



УДК 004.822:514

## РАСПОЗНАВАНИЕ ПОЗЫ ТЕЛА ЧЕЛОВЕКА С ПОМОЩЬЮ 3D КАМЕРЫ С СЕНСОРОМ ГЛУБИНЫ

Шпирко А.А., Дорофеев Н.С., Бобков А.С.

*Волгоградский государственный технический университет,  
г. Волгоград, Российская Федерация*

**alex555@yandex.ru**

**nikita.dorofei@gmail.com**

**bobkovart@gmail.com**

В работе описан метод контроля правильности выполнения физических упражнений лечебной физкультуры с помощью технологии Microsoft Kinect. Формализован процесс идентификации позы тела человека по координатам некоторых «характерных» узлов тела человека с использованием продукционных правил. С помощью нейронной сети – многослойного персептрона происходит классификация упражнений на правильные и неправильные. Обучение нейронной сети происходило на выборке, составленной экспертом в области ЛФК.

**Ключевые слова:** упражнения ЛФК; продукционные правила; нейронная сеть; контроль правильности выполнения упражнений.

### ВВЕДЕНИЕ

Лечебная физическая культура (ЛФК) – это метод лечения, состоящий в применении физических упражнений и естественных факторов природы к больному человеку с лечебно-профилактическими целями. В основе этого метода лежит использование биологической функции организма – движения [Medicinform, 2012].

До недавнего времени обеспечить контроль правильности выполнения физических упражнений мог только персональный тренер или врач-специалист по ЛФК. Самостоятельное выполнение упражнений возможно с использованием специализированной литературы или мультимедийных дисков, с записанными примерами выполнения упражнений. Но в обоих случаях, невозможно проконтролировать правильность выполнения человеком упражнений.

Применение сенсора Microsoft Kinect [Орлова и др., 2011], в разрабатываемой системе, позволит сделать процесс занятий удобным и безопасным. Также разработанные модели и методы применяются в работах по распознаванию эмоциональных реакций человека [Заболеева-Зотова и др., 2011], [Розалиев и др., 2011a], [Розалиев и др., 2011b].

В статье описан метод контроля правильности выполнения физических упражнений (ФУ) лечебной физкультуры с помощью технологии Microsoft Kinect.

### 1. Распознавание движения

Любое физическое упражнения – это движение. Для контроля правильности выполнения ФУ необходимо распознать движение человека.

Выполняемое ФУ снимается безмаркерным способом с помощью сенсора Kinect (3-D камера от Microsoft с сенсором глубины), далее используется система Brekel Kinect, предоставляющая распознанное движение в формате .bvh – файла.

BVH обозначает данные Bio Vision Hierarchical. Этот формат предоставляет возможность представления информации об иерархии каркаса тела человека в добавление к данным о движении [Kinecthacks, 2012].

```

HIERARCHY
ROOT_HIPS
{
  OFFSET 0.0000 0.0000 0.0000
  CHANNELS 6 Xposition Yposition Zposition Zrotation Xrotation Yrotation
  JOINT Chest1
  {
    OFFSET 0.0415 6.2451 0.0203
    CHANNELS 3 Zrotation Xrotation Yrotation
    JOINT Chest2
    {
      OFFSET 0.0118 22.5192 -0.7792
      CHANNELS 3 Zrotation Xrotation Yrotation
      JOINT Leftcollar
      {
        OFFSET -0.3300 25.0202 0.8579
        CHANNELS 3 Zrotation Xrotation Yrotation
        JOINT Leftshoulder
        {
          OFFSET 15.6466 -3.5590 0.4633
          CHANNELS 3 Zrotation Xrotation Yrotation
          JOINT Leftelbow
          {
            OFFSET 6.7754 -28.9612 2.1049
            CHANNELS 3 Zrotation Xrotation Yrotation
            JOINT Leftwrist
            {
              OFFSET 5.9027 -23.6186 1.7304
              CHANNELS 3 Zrotation Xrotation Yrotation
              End Site
              {
                OFFSET 3.1558 -12.6376 1.4914
              }
            }
          }
        }
      }
    }
  }
}
JOINT Rightcollar
{
  OFFSET -0.3300 25.0202 0.8579
  CHANNELS 3 Zrotation Xrotation Yrotation
  JOINT Rightshoulder
}

```

Рисунок 1- Структура BVH-файла

## 2. Идентификация позы человека и физического упражнения

Каждое ФУ характеризуется набором поз тела человека. Каждая поза характеризуется набором координат узлов (суставов) относительно центра тела.

ФУ ЛФК характеризуется набором т.н. «характерных» узлов человеческого тела. Это узлы, на которые направлено его воздействие и взаиморасположение которых в процессе выполнения определяет правильность выполнения упражнения

Идентификация ФУ и характерных узлов человеческого тела, реализована с помощью продукционной модели вида:

$$pm = \langle S, L, A \rightarrow B, C, Q \rangle,$$

Где  $S = \{S_1, S_2, S_3\}$  – множество зон человеческого тела (верхняя, нижняя, центральная);  $L$  - условие применимости ядра продукции. Сигнал о том, что было выбрано упражнение.

$A \rightarrow B, C$  - ядро продукции, интерпретируется фразой:

«Если  $A$ , то  $B$  и  $C$ »,

где  $A = \{a_1, a_2, a_3, \dots, a_n\}$  – множество физических упражнений;

$B = \{\{b_1, b_2, \dots, b_n\}_1, \{b_1, b_2, \dots, b_n\}_2, \dots, \{b_1, b_2, \dots, b_n\}_m\}$  – множество, включающие себя подмножества из набора характерных узлов (суставов) человеческого тела  $b$ , где  $b$  это вектор – строка вида  $(b_x, b_y, b_z)$ .

$C = \{c_1, c_2, \dots, c_n\}$  – множество нейронных сетей, каждая из которых настроена для определения правильности конкретного упражнения.

## 3. Метод проверки правильности выполнения физических упражнений

Для контроля правильности и эффективности выполнения физического упражнения, необходимо определить, какие действия при выполнении упражнения будут считаться правильными, а какие нет. Эта задача ложится на эксперта в области ЛФК. На основании мнения эксперта и основных правил лечебной физкультуры, с помощью системы распознавания движения Brekel Kinect были

записаны упражнения, которые будут считаться правильными. Также были записаны движения, представляющие собой наиболее распространенное неправильное выполнение упражнений.

Под упражнением в данном контексте понимается набор поз которые последовательно принимает человек и перед системой стоит задача определить к какому классу относится в текущий момент времени поза, которую принимает человек: к классу «правильное выполнение» или к классу «неправильное выполнение».

Обучающая выборка состояла из 2000 поз на каждый класс, где каждая поза характеризовалась набором «характерных» узлов для каждого физического упражнения.

Было протестировано десять различных методов, способных решать задачи классификации, на одних и тех же входных данных. По окончании, метод, показавший наилучший результат, был использован для решения задачи классификации правильности выполнения ФУ.

Таблица 1 - Сравнение эффективности методов решения задачи идентификации

Метод	Правильно Идентифицированные образы(%)
Дерево решений: CHAID алгоритм (автоматический обнаружитель взаимодействий)	62,58
Дерево решений: C&RT алгоритм	69,99
Дерево решений: boosting trees алгоритм («растущие деревья»)	81,74
Дерево решений: random forest (алгоритм «случайные леса»)	76,18
Метод опорных векторов (Support Vector Machine)	76,25
Метод К-ближайших соседей (K-Nearest neighbors)	82,11
Байесовский классификатор (Naive Bayes Classifiers)	87,36
Нейронная сеть: вероятностная сеть	94,12
Нейронная сеть: Многослойный перцептрон	96,02
Нейронная сеть: радиально-базисных функций	68,20

### 3. Характеристики нейронной сети

Наилучший результат показала нейронная сеть многослойный персептрон со следующими параметрами:

Тип НС	%успешной классификации и всей выборки	%успешной классификации и тестовой выборки	%успешной классификации и контрольная выборки	Алгоритм обучения	Функция ошибки	Функция активации скрытого слоя	Функция активации внешнего слоя
МРП18-12-2	96,0231	96,0001	95,9887	BFGS5	Кросс-энтропия	Экспоненциальная	Софтмакс

Рисунок 2 – Параметры нейронной сети

Алгоритм обучения – BFGS. В данном алгоритме используется BFGS-схема обновления, названная по первым буквам имен Broyden-Fletcher-Goldfarb-Shanno (если точнее, то эта формула строит не сам гессиан, а обратную к нему матрицу; таким образом, не надо тратить время на её обращение). В случае больших размерностей объем памяти порядка  $N^2$ , требуемый для хранения гессиана, оказывается слишком большой нагрузкой, также как и затраты машинного времени на его обработку. Поэтому вместо использования  $N$  значений градиента для построения гессиана можно обойтись меньшим числом значений, позволяющим использовать объем памяти порядка  $N \cdot M$ . Обычно на практике  $M$  выбирают в промежутке от 3 до 7, в сложных случаях можно увеличить эту константу до 20. В результате такой экономии получается не сам гессиан, а лишь его аппроксимация. Изначально этот метод разрабатывался для оптимизации функций очень большого числа аргументов (сотни и тысячи), поскольку именно в этом случае увеличение числа итераций из-за пониженной точности аппроксимации гессиана полностью окупается снижением накладных расходов на обновление гессиана, однако нет причин, по котором этот метод нельзя применять для задач малой размерности. Основным его достоинством является масштабируемость, поскольку он обеспечивает высокое быстродействие на задачах большой размерности, при этом позволяя решать и задачи малой размерности [Осовский, 2007].

Функция ошибки – кросс-энтропия. Функция, основанная на теоретико-информационных характеристиках. Особенно хорошо подходит для задач классификации. Имеется два варианта: для сетей с одним выходом и для сетей с несколькими выходами; в первом варианте используются логистические функции активации, во втором – функции софтмакс [Боровиков, 2007].

Функция активации внешнего слоя Софтмакс. Это функция активации, специально предназначенная для классификационных сетей с кодированием по методу один-из- $N$ . Вычисляет нормированную экспоненту (т.е. сумма выходов равна единице). В сочетании с кросс-энтропийной функцией ошибок позволяет модифицировать многослойный персептрон для оценки вероятностей принадлежности классам [Хайкин, 2006].

### 4. Пример работы системы

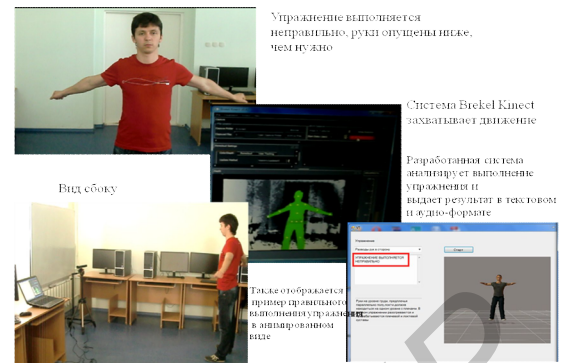


Рисунок 3 – Пример работы системы. Упражнение выполняется неправильно

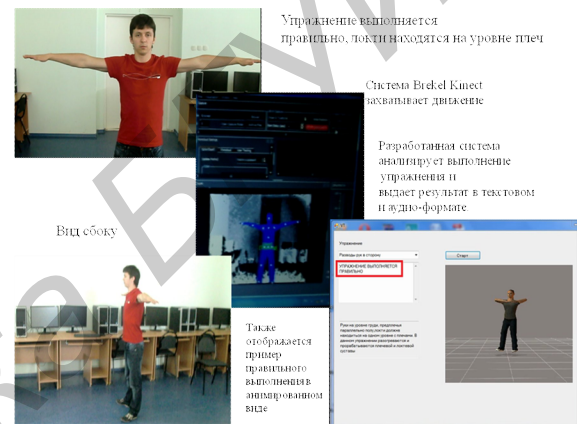


Рисунок 4 – Пример работы системы. Упражнение выполняется неправильно

На рисунках представлены примеры выполнения правильного и неправильного выполнения ФУ. Распознанное движение сохраняется в BVH - формат, подается на вход нейронной сети, которая уже возвращает результат – заключение правильно или неправильно выполняется то или иное ФУ. Одновременно пользователю показывается анимированный пример правильного выполнения и звуковое сопровождение.

### Заключение

В результате работы была автоматизирована проверка правильности выполнения ФУ, разработаны правила идентификации позы человека и ФУ, построена нейронная сеть для проверки правильности выполнения ФУ.

Работа частично поддержана Российским фондом фундаментальных исследований (проекты 12-07-00266, 12-07-00270).

## Библиографический список

[Орлова и др., 2011] Орлова, Ю.А. Обзор современных автоматизированных систем распознавания эмоциональных реакций человека / Ю.А. Орлова, В.Л. Розалиев // Изв. ВолгГТУ. Серия "Актуальные проблемы управления, вычислительной техники и информатики в технических системах". Вып. 10 : межвуз. сб. науч. ст. / ВолгГТУ. - Волгоград, 2011. - № 3. - С. 68-72.

[Medicinform, 2012] Основы ЛФК / Медицинская информационная сеть. – [2012]. – Электронный ресурс. Режим доступа: <http://www.medicinform.net/fizio/lfk/>

[Kinethacks, 2012] Применение Kinect на PC /Kinethacks. – [2012]. – Электронный ресурс. Режим доступа: <http://kinethacks.net/>

[Заболеева-Зотова и др., 2011] Развитие системы автоматизированного определения эмоций и возможные сферы применения / А.В. Заболеева-Зотова, Ю.А. Орлова, В.Л. Розалиев, А.С. Бобков // Открытое образование. - 2011. - № 2. - С. 59-62.

[Розалиев и др., 2011a] Розалиев, В.Л. Моделирование эмоционального состояния человека на основе гибридных методов / В.Л. Розалиев, А.В. Заболеева-Зотова // Программные продукты и системы: международный науч.-практ. журнал. – Тверь, 2010 – Вып.2 (90). – С.141-146.

[Розалиев и др., 2011b] Розалиев, В.Л. Применение нейронных сетей и грануляции при построении автоматизированной системы определения эмоциональных реакций человека / В.Л. Розалиев, А.С. Бобков, О.С. Федоров // Изв. ВолгГТУ. Серия «Актуальные проблемы управления, вычислительной техники и информатики в технических системах»: межвуз. сб. науч. ст. / ВолгГТУ. - Волгоград, 2010. - Вып. 9, № 11. - С. 63-68.

[Хайкин, 2006] Хайкин, С. Нейронные сети: полный курс / С. Хайкин. - М.: Филинь, 2006. - 122 с

[Осовский, 2007] Осовский, С. Нейронные сети для обработки информации /С. Осовский. - М.: Финансы и статистика, 2007.- 344 с.

[Боровиков, 2007] Боровиков, В. П. Нейронные сети STATISTICA Neural Networks: Методология и технология современного анализа данных / В. П. Боровиков. - М. : Статсофт, 2007. - 632 с.

### RECOGNITION OF BODY POSTURES USING A 3D CAMERA WITH THE DEPTH SENSOR

Shpirko A.A., Dorofeev N.S., Bobkov A.S.

*Volgograd State Technical University, Volgograd,  
Russian Federation*

**alex555@yandex.ru**

**nikita.dorofei@gmail.com**

**bobkovart@gmail.com**

A new method of controlling correctness of performing physical exercises of physical therapy with the help of 3D camera with depth sensor is described. Process of identifying human body postures by coordinates of some "specific" nodes of the human body using production rules was formalized. With the help of neural networks - multilayer perceptron exercises are classified on right and wrong. Neural network was trained using samples prepared by the experts in the field of physical therapy.

## INTRODUCTION

Relevance of the work by the fact that to date the only method by which to correctly perform the person exercising physical therapy can be controlled - it is a

personal trainer or a doctor - a specialist gymnastics. With self-perform these exercises, it is likely that people will not carry them out correctly, because not all people who need exercise therapy sessions have sufficient knowledge in this area. While automating the control of correctness, the system checks the correctness of the exercises, provides tips on technique performance visually and in text form.

## MAIN PART

Any physical exercise represents motion. In order to work with any motion, first of all, this motion should be captured. 3D sensor from Microsoft, named Kinect, captures the motion, then motion data is recognized and transmitted into BVH file. BVH means "Bio Vision Hierarchical" and represents a special format where human body joints and their positions are described in human-readable numbers.

Also any physical exercise can be characterized by a set of human's poses, which, in their turn represent a set of "specified" body joints. And the correctness of performing an exercise depends on these joints interposition. Information about joint positions is gathered from BVH file, which is parsed in real-time. Also in real-time information from sensor is written in it. Besides Cartesian coordinates, this file contains angles of rotations of joints relative to each other.

A neural network, trained with the help of medical experts classifies poses, which a person take while performing an exercise, and makes a decision while he is doing wrong or right. The training set for neural network contained several thousands of samples of exercises both in correct and in incorrect type. Several methods, which can solve classification task were tested and the best results showed multilayer perceptron.

## CONCLUSION

As a result of work, the rules identifying human posture and exercise have been developed, neural network to test the correctness of the exercise, was constructed and thus, control of correctness of performing physical exercises was automated.

This work was partially supported by the Russian Foundation for Basic Research (projects 12-07-00266, 12-07-00270).