

оценки стоимости и качества инновационных решений на проектных этапах жизненного цикла высокотехнологичных изделий машиностроительных производств // Известия МГТУ «МАМИ». - 2012, № 2 (14), т. 2, с. 118-124.

Система бесконтактного копирующего H2C-интерфейса для управления антропоморфным роботом

к.ф.-м.н. Идиатуллоев Т.Т., Романцов Н.А., Чабаненко Е.Б.

Университет машиностроения

8 (495) 223-05-23, timid@mami.ru, nikittik@gmail.com, abc437@rambler.ru

Аннотация. В статье рассмотрены проблемы создания и функционирования человекоориентированных систем управления, в которых используются бесконтактные жестовые интерфейсы. Проведен анализ существующих подходов и сделаны предположения о наиболее эффективных подходах к построению таких систем.

Ключевые слова: человекомашинные интерфейсы, распознавание образов, автоматизированные системы управления, управление техническими системами, пользовательский интерфейс, когнитивная психология, юзабилити, искусственный интеллект

Разработка человекомашинных интерфейсов, безусловно, это одно из важных направлений развития современной техники. С момента появления первых паровых и электрических машин и до наших дней необходимо было обеспечить возможность выдать оператору какую-либо информацию о состоянии оборудования и предоставить ему возможность выполнить управляющие действия.

Естественно, что ранние «интерфейсы» были весьма примитивными и требовали специальной и весьма длительной подготовки, а иногда даже и тренировки. Оператор должен был понимать суть происходящих в системе (машине) процессов, чтобы в зависимости от показаний достаточно примитивных индикаторов, таких как стрелочные манометры, управлять ей, переводя бесчисленные рычаги и вентили в нужные позиции. И при этом достаточно часто цена ошибки была весьма велика – ни о каком автоматизированном интеллектуальном контроле, естественно, не могло быть и речи.

В этом контексте, конечно, крайне завораживающе выглядят фантастические «стимпанковские» механизмы, которые с помощью рычагов, часовых механизмов, паровых, пневматических и гидравлических приводов могут выполнять весьма сложные и разнообразные движения. Вершиной инженерной мысли «доэлектронной» эры стали автоматы и, как ни странно, автомобильные коробки передач. И если автоматы по своей сути – это очень сложные часовые механизмы, то гидромеханические автоматические коробки переключения передач способны реагировать на значительное число внешних факторов и могут включать даже системы кратковременной памяти в виде гидроаккумуляторов.

С приходом электрических, а впоследствии и электронных систем, появилась возможность снизить трудоёмкость управления механизмами и машинами. В производство стали внедряться автоматизированные и даже автоматические системы управления. Но по-прежнему консоль управления, основной элемент интерфейса между человеком и машиной, оставалась перегруженной. Оператор должен был понимать принципы функционирования устройства, которым должен был управлять. Именно поэтому возникло большое число специальностей «оператор чего-нибудь». Многие, даже схожие по назначению механизмы, работали и управлялись по-разному, а значит, без дополнительной подготовки правильно управлять ими было практически невозможно.

Развитие теории построения интерфейсов до определённого момента было связано с совершенствованием технологий производства электронных устройств. Появление компакт-

ных электрических лампочек позволило подсвечивать символические изображения действий на пультах управления. Создание сложных переключателей позволяло менять поведение системы в разных режимах. А развитие датчиковой аппаратуры позволило создавать системы, реагирующие на действия пользователя, в том числе бесконтактным способом. Хорошим примером может являться механизм открытия дверей, снабженный инфракрасным датчиком движения – приближение человека приводит к срабатыванию электропривода раскрытия дверей.

Однако настоящий прорыв наступил с появлением цифровой микропроцессорной техники. Способность синтезировать изображение произвольных объектов на графических мониторах дает возможность формировать интерфейсы практически для любых ситуаций. Появилась даже целая отрасль – проектирование пользовательских интерфейсов.

Высокая вычислительная мощность компьютеров позволяет создавать достаточно отзывчивые интерфейсы, которые хорошо сочетаются с сенсорными устройствами ввода. Кроме того, не представляет сложности добавить акселерометры, гироскопы и прочие устройства ввода информации, не считая различных клавиатур и позиционных указателей.

Рассмотрим, как сейчас выглядит типичный процесс управления технической системой, снабжённой компьютеризированным интерфейсом. Но взглянем на него в разрезе этапов, формирующих модель поведения оператора.

Первый и самый неочевидный для многих этап – это инициирование самого взаимодействия человека и вычислительной системы. «Неочевидный» потому, что, как правило, компьютеры большую часть своего времени занимаются тем, что ожидают действий пользователя. При этом обычно устройства вывода (чаще всего монитор) формируют «приглашение» в виде какого-нибудь интерфейса пользователя. И это означает, что именно проектировщик пользовательского интерфейса влияет на возможные сценарии поведения оператора на первом этапе.

Именно синтетические объекты, из которых состоит пользовательский интерфейс, а также доступные устройства ввода, определяют, какие действия может выполнить пользователь. И грамотно разработанная система должна представить ему однозначно интерпретируемые элементы управления. Например, нарисованные на экране кнопки или переключатели, скорее всего, служат для запуска каких-либо действий или для переключения состояний.

Правильно подобранные ассоциации с реальными объектами позволяют разработчику программного обеспечения обрабатывать только ожидаемые действия пользователя над соответствующими виртуальными образами, а не изобретать способы реагирования на его нешаблонное поведение. И подобная детерминированность общения, конечно, значительно упрощает общую логику взаимодействия оператора и вычислительной системы. Но, что очень важно, работать ожидаемым образом данная система будет только если взаимодействие инициировано со стороны вычислительной системы – именно она предлагает человеку выбрать действие из предложенных вариантов. Такую модель передачи управления можно обозначить как «компьютер-человек-компьютер» (C-to-H-to-C, C2H2C).

Конечно, в очень большом числе случаев информационные системы не предполагают ответных действий пользователя. К примеру, различные мультимедийные панели или световые экраны просто генерируют информацию и передают её человеку-потребителю. Это модели «компьютер-человек» (C-to-H, C2H).

Из повсеместного распространения этих моделей существует одно «технологическое» следствие – человек не может отдать «неожиданные» команды, то есть такие, для которых не предусмотрен какой-нибудь шаблон их ввода вроде «тычка» (touch) или «смахивания» (swipe). Это значит, что возможность ввести неустановленную команду связана исключительно с посредственной проработкой логики пользовательского интерфейса, в том числе взаимодействия «управляемых» объектов. Чаще всего подобное происходит тогда, когда разработчик использует в интерфейсе какой-нибудь заимствованный элемент управления

(control), часть функциональности которого ему неизвестна. Например, это может быть кнопка, которую можно «перетаскивать», но об этом не знает наш разработчик, благодаря чему пользователь сможет «модифицировать» интерфейс, «утащив» кнопку за край экрана.

Разумный разработчик будет выбирать наиболее простые «устройства ввода» - такую концепцию интерфейса, которая будет допускать как можно меньше свободы оператора. Желательно, чтобы пользователь сразу «понимал» ожидаемое действие, например натягивал рогадку для выбора стартовой скорости снаряда.

Именно простота восприятия принципов управления привела к тому, что в настоящий момент графические сенсорные экраны, симулирующие интерфейсы типа «touch, drag and swipe», практически вытеснили все остальные универсальные устройства ввода-вывода.

Однако остаётся значительное количество ситуаций, когда описанные модели «не работают». Например, когда примитивный интерфейс предполагает переусложнённую схему действий из десятков последовательных элементарных шагов, таких как адресные книги в мобильных телефонах, которые требуют много «нажатий» для поиска нужного контакта. Или ситуации, когда нужная операция оказывается неочевидной, или нет технической возможности использовать «удобный для разработчика» интерфейс. Хороший пример – телефонные автоинформаторы со словами: «Нажмите единицу, если Вы хотите дождаться ответа оператора».

Такие схемы взаимодействия, в которых инициатором взаимодействия выступает именно человек, когда компьютер либо не может синтезировать понятные объекты управления, либо их использование не отвечает задачам пользователя, можно отнести к моделям «человек-компьютер-человек» (H-to-C-to-H, H2C2H) и «человек-компьютер» (H-to-C, H2C).

		инициатор	
		человек (оператор)	компьютер
контролер	человек (оператор)	H2C	C2H2C
	компьютер	H2C2H	C2H

Рисунок 1. Квадрат моделей информационных потоков

Эти четыре способа организации взаимодействия оператора и вычислительной машины образуют квадрат моделей информационных потоков (рисунок 1). В пространстве вовлечённости участников в организацию и поддержание взаимодействия модели представляют собой направленные отрезки или ломанные линии, положение узловых точек которых зависит, к примеру, от степени абстракции передаваемой информации (рисунок 2). Здесь, безусловно, открывается широкое поле для исследований в области когнитивистики и юзабилити.

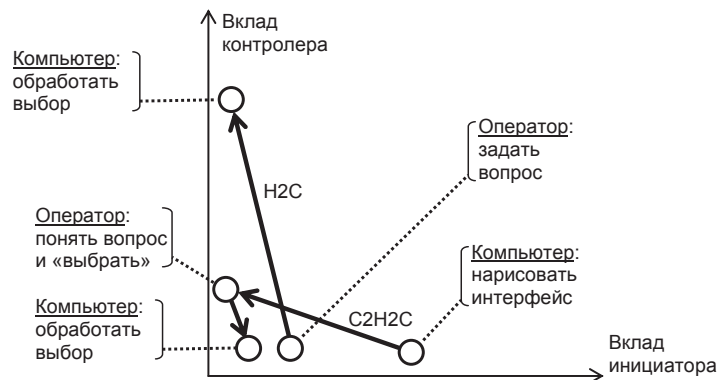


Рисунок 2. Пространство вовлечённости (вклада в осуществление взаимодействия)

Конечно, следует понимать, что внедрение бесконтактных и голосовых интерфейсов

всё-таки не переводит информационную систему в разряд поддерживающих модели H2C2H и H2C. Требуется модификация логики взаимодействия системы и потенциального пользователя. Вычислительная система должна научиться определять желание человека начать взаимодействие, а затем отреагировать на его управляющие действия. И только если потребуются уточнение команды, может быть использован синтезированный интерфейс.

При этом ничто не мешает компьютеризированной системе управления демонстрировать ожидание управления от потенциального пользователя просто как «готовность к общению». Но желательно, чтобы интерфейс непрерывно демонстрировал результаты текущего распознавания действий человека, пусть даже как спектрограмму поступающих акустических данных. Примерно в таком ключе функционируют современные системы голосового управления, такие как Apple Siri и Google Now.

Однако голосовой интерфейс не исчерпывает все возможные подходы к построению H2C и H2C2H систем.

С развитием робототехники и, в частности, с появлением интеллектуальных систем управления всё больший интерес начинает привлекать возможность использовать естественные для человека невербальные действия. Ведь применение роботов, к примеру, в домашнем хозяйстве предполагает ориентацию на крайне низкий уровень специальной технической подготовки пользователей. В идеале собственник робота вообще не должен изучать какие-либо «методы управления». А естественными для человека являются голосовые и жестовые методы коммуникации.

Однозначную полезность жестов для передачи информации и управления демонстрируют флажковая флотская азбука, сурдоперевод, жестикуляция строителей и спасателей. Существует множество областей жизнедеятельности, для которых люди изобретали свои системы жестикуляции, которые с некоторой степенью приближения можно интерпретировать как языки передачи команд.

Естественными препятствиями на пути внедрения жестовых методов управления являются как неоднозначность значительного числа ситуативных движений человека, непосредственно не связанных с отдачей команд, так и задача детектирования самих движений. И если первая требует специальных исследований, то для преодоления второй существует два пути. Первый (контактный) – создание экзоскелетов или каких-нибудь иных наборов носимых датчиков, а второй (бесконтактный) – анализ видеоизображения оператора с целью сопоставления его с виртуальной моделью скелета человека.

Второй подход нашёл техническое воплощение в проекте Kinect компании Microsoft, который привёл к созданию одноимённого контроллера для игровой приставки XBOX 360 (XBOX ONE). К настоящему времени компания распространяет специальный комплект разработчика, позволяющий использовать этот контроллер с компьютерами под управлением операционной системы Windows. Для большей определённости, в технических описаниях весь комплекс принято разделять на сенсор Kinect, систему анализа сцены и систему трекинга фигур.

Сам по себе сенсор Kinect является устройством для прямого дистанционного трехмерного измерения объектов и сцен, совмещенным с цветной видеокамерой. В процессе работы сенсор строит «карту глубины» для наблюдаемой сцены (рисунок 3).

Kinect создает карту глубины в два этапа. Чтобы получать информацию о трехмерном представлении пространства, Kinect проецирует перед собой сетку из инфракрасных точек. Далее, проводится считывание построенной сетки с помощью встроенной инфракрасной камеры. За счёт того что эмиттер сетки и камера разнесены на некоторое расстояние, считанная сетка имеет искажения, связанные с удалённостью освещаемых объектов. В дальнейшем с помощью метода триангуляции смещения преобразуются в оценку удалённости отражающей поверхности от сенсора (рисунок 3).

На основании полученных данных программное обеспечение Kinect строит карту глу-

бины. Затем на основе этой карты можно может быть выполнено распознавание позы человеческого тела как набора точек виртуального скелета оператора (рисунок 4).

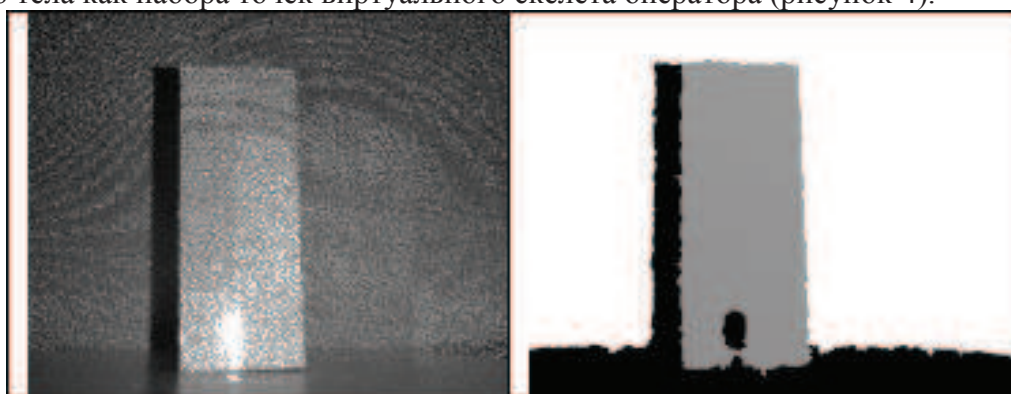


Рисунок 3. Пример анализа сцены с построением карты глубины

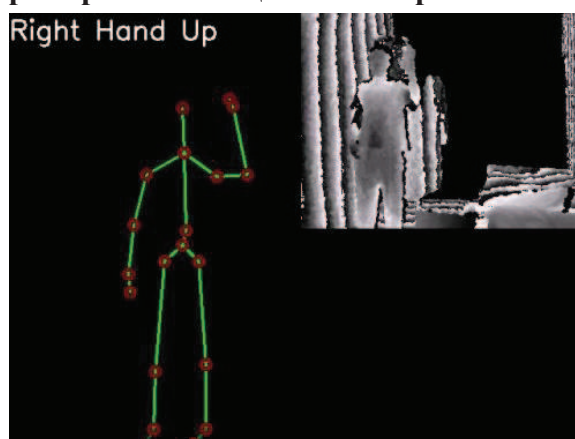


Рисунок 4. Трекинг (отслеживание) виртуального скелета (фигуры) оператора

Результаты трекинга фигуры представляют собой массив трёхмерных координат узлов скелета, которые могут быть использованы для анализа поз и жестов оператора.

Использование данных трекинга скелета возможно, к примеру, в задаче жестового управления летающей мобильной платформой – квадрокоптером.

Очень часто в реальной практике авиации возникают ситуации, когда для безопасной посадки необходимо взаимодействие пилота и наземного оператора, а в некоторых случаях передача управления вообще могла бы считаться оптимальным решением. И поскольку использование специального пульта для телеуправления не всегда возможно, то распознавание жестов бортовой вычислительной системой могло бы решить проблему передачи команд.

Например, экспериментальная установка в режиме телеуправления интерпретирует движения рук оператора, стоящего напротив сенсора Kinect, как команды движения платформы, её поворота и тангажа. Например, подъем двух рук вверх приводит к взлёту коптера, а противоположное движение – к посадке на поверхность. Схема работы данной системы и фрагмент интерфейса приведены на рисунке 5.

Подобная методика управления, предусматривающая интерпретацию достаточно простых жестов в набор элементарных команд, может быть адаптирована практически под все типы подвижных платформ, например, для управления автомобилем или башенным краном.

В простейшем случае «язык тела» оператора будет состоять даже не из жестов, а из статических поз, которые будут различаться таким программным обеспечением, как нажатие клавиши условного пульта управления. И таких клавиш может быть тем больше, чем точнее удастся различать расположение различных частей тела оператора и чем больше их (частей тела) будет использовано.

Как правило, программная реализация системы управления предполагает регулярный «опрос» модуля распознавания позы оператора с целью определения момента её смены. Как

только поза станет устойчиво детектироваться и совпадёт с одним из шаблонов, в систему будет передана команда управления.

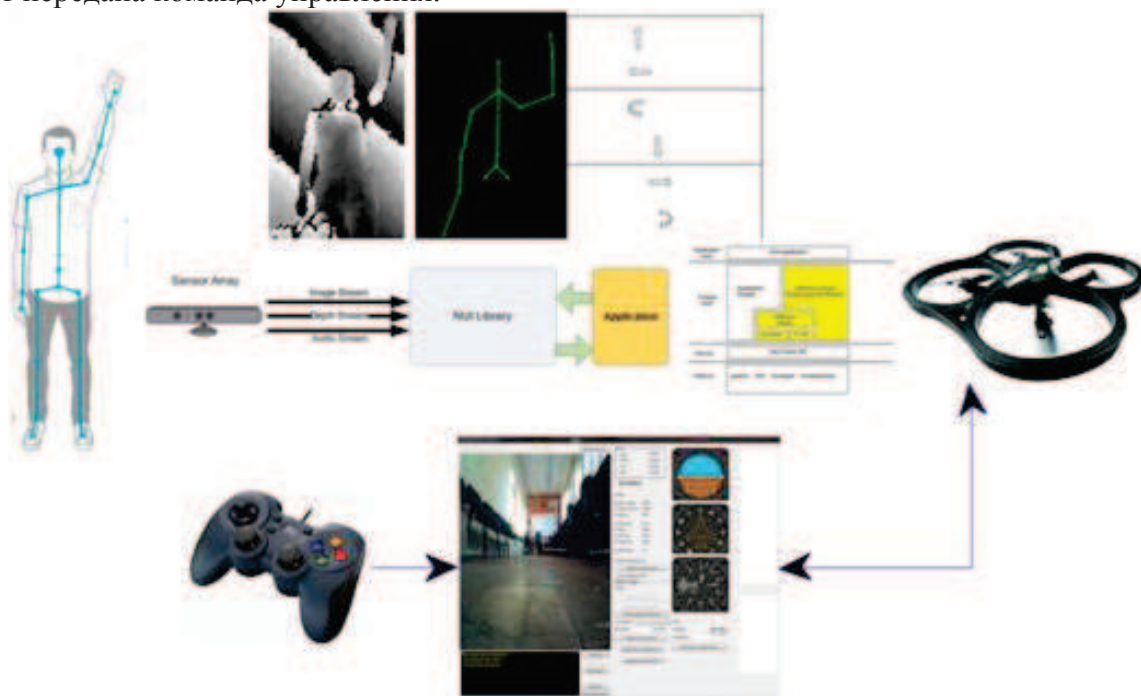


Рисунок 5. Схема управления и интерфейс системы полётного контроля

Для устойчивого и надёжного функционирования система шаблонов распознавания должна обладать двумя характеристиками:

- «управляющие» позы должны хорошо различаться, несмотря на то, что человек не может достаточно точно воспроизводить положение частей тела и, главное, сохранять их неподвижность продолжительное время;
- должна существовать возможность «выключить» управляющее воздействие, сознательно «сломав» позу, например опустив руку, причём не вызвав при этом ложное срабатывание какой-нибудь другой команды.

Очевидный вариант построения такой системы – это введение нескольких активных зон вокруг силуэта оператора, заметно отстоящих друг от друга, и при этом достаточно крупных, чтобы удовлетворить условиям надёжного управления.

Экспериментальная проверка показывает, что при «силуэтном» распознавании, то есть при проецировании фигуры оператора на плоскость, перпендикулярную лучу зрения, активных зон всего восемь. При этом «верхняя» и «нижняя» зоны, то есть «руки над головой» и «руки вдоль тела» являются неустойчивыми по отношению к зеркальной симметрии – различить, какая из рук поднята, достаточно сложно. А следовательно, «цепные» инструкции, связывающие отдельные команды, должны строиться с учётом процесса смены поз с «блокировкой» интерпретации позы до остановки движения оператора.

Более прогрессивным является принцип распознавания именно жестов, то есть «рисунка» движения конечностей оператора. При этом «стартовая поза» может быть даже общей (например, руки в стороны), а характер движения определит реакцию системы. Скажем, движение руками вверх – набор высоты. Подобный подход позволит серьезно увеличить количество возможных команд.

Но для управления антропоморфными роботами этого может оказаться недостаточным. Дело в том, что человек интуитивно ассоциирует элементы структуры робота с частями человеческого тела. Поэтому для подобного класса роботов может казаться более эффективным применение систем управления копирующего типа.

Бесконтактными копирующими интерфейсами можно считать определенное семейство

Н2С-интерфейсов, которые предполагают формирование инструкций по перемещению активных зон устройства через демонстрацию таких движений человеком-оператором. В зависимости от конструкций, роботизированная платформа может либо воспроизводить движения человека, либо определять «ожидаемые пользователем действия», которые необходимо выполнить.

Для построения таких систем в последовательность обработки данных добавляется несколько этапов. На следующем этапе, после «считывания» скелета, массив координат узлов «виртуального скелета» оператора накладывается на массив точек, ассоциированный с опорной поверхностью (полом), и на основании этого производится расчёт базисных матриц и матриц поворота зоны опоры и торсового узла.

Завершающей фазой обработки является «вписание» виртуальной модели антропоморфного робота в имеющуюся систему узлов с учетом возможностей его кинематики и необходимостью правильного размещения опорных поверхностей (стоп) для устойчивого стояния.

Выбранный в качестве исполнительного механизма робот *Bioid* в модификации *C* имеет такое расположение осей вращения узлов, что прямая кинематическая задача для конечностей имеет простое решение. После соответствующей математической обработки получается массив углов поворотов осей сервоприводов, которые обеспечивают принятие роботом нужной позы. Таким образом, робот как бы «зеркалит» оператора, копируя положение его тела.

Разработанный метод может быть использован при создании систем программного управления роботизированных систем, чтобы дать возможность оператору задавать движения робота через их непосредственный «показ». Это позволит экономить много времени при настройке роботизированной системы на определенный характер движения.

А также при решении задачи динамической стабилизации робота при быстрой смене поз можно реализовать систему непосредственного управления движением робота путем копирования движения оператора.

По сравнению с подходом, основанном на использовании специального сенсорного костюма управления, достоинством предложенного метода является реализация системы бесконтактного управления антропоморфным роботом. При необходимости считывание поз оператора может производиться камерой, размещённой на самом роботе, что позволяет сделать робота достаточно автономным.

В дальнейшем планируется внедрить в программное обеспечение робота возможность полноценного управления всеми перемещениями с помощью различных жестов и наклонов корпуса, а также продолжить исследования на других типах робототехнических систем.

Литература

1. Войцеховский Я. Дистанционное управление моделями. Пособие моделиста и радиолюбителя. Пер. с польск. / Под ред. А.П. Павлова и Н.Н. Пуяткина. – М.: Связь, 1977, 432 с.
2. Готшалк О.А. Системы автоматизации и управления. Конспект лекций. – С.-Пб.: СЗПИ, 1998, 35 с.
3. Stephane Piskorski, Nicolas Brulez, Pierre Eline. AR.Drone Developer Guide. Parrot S.A. 2012, 107 с.
4. US Patent Application № 20120278904, Microsoft corporation, Redmond, November 1, 2012.

Перспективы развития обкатного зубофрезерования цилиндрических колёс

д.т.н. Калашников А.С., к.т.н. Моргунов Ю.А., к.т.н. Калашников П.А., Котькина Т.В.

Университет машиностроения
8-916-3768356, morgunov@mami.ru

Аннотация. Рассмотрены условия зацепления червячной фрезы с профилем