

УДК 004.5:004.81

Т. Г. Петренко, О. С. Тимчук, кандидаты техн. наук

МОДЕЛЬ ЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ПОЛЬЗОВАТЕЛЯ ЧЕЛОВЕКО-МАШИННЫХ СИСТЕМ

Аннотация. В статье предложена новая модель оценки эмоционального состояния пользователя человеко-машинных систем с естественным пользовательским интерфейсом (NUI). Модель позволяет учитывать невербальную информацию, получаемую от NUI Kinect. Проблемы неопределенности, возникающие при обработке невербальной информации, решаются с помощью методов теории дискретных интервальных нечетких множеств и систем второго типа.

Ключевые слова: распознавание эмоций, иммерсивность, NUI, Kinect, дискретная интервальная система, нечеткая логическая система, система второго типа

T. Petrenko, PhD., O. Tymchuk, PhD.

HUMAN-COMPUTER SYSTEMS USER EMOTION STATE MODEL

Abstract. The paper proposes a new emotion recognition model of user of human-computer systems with a natural user interface (NUI). Model allows to take into account the non-verbal information from NUI Kinect. The problems of uncertainty occurring in processing of non-verbal information are solved by the methods of the theory of discrete interval type-2 fuzzy sets and systems.

Keywords: emotion recognition, immersive, NUI, Kinect, discrete interval type-2 fuzzy logic system

Т. Г. Петренко, О. С. Тимчук, канд. техн. наук

МОДЕЛЬ ЭМОЦИОНАЛЬНОГО СТАНУ КОРИСТУВАЧА ЛЮДИНО-МАШИННИХ СИСТЕМ

Анотація. У статті запропоновано нову модель оцінки емоційного стану користувача людино-машинних систем з природним інтерфейсом користувача (NUI). Модель дозволяє враховувати невербальну інформацію, що отримується від NUI Kinect. Проблеми невизначеності, що виникають при обробці невербальної інформації, вирішуються за допомогою методів теорії дискретних інтервальних нечітких множин і систем другого типу.

Ключові слова: розпізнавання емоцій, імерсивність, NUI, Kinect, дискретна інтервальна система, система другого типу, нечітка логічна система

Введение. Естественный пользовательский интерфейс (NUI) [1] является важной составляющей современных человеко-машинных систем и обеспечивает получение пользователем иммерсивных ощущений. Такие ощущения формируются у пользователя при работе с мультисенсорными устройствами за счет более высокой интерактивности взаимодействия пользователя и программной системы. Примерами NUI являются 3D Sensor Kinect, Creative Senze3D. Современные NUIs способны поддерживать не только вербальные, но и невербальные коммуникации пользователя с программной системой. Невербальное общение всегда проявляется при естественном человеческом взаимодействии посредством когнитивных параметров [2, 3]. Основным, внешне проявляемым, когнитивным параметром является эмоциональное состояние человека. Существующие

модели эмоционального состояния пользователя человеко-машинных систем [4, 5] не позволяют в полной мере оценить состояние человека по следующим причинам:

- в моделях слабо поддерживается фиксация невербальной информации, получаемой от NUIs;
- среди специалистов отсутствует однозначная интерпретация невербальной информации, получаемой от NUIs;
- невербальная информация обычно представлена только словестным описанием;
- невербальная информация носит непостоянный характер в течение определенных интервалов времени;
- невербальная информация может описываться значительным количеством параметров;
- временные ряды невербальной информации обладают нелинейной структурой.

© Петренко Т.Г., Тимчук О.С., 2014

Перечисленные причины позволяют выделить основные характеристики информации, с которой должны работать человеко-машинные системы, оборудованные NUIs. Это – сложность, неполнота, неопределенность, высокая размерность и динамичность. В данной работе предложена модель оценки эмоционального состояния пользователя с учетом возможностей NUI и с помощью методов теории дискретных интервальных нечетких множеств и систем второго типа (DIT2FSs и DIT2FLSs соответственно) [6]. В качестве NUI рассматривается 3D Sensor Kinect [7].

Основные области применения разработанной модели: индустрия видеоигр и развлечений, робототехнические комплексы, операторские станции и т.д.

Изложение основного материала. Разделим обработку информации об эмоциональном состоянии пользователя на два этапа.

Первый этап – обработка данных, получаемых от SDK Kinect [8] (координаты опорных точек тела пользователя; угол наклона головы пользователя; отклонение формы рта, бровей и глаз от нейтральной формы лица пользователя). Анализ данной информации позволяет определить базовые признаки эмоции пользователя.

Второй этап обеспечивает нечеткую интерпретацию признаков эмоции с помощью DIT2FLS.

В работе, с целью упрощения изложения материала, оценка эмоции пользователя представлена на базе информации, получаемой из одного кадра. При построении динамической человеко-машинной системы выполняется анализ серии кадров, получаемых от SDK Kinect.

В соответствии с положениями теории DIT2FSs и DIT2FLSs представим модель эмоционального состояния пользователя

$$\begin{aligned} \tilde{A}(t) &= F^*(IN(t), LI, LO, R), \\ IN(t) &= \langle in_n \rangle, n = \overline{1, N}, \\ LI &= \langle li_n \rangle, \\ R &= \langle r_p \rangle, p = \overline{1, P}, \end{aligned}$$

где $\tilde{A}(t)$ – DIT2FS, описывающее эмоциональное состояние пользователя в момент

времени t ; $IN(t)$ – набор четких входных значений (признаки эмоции); N – количество входных значений; LI – набор входных лингвистических переменных, которые описывают признаки эмоционального состояния пользователя; LO – результирующая лингвистическая переменная, описывающая эмоциональное состояние пользователя; R – набор нечетких правил, на основе которых оценивается эмоциональное состояние пользователя; P – количество нечетких правил; F^* – операция нечеткого логического вывода (алгоритм Мамдани).

SDK Kinect позволяет определить три базовых признака эмоции

$$IN(t) = \langle in_1, in_2, in_3 \rangle,$$

где in_1 – оценка угла наклона головы пользователя; in_2 – оценка отклонения параметров лица пользователя от нейтральной формы лица; in_3 – оценка отклонения опорных точек тела пользователя от нейтральной формы положения тела в пространстве.

Значения набора $IN(t)$ приводятся к нормированному интервалу $[0; 1]$.

Для оценки угла наклона головы пользователя в модели используется угол наклона головы влево /вправо (RA), значение которого определяет SDK Kinect

$$in_1 = \frac{RA - RA_{MIN}}{RA_{MAX} - RA_{MIN}},$$

$$RA \in [-90^\circ; 90^\circ], RA_{MIN} = -90, RA_{MAX} = 90.$$

Отклонения параметров лица пользователя (отклонение формы рта, бровей, глаз) от нейтральной формы лица характеризуют общее выражение лица пользователя [3] (нейтральное, позитивное, негативное). SDK Kinect определяет 7 типов отклонений, значение которых выражается в диапазоне $[-1; 1]$ (табл. 1).

Оценка отклонения параметров лица пользователя in_2 в работе выполняется с помощью правил типа IF-THEN (система правил представлена в виде дерева решений, (рис. 1). Построение правил в работе осуществляется в соответствии с исследованиями в области когнитивной психологии. Пример правила: **IF AU6_0 AND AU5_0 AND AU4_2 AND AU3_1 THEN $in_2=0,3$** .

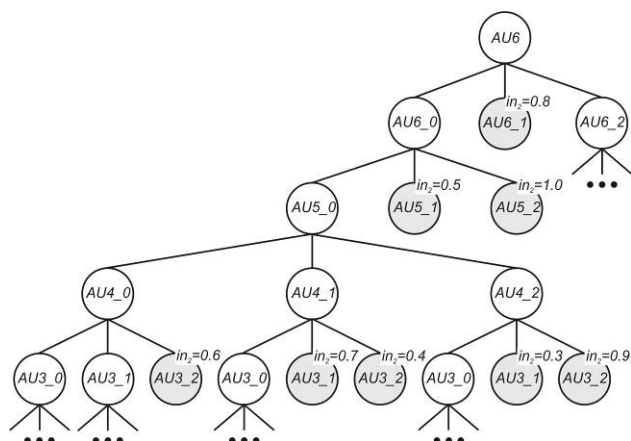


Рис. 1. Фрагмент дерева определения оценки отклонения лица пользователя от нейтральной формы лица

Отклонения опорных точек тела (суставов) пользователя от нейтральной формы положения тела в пространстве в данной модели рассматриваются как жест, совершаемый пользователем в момент времени t . SDK Kinect определяет координаты 20 суставов пользователя.

Анализ координат суставов позволяет с определенной уверенностью определить жест (оценка in_3), совершаемый пользователем во время взаимодействия с человеко-машинной системой. Согласно данным из когнитивной психологии, жесты характеризуют внутреннее состояние пользователя. На рис. 2 представлены составляющие жеста (тело пользователя на рисунке отображено зеркально).

1. Параметры лица пользователя, определяемые SDK Kinect

Параметр лица пользователя	Иллюстрация	Интерпретация параметра
AU0 – нейтральное лицо		AU1=0, AU2=0, AU3=0, AU4=0, AU5=0, AU6=0.
AU1 – верхняя губа		AU1_0=[-0,5; 0,5] – нейтральное положение; AU1_1=[-1; -0,5] – верхняя губа опущена вниз; AU1_2=(0,5; 1] – зубы видны полностью.
AU2 – нижняя челюсть		AU2_0=[-0,5; 0,5] – челюсть закрыта; AU2_1=[-1; -0,5] – челюсть закрыта; AU2_2=(0,5; 1] – челюсть полностью открыта.
AU3 – разрез губ		AU3_0=[-0,5; 0,5] – нейтральный разрез; AU3_1=[-1; -0,5] – губы округлены; AU3_2=(0,5; 1] – губы вытянуты.
AU4 – нижние брови		AU4_0=[-0,5; 0,5] – нейтральные; AU4_1=[-1; -0,5] – брови подняты полностью; AU4_2=(0,5; 1] – брови опущены полностью.
AU5 – угол наклона губ		AU5_0=[-0,5; 0,5] – нейтральный; AU5_1=[-1; -0,5] – счастливая улыбка; AU5_2=(0,5; 1] – нахмуренное лицо.
AU6 – верхние брови		AU6_0=[-0,5; 0,5] – нейтральные; AU6_1=[-1; -0,5] – брови опущены полностью; AU6_2=(0,5; 1] – брови подняты полностью.

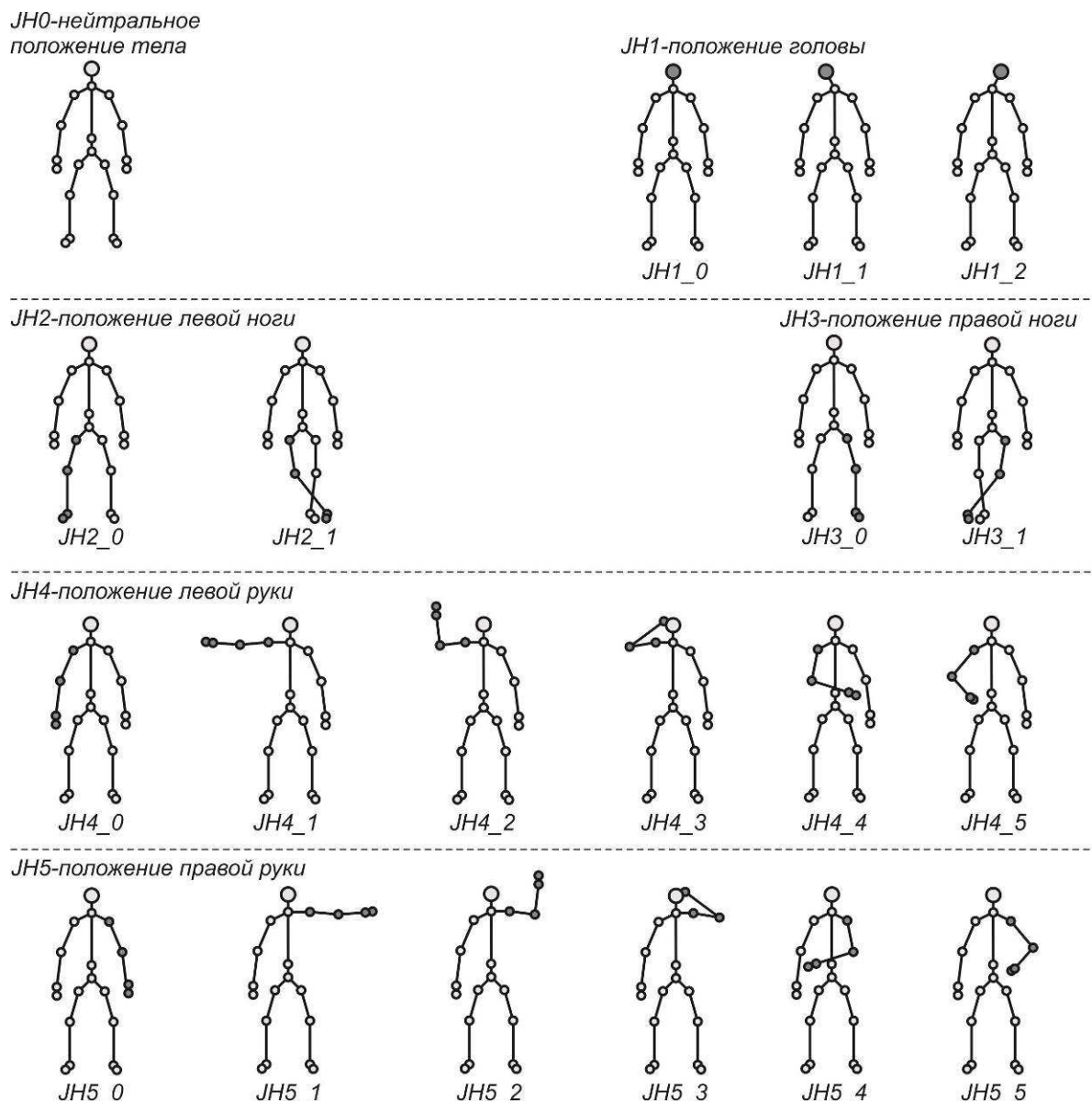


Рис. 2. Составляющие жеста (анализ данных SDK Kinect)

Жест как комбинация составляющих в работе определяется с помощью правил типа IF-THEN (система правил представлена в виде дерева решений, рис. 3). Построение правил в работе осуществляется в соответствии с исследованиями в области когнитивной психологии. Пример правила: **IF JH1_1 AND JH2_0 AND JH3_0 AND JH4_3 THEN $in_3=0,3$** .

Набор LI содержит 3 лингвистических переменных

$$LI = \langle li_1, li_2, li_3 \rangle,$$

где li_1 – лингвистическая переменная, описывающая угол наклона головы пользователя; содержит 3 терма; определена на

универсальном множестве $X_1 = [0; 1]$ (рис. 4, а); li_2 – лингвистическая переменная, описывающая выражение лица пользователя; содержит 3 терма; определена на универсальном множестве $X_2 = [0; 1]$ (рис. 4, б); li_3 – лингвистическая переменная, описывающая внутреннее состояние пользователя; содержит 5 термов; определена на универсальном множестве $X_3 = [0; 1]$ (рис. 4, в).

Результирующая лингвистическая переменная LO содержит 6 термов (базовые классы внешнего проявления эмоционального состояния пользователя [3]). LO

определена на универсальном множестве $X_{res} = [0; 1]$ (рис. 4, г).

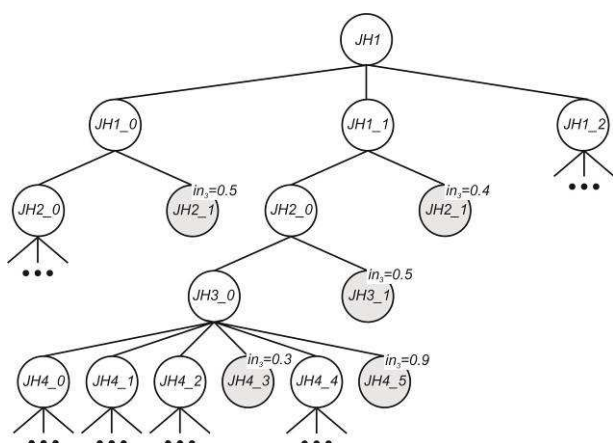


Рис. 3. Фрагмент дерева оценки отклонения опорных точек тела пользователя от нейтральной формы положения тела в пространстве

Набор правил R содержит 45 нечетких правил типа IF-THEN. Для каждого правила экспертом установлена степень доверия к правилу. Пример нечеткого правила из набора R : **IF** in_1 *is* «средний угол наклона головы» **AND** in_2 *is* «позитивное выражение лица» **AND** in_3 *is* «внутреннее состояние – размышление» **THEN** «эмоциональное состояние пользователя – радость» (доверие к правилу – 1,0).

В работе выполнена программная реализация разработанной модели. Программная реализация представлена в виде набора динамических библиотек, архитектура которых имеет прозрачную структуру, содержит унифицированные структуры данных, является расширяемой и полной. Перечисленные качества позволяют подключать разработанный набор библиотек к уже готовым информационным технологиям организации человеко-машинного интерфейса. Для обработки DIT2FLS использован пакет библиотек поддержки DIT2FLS (DIT2FLS Toolbox) [9]. DIT2FLS Toolbox содержит средства для проектирования и моделирования DIT2FLS, реализует все фазы нечеткого логического вывода.

Эффективность предложенной в статье модели подтверждена успешной реализацией модифицированной видеоигры «Pac-Man», где стратегия поведения компьютерных персонажей динамически изменяется в соответствии с эмоциональным состоянием человека-игрока [10]. Учет оценки эмоционального состояния человека-игрока позволил организовать интерактивное индивидуальное взаимодействие человека-игрока с объектами видеоигры и сформировать у пользователя иммерсивные ощущения.

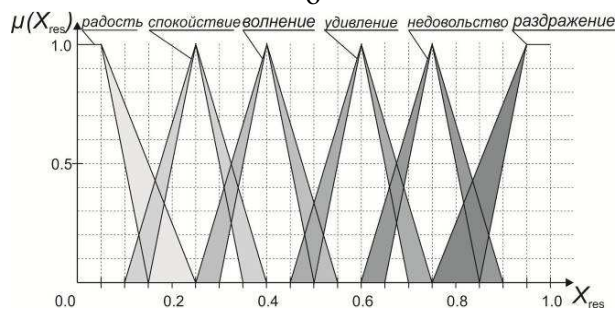
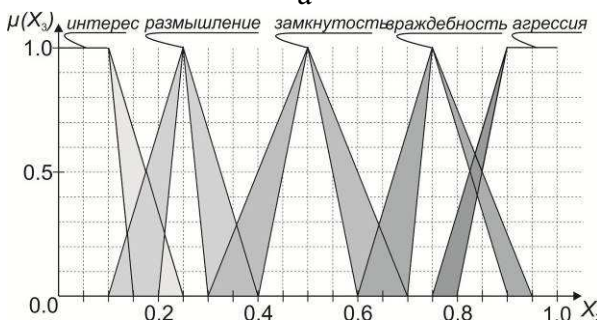
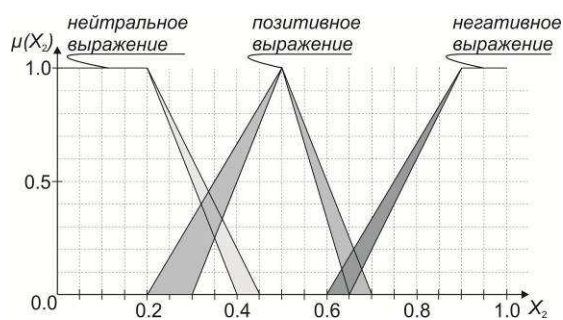
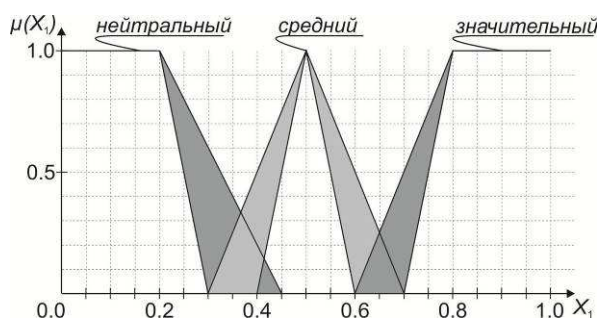


Рис. 4. Лингвистические переменные модели эмоционального состояния пользователя. (а), (б), (в) – входные, (г) – результирующая

Эксперимент. Пусть в момент времени t от SDK Kinect были получены следующие данные:

1) Значение угла наклона головы влево /вправо (RA) – -27° .

2) Отклонения параметров лица пользователя от нейтральной формы лица:

- верхние брови (AU6) – $-0,1$;
- угол наклона губ (AU5) – $0,2$;
- нижние брови (AU4) – $0,7$;
- разрез губ (AU3) – $-0,8$;
- нижняя челюсть (AU2_0) – $0,15$;
- верхняя губа (AU1_0) – $0,2$.

3) Положение пользователя в пространстве. В работе рассматривается двумерное пространство. Каждый сустав пользователя задается декартовыми координатами (табл. 2). Начало координат – тазобедренный сустав. Положительное направление осей – общепринятое для декартовой системы координат.

2. Координаты 20 суставов пользователя

Название сустава	Координаты
Щиколотка левая	(-0,140193; -0,7792207)
Щиколотка правая	(0,00115087; -0,75275)
Локоть левый	(-0,418351; 0,4164155)
Локоть правый	(0,2078099; 0,1633845)
Стопа левая	(-0,158239; -0,8299173)
Стопа правая	(-0,0238999; -0,803163)
Рука левая	(-0,270471; 0,6782582)
Рука правая	(0,2303156; -0,232154)
Голова	(-0,0900773; 0,745175)
Тазобедренный сустав	(0,0573179; 0,1352181)
Бедро левое	(-0,138075; 0,0621909)
Бедро правое	(0,0119799; 0,0586523)
Колено левое	(-0,138051; -0,4264791)
Колено правое	(0,0344744; -0,411151)
Шея	(-0,0257272; 0,557311)
плечо левое	(-0,211146; 0,4783091)
Плечо правое	(0,1253217; 0,4436179)
Позвоночник	(-0,043089; 0,1924895)
Запястье левое	(-0,311998; 0,6251581)
Запястье правое	(0,236292; -0,1146519)

На рис. 5 представлено положение тела пользователя в момент времени t в соответствии с данными из табл. 2.



Рис. 5. Положение тела пользователя в момент времени t

На первом этапе работы модели выполняется обработка данных от SDK Kinect с целью определения базовых признаков эмоции пользователя – набора четких входных значений $IN(t)$:

1) *Определение экспертной оценки*

$$in_1 = \frac{-27 - (-90)}{90 - (-90)} = 0,35.$$

2) *Определение экспертной оценки in_2 .*

Значения и интерпретация параметров лица пользователя, полученные от SDK Kinect в момент времени t , представлены в табл. 3.

3. Значения и интерпретация параметров лица пользователя

Значение параметра	Интерпретация параметра
AU6 = $-0,1$	Нейтральное положение верхних бровей (AU6_0)
AU5 = $0,2$	Нейтральный угол наклона губ (AU5_0)
AU4 = $0,7$	Нижние брови опущены полностью (AU4_2)
AU3 = $-0,8$	Губы округлены (AU3_1)
AU2 = $0,15$	Челюсть закрыта (AU2_0)
AU1 = $0,2$	Нейтральное положение верхней губы (AU1_0)

Параметры лица пользователя, приведенные в табл. 3, соответствует правилу **IF AU6_0 AND AU5_0 AND AU4_2 AND AU3_1 THEN $in_2=0,3$** (рис. 1).

Отсюда следует, что экспертная оценка in_2 равняется $0,3$.

3) *Определение экспертной оценки in_3 .*

Анализ координат суставов пользователя позволяет выделить жест (рисунки 2 и 3),

совершенный пользователем в момент времени t ($JH1_1, JH4_3$).

Выделенный набор жестов соответствует правилу ***IF JH1_1 AND JH2_0 AND JH3_0 AND JH4_3 THEN $in_3 = 0,3$*** (рис. 3).

Отсюда следует, что экспертная оценка in_3 равняется 0,3.

В результате обработки данных от SDK Kinect сформирован набор четких значений $IN(t)$

$$IN(t) = \langle 0,35; 0,3; 0,3 \rangle.$$

На втором этапе работы модели выполняется нечеткая интерпретация признаков эмоции с целью оценки эмоционального состояния пользователя.

Заданный набор четких входных значений $IN(t)$ активирует 8 нечетких правил (табл. 4).

4. Набор активированных правил нечеткой модели эмоционального состояния пользователя

li_1	li_2	li_3	LO	Уверенность
Нейтральный	Нейтральное	Размышление	Волнение	0,7
Нейтральный	Нейтральное	Замкнутость	Спокойствие	0,8
Нейтральный	Позитивное	Размышление	Интерес	0,9
Нейтральный	Позитивное	Замкнутость	Волнение	0,9
Средний	Нейтральное	Размышление	Волнение	0,9
Средний	Нейтральное	Замкнутость	Волнение	1,0
Средний	Позитивное	Размышление	Радость	1,0
Средний	Позитивное	Замкнутость	Удивление	0,6

Результатом работы DIT2FLS (алгоритм операции нечеткого логического вывода – алгоритм Mamdani, алгоритм понижения типа – ЕКМ [6], метод дефаззификации – Centroid) является значение, равное 0,28, т.е. эмоциональное состояние пользователя в момент времени t – спокойствие.

Выводы. В работе построена двухуровневая модель эмоционального состояния пользователя на базе теории DIT2FSS и

DIT2FLSs. На первом этапе работы модели данные от SDK Kinect интерпретируются как положения частей тела и зон лица пользователя. На втором этапе работы модели выполняется нечеткая интерпретация на основе DIT2FLS положений частей тела и состояний зон лица как жестов и мимических реакций, значимых для системы человеко-машинного взаимодействия.

Аппарат DIT2FSS и DIT2FLSs позволяет учесть индивидуальные особенности пользователя. Индивидуализация человеко-машинного интерфейса характеризуется эффективной обратной связью, своевременной и соответствующей реакцией на изменение состояния пользователя.

Список использованной литературы

1. Wigdor D. Brave NUI World: Designing Natural User Interfaces for Touch and Gesture [Текст] / D. Wigdor, D. Wixon. – 1st ed. – Morgan Kaufmann, 2011. – 264 p.
2. Солсо Р. Когнитивная психология [Текст] / Р. Солсо. – Изд. 6-е. – СПб.: Питер, 2002. – 589 с.
3. Solso R. L. Cognitive Psychology [Текст] / R. L. Solso, M. K. MacLin, O. H. MacLin – 8th ed. – Pearson, 2005. – 592 p.
4. Ekman P. Facial Expression and Emotion [Электронный ресурс]. – Режим доступа: <http://www.paulekman.com/wp-content/uploads/2013/07/Facial-Expression-Of-Emotion1.pdf>. – Дата доступа (04.03.2014).
5. Halder A. General and Interval Type-2 Fuzzy Face-Space Approach to Emotion Recognition [Текст] / [A. Halder, A. Konar, R. Mandal and others] // IEEE Transactions on Systems, Man, and Cybernetics: Systems. – 2013. – Vol. 43, No. 3. – Pp. 587 – 605.
6. Kolakowska A. Emotion Recognition and its Application in Software Engineering [Текст] / [A. Kolakowska, A. Landowska, M. Szwoch and others] // Proceedings of the 6th International Conference on Human System Interaction. – 6–8 June 2013. – Pp. 532 – 539.
7. Sadeghian A. Advances in Type-2 Fuzzy Sets and Systems: Theory and Applications [Текст] / A. Sadeghian, J. Mendel, H. Tahayori. – Springer. Pp.2013. – 262.

8. Borenstein G. Making Things See: 3D vision with Kinect, Processing, Arduino, and MakerBot [Текст] / G. Borenstein. – Make, 2012. – 440 p.

9. Kinect for Windows. Developer Downloads [Электронный ресурс]. – Режим доступа: <http://www.microsoft.com/en-us/kinectforwindows/develop/developer-downloads.aspx>. – Дата доступа (04.03.2014).

10. Petrenko T. Package Library and Toolbox for Discrete Interval Type-2 Fuzzy Logic Systems [Текст] / Т. Petrenko, О. Tymchuk // Proceedings of the 18th International Conference on Soft Computing (MENDEL), Brno, Czech Republic. – 27–29 June 2012. – Pp. 233 – 238.

11. Petrenko T. Adaptive Behavior Control Model of Non Player Character [Текст] / Т. Petrenko, О. Tymchuk // Proceedings of the 15th International Conference on Computer Modelling and Simulation (UKSim-AMSS), Cambridge, United Kingdom. – 10–12 Apr. 2013. – Pp. 39 – 44.

Получено 02.03.2014

References

1. Wigdor D., and Wixon D., (2011), *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*, 1st ed., Morgan Kaufmann (In English).

2. Solso R.L., MacLin, M.K., and MacLin O.H., (2005), *Cognitive Psychology, 8th ed.*, Pearson (In English).

3. Keltner D., and Ekman P., (2003), Facial Expression of Emotion (In English), available at: <http://www.paulekman.com/wp-content/uploads/2013/07/Facial-Expression-Of-Emotion1.pdf> (accessed 4 March 2014).

4. Halder A., Konar A., Mandal R., and others., (2013), General and Interval Type-2 Fuzzy Face-Space Approach to Emotion Recognition, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 43, No. 3, pp. 587 – 605 (In English).

5. Kolakowska A., Landowska A., Szwoch M., and others., (2013), Emotion Recognition and its Application in Software Engineering, *Proceedings of the 6th International Conference on Human System*

Interaction (HSI), June 6–8, 2013, Sopot, Poland, pp. 532 – 539 (In English).

6. Sadeghian A., Mendel J., and Tahayori H., (2013), *Advances in Type-2 Fuzzy Sets and Systems: Theory and Applications*, Springer (In English).

7. Borenstein G. (2012), Making Things See: 3D Vision with Kinect, Processing, Arduino, and MakerBot, Make (In English).

8. Kinect for Windows. Developer downloads (In English), available at: <http://www.microsoft.com/en-us/kinectforwindows/develop/developer-downloads.aspx> (accessed 4 March 2014).

9. Petrenko T., and Tymchuk O. Package Library and Toolbox for Discrete Interval Type-2 Fuzzy Logic Systems, (2012), *Proceedings of the 18th International Conference on Soft Computing (MENDEL)*, June 27–29, 2012, Brno, Czech Republic, pp. 233 – 238 (In English).

10. Petrenko T., and Tymchuk O. Adaptive Behavior Control Model of Non Player Character, (2013), *Proceedings of the 15th International Conference on Computer Modelling and Simulation (UKSim-AMSS)*, April 10–12, 2013, Cambridge, United Kingdom, pp. 39 – 44 (In English).



Петренко Татьяна Григорьевна,
к.т.н., доцент каф. компьютерных технологий
Донецкого национального университета,
тел.050-3487748,
e-mail:
petrenkotg@ieee.org



Тимчук Олег Сергеевич,
к.т.н., старший преподаватель каф. компьютерных технологий
Донецкого национального университета,
тел. 099-0188285,
e-mail:
o.s.timchuk@gmail.com.