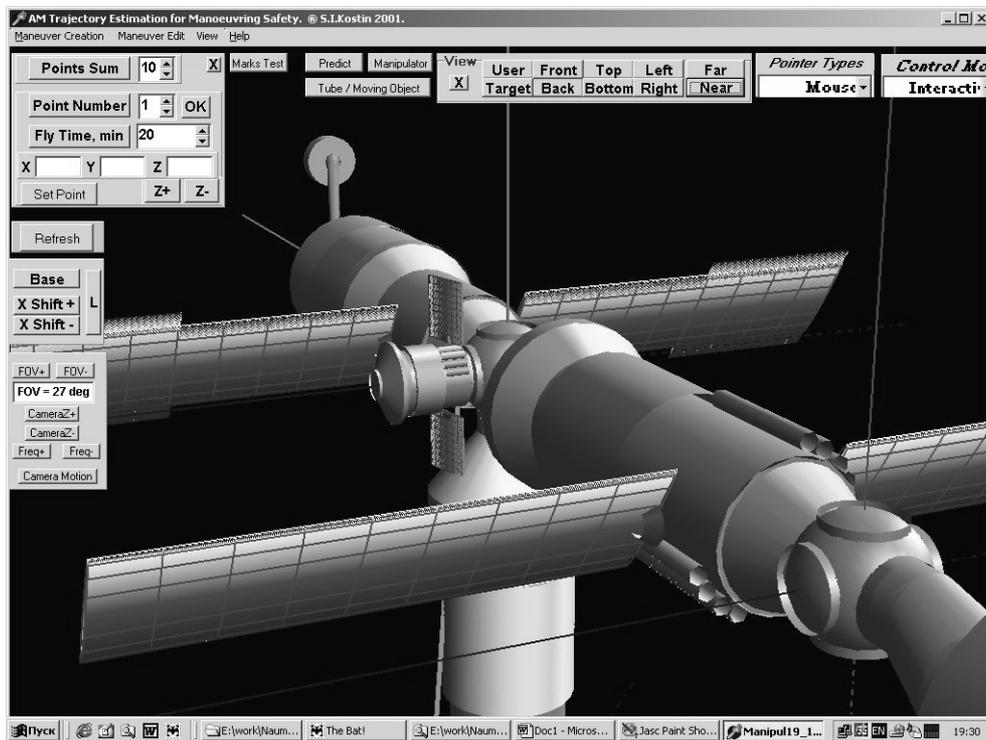


Рис. 2а. Трехмерное изображение геометрической модели орбитальной станции (ракурс: сзади — сверху — правый бок).

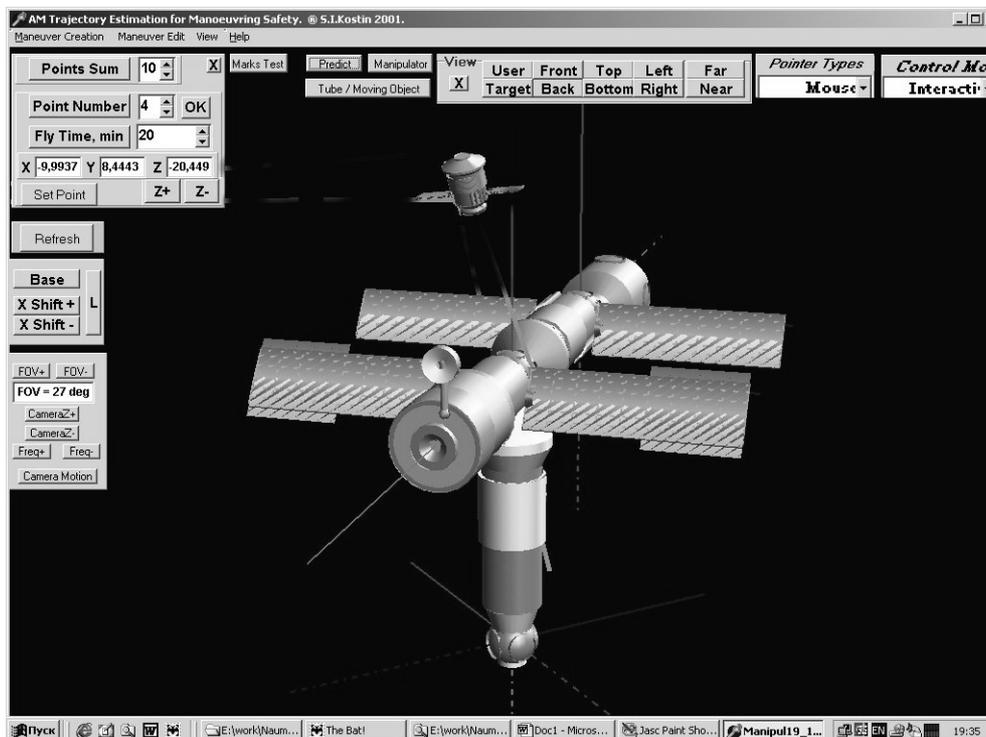


Рис. 2б. Трехмерное изображение геометрической модели орбитальной станции (ракурс: спереди — сверху — правый бок).

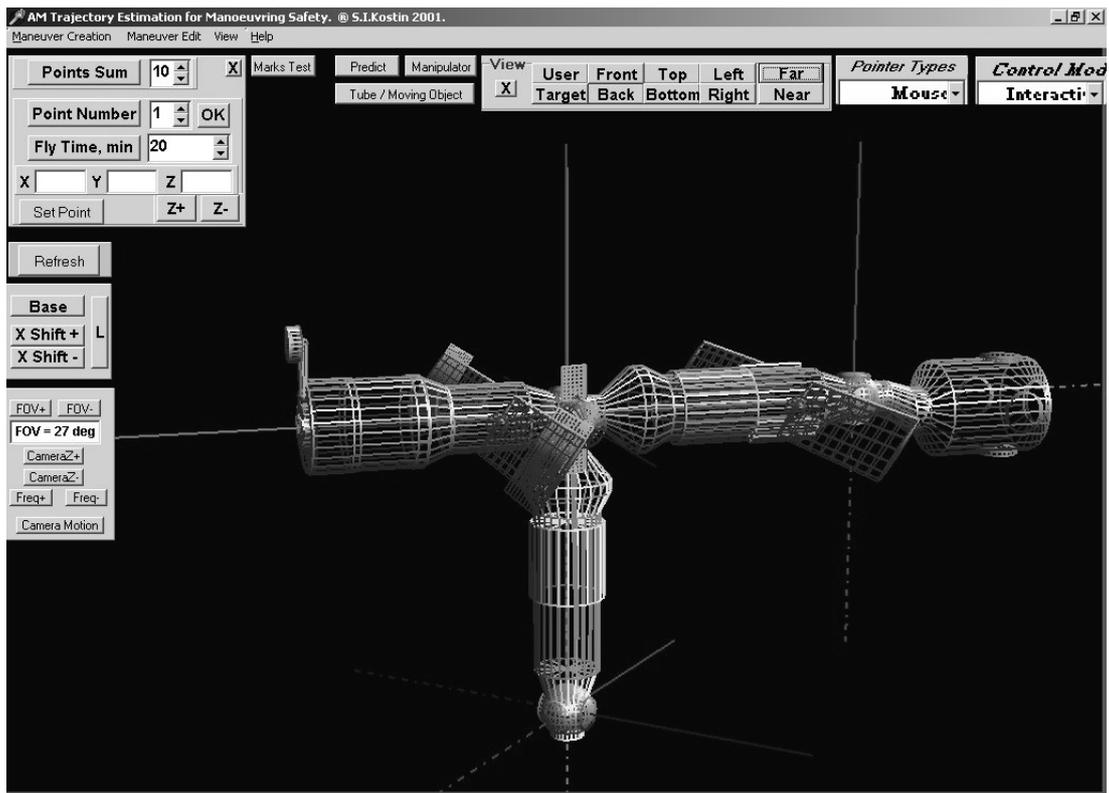


Рис. 2с. Изображение проволочной модели орбитальной станции.



Рис. 3а. 3D изображение геометрической модели сборочного участка орбитальной станции с виртуальным манипулятором (ракурс: сверху — сзади — правый бок).

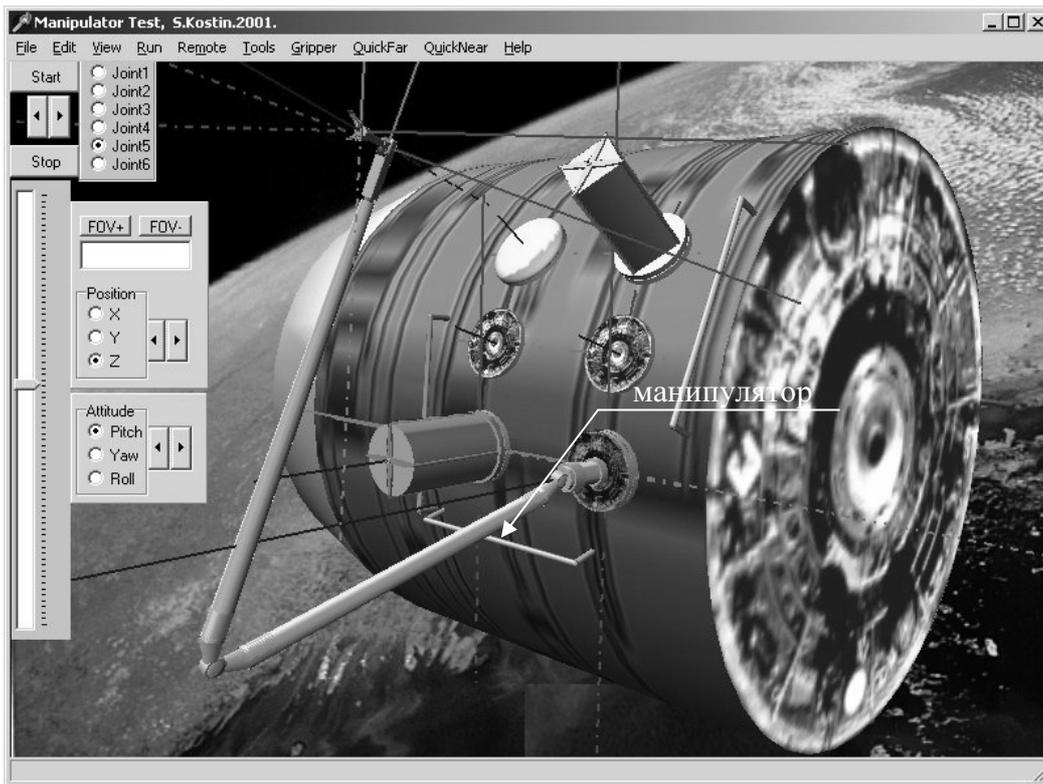


Рис. 36. 3D изображение геометрической модели повернутого на 80° сборочного участка орбитальной станции с виртуальным манипулятором (ракурс: сверху — сзади — правый бок).

Был изготовлен макет сборочного участка орбитальной станции, который использовался для получения видео изображения реальной внешней среды. Его видеоизображение представлено на рис. 4.

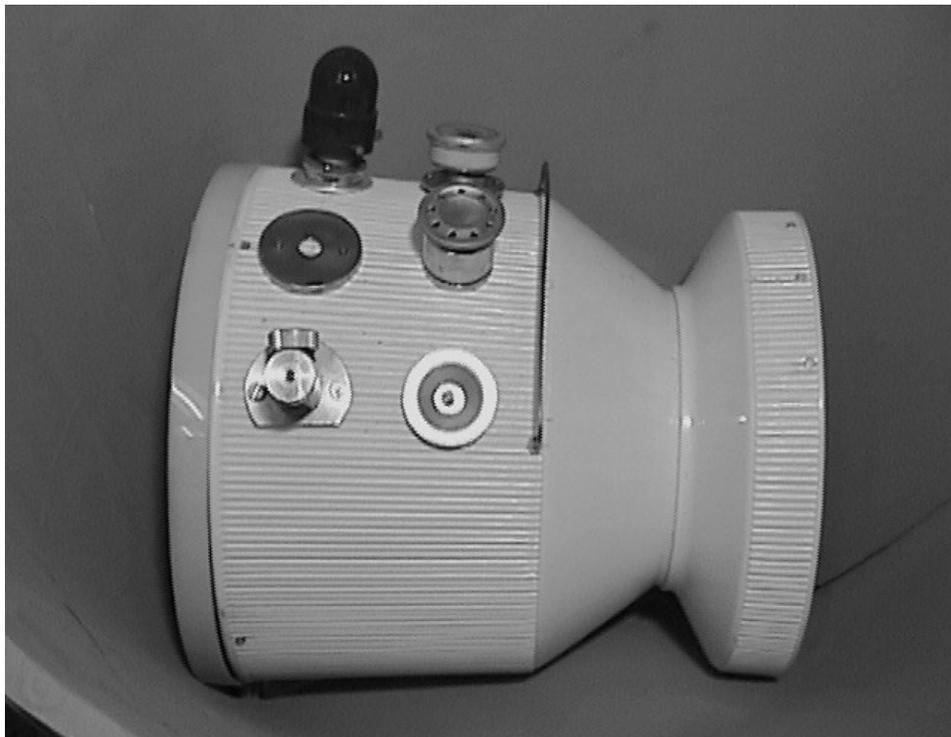


Рис. 4. Видеоизображение макета сборочного участка орбитальной станции.

Эксперименты, проводимые на созданном аппаратно-программном экспериментальном комплексе при использовании весьма недорогих графических станций, характеристики которых были приведены в [1], раздел 4, дали ниже следующие результаты.

В случае реальной внешней среды средней сложности и дополняющего ее виртуального объекта (манипулятора), представленных на рис. 1, имеют место ниже следующие затраты времени графическими станциями, идущие на выполнение третьего из вышеупомянутых процессов по формированию кадра изображения дополненной реальности (Таблица 1).

Таблица 1. Затраты времени на формирование кадра видеоизображения

| Затраты времени | Тип видеокарты | |
|--|----------------|-------|
| | AGP | AGP×2 |
| Суммарное время, мсек. | 39,5 | 35,1 |
| Время генерации и вывода на дисплей виртуального объекта, мсек. | 14,7 | 13,0 |
| Время захвата видеоизображения реальной внешней среды генерации ее геометрической модели и вывода на дисплей изображения дополненной реальности, мсек. | 20,7 | 18,0 |
| Остальные затраты времени мсек. | 4,1 | 4,1 |

Таким образом, проведенные эксперименты подтвердили, что разработанная технология обеспечивает возможность достижения частоты смены кадров, достаточной для получения непрерывного восприятия дополненной реальности. Причем время выполнения процесса в значительной степени зависит от типа используемой видеокарты вывода изображения и может быть существенно сокращено при использовании более совершенных видеокарт.

При использовании таких карт также может быть существенно усложнена внешняя среда и виртуальный объект.

Проверка точности совмещения изображений реальной внешней среды с ее геометрической моделью выявили наличие ошибок совмещения двух видов - статических и динамических. Статические ошибки появляются даже в случае, если позиция и ориентация наблюдателя и/или объекта наблюдения остаются неизменными. Динамическая ошибка не появляется, пока наблюдатель или объект наблюдения не движутся. Если человек использует для наблюдения дополненной реальности, смонтированные в головной шлем дисплеи, то динамическая ошибка является главной компонентой суммарной ошибки совмещения.

Главными причинами статической ошибки являлись:

- нелинейные искажения TV-камеры,
- отличие позиции/ориентации реальной TV-камеры, отслеживающей пространственное положение головы от позиции/ориентации виртуальной камеры, формирующей изображение геометрической модели внешней среды,
- ошибки формирования геометрической модели внешней среды,
- различие углов обзора (FOV) реальной камеры и виртуальной камеры, позиция/ориентация которой определяет ракурс и масштаб изображения геометрической модели внешней среды и виртуального объекта.

Результирующая статическая ошибка составляла 15 пиксел, получаемая при использовании TV-камер с FOV $65^\circ \times 65^\circ$. Замена этой камеры на камеру с

FOV $40^\circ \times 40^\circ$ привело к уменьшению ошибки до 7–10 пиксел, что явилось следствием уменьшения нелинейных искажений новой TV-камеры.

Для уменьшения отличий в позиции/ориентации реальной и виртуальной TV-камер, вызванных неточностью измерения позиции/ориентации головы, был изменен источник входных данных о положении реальной TV-камеры. Координаты позиции/ориентации для виртуальной TV-камеры ранее формируемые компьютером системы отслеживания позиции/ориентации головы, теперь вводятся непосредственно с роботоподобного устройства, перемещающего реальные TV-камеры, что исключает возможные ошибки, обусловленные неточностью отслеживания TV-камерами позиции/ориентации головы. Входными данными о позиции/ориентации являются матрицы T позиции/ориентации каждой из реальных TV-камер обзора внешней среды. Матрицы вычисляются как известные функции суставных координат устройства, измеряемых суставными сенсорами.

Однако в этом случае весьма вероятна систематическая ошибка определения позиции/ориентации TV-камер, обусловленная неточностью математической модели роботоподобного устройства, устанавливающей зависимость элементов матрицы T^0 позиции/ориентации платформы с реальными камерами от суставных координат устройства вследствие чего положение/ориентация наблюдателя реальной внешней среды и наблюдателя модели не совпадают.

В связи с этим была разработана специальная технология калибровки для устранения этой ошибки. Она позволяет определить так называемую корректирующую матрицу $T_{кор}$, с помощью которой вычисляется матрица T^0 действительного положения/ориентации систем координат, связанной с каждой из TV-камер таким образом, что ее начало координат является оптическим центром камеры и ось X_3 является оптической осью.

Матрица T^0 может быть представлена в следующем виде:

$$T^0 = TT_{кор}. \quad (1)$$

Предложенная технология калибровки базируется на определении позиции и ориентации реальных TV-камер в базовой системе координат по TV-изображению некоторой известной простой конструкции, позиция/ориентация которой в этой базовой системе координат точно известна.

Например, такой конструкцией может быть конструкция, изображенная на рис. 5. Тогда, используя известные формулы оптического преобразования

$$X_1^{(j)} = -k_x f \frac{T_{1.}^0 X^{(i)}}{T_{3.}^0 X^{(i)}}, \quad i = 1, 2, \dots, n;$$

$$X_2^{(j)} = -k_y f \frac{T_{2.}^0 X^{(i)}}{T_{3.}^0 X^{(i)}}, \quad i = 1, 2, \dots, n, \quad (2)$$

где $X_1^{(i)} X_2^{(i)}$ — координаты изображения i -х характерных точек конструкции на чувствительной плоскости в координатной системе камер,

$X^{(i)} = (X_1, X_2, X_3)$ — вектор позиции i -ой характерной точки конструкции в базовой системе координат,

T_1^0, T_2^0, T_3^0 — первая, вторая и третья строки искомой матрицы T^0 позиции/ориентации системы координат камеры в базовой системе координат объекта внешней среды, знание которой необходимо для формирования изображения виртуального объекта и геометрической модели внешней среды.

Очевидно, в общем случае достаточно 12 уравнений (2), чтобы определить все элементы искомой матрицы и еще две, чтобы найти неизвестные величины $k_x f$ и $k_y f$.

Таким образом, в общем случае, достаточно использовать 7 характерных точек, чтобы решить задачу нахождения искомой матрицы T^0 определения путем разрешения системы (2).

На практике задачу можно упростить, если задать в базовой системе координат позицию/ориентацию известной конструкции, таким образом, чтобы характерные точки имели одну или две нулевые компоненты, а также используя симметричные относительно осей координат точки. При этом система уравнений (2) упрощается.

Еще более просто определить матрицу $T_{кор}$, если матрицу T^0 не вычислять, а задавать. Причем выбрать такое ее значение, которому соответствует позиция/ориентация TV-камеры, легко достижимые путем управления роботоподобным устройством вручную, например, путем последовательных поступательных перемещений вдоль осей координат базовой системы, а затем, вокруг осей.

Такой матрицей может быть матрица вида

$$T^0 = \left\| \begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & X_3 \\ \hline 0 & 0 & 0 & 1 \end{array} \right\|.$$

Если в качестве простой наблюдаемой с помощью TV-камеры конструкции взять конструкцию, изображенную на рис. 5, то, очевидно, изображение ее в позиции TV-камеры, соответствующей матрице T^0 , будет иметь вид, представленный на рис. 5.

Правда, при любом расстоянии X_3 от конструкции, этот вид будет одинаков и будет только отличаться масштабом. Для определения неизвестного X_3 целесообразно воспользоваться следующим приемом. Измерим расстояние d_{AB} между изображениями каких-либо характерных точек конструкции, например, между точками A и B , лежащими на расстоянии d_{AB} по оси X_2 системы координат экрана в случае, когда TV-камера занимает позицию $(0,0,X_3)$. Затем переместим камеру вдоль оси X_3 на расстояние Δ и снова измерим расстояние d'_{AB} между изображениями этих точек. Тогда из уравнений оптического преобразования имеем:

$$d_{AB} = K_x f \frac{D_{AB}}{X_3},$$

$$d'_{AB} = K_x f \frac{D_{AB}}{X_3 + \Delta}.$$

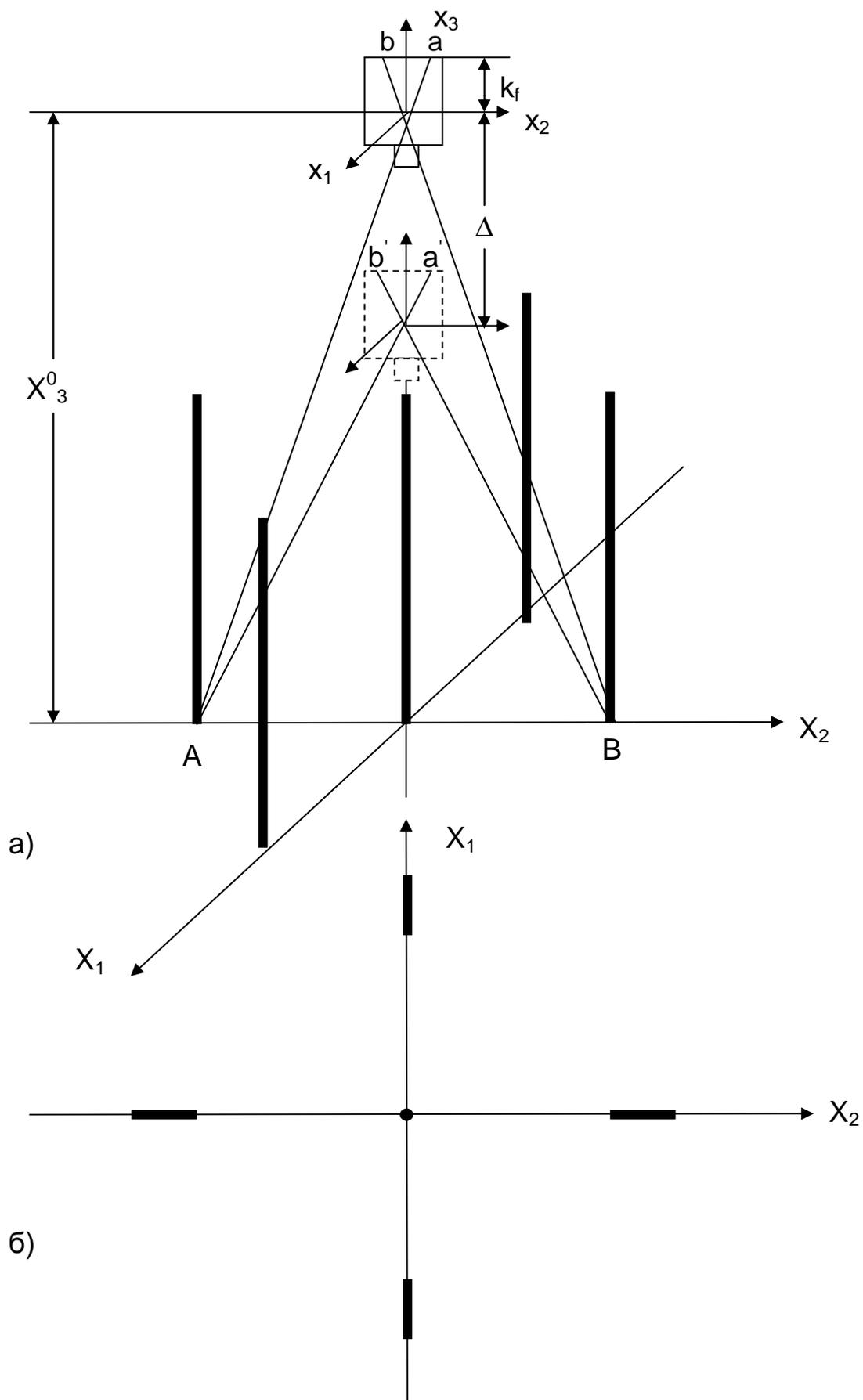


Рис. 5. Специальная конструкция (макет декартовой системы координат) для определения корректирующей матрицы $T_{кор}$ (а) и TV-изображения специальной конструкции, соответствующее позиции/ориентации TV-камеры, определенной матрицей T^0 (б).

Искомое X_3 будет определяться как $X_3 = \frac{\Delta \cdot d'_{AB}}{d_{AB} - d'_{AB}}$

Определив таким образом T^0 и зная матрицу T , определенную по известным значениям суставных координат роботоподобного устройства, можно с помощью (1) определить значение $T_{кор}$ как $T_{кор} = T^{-1}T^0$. На эту матрицу необходимо домножать справа вычисленную текущую матрицу T , чтобы получить текущую матрицу позиции/ориентации виртуальной TV-камеры.

Ошибка, обусловленная разницей углов зрения реальной и виртуальной TV-камер, устраняется путем ручной настройки угла зрения виртуальной TV-камеры, которая устанавливается вышеупомянутым способом в заданную скорректированную позицию, идентичную позиции реальной TV-камеры.

Эта настройка предусматривается при разработке ПО генерации добавленной реальности и осуществляется с помощью клавиатуры и мыши. С этой целью на видео изображение вышеупомянутой реальной простой калибровочной конструкции накладывается виртуальное изображение этой же виртуальной конструкции, полученное для тех же самых позиции и ориентации, что и реальное. Ошибка совмещения в этом случае обусловлена лишь отличием углов зрения виртуальной и реальной TV-камер.

В процессе экспериментов было установлено, что статическая ошибка совмещения изображения реальной внешней среды, полученного с помощью мобильных TV-камер, с изображением ее компьютерно синтезированной геометрической модели, полученным с помощью виртуальных TV-камер, не превышает 2-3 пикселя, при условии компенсации нелинейных искажений реальных TV-камер, а также проведения тщательной калибровки позиции/ориентации этих TV-камер с целью уменьшения ошибки позиционирования до 0,1 мм и угловой ошибки до 0,005 рад, а также уточнения величины FOV реальных TV-камер для согласования этой FOV с FOV виртуальных TV-камер.

Ниже приводятся рисунки, которые иллюстрируют различные варианты погружения (добавления) изображения виртуального манипулятора к изображению реальной модели внешней среды, добытой с помощью TV-камер.

Рис. 6 и 7 иллюстрируют внешнюю среду, которой является макет сборочной части орбитальной станции, и реальную внешнюю среду, дополненную виртуальным манипулятором. Рис. 8–11 иллюстрируют изображение добавленной реальности, когда виртуальный объект движется. Рис. 12–15 иллюстрируют изображение добавленной реальности, когда человек-наблюдатель приближается к макету сборочной части станции. Рис. 16 иллюстрирует сложный эффект загороживания манипулятором сцены и, наоборот, сценой — манипулятора.

Причиной появления динамической ошибки при перемещении головы наблюдателя или возможных перемещениях объектов внешней среды является возникающее в этом случае отличие изображения реальной внешней среды, полученного с помощью видеокамеры, от изображения ее геометрической модели, полученного с помощью компьютерного синтеза. Это отличие вызвано: (1) появляющейся разницей в позиции/ориентации реальной TV-камеры, при котором формируется кадр видеоизображения, от позиции/ориентации виртуальной камеры, при котором синтезируется изображение геометрической модели; (2) «размытостью» (нечеткостью контура) видеоизображения, обусловленным смещением объектов внешней среды относительно TV-камеры за период экспозиции.

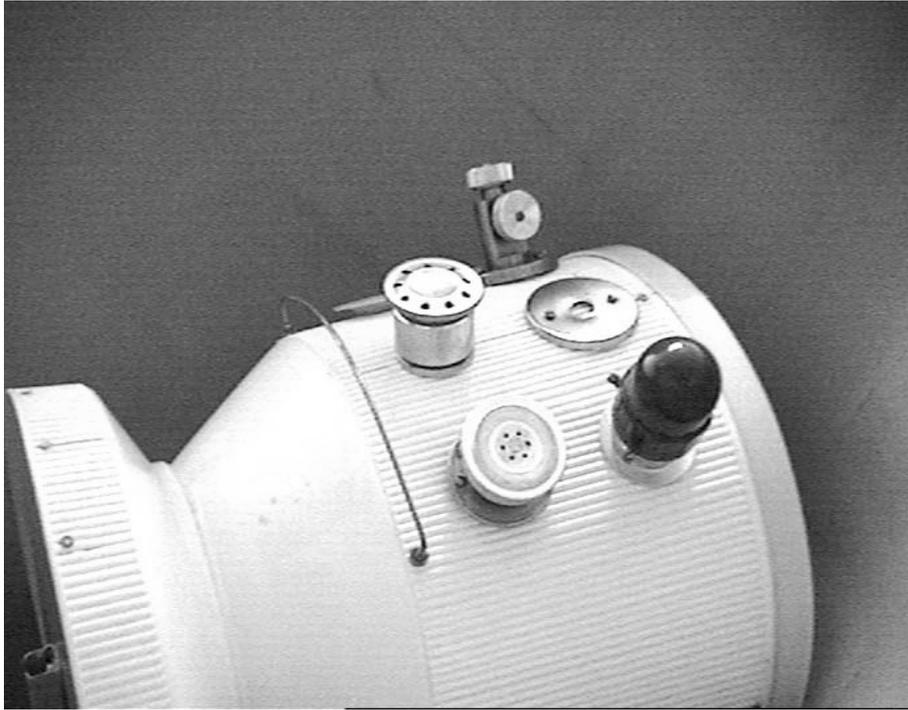


Рис. 6. TV изображение макета реальной внешней среды сборочной части орбитальной космической станции.



Рис. 7. Реальная внешняя среда, дополненная виртуальным манипулятором.



Рис. 8. Изображение виртуального объекта (I фаза).

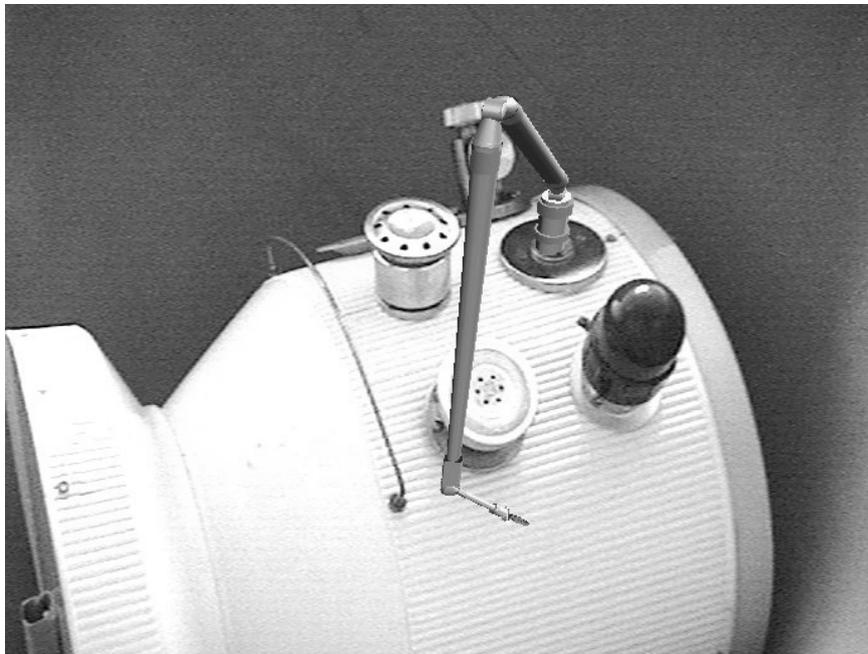


Рис. 9. Изображение виртуального объекта (II фаза).

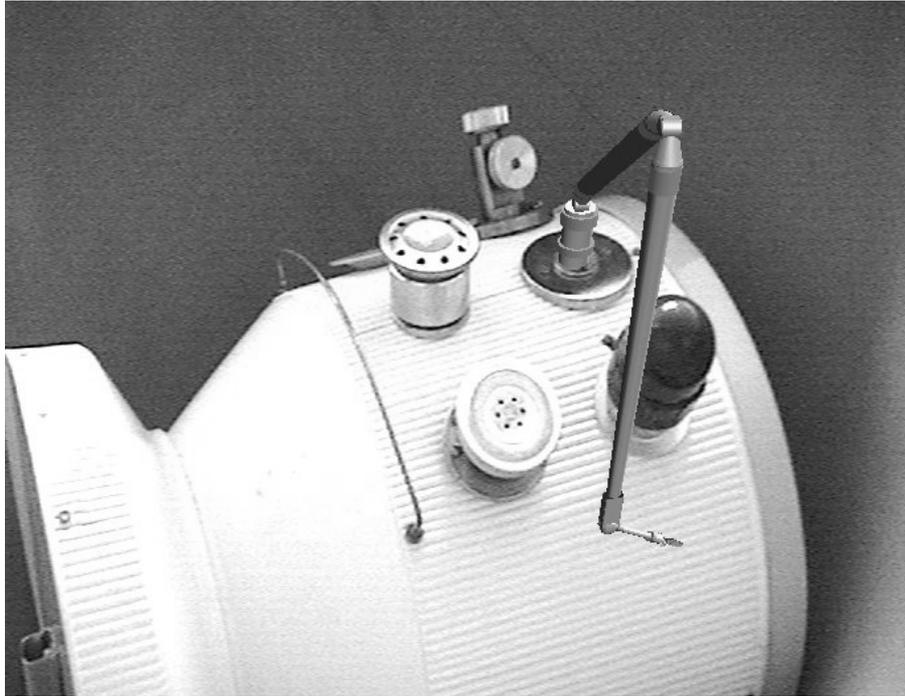


Рис. 10. Изображение виртуального объекта (III фаза).

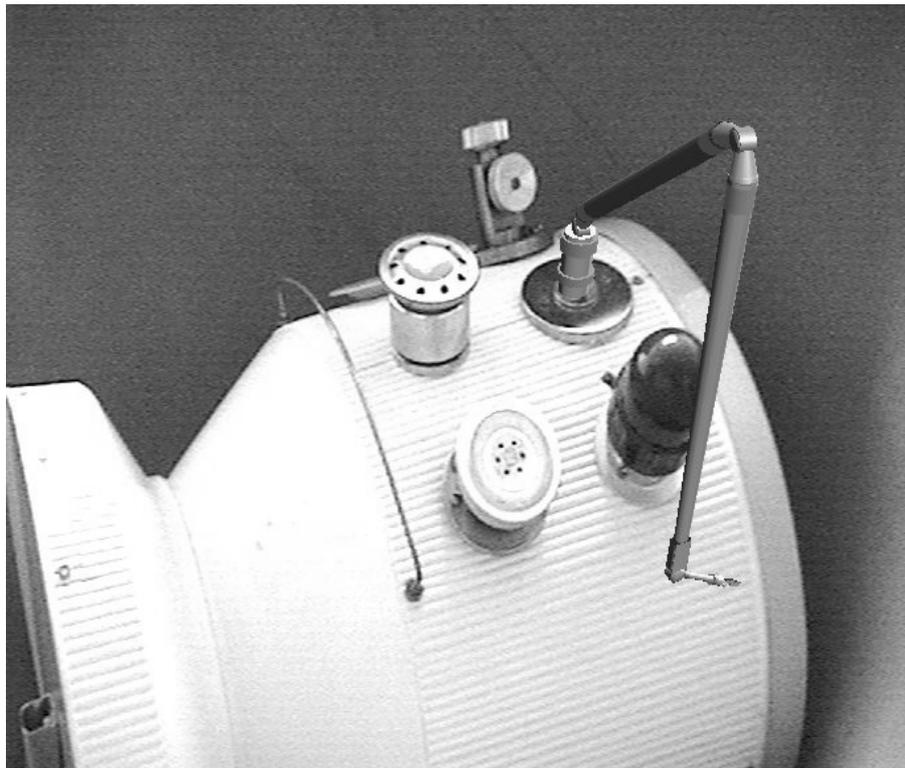


Рис. 11. Изображение виртуального объекта (IV фаза).

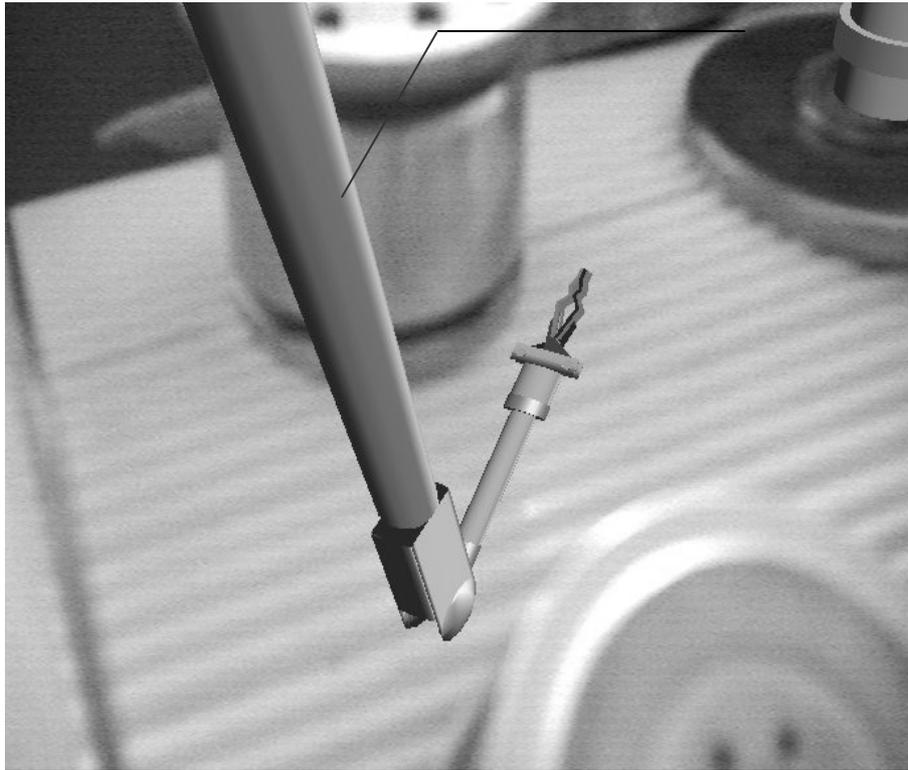


Рис. 12. Иллюстрация добавленной реальности, когда человек-наблюдатель удаляется (I фаза движения).



Рис. 13. Иллюстрация добавленной реальности, когда человек-наблюдатель удаляется (II фаза).

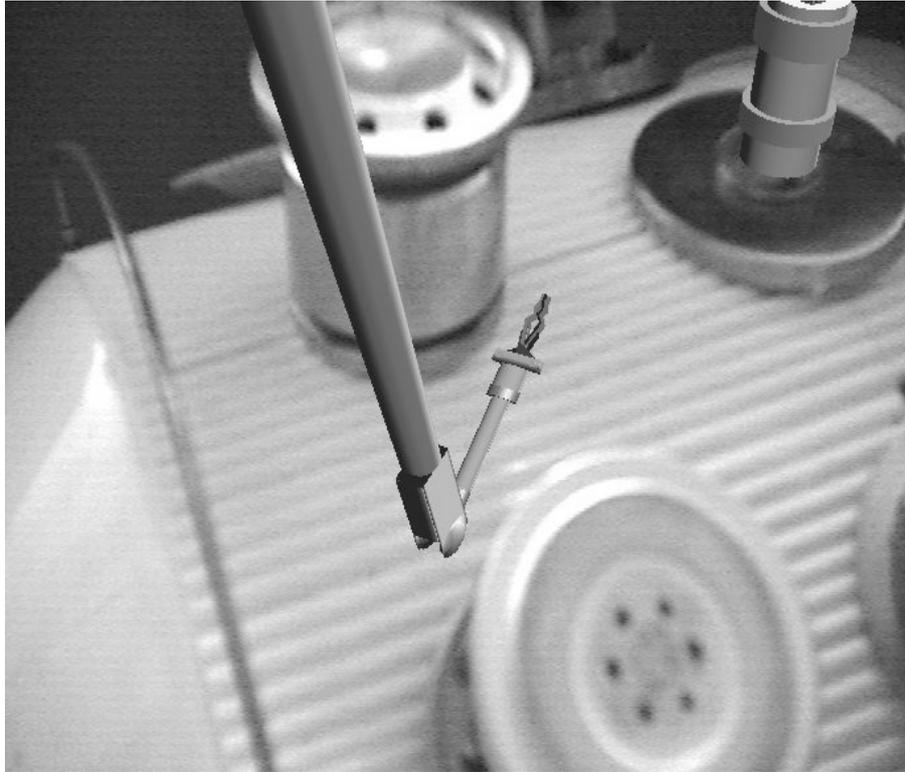


Рис. 14. Иллюстрация добавленной реальности, когда человек-наблюдатель удаляется (III фаза).

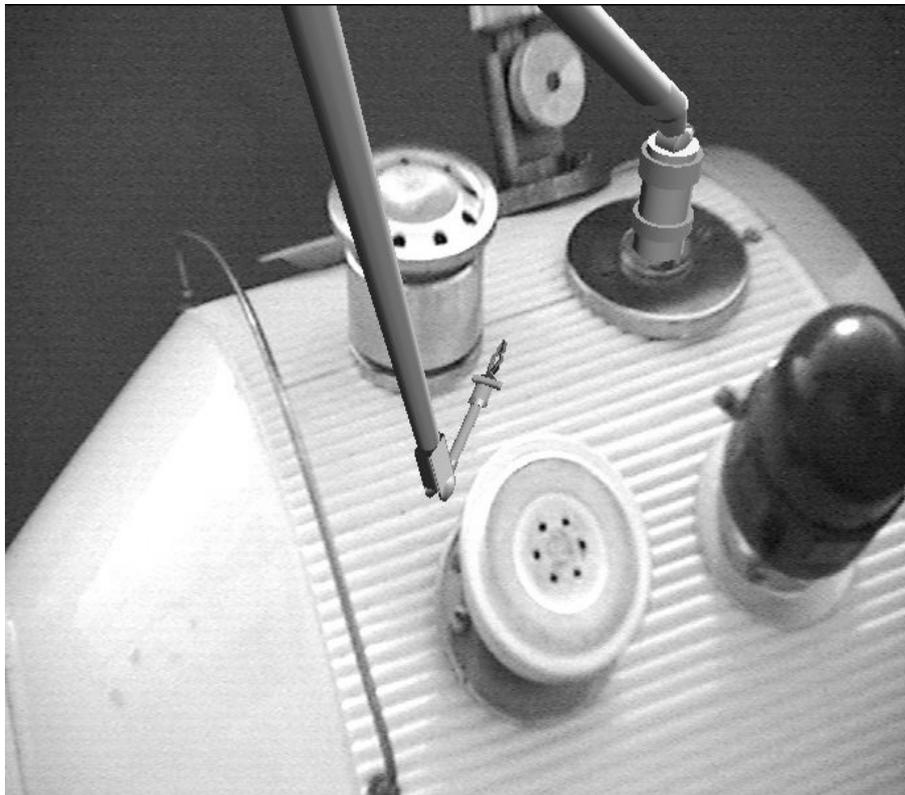


Рис. 15. Иллюстрация добавленной реальности, когда человек-наблюдатель удаляется (IV фаза).



Рис. 16. Иллюстрация изображения добавленной реальности, когда имеет место "сложный" эффект загромождения.

Как показали эксперименты, определяющей причиной разницы в позиции/ориентации реальной и виртуальной камер при перемещении головы оператора, является динамическая ошибка при отслеживании позиции/ориентации головы платформой с камерами роботоподобного устройства. Она обусловлена инерционностью устройства и несовершенством системы управления.

Используемая следящая система управления устройства СФЕРА-36 является весьма быстродействующей. Она отслеживает ступенчатый входной сигнал экспоненциально, постоянная времени экспоненты составляет 0,07 сек. Однако, такие динамические характеристики системы оказываются недостаточными. Действительно, при такой постоянной времени при развороте головы со скоростью 40 град/сек установившаяся ошибка отслеживания будет составлять 2,8 град. При FOV камеры 40 град. Такая ошибка соответствует огромному сдвигу виртуального изображения относительно реального. Этот сдвиг составляет 7 % относительно размера экрана, что является недопустимым, если еще принять во внимание, что входной сигнал для системы управления формируется компьютером обработки данных блока 3 (рис. 1) с запаздыванием порядка 40 мсек, определяемым затратами времени на вычисление.

Возможным способом уменьшения разницы между позицией/ориентацией виртуальной и реальной камер, является улучшение динамических свойств системы управления, однако эксперимент показал, что таким образом можно уменьшить ошибку не более чем в два раза, что недостаточно.

Более кардинальный путь уменьшения разницы в позиции/ориентации реальной и виртуальной камер предполагает использование для синтеза изображения геометрической модели внешней среды не координаты позиции/ориентации головы, полученные с помощью HTS, а координаты текущего положения/ориентации реальной TV-камеры обзора внешней среды, которые вычисляются по измеренным текущим значениям суставных координат роботоподобного устройства, так как функциональная зависимость матрицы пози-

ции/ориентации платформы с камерами от суставных координат используемого роботоподобного устройства всегда известна.

Именно такой путь был избран при разработке прототипа программно-аппаратного комплекса для тестирования предложенной технологии погружения виртуального тела в реальный мир.

Правда при таком подходе не устраняется запаздывание в отслеживании платформой с камерами позиции/ориентации головы, вызванное затратами времени на вычисление координат головы компьютером блока 3 (рис. 1) и обработку этих координат системой управления камерами (блок 8). Однако, как показали эксперименты, человек практически не ощущает этого запаздывания. Размытость видеоизображения тем большая, чем больше период экспозиции, также существенно ухудшает реалистичность визуального восприятия «эффекта» погружения виртуального тела в реальный мир.

Заметим, что при скорости вращения головы 40 град/сек для периода экспозиции 20 мсек (50 % времени формирования кадра) угол поворота составляет 0,8 град, т.е. размытость контура изображения равна 2 % от размера экрана.

Это вызвано, во-первых, тем, что компьютерно-синтезированное изображение имеет четкие границы и поэтому точное совмещение видеоизображений, имеющего нечеткие границы с компьютерно-синтезированным изображением, в принципе, невозможно. Во-вторых, в этом случае оказывается затруднительно определить оптимальные значения координат позиции/ориентации виртуальной камеры, которым должно соответствовать синтезируемое изображение, т.к. время экспозиции может быть значительным, изменяясь от кадра к кадру в зависимости от освещенности, и в пределе достигать времени формирования кадра изображения, т.е. 40 мсек.

Лучше всего выбирать момент середины периода экспозиции и использовать значения координат позиции/ориентации реальной камеры для синтеза компьютерного изображения, соответствующего именно этому моменту времени. Однако процесс измерения суставных координат роботоподобного устройства, формирование по ним координат позиции/ориентации TV-камеры и передача их на графическую станцию для формирования компьютерного изображения, занимают определенное время, поэтому задача оптимального выбора координат позиции/ориентации виртуальной камеры для формирования компьютерного изображения является весьма сложной. А главное, решение ее полностью не разрешает проблему точного совмещения виртуального и реального изображений при перемещении наблюдателя.

Эффективный путь уменьшения размытости основан на использовании специальных TV камер с очень малым временем экспозиции, что в принципе, открывает возможность кардинально уменьшить размытость изображения и более точно определить позицию реальной а, следовательно, и виртуальной камеры для формирования компьютерного изображения.

Однако практически реализация этого пути требует значительных технических усовершенствований аппаратно-программного комплекса, реализующего технологию погружения, блок-схема которого представлена на рис. 1.

Основное усовершенствование состоит в обеспечении синхронизации моментов начала формирования кадров изображения TV-камер обзора внешней среды (блок 1) с моментом начала формирования кадров TV кадров изображения маркерного устройства, а также с моментом начала вычисления кадра изображения геометрической модели внешней среды, производимого графическими станциями (блоки 7 и 8), и, наконец, с моментом начала вычисления матри-

цы позиции/ориентации виртуальной TV камеры, осуществляемой блоком 2 (центральный процессор устройства СФЕРА-36). При этом, очевидно, интервалы циклической работы алгоритмов, реализуемых блоками 2 и блоками 7 и 8 должны быть равны периоду формирования TV кадров видеоизображения.

Это усовершенствование позволяет обеспечить жесткую очередность работы алгоритмов, выполняемых на разных блоках, относительно друг друга и сделать эту очередность идентичной на каждом из одинаковых временных интервалов, равном периоду формирования кадра. Это в свою очередь облегчает экстраполяцию позиции/ориентации головы и TV-камеры, которая необходима, т.к. интервалы формирования позиций/ориентаций занимают время, сравнимое с интервалом формирования TV изображения, и в течение этого времени возможны существенные приращения позиции/ориентации TV-камеры.

На рис. 17 представлена предлагаемая нами укрупненная временная диаграмма работы каждого из блоков, поясняющая возможность компенсации динамической ошибки при совмещении TV изображения внешней среды с компьютерно-синтезированным изображением его геометрической модели.

В момент t_k мобильные TV-камеры обзора внешней среды (блок 1) и камер формирования изображения маркеров, входящая в блок 3, начинают формировать соответствующие изображения. Причем позиция/ориентация камер точно соответствуют моменту времени t_k , т.к. время экспозиции считается пренебрежимо малым по сравнению с интервалом формирования кадра.

Компьютер блока 3 формирует значение матрицы $T^H(t_k)$ позиции/ориентации головы для момента времени t_k и вычисляет желаемое значение матрицы позиции/ориентации TV-камеры $T_d^C(t_{k+1})$ перемещаемой роботоподобным устройством для момента t_{k+1} , путем экстраполяции значения матрицы $T^H(t_k)$ на момент t_{k+1} . Это значение передается блоку 2 системы управления мобильными TV-камерами и, является заданием для него, которое должно отслеживаться платформой с TV-камерами.

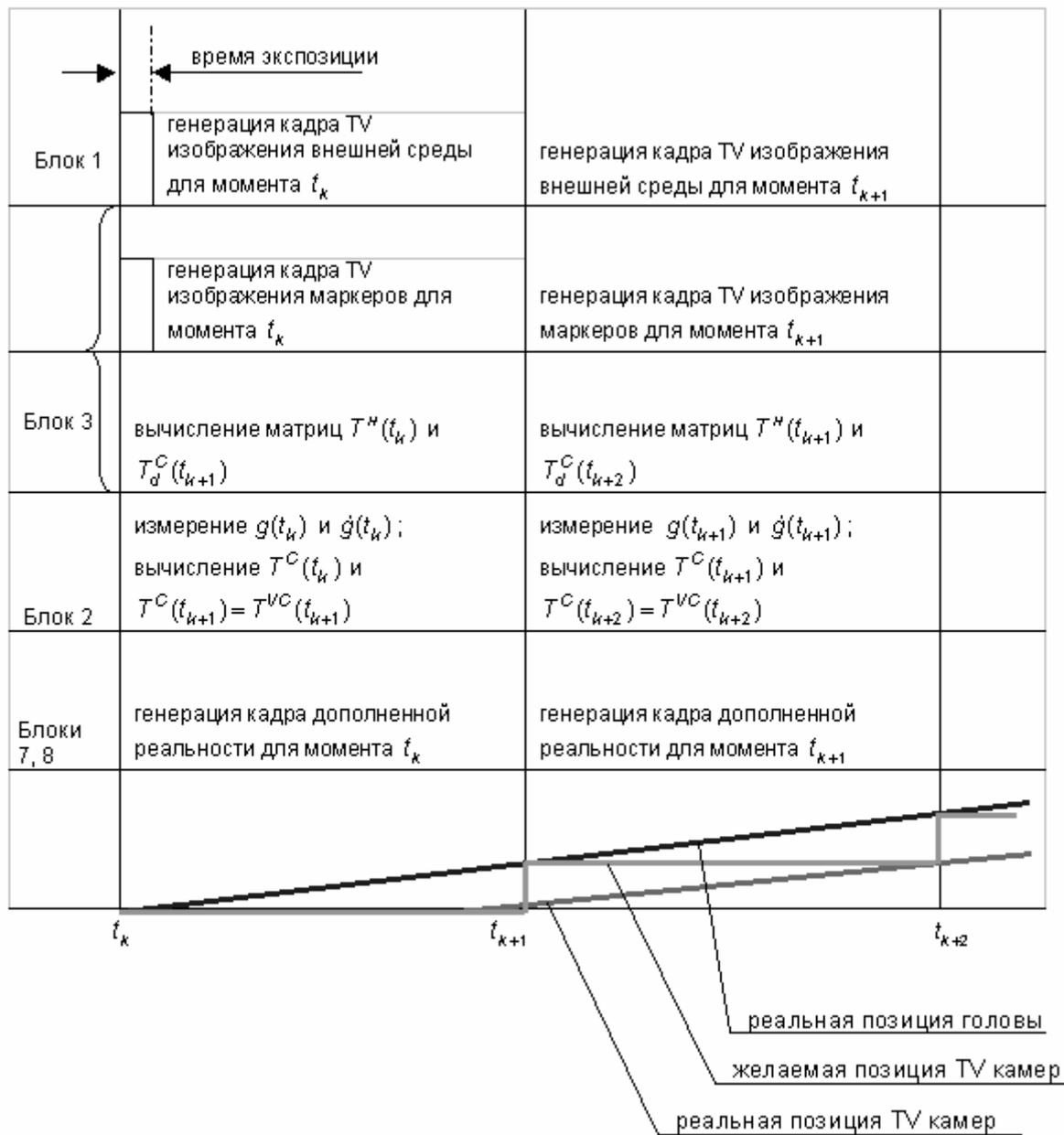
Так как величина $t_{k+1} - t_k = \Delta t$ при предлагаемом подходе известна, то ожидаемое на момент t_{k+1} значение матрицы T^C , определяется по формуле:

$$T_d^C(t_{k+1}) = T^H(t_k) \delta T_d^H(\Delta t),$$

$$\text{где } \delta T_d^H(\Delta t) = \begin{bmatrix} 1 & -\delta_3 & \delta_2 & \Delta_1 \\ \delta_3 & 1 & -\delta_1 & \Delta_2 \\ -\delta_2 & \delta_1 & 1 & \Delta_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \text{ — приращения матрицы,}$$

$\Delta_1, \Delta_2, \Delta_3, \delta_1, \delta_2, \delta_3$ — известные малые смещения и повороты системы координат камеры относительно осей базовой системы координат за истекший период времени $t_k - t_{k-1} = \Delta t$, который равен периоду $t_{k+1} - t_k = \Delta t$.

Предполагается также, что вычислительные мощности компьютера позволяют в течение интервала Δt произвести все необходимые вычисления.



Р

Рис. 17. Временные диаграммы работы основных блоков экспериментального аппаратно-программного комплекса.

Блок 2 обеспечивает опрос в момент t_k датчиков суставных координат $g(t_k)$ и суставных скоростей $\dot{g}(t_k)$ (предполагается, что временной интервал опроса достаточно мал) и по этим значениям путем решения прямой задачи геометрии и кинематики вычисляет матрицу $T^C(t_k)$ положения/ориентации TV-камеры для момента t_k и матрицу $\dot{T}^C(t_k)$, недиагональные элементы которой являются скоростями изменений элементов матрицы $T^C(t_k)$ в момент времени t_k , а диагональные элементы — единицы, а также этот блок вычисляет матрицу $T^C(t_{k+1})$ — ожидаемых значений матрицы T^C в момент времени t_{k+1} .

Матрица $T^C(t_{k+1})$ должна быть равна матрице позиции/ориентации виртуальной камеры $T^{VC}(t_{k+1})$, которой должен соответствовать формируемый для момента t_{k+1} блоками 7 и 8 кадра, дополненного виртуальным манипулятором реального изображения внешней среды.

Матрица $T^C(t_{k+1})$ вычисляется по формуле:

$$T^C(t_{k+1}) = T^{VC}(t_{k+1}) = T^C(t_k) \dot{T}^C(t_k) \Delta t,$$

которая предполагает реалистичную гипотезу: скорость $\dot{g}(t_k)$ в течение интервала Δt неизменна.

Таким образом, в момент t_{k+1} как впрочем и любой другой момент, начала формирования TV кадра изображения реальной внешней среды начинает синтезироваться кадр компьютерного изображения соответствующего геометрической модели внешней среды. Масштаб и ракурс его соответствует позиции/ориентации реальной TV с точностью до используемой правдоподобной гипотезы о неизменности скорости перемещения TV-камеры на интервале $[t_k, t_{k+1}]$. Это позволяет свести динамическую ошибку наложения к минимуму.

Очевидным недостатком предлагаемого подхода является требование очень сильной освещенности рабочей зоны, без которой невозможно уменьшить экспозицию.

Возможным выходом являются использование импульсных источников освещения, момент включения которых синхронизирован с моментом начала развертки кадра TV кадра.

К сожалению, нам не удалось практически реализовать на созданном аппаратно-программном комплексе предложенный подход уменьшения динамической ошибки совмещения, порожденный размытостью видеоизображения. Это объясняется следующим: цикличность алгоритма блока 2 управления перемещениями TV-камерами обзора внешней среды, равная 32 мсек, не совпадает с периодичностью формирования TV-изображений внешней среды и маркеров HTS. Это не позволило синхронизировать функционирование циклического алгоритма блока 2 с алгоритмом компьютера формирования координат позиции/ориентации головы, входящего в блок 3, что сделало бессмысленным использование TV-камер с малым временем экспозиции.

В существующем варианте созданного аппаратно-программного комплекса проверки предложенной технологии «погружения» для достижения приемлемой динамической ошибки совмещения компьютерного и TV изображения внешней среды скорость вращения головы наблюдателя не должна превышать 3–4 град/сек. При этом динамическая ошибка совмещения не превышает 5–8 пикселя.

5.2. Тактильно-силовой эффект погружения

Эффект погружения виртуального объекта в реальный мир экспериментально исследовался, как для случая, когда виртуальным объектом являлось твердое тело, так и для случая, когда виртуальным объектом являлся телеуправляемый манипулятор.

В первом случае для исследования использовались следующие устройства:

К ним относятся:

- роботоподобный механизм с шестью степенями свободы, оснащенный 6-ю приводами, захватным устройством, в которое помещаются сменные муляжи виртуального тела, а также 6-ти компонентным датчиком, измеряющим векторы момента и силы, прикладываемые к муляжу рукой человека;
- устройство управления приводами роботоподобного механизма, программное обеспечение которого формирует управление приводами, обеспечивающее такое перемещение муляжа тела в пространстве, как будто это реальное тело, имеющее заданную массу и матрицу инерции.

Для проверки реалистичности тактильно-силового взаимодействия с имитатором виртуального тела было использовано тело в форме додекаэдра. Он должен был перемещаться под действием приложенных к нему сил.

Было проверено взаимодействие с телами массой 0,05; 0,2; 0,6; 1,2; 4,0 тонны. Матрицы инерции были выбраны такими, что вышеупомянутые массы соответствовали додекаэдрам, сделанным из алюминия. Рис. 18 показывает процесс взаимодействия рук человека с имитатором тела, сделанном из пенопласта.

Эксперименты показали высокий реализм взаимодействия, тем больший, чем больше масса тела. В первую очередь это объясняется весьма малым интервалом времени формирования управляющих заданий приводам роботоподобного устройства, включающим затраты на измерение силы и момента, приложенного к имитатору, на интегрирование уравнений динамики с целью получения векторов угловых и линейных скоростей тела, на определение вектора суставных координат роботоподобного устройства и ввод этих величин в качестве заданий для сервоприводов. Этот интервал не превышает 10 мсек., поэтому человек воспринимает процесс взаимодействия как непрерывный процесс. Проверка ошибки силового взаимодействия осуществляется следующим способом.



Рис. 18. Взаимодействие руки человека с имитатором виртуального тела.

Человек прикладывает силу и момент к имитатору тела, закрепленному на конце роботоподобного устройства (рис. 18). Начальная скорость тела равнялась нулю. Тело ускорялось под действием силы, затем оно двигалось с постоянной скоростью после окончания приложения силы. После чего человек прикладывает силу, чтобы погасить скорость тела до нуля. Интервал движения тела был равен 15 сек. Время приложения человеком силы к телу было меньше 1 сек.

Ошибка физической имитации движения тела определялась как отклонение измеренной длины реальной траектории движения тела от длины идеальной траектории, отнесенных к длине идеальной траектории.

Идеальная траектория вычислялась путем численного интегрирования уравнений движения тела.

Эксперименты были проведены для различных вариантов массоинерционных характеристик тел. График, приведенный на рис. 19, иллюстрирует вышеуказанную ошибку имитации как функцию массы тела. Ошибка уменьшается с ростом массы и составляет 2,5 % для тела массой 1,2 тонны и 1,1 % для тела массой 4 тонны.

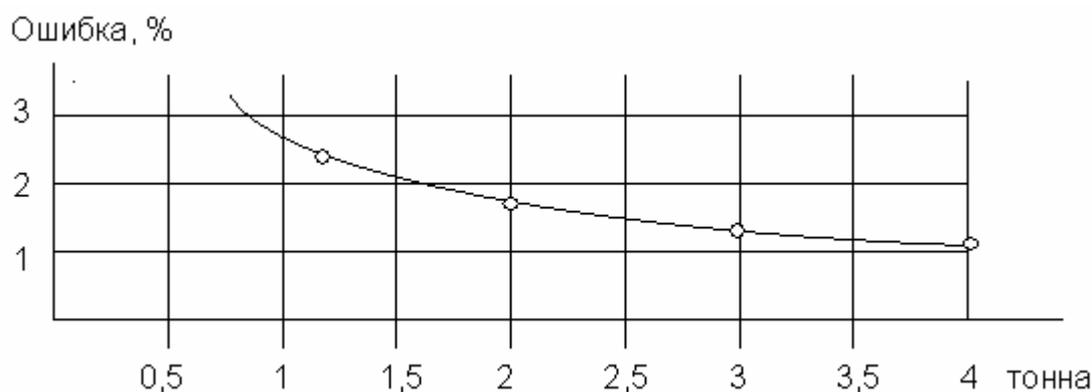


Рис. 19. Ошибка имитации движения тела как функция его массы.

Кроме того, было выявлено следующее:

- вышеуказанная ошибка имитации имеет место только в случае кратковременного приложения силы (не более 1,5 сек.). В случае более длительного приложения силы возникает нарушение процесса моделирования силового взаимодействия, которое обусловлено динамическими свойствами системы управления роботоподобным устройством, которые в этом случае начинают проявляться;
- имитация движения тела малой массы затруднительна, так как тело движется очень быстро и выходит за пределы зоны возможных положений тела, обусловленной конструкцией роботоподобного устройства.

В случае, когда в качестве виртуального объекта использовался манипулятор с копирующим дистанционным управлением, целью эксперимента являлось выяснение качества переходных процессов установления силы, возникающей на задающей руке при столкновении виртуального манипулятора с твердой поверхностью. Полагалось, что в этом случае к рабочему инструменту виртуального манипулятора прикладывалась сила, являющаяся ступенчатой функцией времени. Эта сила должна быть «повторена» на рукоятке задающей руки, которую схватил и жестко удерживает человек-оператор.

Чем точнее «повторение», тем реалистичнее воспроизводится эффект погружения.

В идеале функциональная зависимость этой силы от времени также должна быть ступенчатой функцией, идентичной той, какой удовлетворяет сила приложения к рабочему инструменту виртуального манипулятора.

Для проведения эксперимента использовались устройства, обозначенные на рис. 1 блоком 3. Технические характеристики этих устройств были приведены выше.

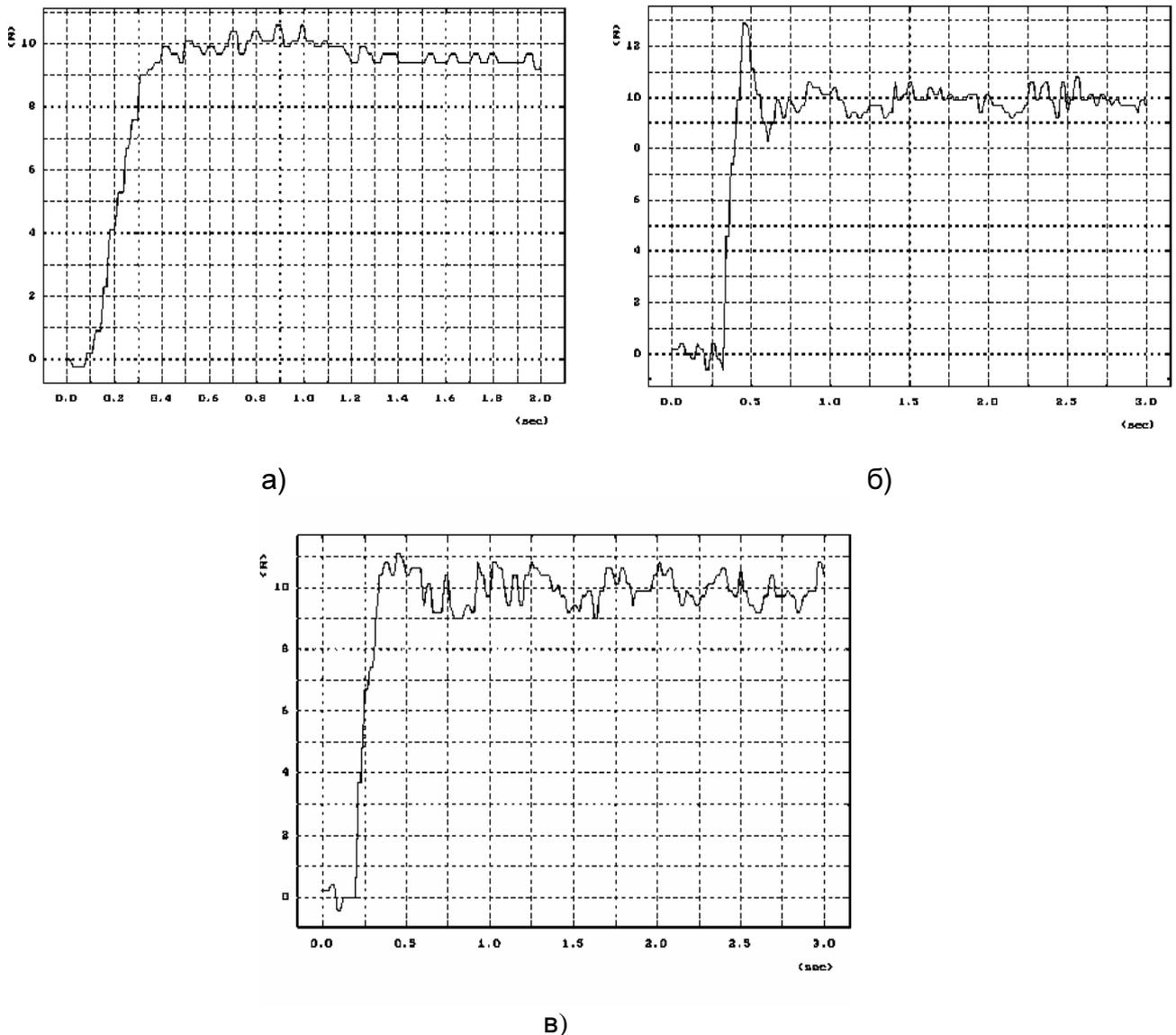


Рис. 20. Переходные процессы установления контактной силы; желаемая величина силы $G_d = 10H$: а) матрица усиления $k_1 = k_2 = 10^4$, б) матрица усиления $k_1 = 1,5k_2 = 10^4$, в) матрица усиления $k_1 = 0,5k_2 = 10^4$.

Система управления задающей руки, реализованная с помощью вычислительных средств устройства СФЕРА-36, построена в соответствии с методом, подробно описанным в [2].

Задающая рука была снабжена 6-мерным запястным силомоментным датчиком. Матрица жесткости датчика диагональная $C_W = \{2 \cdot 10^3, 2 \cdot 10^3, 2 \cdot 10^3, 2 \cdot 10^4, 2 \cdot 10^4, 2 \cdot 10^4\}$.

Жесткость трансмиссии, звеньев руки являются такими, что трансмиссии звеньев можно считать абсолютно жесткими по сравнению с запястьем манипулятора.

Уравнения связи имеют форму (4) в [2], где $K = n = \|010000\|$ — (1×6) –матрица связей, закон управления:

$$U = k_1 J^T n^T (x_e - x_{de}) + k_2 J^T (I - n^T n) (x_{dr} - x_r),$$

где $x_{dr} = const = 0$; $x_{de} = L_W^T C_W^T L_W G_d$; $G_d = (0, 1, 0, 0, 0, 0)$; $x_{de} = (0, 10^{-2}, 0, 0, 0, 0)$.

Рис. 20а, 20б, 20в представляют переходный процесс установления контактной силы в случае, если желаемой величиной этой силы является ступенчатая функция G_d . Были использованы нижеследующие величины матрицы коэффициентов усиления $k_1 = k_2 = 10^4$; $k_1 = 1,5k_2 = 10^4$; $k_1 = 0,5k_2 = 10^4$.

Эксперименты показали, что вместо ступенчатой функции имеет место экспонента с постоянной времени 0,6 сек. Человек воспринимает это взаимодействие как ступенчатое.

6. Заключение

В результате проведенного исследования было сделано следующее:

1. Предложен перспективный подход к решению проблемы погружения виртуального объекта в реальный мир. Этот подход позволяет разработать информационную технологию, позволяющую обеспечить комбинацию как визуального восприятия погружения, так и тактильно-силового восприятия.
2. Предложенный подход наиболее адаптирован к ситуации, когда реальная внешняя среда находится на значительной дистанции от человека, а в качестве виртуального объекта используется либо твердое тело, с которым должны взаимодействовать человеческие руки, либо телеуправляемый манипулятор.
3. Были разработаны методы, алгоритмы, а также аппаратно-программные средства, реализующие предложенный подход с меньшими затратами по сравнению с известными решениями и обеспечением более реалистичного визуального и тактильно-силового эффектов восприятия.
4. Был изготовлен прототип экспериментального комплекса, который позволяет проверить и оценить степень реальности визуального и тактильно-силового погружения виртуального тела в реальный мир.
5. Была проведена экспериментальная проверка предложенного подхода, которая показала нижеследующее:
 - а) При предложенном подходе, базирующемся на наблюдении внешней среды с помощью видео камеры, легко достигается стандартная частота смены кадров генерированной дополненной реальности (Augmented Reality), что обеспечивает «непрерывность» визуального восприятия; причем возможно значительное усложнение графики;

- b) Статическая точность совмещения (registration) изображения реальной внешней среды с геометрической моделью внешней среды не превышает 2–3 пикселя, при условии компенсации нелинейных искажений TV-камеры, а также приведения тщательной калибровки позиции/ориентации реальных TV-камер с целью уменьшения ошибки позиционирования до 0,1 мм и угловой ошибки до 0,005 рад., а также уточнения угла обзора (FOV).
- c) Динамическая ошибка при используемой аппаратуре экспериментального комплекса достаточно высока; она составляет 4–5 пиксел при угловой скорости вращения головы, не превышающей 4 град/сек; было показано, что уменьшение динамической ошибки требует использования TV-камер с малой экспозицией, а также синхронизации работы циклических алгоритмов на всех параллельно работающих компьютерах, реализующих предложенный подход; частота синхронизации должна быть равна частоте формирования TV кадров.
- d) тактильно-кинестетический эффект имеет высокую степень реалистичности как в случае взаимодействия человека с виртуальным телом, так и в случае копирующего управления виртуальным манипулятором.

Литература

- [1] Кулаков Ф. М. , Смирнов Е. Н. , Липатов А. Е. , Латыпов В. Н. . Информационная технология добавления виртуального объекта в реальный мир // Труды СПИИРАН. Вып. 2 — СПб: СПИИРАН, 2004.
- [2] Кулаков Ф. М. Робастное управление податливым движением роботов с упругими элементами // М: Известия РАН, "Теория и системы управления", N 4, 2000.