

синтаксического уровня мышления.

Понятно, что для тех задач, которые имеют не единственный вариант решения, организация самого пространства и навигация существенно усложняются – в этом пространстве появляются дополнительные измерения, учитывающие разнообразие вариантов решений. Именно такой случай представляет большой интерес, так как он позволяет тестировать и развивать более высокие этажи семантического мышления.

Используемые в настоящее время в школах тестирующие компьютерные программы являются пассивными фиксаторами успеха и неудач при ответах «да» и «нет» на простые вопросы из заданного множества, то есть оценивают в какой-то степени уровень знаний, но не мыслительность учащихся и даже не уровень их понимания предмета. Чтобы преодолеть этот барьер «формальной проверки знаний», необходима разработка более совершенных тестов. Нам представляется, что описанные выше компьютерные навигаторы решения типовых задач должны стать новой интерактивной формой автоматизации образовательного процесса. Такие навигаторы пока еще не созданы, но возможность их создания не представляется чем-то фантастичной. Как показал наш предварительный опыт, в этом направлении необходимы дополнительные исследования и разработки как синтаксиса, так и семантики пространства «типовых задач». Иначе говоря, в этой области искусственного и естественно-интеллекта пока еще не наработано необходимое понимание системной организации интерактивных решателей задач. Нарботка такого опыта необходима не только для создания учебных навигаторов решателей задач, но и для решения фундаментальной проблемы «как измерить мысль».

Трансформация мимики произвольного человеческого лица для задач интеллектуального человеко-машинного интерфейса

Е.В. Тимофеев

Рассматриваются возможные подходы к конструированию мимических систем и вводится понятие интеллектуальной мимической системы. Формулируются задачи, стоящие перед разработчиком интеллектуальной мимической системы, и предлагаются пути их решения. Приводятся примеры формул для преобразования изображений человеческого лица. Обсуждаются перспективы внедрения такой системы в специализированные программные среды.

1. Интеллектуальная мимическая система

Включение мимических сигнализационных стимулов в человеко-машинный интерфейс – задача, реально стоящая перед разработчиками современных мультимедийных программных сред. Для того, чтобы идея такого усовершенствованного интерфейса получила широкое применение, необходимо соблюдение ряда условий. Во-первых, затраты на создание мимической системы должны быть невелики. К примеру, если создавать масштабную мимическую систему, то услуги художника, который должен будет нарисовать все мимические состояния для одного или нескольких лиц, обойдутся разработчику недешево. Во-вторых, мимические стимулы должны быть разнообразными, то есть если в программе будет использовано, к

примеру, только два мимических изображения, то в течение короткого времени реакции оператора на данные стимулы притупятся, а сами стимулы станут вызывать раздражение. В-третьих, визуализация мимики должна носить естественный характер: смена эмоций должна производиться плавно, а мимические конструкции должны соответствовать реальной человеческой мимике. Использование графической базы для визуализации мимики не удовлетворяет двум из трех перечисленных условий: либо эмоций будет мало, либо система окажется дорогой, а кроме того невозможно будет добиться плавной трансформации мимики – для этого понадобится прорисовка всех промежуточных состояний. Задание исходного изображения лица с помощью графического инструментария некоторого языка программирования было бы не эффективно, так как данный подход требует большого объема работ по каждому отдельному лицу, а от повышенной трудоемкости страдают и полнота, и правдоподобие мимической системы.

Подходящей альтернативой этим двум подходам автору видится так называемая интеллектуальная мимическая система, позволяющая распознавать элементы лица на введенной произвольной фотографии (или рисунке) и затем трансформировать исходное изображение в соответствии с некоторыми правилами.

При создании интеллектуальной мимической системы отпадает необходимость задания исходного изображения (исходных графических данных) – его может ввести пользователь в виде графического файла. Эта возможность решает также и проблему монотонности стимулов – пользователь может вводить много разных портретов. Наконец правила трансформации позволяют делать переход от одной эмоции к другой плавным и естественным; это также могло бы помочь в решении проблемы отображения смешанных эмоций.

Таким образом, перед разработчиком интеллектуальной мимической системы стоит задача создания следующих процедур: распознавание изображения, описание лицевых элементов и их преобразование. Трансформация мимики происходит по определенной шкале с некоторыми крайними значениями, соответственно может быть

шкалирована и поступающая в мимическую систему извне информация. То есть на вход системы подается указатель на эмоциональную шкалу и значение на шкале, соответствующее силе эмоции. На выход же подается изображение лица, соответствующим образом трансформированного, выводимое на экран монитора.

2. Задача распознавания изображения человеческого лица

Возможность распознать лицо человека на фотографии, выделить в качестве отдельных объектов лицевые элементы (рот, глаза, брови и т.д.) дает нам фактически ключ к преобразованию такого изображения с целью визуализации той или иной мимики. Разработать полностью автоматизированную процедуру распознавания лица вполне реально, однако на такое распознавание накладывался бы ряд существенных ограничений. Во-первых, пришлось бы делать допущения примерного расположения лицевых элементов относительно границ изображения. Во-вторых, выделение объектов по контрастности изображения дает эффект только если цвет кожи лица на фотографии имеет явно светлый тон в сравнении с цветом губ или бровей, что далеко не всегда так. (Расознавание же собственно по цвету подошло бы только для работы с фотографиями одного контрастного человека, чьи цвета заранее известны.) В-третьих, даже такое распознавание, в случае его удачного осуществления, дает нам только приблизительные очертания объектов, что хорошо иллюстрируется на примере глаза – очень неоднородного в плане контрастности объекта (это верно и для остальных лицевых элементов). Для проведения же трансформации мимики важно довольно точно определить реальные границы лицевых элементов.

Возможно было бы привлечь к решению задачи распознавания оператора, предоставляя ему указывать контуры объектов с помощью манипулятора прямо на экране. Такое распознавание «вручную» работа весьма кропотлива; имеется прямая зависимость между

объемом данной работы и точностью описания объектов, то есть для достижения максимальной точности понадобится выделять каждый пиксел в контуре. Более точное выделение объектов дает полуавтоматическая система распознавания, когда оператор подготавливает второй графический файл с обрисованными контурами лицевых элементов, а программа осуществляет непосредственную их идентификацию. При таком подходе области нахождения объектов программа может определять самостоятельно, проверяя полученные контуры на замкнутость и размер.

В случае, если распознаваемый объект является сложным, то есть в нем содержатся некоторые внутренние элементы, существующие для визуализации мимики, следует использовать специальные алгоритмы их распознавания дополнительно к распознаванию контуров. Примерами таких внутренних элементов являются раздельная линия губ, радужная оболочка глаза и т.п. Возможен подход полностью автоматизированного распознавания этих элементов, основанный на принципе выделения объектов по контрастности изображения.

Целесообразно, на взгляд автора, выделение 6 лицевых элементов: бровей, глаз, рта и носа. Именно эти части лица ответственны за конструирование мимики, что же касается носа, то его координаты были бы полезны для определения верхней границы для трансформаций рта. В программной реализации данные о контурах объектов могут храниться в двумерном массиве в виде пар $[x; y]$.

3. Модель лица

В разработанной автором интеллектуальной мимической системе описание лицевых элементов с введенной в систему фотографии (или, по-другому, модель лица) формируется набором двумерных массивов, задающих контуры каждого из объектов. Номер строки в таком массиве соответствует значению координаты на оси X , сама же строка содержит Y -значения координат пикселов, составляющих контур объекта.

Для начала нужно определить крайние значения координат контура объекта — они помещаются в отдельный массив. После этого производим заполнение первого столбца в двумерном массиве Y -значений — координаты берутся из исходного массива, полученного при распознавании объекта. Так описывается верхний отрезок контура. Затем заполняется второй столбец — нижний отрезок: если значение в первом столбце отлично от нулевого, то значение помещается во второй столбец. В результате мы получаем массив с координатами верхнего и нижнего отрезков контура, с его помощью можно теперь адресоваться к любому пикселу изображения. Именно по этой причине возможное наличие строго вертикальных отрезков контура (например, бровей) не требует явного описания их координат. Таким образом, значения координат произвольного пиксела $(x; y)$ изображения в выделенном нами объекте находятся в следующих пределах:

$$\text{Массив } Y\text{-значений } [x][0] \leq y \leq \text{Массив } Y\text{-значений } [x][1],$$

где x попадает в промежуток между крайним левым и крайним правым значениями координат контура объекта.

Данная схема является универсальной, в соответствии с ней могут описываться все объекты. В разработанной автором системе включениями являются рот и нос. Массив Y -значений для первого содержит дополнительный столбец для задания раздельной линии губ (подобно отрезкам контура рта), а в модели второго нас интересуют только крайние значения контура.

4. Трансформация лицевых элементов

Преобразование лицевых элементов производится попиксельно, то есть берутся координаты каждого пиксела изображения из области трансформируемого объекта и к ним прибавляется вычисляемый коэффициент преобразования. Цветовое значение пиксела переносится на новое место, старое же его положение либо автоматически заполняется новым пикселом, перенесенным в изображение

как раз таким способом, либо принудительно заполняется с помощью процедур цветовой экстраполяции и цветогенерации. Если новое изображение лицевого элемента должно быть больше исходного (например, растягивание губ при улыбке), применяется техника дублирования пикселей, определяемых по коэффициенту экстраполяции, который, в свою очередь, зависит от силы визуализируемой эмоции.

На сегодняшний день разработанная автором система может производить следующие преобразования лицевых элементов. Для рта улыбка с расходящимися и сомкнутыми губами, опускание углов рта, открытие рта и сжатие губ. Для бровей: поднятие бровей целиком, внутренних и внешних углов бровей по-отдельности, опускание внешних концов бровей и нахмуривание бровей. Для глаз: прищуривание и смыкание век, поднятие верхнего и нижнего век.

Приведем примеры конкретных процедур преобразования выделенных объектов.

Визуализация улыбки с сомкнутыми губами — это процедура поднятия концов губ вверх при одновременном утончении нижней губы и расхождении концов губ по горизонтали. Новая координата для X высчитывается по следующей формуле:

$$c + a\left(\frac{g}{Z} + 1\right),$$

где c — x -значение середины объекта, a — расстояние от c до x -значения пиксела, g — величина, задающая силу эмоции, Z — обратный коэффициент горизонтальной растяжки. Если расстояние между двумя новыми пикселями, соответствующими двум соседним старым, равно 2, то один из пикселей дублируется для заполнения пробела.

Новая координата по Y для верхней губы высчитывается по следующей формуле:

$$m - \frac{a^2g}{C},$$

где m — старое y -значение пиксела, C — константа, а остальные переменные уже были описаны выше. Для нижней губы формула будет

выглядеть следующим образом:

$$L + \frac{(m - L)}{(g/S + 1)} - \frac{a^2g}{C},$$

где L — y -значение разделительной линии губ для пиксела с x -значением, равным $a + c$, S — обратный коэффициент редукции нижней губы, а остальные переменные заданы выше.

Визуализация улыбки с расхождением губ — это процедура поднятия концов губ вверх с открытием рта за счет поднятия верхней губы, одновременным утончением нижней губы и расхождением концов губ по горизонтали. Новая координата для X высчитывается по такой же формуле, как и для улыбки с сомкнутыми губами. Y -координата для верхней губы высчитывается по следующей формуле:

$$m - \frac{a^2g}{T} - t,$$

где T — константа, а t — смещение верхней губы по вертикали, зависящее от величины переменной g . Для нижней губы формула выглядит следующим образом:

$$L - \frac{(L - m)}{(g/S + 1)} - \frac{a^2g}{((X - c)^2gT / (tT + (X - c)^2g))},$$

где L — y -значение нижней отрезка контура губ для пиксела с x -значением, равным $a + c$, X — крайнее левое x -значение контура губ. Промежуток между верхней и нижней губой может заполняться некоторым однородным цветом, значение которого является производным от цветовой гаммы всего изображения.

5. Визуализация преобразований лица в целом

Для создания на экране правдоподобного изображения трансформированного лица необходимо сочетать изменения всех основ-

ных лицевых элементов в соответствии с конкретной визуализируемой эмоцией. Поскольку сила эмоции задается одним числом, то этот показатель для удобства должен быть общим для всех процедур трансформации различных объектов. Если же динамика изменений одного объекта не будет соответствовать динамике изменений другого, то их всегда можно откорректировать, искажая значения показателя с помощью определенной функции.

Перед выводом на экран трансформированных объектов нужно осуществить цветозаполнение их областей, чтобы в случае изменения формы объектов на исходном месте не оставалось ни старого изображения, ни незаполненных участков. Цветозаполнение производится на основе цветových значений, смежных с краем объекта пикселей. Далее необходимо, чтобы объекты трансформировались в естественных рамках, не наезжали друг на друга и за пределы лица. Именно для этого нам нужны координаты изображения носителя.

При визуализации объектов важна синхронность видимых и экране преобразований. Поэтому вывод на экран целесообразно производить тогда, когда трансформированное изображение уже находится в памяти компьютера. При переходе к новому состоянию изображения лица необходимо перерисовать его исходное состояние, чтобы новое изображение не накладывалось на старое.

6. Перспективы внедрения интеллектуальных мимических систем в специализированные программные среды

Интеграция интеллектуальных мимических систем в диалоговые среды в целях моделирования участника диалога является эффективным средством совершенствования человеко-машинного взаимодействия. Внедрение таких систем служит целям более быстрого и эффективного получения информации пользователем той или иной программы. Поскольку взаимодействие данного инструмента актуально в первую очередь в тех случаях, которые осуществляют выдачу актуальной и

формации, важной для оператора. К их числу следует отнести системы контроля данных, коммуникационные программы, обучающие и развивающие, а также справочные системы. Не следует забывать и о многофункциональных программных средах, например, операционных системах.

Обучающие системы являются одним из наиболее ярких примеров программных сред, которым усовершенствование интерфейса посредством включения в программу мимической интеллектуальной системы принесло бы значительное повышение эффективности ее работы - эффективности процесса обучения. Такое решение позволило бы более успешно моделировать преподавателя и его активность, а также усовершенствовало бы систему оценивания знаний обучающегося, включая в нее новые стимулы. И именно использование интеллектуальной мимической системы, позволяющей работать с произвольными фотопортретами, делает данное решение для обучающих программ особенно полезным, поскольку привносит разнообразие поступающих пользователю стимулов и оживляет процесс обучения.

С развитием компьютерных и сетевых технологий, по мере их удешевления все больше пользователей начинают использовать возможности компьютерного обучения. Поэтому современные обучающие программы должны рассчитываться именно на массового пользователя, а следовательно должны использовать все преимущества, предоставляемые компьютером для моделирования органического процесса обучения. В том числе и возможности передачи информации, предоставляемые мимическими системами. В этой связи следует ожидать в ближайшем будущем рост интереса к исследованному, связанному с мимическими системами, наиболее приоритетной из которых является их интеллектуализация.

В заключение автор выражает благодарность профессору А.С. Стрельцову за внимание к работе и полезные замечания при ее обдумывании.

Тимофеев Е.В. Компьютерные обучающие системы / М.: Мир, 1991. - 120 с.

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] Куприянов В.В. Роль мимики в дифференциации личности // Сборник трудов научно-исследовательского института психиатрии им. Асагиани. Тбилиси, 1974.
- [3] Куприянов В.В., Стовичек Г.В. Лицо человека: анатомия, мимика. М.: Медицина, 1988.
- [4] Nummenmaa Tapio. Pure and blended emotion in the human face: Psychometric experiments. Helsinki: Suomalainen tiedeakatemia, 1992.
- [5] Биркенбил В. Язык интонации, мимики, жестов. С-Пб.: Питер, 1997.
- [6] East-West conference on artificial intelligence. From theory to practice. Moscow: Association for Artificial Intelligence of Russia, 1993.
- [7] Поспелов Д.А. Десять горячих точек в исследованиях по искусственному интеллекту // Интеллектуальные системы. Т. 1. Вып. 1-4. 1996.
- [8] Баула В.Г., Зырянов М.И. Интеллектуальные обучающие системы и их применение в дистанционном образовании // Интеллектуальные системы. Т. 1. Вып. 1-4. 1996.
- [9] Федоров Б.И., Джалишвили З.О. Логика компьютерного диалога. М.: Онега, 1994.
- [10] Вашик К., Кудрявцев В.Б., Строгалов А.С. Проект IDEA. Dortmund: Link Link Software GmbH, 1995.

- [11] Психология эмоций. Тексты. М.: изд-во Моск. ун-та, 1993.
- [12] Ekman Paul. Telling lies: Clues to deceit in marketplace, politics and marriage. New York, London: Norton, 1985.
- [13] Izard Carroll E. The psychology of emotions. New York, London: Plenum press, 1991.
- [14] Брусиловский П.Л. Интеллектуальные обучающие системы // Информатика. Научно-технический сб. Сер. Информационные технологии. Средства и системы. Вып. 2. М.: ВНИИМИ, 1990.
- [15] Брусиловский П.Л. Модели обучаемого в интеллектуальных обучающих системах // Управляющие системы и машины. Киев, 1992. №7-8.
- [16] Брусиловский П.Л. Построение и использование моделей обучаемого в интеллектуальных обучающих системах // Техническая кибернетика. 1992. №5.
- [17] Довгялло А.М., Ющенко Е.Л. Обучающие системы нового поколения // Управляющие системы и машины. 1988, №1.
- [18] Петрушин В.А. Экспертно-обучающие системы. Киев: Наукова думка, 1992.
- [19] Кудрявцев В.Б., Вашик К., Алисейчик П.А., Перетрухин В.В., Строгалов А.С. Компьютерные интеллектуальные системы обучения // Компьютерные аспекты в научных исследованиях и учебном процессе. М.: изд-во Моск. ун-та, 1996.
- [20] Кудрявцев В.Б., Вашик К., Алисейчик П.А., Перетрухин В.В., Строгалов А.С. Об автоматном моделировании обучения // Дискретная математика. 1996. Т. 8. Вып. 4.
- [21] Строгалов А.С. Компьютерные обучающие системы: некоторые проблемы их разработок // Вузовская педагогика в информационном обществе. М.: Изд-во РГГУ.

