

Интеграция языковых средств мимики в интеллектуальный человеко-машинный интерфейс

Е.В. Тимофеев

Рассматриваются возможности языка мимики по передаче информации в процессе диалога и преимущества использования мимических языковых средств в человеко-машинном интерфейсе. Формулируются принципы моделирования мимических систем и предлагаются варианты конструирования диалоговой системы, использующей мимическую сигнализацию. Обсуждаются возможности включения невербальной диалоговой компоненты в информационные системы.

1 Язык мимики как важнейший элемент невербального общения

Коммуникация между людьми - процесс сложносоставной и не ограничивается только речевым (или вербальным) способом передачи информации. Более того, в условиях реального общения получение субъектом некоторой вербальной информации приводит к появлению у него определенной информационной потребности, направленной на подтверждение истинности и полноты услышанного, определение отношения собеседника к данной информации и, собственно, к субъекту, а также выявление текущего эмоционального состояния собеседника. Передача информации в вербальной форме как правило подконтрольна человеку, на словах человек сообщает не более того, что он хочет сообщить. Необходимую же дополнительную информацию реципиент может черпать из невербальных сообщений, содержащихся в интонации, мимике, жестах. С другой стороны для полноценного коммуникативного акта (естественного в полноте сообщений, открытого с психологической точки зрения) необходимо задействование как вербальной, так и невербальной компоненты общения. В процессе диалога каждая из сторон имеет свои цели (что-либо объяснить, доказать, развеселить, оскорбить, навязать свою точку зрения и

т.д.) и использует коммуникативный акт для их достижения. Чем полнее коммуникативный акт в смысловыразительном плане, тем с большим успехом достигаются цели участников диалога. Это особенно важно в таком деле как учебный процесс, и на сегодняшний день обучающие системы, в том числе и главным образом компьютерные, не могут являться равноценной по эффективности заменой живого преподавателя, в частности, по игнорированию разработчиками таких систем невербальных компонент образовательного процесса.

Пожалуй, самой важной компонентой невербального общения следует считать мимику. Мимические сигналы легко воспринимаются реципиентом сообщения, поскольку человек полнее усваивает зрительную информацию, их также отнюдь не просто контролировать, по крайней мере в комплексе: "мимическая выразительность воспроизводится как безусловный рефлекс-[3]. Поэтому человек больше склонен обращать внимание на мимику, доверять той информации, которую она несет, и, в случае ведения психологически естественного диалога, сам возлагать на мимику значительную информативную нагрузку. "Лицо представляет собой лишь небольшой отдел головы человека, но является, пожалуй, главной характеристикой его внешности-[4]. Таким образом при организации (моделировании) полноценного коммуникативного процесса необходимо особое внимание уделять мимической сигнализации.

2 Передача информации с помощью языка мимики

Язык мимики довольно сложен и способен нести большую информативную нагрузку. "Существуют тысячи разновидностей выражения лица, каждое из которых отлично от других-[13]. Мимика служит для выражения не только эмоций, которые являются первоисточником информации при генерации мимических сигналов, но и диалоговых сигналов (conversational signals), типа вопросительного выражения лица, мимических символов (facial emblems), например, подмигивание, а также лицевых операторов (facial manipulators), таких, как кусание губ или надутые щеки [13].

Традиционно исследования, посвященные мимике, как правило тесно связаны с работами по психологии эмоций; сама же мимика воспринимается в первую очередь как средство отображения эмоций. Рассматривая различные точки зрения на проблему условий возникновения эмоций можно констатировать их "двойную обусловленность: с одной стороны, потребностями (мотивацией), с другой - особенностями воздействий (ситуации)-[12]. Обобщив данное

положение, мы получаем два основных информационных источника экспрессии лица (имеются в виду не только эмоции): внешний и внутренний стимул. С первым связано реагирование субъекта на некоторую ситуацию, полученную информацию - мимика выражает его отношение к этой информации. С внутренним стимулом связана, во-первых, искусственная мимика, направленная на достижение некоторой цели, сознательно преследуемой субъектом, во-вторых, мимика естественная, чье проявление направлено на достижение неосознанных целей, и, в-третьих, естественная мимика, отражающая психическое, эмоциональное состояние субъекта. Обычное явление - когда на мимику влияют сразу два стимула: и внешний, и внутренний. В этом случае выражение лица несет информацию как об отношении ко внешней ситуации, так и, в неявном виде, о мотивации субъекта, а также о его текущем эмоциональном настрое.

Смысловыразительным свойством обладают не только слова языка мимики в целом (общее выражение лица), но и элементарные мимические составляющие (ЭМС) - улыбка, раскрытый рот, поднятая бровь и т.д. - буквы мимического алфавита. Последние имеют, как правило, некие обобщенные значения; так, приподнятые верхние веки, характерные для выражения страха, возмущения, ненависти и восторга, означают в общем случае сильное волнение. Необходимость дифференцированного восприятия мимики обусловлена и тем, что "изменения лица происходят в трех автономных зонах: "лоб-брови" "глаза" и "рот-губы". Наиболее точно идентифицируются мимические проявления в области рта, наименее точно - в области глаз-[1]. Таким образом, основная информационная нагрузка (имеется в виду фактическая смысловыразительная нагрузка) в мимике приходится на область рта.

3 Искусственные мимические системы

Современное состояние дел в работах по усовершенствованию человеко-машинного интерфейса для компьютерных сред позволяет говорить о моделирующих человеческое поведение системах как о перспективном направлении данной отрасли информатики. Интеграция искусственных мимических систем в диалоговые среды для задач моделирования собеседника является, пожалуй, одним из наиболее эффективных для восприятия и эффективных с информационной точки зрения примеров усовершенствования человеко-машинного интерфейса. Основной задачей такого проекта становится увеличение скорости и полноты усвоения информации пользователем некоторой компьютерной среды. Примерами такого рода сред могут быть системы контроля данных, коммуникационные программы, операционные системы, обучающие,

развивающие и справочные системы.

Повышение скорости усвоения информации в средах, включающих в себя искусственную мимическую систему, возможно за счет того, что мимические сигналы являются графическими образами, воспринимаемыми реципиентом фактически мгновенно. Кроме того эти сигналы обрабатываются на подсознательном уровне, что не требует усилий по их расшифровке. Факторы, отрицательно влияющие на восприятие информации (такие как усталость, монотонность поступающих стимулов), в случае с передачей информации посредством мимики в меньшей степени сказываются на способности достаточно полно усваивать информацию. На полноту ее усвоения также влияет и тот факт, что мимическая система производит эффект "одушевления" компьютера: диалог со средой становится более естественным, оператор получает новизну стимулов, что положительно сказывается на его эмоциональном настрое и соответственно влияет на его работу.

4 Моделирование мимических систем

При моделировании мимической системы человека в первую очередь необходимо определиться с тем, каким образом будут описываться мимические конструкции. Наиболее простой способ - создание графической базы данных, элементы которой будут использованы при визуализации мимики. Другой способ, более сложный для реализации, - задание исходного изображения человеческого лица в терминах некоторого языка программирования или математических функций, что позволило бы затем трансформировать графический вывод мимики по набору правил. Такой путь позволил бы значительно сократить размеры всей системы, однако в остальном его эффективность (полнота, правдоподобие, четкость отображения эмоций) была бы такой же или меньшей, чем при первом способе. Наконец, самым сложным для реализации, но в то же время вызывающим интерес со стороны разработчиков информационных систем, является вариант, когда исходное изображение лица является произвольным - его нужно распознать, чтобы иметь возможность трансформировать для придания нужного выражения. В данной статье предлагается подробнее остановиться на первом способе описания мимики и обсудить проблемы третьего пути.

Сразу отметим, что возможно использование различных вариантов графических баз и, соответственно, механизмов построения мимических конструкций; можно использовать и несколько схем перехода от одного выражения лица к другому. В общем случае система должна

иметь входной информационный канал, по которому поступали бы сообщения на некотором языке (в данном случае актуальным является ограниченный ЕЯ, либо некоторый специальный язык). Система должна обработать эти сообщения и выделить их смысловое содержание; в соответствии с ним на выход подается определенная мимическая конструкция - семантическая проекция полученной информации на языке мимики.

В качестве мимического алфавита можно использовать, во-первых, набор элементарных мимических составляющих (ЭМС) (о них упоминалось выше), во-вторых, множество триад тех же ЭМС, определяющих мимику трех основных лицевых областей ("лоб-брови "глаза "рот-губы"), в-третьих, множество цельных мимических представлений, соответствующих конкретным выражениям лица - своеобразный иероглифический метод. Первый подход предоставляет компактный конечный мимический алфавит, что уменьшает размер графической базы, а также дает большую гибкость в работе с несимметричной мимикой (типа кривой ухмылки или подмигивания одним глазом), упрощая независимый мимический вывод для правой и левой половин лица. Однако в этом случае потребуются сложный механизм визуализации; кроме того, как показала практика, для достижения высокого уровня соответствия между реальной мимикой человека и ее графической репрезентацией понадобится детальная проработка ЭМС для каждой конкретной мимической конструкции, что закономерно приведет к увеличению количества связей в системе и к общему ее усложнению.

Второй подход позволяет формализовать структуру мимической репрезентации, что упрощает систему в целом. Также в данном случае упрощается и механизм визуализации, поскольку в каждом случае вывода мимической конструкции производится вывод только трех легко взаимопозиционируемых компонент. Однако, при том, что размер графической базы вырастает по сравнению с первым случаем, система все также нуждается в механизме комбинирования изображения и таком отношении эмоция-ЭМС, как один ко многим (мимическая составляющая здесь равнозначна элементу графической базы). При третьем подходе это отношение заменяется на один к одному, а визуализация мимической конструкции производится наиболее простым способом. Два главных недостатка этого подхода - большая по объему графическая база и более трудоемкий процесс ее создания.

Мимический словарь может быть двух типов: в качестве заголовков словарных статей могут быть использованы либо эмоции, либо отдельные ЭМС. В первом случае словарная статья должна содержать соответствующие определенной эмоции ЭМС, во втором статью составят те эмоции, для проявления которых характерна данная ЭМС. Для

задач отображения эмоциональной экспрессии больше подходит первый вариант словаря, поскольку при построении той или иной мимической конструкции логичнее идти от требующейся эмоции к ее полному мимическому описанию. Второй вариант словаря больше подходит для задач распознавания мимики, представленной на графическом языке, когда по отдельной мимической составляющей требуется определить, каким эмоциям она соответствует.

Выделение смысла входного текста - отдельная сложная задача. В простейшем случае соответствие между текстом и графической репрезентацией эмоции может задаваться словарем эмоционально окрашенных слов, который может быть расширен путем включения в него устойчивых выражений. При ведении диалога с системой, оснащенной словарем эмоционально окрашенных слов, могут возникать ситуации, когда текстовая конструкция несет смешанную эмоциональную окраску. Один из вариантов решения этой проблемы - использование мимических соответствий смешанных эмоций (например, удивление + ненависть и т.п.). При этом с текстом должна быть ассоциирована лишь некоторая часть эмоций - базовый набор, каждой паре эмоций из которого будет соответствовать своя мимическая конструкция. В подобной системе увеличение базового набора на несколько единиц приведет к расширению всего мимического словаря на десятки статей (при 10 базовых эмоциях потребуется 100 мимических репрезентаций их сочетаний). Это сильно ограничивает возможности разработчика. Кроме того, в полученной текстовой конструкции смешанная эмоциональная окраска должна задаваться не более чем двумя эмоциями, что тоже накладывает свои ограничения. Также следует учитывать, что "основные эмоции идентифицируются легче и точнее составных-[1].

Другой вариант - использовать механизм поглощения эмоций, когда, например, две эмоции (или больше) из мимического словаря преобразуются в третью из того же множества эмоций (необходимость поглощения трех и большего количества эмоций маловероятна). При этом поглощение должно производиться по некоторому набору правил с учетом силы тех или иных эмоций, их знака и т.д. При таком подходе, в отличие от первого случая, увеличение мимического словаря является желательным и, более того, должна быть предусмотрена возможность пополнять эмоциями уже сформированную систему.

Рассматривая работу мимической системы в целом, можно говорить о том, что информация, поступающая из входного канала, представляет собой последовательность сообщений, различных с точки зрения информативной нагрузки и одинаковых по той реакции, на которую они рассчитаны - графический вывод определенной эмоции. При этом обыкновенная смена текущей эмоции на новую является

слишком упрощенным решением, не соответствующим процессам и правилам функционирования человеческой мимической системы, в которой проявляемые эмоции коррелируют с текущим эмоциональным состоянием субъекта. Последнее может определяться как ранее полученными сообщениями, так и эмоциональным настроением данного субъекта, зависящим от некоторых внешних условий, от цикличности естественных эмоциональных состояний, а также и от случайного фактора. В общем случае логично было бы предложить схему эмоциональных переходов по задаваемому конечному автомату.

5 Возможные направления исследований по интеграции невербальных языковых средств в человеко-машинный интерфейс

Одно из наиболее интересных направлений исследований по рассматриваемой нами тематике - это реализация возможности трансформировать мимику произвольного человеческого лица, вводимого в систему в качестве графического объекта непосредственно в процессе ее работы. Такая возможность позволила бы значительно разнообразить визуальные стимулы диалога, что решало бы проблему монотонности стимулов. Также автоматически решалась бы и проблема исходного изображения - оно могло бы быть произвольным, система брала бы на себя заботу распознать элементы лица. Автоматизация распознавания лица - задача нетривиальная, и для ее упрощения можно было бы привлечь к процессу оператора, хотя его участие должно быть по возможности минимально. Трансформация черт распознанного лица осложняется в частности тем, что для каждого лицевого элемента (рот, глаза, брови и т.д.) должны существовать алгоритмы, основанные на различных принципах. Это хорошо проиллюстрировать на примере глаз, при трансформации которых может понадобиться визуализация отсутствующих на исходном изображении объектов, например, полной окружности радужной оболочки при широко раскрытых глазах. Проблемы возникают и при работе с цветным изображением или даже использующим оттенки серого, поскольку требуется экстраполяция участков поверхности лица в смежные области.

Также может представлять интерес задача расширения невербальных стимулов диалога - включение в информационную систему, в первую очередь, звуковых и интонационных реакций на внешние ситуации. Для задания интонаций подошло бы озвучивание междометий и, к примеру, вводных слов, записанных в аудио-файлы. Возможно использование эмоционально окрашенных мелодий для создания определенной

психологической атмосферы или отражения эмоциональной нагрузки ситуации.

Наиболее актуальным на сегодня применением невербальной диалоговой компоненты являются обучающие системы. К примеру, мимическая система могла бы быть эффективно использована для задач оценивания успеваемости обучающегося. В то же время мимика может быть использована и для корректировки поведения ученика в ходе процесса обучения. Главной же целью использования невербальной компоненты в обучающих системах должно стать моделирование некоторых функций "учителя без чего сложно рассчитывать на эффективность обучения с помощью компьютерных обучающих систем.

Список литературы

- [1] Барабанщиков В.А., Малкова Т.Н. Идентификация экспрессии человеческого лица // Проблемы диагностики и управления состоянием человека оператора. М., Общество психологов СССР при Президиуме АН СССР, 1984.
- [2] Руднев В.А. Вопросы клинической генетики и вегетологии. Красноярск, 1981.
- [3] Куприянов В.В. Роль мимики в дифференциации личности // Сборник трудов научно-исследовательского института психиатрии им. Асатиани. Тбилиси, 1974.
- [4] Куприянов В.В., Стовичек Г.В. Лицо человека: анатомия, мимика. М., Медицина, 1988.
- [5] Nummenmaa Tapio. Pure and blended emotion in the human face: Psychometric experiments. Helsinki, Suomalainen tiedeakatemia, 1992.
- [6] Биркенбил В. Язык интонации, мимики, жестов. С-Пб., Питер, 1997.
- [7] East-West conference on artificial intelligence. From theory to practice. Moscow, Association for Artificial Intelligence of Russia, 1993.
- [8] Поспелов Д.А. Десять горячих точек в исследованиях по искусственному интеллекту // В сб. Интеллектуальные системы. М., том 1, вып. 1-4, 1996.
- [9] Баула В.Г., Зырянов М.И. Интеллектуальные обучающие системы и их применение в дистанционном образовании // В сб. Интеллектуальные системы. М., том 1, вып. 1-4, 1996.
- [10] Федоров Б.И., Джалиашвили З.О. Логика компьютерного диалога. М., Онега, 1994.
- [11] Вашик К., Кудрявцев В.Б., Строгалов А.С. Проект IDEA. Dortmund, Link @ Link Software GmbH, 1995.

- [12] Психология эмоций. Тексты. М., Издательство МГУ, 1993.
- [13] Ekman Paul. Telling lies: Clues to deceit in marketplace, politics and marriage. New York - London, Norton, 1985.
- [14] Izard Carroll E. The psychology of emotions. New York - London, Plenum press, 1991.