

УДК 004.9

Кудубаева С.А.^{1*}, Рюмин Д.А.^{1.}, Калжанов М.У.²

¹Костанайский государственный университет имени А. Байтурсынова, Республика Казахстан, г. Костанай

²Костанайский государственный педагогический институт, Республика Казахстан, г. Костанай
*E-mail: saule.kudubayeva@gmail.com

Метод опорных векторов для распознавания жестовой речи с использованием сенсора KINECT

Для распознавания жестов использованы различные методы, среди них: метод опорных векторов, скрытые марковские модели, нечёткие модели, модель искусственного нейрона, метод разности изображения. Задачи по распознаванию жестовой речи любого языка характеризуются многими параметрами, в первую очередь, это характеристики канала передачи жестовой речи, размер словаря распознавания, вариативность жестов и т.д. Помимо существующих технико-экономических сложностей, на развитие казахских речевых технологий, которые включают в себя и распознавание жестового языка, в первую очередь, влияют особенности казахского языка и речи, вызывающие сложности в процессе автоматической обработки: множество правил словообразования, наличие семи падежей в образовании существительных, вариативность казахского языка и речи в виду наличия нескольких областей. Разработка базы данных для казахского жестового языка, состоящей из дактильной азбуки 42 жестов, является начальной ступенью при создании системы автоматического распознавания отдельных жестов рук. Для распознавания жестов использован сенсор Kinect, координаты скелета руки полученные от него и ключевые характеристики обработаны через XML файлы при помощи инструментов и математических вычислений в MATLAB. Объектом распознавания жестов в данной работе являются буквы и цифры и некоторые жесты в статичной форме.

Ключевые слова: казахский жестовый язык, дактиль, распознавание образов, сегментация, видео файл.

Kudubayeva S.A., Ryumin D.A., Kalzhanov M.U.
**The method of basis vectors for recognition sign language
by using sensor KINECT**

For signs recognition different methods, such as the method of reference vectors, hidden Markov models, fuzzy models, the model of artificial neuron, image difference method are used. The tasks of recognition of gestural speech of any language are characterized by many parameters. First of all this is characteristics of the transmission channel of gestural speech, recognition vocabulary scope, the variability of gestures, etc. Besides technical and economic problems Kazakh speech technologies development, among which there is Sign language recognition, are firstly influenced by peculiarities of the Kazakh Language and speech (many rules of word formation, seven cases of nouns, territorial varieties of the Kazakh language and speech) causing difficulties in automatic processing. The development of Kazakh Sign Language database consisting of 42 gestures is the first step in designing the system of automatic recognition of separate hand gestures. For gestures recognition Kinect sensor, hand skeleton coordinates obtained via it and key characteristics processed through XML files via tools and mathematic calculations in MATLAB are used. Gestures recognition objects in this article are letters, digits and some gestures in static form.

Keywords: Kazakh Sign Language, dactyl, image recognition, segmentation, video file.

Кудубаева С.А., Рюмин Д.А., Калжанов М.У.
**KINECT сенсорын пайдалана отырып ым тілін танып білудегі
сүйеніш вектор әдісі**

Мақалада есту мүмкіндіктері шектеулі адамдардың ым тілін танып білуде сүйеніш вектор әдісі туралы қарастырылған. Қазақша ым тілі үшін деректер қорын құрастыру, дактильді азбука 42 әріпті ым тілінен тұратын, жеке ым қолдарын автоматты түрде танып білу жүйесін құру бастапқы қадамын жасау. Ым-ишара тілін танып білуде Kinect сенсоры қолданылды, қол сүйегінің координатасы алынады, кілттік мінездеме XML файл инструмент арқылы және MATLAB математикалық есептеулердің көмегімен өңделді. Ым тілін танып білу объектісі берілген жұмыста әріптер, сандар және статикалық формадағы ым ишаралары болып табылады. Ымдарды тануға арланған әртүрлі әдістер қолданылады, олардың ішінде: тірек векторлар әдісі, жасырын марков модельдері, анық емес модельдер, жасанды нейрон моделі, бейнелердің әртүрлілігі әдісі. Кез келген тілдің ыммен сөйлеуді тану бойынша міндеттері көптеген параметрлермен сипатталады, ең алдымен ыммен сөйлеуді жиберу каналдарының сипаттамасы, тану сөздігінің өлшемі, ымдау вариативтілігі және т.с.с. 42 ым дактильді әліпбиден тұратын қазақша ым тілінің мәліметтер қорын әзірлеу автоматтандырылған жеке қол қимылдарын тану жүйесінің алғашқы кезеңі болып табылады. Ымдарды тану үшін Kinect сенсоры қолданылған, қол қаңқасының координаттары алынған және кілттік сипаттамасы MATLAB бағдарламасының құралдары мен математикалық есептеулері көмегімен XML файлы арқылы өңделген. Берілген жұмыста ымдарды тану объектісі ретінде әріптер мен сандар және кейбір статикалық формадағы ымдар болып табылады.

Түйін сөздер: қазақша ымдау тілі, дактиль, бейнелерді тану, сегментация, бейне файл.

1 Введение

На текущий момент в большинстве случаев, задача по распознаванию потоков сводится к делению некоторых наборов данных на множество уже определённых сегментов, кластеров. Используя скрытые марковские методы (СММ) наборы данных можно классифицировать. Кроме СММ, распознавание потоков производится методом опорных векторов, который в свою очередь также является классификатором [1]. Идея метода заключается в том, чтобы все входные данные представлять в многомерном пространстве и сегментировать с помощью гиперплоскостей. Распознавание жеста на основе метода опорных векторов заключается в представлении жеста определённой последовательностью ключевых характеристик, которые впоследствии будут подлежать сравнению и сегментации. На начальном этапе производится разбиение входных данных на пересекающиеся отрезки - фреймы. Затем из каждого фрейма извлекаются характеристики (признаки). Полученные характеристики в свою очередь объединяются в вектора, определяющие жесты. На основе сформированных векторных значений осуществляется классификация жеста на основе метода опорных векторов. Метод опорных векторов (SVM) и СММ, показывают достаточно низкий процент ошибок распознавания.

2 Экспериментальная часть

Исследованиями в области распознавания образов занимаются Крак Ю.В., Heloir A., N. Courty, S. Gibet, и другие. Ранее работы о проблемах казахской жестовой речи, способах и методах распознавания жестов выполнены в работах [2-8]. По типу показа жесты делятся на статические и динамические. При воспроизведении статического жеста не возникает необходимости в движении рук, так как положение кисти и пальцев

рук стационарно в пространстве в течение рассматриваемого промежутка времени. Динамические жесты в свою очередь воспроизводятся путём движения какой-либо части человеческого тела, в большинстве случаев кистью во времени и пространстве. Границы слов в потоке жестовой речи могут быть определены лишь в процессе распознавания (декодирования знаков), посредством подбора оптимальной последовательности жестов, наилучшим образом согласующейся с входным потоком жестов по математическим моделям.

Точность работы систем автоматического распознавания жестового языка существенно улучшается с увеличением размера словаря жестов, который может содержать в себе множество повторений одного и того же жеста. Малый словарь содержит единицы и десятки слов. Такой словарь подходит для распознавания порядка цифр, которые входят в номера телефонов, показаний приборов, а также систем управления подвижными техническими объектами, например, такими как машина, вертолёт и систем управления различным оборудованием, например, бытовой техникой.

Разработка системы автоматического распознавания жестов глухих потребовала для начала создания базы данных для казахского жестового языка – словаря жестов. Начальной ступенью в данном направлении явилось создание базы состоящей из дактильной азбуки сорока двух жестов, показанных на рисунке 1, а также цифр, чисел.

							
							
							
							
							
							

Рисунок 1 – Дактильная азбука казахского жестового языка

Для решения глобальной проблемы распознавания информации (жестов, объектов и т.д.) использованы средства компьютерного зрения: анализ и кластеризация информации. В связи с тем, что даже один человек никогда не покажет один и тот же жест подобно показанному ранее, выявлено, что проведение анализа потока информации намного сложнее, чем её синтез из-за большой вариативности сигналов как по времени, так и по содержанию. Задачи по распознаванию как одномерных сигналов (речи), так и

двумерных сигналов (текста) и трёхмерных сигналов (объемных жестов) еще далеко не решены. Однако стоит заметить, что существуют вероятностные модели (СММ, нечёткая модель, модель искусственного нейрона и т.д.), которые направлены на генерацию и классификацию гипотез распознавания сигналов.

Метод опорных векторов (SVM) – метод классификации, использующий теорию статистического обучения. В условии ограниченного набора обучающих данных, метод опорных векторов обычно превосходит традиционные методы оценки параметров, построенные на законе больших чисел. Первоначально, метод опорных векторов использовался для решения задач бинарной классификации, то есть для поиска максимально удалённой от точек двух классов гиперплоскости, разделяющей эти два класса. В нашем случае, движение, регистрируемое камерой, дискретно измеряющим ускорение по трём осям, можно представить в таком виде: $G = (a_x, a_y, a_z)$ где $a_T = (a_T^0, a_T^1, \dots, a_T^{L-1})$, $T = x, y, z$ – вектор ускорения оси, а L – длина последовательности, количество измерений.

Для распознавания жеста, разделим его на $N + 1$ сегментов одинаковой длины. Каждые два соседних сегмента будут образовывать фрейм, который представим в виде:

$$R_k = (r_{xk}, r_{yk}, r_{zk}), k = 0, 1, \dots, N - 1, r_{Tk} = (r_{Tk}^0, r_{Tk}^1, \dots, r_{Tk}^{2l_s-1}), T = x, y, z$$

Пусть два типа жестов $G1$ и $G2$ должны быть классифицированы. Определим выборку из n тренировочных данных: $\{(\varphi_i, \xi_i)\}, i = 1, \dots, n$ где $\varphi_i \in R^d$ – вектор характеристик и

$$\xi_i = \begin{cases} +1 & \text{если } \varphi_i \in G1 \\ -1 & \text{если } \varphi_i \in G2 \end{cases}$$

Разделяющая гиперплоскость записывается так: $\alpha * \varphi + \beta = 0$ Тогда классифицирующая функция будет иметь вид: $f(x) = \text{sign}(h(x))$, где $h(x) = \sum_{i=1}^n \lambda^* \varphi_i K(\xi_i, \xi_j) + \mu^*$. Где λ^*, μ^* – оптимальное решение.

Однако распознавание жестов подразумевает распознавание более двух жестов, сравнивались. Использование метода опорных векторов показало, что необходимо оптимально подобрать ядро функции-классификатора. Результаты показали, что данный метод не требует больших вычислительных мощностей для непосредственной работы алгоритмов.

Однако распознавание жестов подразумевает распознавание более двух жестов, поэтому сравнивались более 30 изображений с классом идеальных жестов – эталонами. В нашем случае класс $G1$ – это класс идеально воспроизведенных изображений статичных жестов, а класс $G2$ – это все другие, сравниваемые. В казахском жестовом языке цифры 1-10 являются статичными, поэтому вся работа по распознаванию жестов начата с этих жестов. Относительно специфичных букв "ң, ғ, қ, ұ, ә" и других.

В реальных условиях использование метода моментов изображения для распознавания рук часто оказывается недостаточно из-за неоднородности фона. В связи с этим необходимо также использование метода разности изображений, который, в свою очередь, призван анализировать входной видеопоток данных с выявлением разности изображений, через нахождение центра масс при движении руки человека [9], в соответствии с рисунком 2. Благодаря данному методу отпадает необходимость в обязательном наличии однородного фона в потоке трёхмерных сигналов, следствием чего является более надежное функционирование программного обеспечения в режиме реального времени.

Запись элементов казахского жестового языка производилась, при помощи сенсора Microsoft Kinect 2.0, на расстоянии от 1,5 до 2 метров, что является самым оптимальным расстоянием. Задний фон состоял из стены с однородным светлым фоном. Захват потока видео осуществлялся с разрешением 956 на 824 пикселя и частотой 25 кадров в секунду.



Рисунок 2 – Вычисление разности изображений

Корневой каталог базы данных состоит из 60 подкаталогов, 10 из которых содержит информацию о жестах, показывающих числа в промежутке от 1 до 10. Остальные каталоги хранят информацию о дактиле казахского языка жестов, состоящего из 42 букв и некоторых слов. Отдельно взятый подкаталог включает в себя 50 видео файлов с записанным одним и тем же жестом, такое же количество текстовых файлов с координатами скелета найденного человека в потоке кадров при помощи сенсора Kinect, разделенного на 25 суставов. Каждая определенная точка - это пересечение двух осей (X, Y) на координатной плоскости и дополнительное значение Z с двойной точностью, обозначающее глубину точки, которая измеряется расстоянием от сенсора до точки объекта в диапазоне от 0 до 1. Основные характеристики базы данных представлены в таблице 1.

Таблица 1– Основные характеристики базы данных

Сенсор	Microsoft Kinect 2.0
Задний фон	однородная светлая стена
Расстояние до камеры	$\approx 1,5 - 2$ метра
Разрешение захвата видео потока	956×824
Частота кадров	25 к/с
Формат видео	mp4
Общий объем базы данных	≈ 10 Гб
Общая длительность видео	≈ 2 часа

Кроме описанных файлов также имеется текстовый файл, хранимый служебную информацию о жесте. Среднее время одного видео файла составляет $\approx 4 - 5$ секунд. Для обучения автоматизированной системы распознавания отдельных жестов, использовалось 10 видео файлов с привязанными к ним текстовыми файлами с координатами о необходимых скелетных точках. Данные файлы принято считать правильными эталонами показа жеста, а остальные используются, как тестовые данные.

Для вывода информации о жестах в диалоговой системе, создана трёхмерная (3D) модель руки человека, которая представлена на рисунке 3.



Рисунок 3 – Трёхмерная модель руки человека

Скелет модели руки с наложением текстур спроектирован в Autodesk 3ds Max. Далее произведено разбиение модели на кластеры, которые в свою очередь состоят из множества градусных координат осей X, Y, Z трёхмерного пространства. Данные кластеры формируют XML файлы. На следующем этапе осуществлен автоматический программный парсинг созданных XML файлов через Microsoft Visual Studio в связке с языком программирования $C\#$ и произведено динамическое извлечение полученных градусных координат модели. После парсинга спроецирована 3D модель на виртуальной плоскости координат через векторную систему для построения клиентских приложений WPF, которая в первую очередь направлена на адаптивность модели. Данные процедуры предназначены не только для создания трёхмерной модели, но и для наполнения базы данных.



Рисунок 4 – Изменённая конфигурация трёхмерной модели руки человека

Реализация системы в целом производилась при помощи инструмента математических вычислений MATLAB [10], который позволил исследовать и спроектировать модели с альтернативными подходами к решению. Также применялись вспомогательные программные средства и библиотеки OpenGL [11], OpenCV [12]. Официальный SDK [13] от компании Microsoft предоставил возможность получения необходимых данных с сенсора Kinect. Тестирование автоматизированной системы производилось на вычислительных машинах с разной производительностью. Предложенная система распознавания некоторых статичных жестов рук, протестирована на базе данных, которая включает в общей сложности 1560 жестов двух человек. Каждый класс включает жест, показанный 30 раз. Эталоны каждого жеста выбраны в случайном порядке в количестве 5 образцов для каждого класса. Остальные 25 образцов являются тестовыми. Качество и полнота распознавания оказалась в промежутке 0,63 до 0,91, средняя полнота составляет - 0,70, а точность 0,87. Данные результаты получены на базе данных, которая предварительно

подверглась нарезке на фрагменты, без дополнительных обработок. Характеристики качества распознавания цифр и некоторых букв представлены, в соответствии с таблицей 2.

Таблица 2– Характеристика качества распознавания отдельных жестов

Жест	Полнота	Точность	Жест	Полнота	Точность
1	0,63	0,89	Ә	0,74	0,86
2	0,62	0,90	Ғ	0,71	0,84
3	0,71	0,85	Қ	0,72	0,84
4	0,72	0,88	Ң	0,67	0,88
5	0,75	0,87	Ө	0,73	0,86
6	0,72	0,87	Ү	0,61	0,83
7	0,72	0,85	У	0,63	0,84
8	0,68	0,86	h	0,75	0,90
9	0,70	0,87	i	0,76	0,87
10	0,67	0,84			

Из полученных результатов видно (таблица 1), что средняя точность распознавания, вычисляемая по формуле (1) составляет 0,87 (87%).

$$x_c = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (1)$$

где n – количество жестов; x_n – точность распознавания жеста. Данные результаты получены на записанной базе данных.

Наиболее низкую точность распознавания показали жесты, при демонстрации которых необходимо располагать пальцы на относительно близком расстоянии друг от друга под разными углами. Повысить точность в таком случае возможно за счет увеличения количества эталонов, а также разработке алгоритма способного осуществлять процессы размыкания и замыкания не только бинарного, но и цветного изображения. Это позволит гибко управлять цветовыми параметрами и определять не только более точно координаты объектов, но и их дескрипторы. Из полученных результатов видно, что средняя полнота составляет – 0,70, а точность 0,87. Данные результаты получены на базе данных, которая предварительно подверглась нарезке на фрагменты, без дополнительных обработок. Проведены три типа экспериментов, направленные на проверку способности к классификации метода опорных векторов (SVM). Первый эксперимент оценивал способность алгоритма SVM с предварительно сегментированными жестами из набора номера-форм данных. Второй эксперимент оценивал классификацию способности алгоритма относительно минимального количества обучающих данных. Третий эксперимент испытания изучал способность SVM алгоритма правильно классифицировать данные из данных номеров-форм в виде непрерывного потока данных, который также содержал ряд нулевых жестов. Это очень важный тест для данного исследования, так, если модель может достичь такого же хорошего результата классификации по 5 обучающим примерам, как и по 50 – ти примерам обучения, то исполнитель может

экономить время сбора данных обучения, а также в подготовке модели. С помощью всего лишь 3 примеров функции обучения SVM достигается значение неточности 3,61%. На 10 примеров функции обучения SVM достигли практического результата распознавания на 81,67%, однако стандартное отклонение во всех 10-ти примерах по-прежнему высоко - около 13,35%. Для 13 примеров функции обучения SVM достигли превосходное значение точности на 81,3% с умеренным стандартным отклонением 6,81%. Используя 20 примеров обучения, SVM достигает 81,76% с очень небольшим стандартным отклонением 1,84%. Результаты этого теста позволяют предположить, что по крайней мере 11 примеров обучения требуются на жест, если пользователь хочет достичь надежного результата классификации более 80%.

Одно интересное наблюдение было получено, глядя на лог-файлы, созданные во время эксперимента. Файлы журнала показали, что для жестов, которые были неправильно классифицированы как нулевой жест, апостериорная вероятность каждого такого жеста, имела очень малое значение (обычно ниже 0,1).

Это противоположно случаю, когда были жесты правильно классифицированы, с k – ми жестами, апостериорная вероятность обычно имеющей значение больше 0,7 а суммирование всех остальных жестов около 0,3. Другими словами, когда SVM сделал правильную классификацию вероятность того, что жест имел место, была высокой, в то время, когда SVM не в состоянии сделать правильную классификацию вероятность того, что не был допущен жест (Неправильно) – низкая.

Эти результаты свидетельствуют о том, что SVM будет плохим выбором для распознавания многомерных жестов из непрерывного потока данных, даже несмотря на достигнутые SVM отличные результаты с тем же набором данных, когда жесты были предварительно сегментированы.

В своем нынешнем состоянии, полученные данные по алгоритму SVM не могут быть использованы в качестве алгоритма распознавания в режиме реального времени для жестов из непрерывного потока данных, например, во время разговора жестами глухих, чтобы автоматически классифицировать жест с любой большой точностью. Дальнейшие исследования будут продолжены с комбинацией других методов, например DTW-алгоритм динамической трансформации временной шкалы. Изучив оба CMM и SVM как потенциальные методы распознавания жестов из непрерывного потока данных, был сделан вывод о том, что новый подход будет необходим с изучением точных порогов выбраковки.

Использование алгоритма опорных векторов является очень мощным алгоритмом распознавания образов, который имеет следующие преимущества для распознавания многомерных жестов:

1. Можно легко решить нелинейную задачу классификации с использованием трюка ядра.
2. Определение параметров модели SVM соответствует выпуклой оптимизации проблема и поэтому любое локальное решение также является глобальным решением (в отличие от например искусственной нейронной сети, которая обычно может застрять в локальной минимальной, не найдя глобального минимума)
3. SVM может легко работать с очень большими размерами векторов признаков

4. Даже при большом количестве обучающих данных с высокой размерностью векторов признаков, алгоритм SVM может обучить модель в очень короткий промежуток времени обучения

Недостатки метода. Метод опорных векторов неустойчив по отношению к шуму в исходных данных. Если обучающая выборка содержит шумовые выбросы, они будут существенным образом учтены при построении разделяющей гиперплоскости.

До сих пор не разработаны общие методы построения спрямляющих пространств или ядер, наиболее подходящих для конкретной задачи. Построение адекватного ядра является искусством и, как правило, опирается на априорные знания о предметной области.

3 Заключение

Наиболее низкую точность распознавания показали жесты, при демонстрации которых необходимо располагать пальцы на относительно близком расстоянии друг от друга под разными углами. Повысить точность в таком случае возможно за счет увеличения количества эталонов, а также разработке алгоритма способного осуществлять процессы размыкания и замыкания не только бинарного, но и цветного изображения. Это позволит гибко управлять цветовыми параметрами и определять не только более точно координаты объектов, но и их дескрипторы.

Таким образом, алгоритм выделения рук принимает на вход видеофайл, затем циклически извлекает кадры. Для каждого кадра производится пороговое отсечение яркости, до тех пор, пока не будет определен оптимальный порог. Далее осуществляется заполнение пустот объектов, с последующим поиском на изображении. После удаления шумов, подсчитывается количество найденных объектов на основе вычисления их площадей и определяются руки человека. Из найденных объектов - рук, извлекаются информативные признаки необходимые для распознавания на основе метода опорных векторов.

4 Дискуссия

Авторы статьи являются разработчиками сайта [14] для людей с ограниченными возможностями по слуху. Многие вопросы, связанные с распознаванием дактиля и жестовой речи, изучены пока недостаточно. Традиционные многочисленные методы и алгоритмы цифровой обработки сигналов основаны на использовании фильтрации сигналов, корреляционных и спектральных анализах, вейвлет преобразованиях. Данные приёмы позволяют распознавать визуальные образы с использованием методов сегментации, кластеризации, классификации с использованием вероятностного моделирования и распознавания видеосигналов на основе различных типов искусственных нейронных сетей, скрытых цепей Маркова, динамического программирования, пиксельной обработки видеопотока. Все устройства, которые применяются для распознавания жестов, делятся на две группы: устройства, основанные на видеоданных и устройства, которые используют датчики: акселерометр, гироскоп и магнетометр. Лучший результат дают системы, которые помимо RGB сенсора имеют IR Depth сенсор (сенсор глубины) для распознавания жестов на небольшом расстоянии. В связи с этим авторы данной работы

использовали инструменты математических вычислений MATLAB для обработки данных от сенсора Kinect. Основная проблема распознавания жестовой речи состоит в том, что распознать отдельные жесты недостаточно. Необходимо распознать непрерывную жестовую речь, ее смысл, в этом направлении продолжены наши исследования.

Литература

- [1] *Котельников Е.В.* Параллельная реализация машины опорных векторов с использованием методов кластеризации / Е.В. Котельников, А.В. Козволина // Материалы международной научной конференции "Параллельные вычислительные технологии". – Санкт-Петербург, 2008.
- [2] *Kudubayeva S., Yermagambetova G.* "Parametric representation of the Kazakh gestural speech" Lectures Notes in Artificial Intelligence in Bioinformatics, SPECOM-2014.
- [3] *Krak Iu., Kudubayeva S.* "System for Sign Language Modeling and Learning" 8Th IEEE International Conference on Application of Information Technology, Astana, 2014.
- [4] *Фархадов М.П., Кудубаева С.А., Ермагамбетова Г.Н.* Теория скрытых Марковских моделей и ее применение для распознавания речи // Научный журнал "Вестник ЕНУ им. Л.Н. Гумилева", Астана, ноябрь 2013.
- [5] *Farkhadov M.P., Kudubayeva S.A., Ermagambetova G.N.* Acoustic simulator for the hearing impaired people // American International Journal of Contemporary Research, USA, Vol. 3 No. 11; November 2013.
- [6] *Kudubayeva S.A., Ermagambetova G.N.* To the problem of speech recognition // American International Journal of Contemporary Research, USA, Vol. 3 No. 11; November 2013.
- [7] *Кудубаева С.А., Ермагамбетова Г.Н.* Задача распознавания казахской жестовой речи // Международные Сатпаевские чтения – 2014 "Роль и место молодых ученых в реализации стратегии" Казахстан-2050", посвященные 80-летию КазНТУ имени К.И. Сатпаева, Алматы, 11-12 апреля 2014 г.
- [8] *Крак Ю., Кудубаева С., Ермагамбетова Г.* Жестовый язык и проблемы его распознавания // IV Международной научно-теоретической конференции "Роль физико-математических наук в современном образовательном пространстве", Атырауский государственный университет имени Х. Досмухамедова, Атырау, 24-25 апреля 2014 г.
- [9] *Фисенко В.Т., Фисенко Т.Ю.* Компьютерная обработка и распознавание изображений / В.Т. Фисенко, Т.Ю. Фисенко // Учебное пособие, СПб НИУ ИТМО, 2008. – С. 68-71. 192 с.
- [10] *MathWorks – Makers of MATLAB and Simulink [Электронный ресурс]* – Режим доступа: <http://www.mathworks.com> (дата обращения: 20.03.2016).
- [11] *OpenGL – The Industry Standard for High Performance Graphics [Электронный ресурс]* – Режим доступа: <https://www.opengl.org> (дата обращения: 20.02.2016).
- [12] *OpenCV [Электронный ресурс]* – Режим доступа: <http://opencv.org> (дата обращения: 20.04.2016).
- [13] *Kinect for Windows SDK 2.0 [Электронный ресурс]* – Режим доступа: <https://www.microsoft.com/en-us/download/details.aspx?id=44561> (дата обращения: 20.01.2016).
- [14] *www.surdo.kz*

References

- [1] *Kotelnikov Ye. V.* Parallel Implementation of Reference Vectors Machine Using Clustering Method / Ye. V. Kotelnikov, A. V. Kozvonina // Proceedings of the international scientific conference "Parallel Computing Technologies". – St. Petersburg , 2008.
- [2] *Kudubayeva S., Yermagambetova G.* "Parametric representation of the Kazakh gestural speech" Lectures Notes in Artificial Intelligence in Bioinformatics, SPECOM-2014.
- [3] *Krak Iu., Kudubayeva S.* "System for Sign Language Modeling and Learning" 8Th IEEE International Conference on Application of Information Technology, Astana, 2014.

-
- [4] *Farkhadov M.P., Kudubayeva S.A., Ermagambetova G.N.* Theory of Hidden Markov Models and its Implementation for Speech Recognition // science journal "Herald of L.N Gumilyov ENU", Astana, November 2013.
- [5] *Farkhadov M.P., Kudubayeva S.A., Ermagambetova G.N.* Acoustic simulator for the hearing impaired people // American International Journal of Contemporary Research, USA, Vol. 3 No. 11; November 2013.
- [6] *Kudubayeva S.A., Ermagambetova G.N.* To the problem of speech recognition // American International Journal of Contemporary Research, USA, Vol. 3 No. 11; November 2013.
- [7] *Kudubayeva S.A., Ermagambetova G.N.* Problem of Kazakh Gestural Speech Recognition // International Satpaev Readings – 2014 "Role and Place of Young Scientists in Implementation of "Kazakhstan-2050" Strategy" dedicated to 80 th anniversary of K.I. Satpaev KazNTU, Almaty, April 11-12, 2014.
- [8] *Krak Yu., Kudubayeva S., Ermagambetova G.* Sing Language and Problems of its Recognition // Proceedings of IV International scientific theoretical conference "The Role of Physical Mathematical Sciences in Modern Educational Space", Kh. Dosmukhamedov Atyrau state university, Atyrau, April 24-25, 2014.
- [9] *Fissenko V.T., Fissenko T.Yu.* Computer Processing and Image Recognition / V.T. Fissenko, T.Yu. Fissenko // Manual, Spb SRU ITMO, 2008. – P. 68-71. 192 p.
- [10] *MathWorks - Makers of MATLAB and Simulink [Electronic resource]* – Access mode: <http://www.mathworks.com> (reference date: 20.03.2016).
- [11] *OpenGL - The Industry Standard for High Performance Graphics [Electronic resource]* – Access mode: <https://www.opengl.org> (reference date: 20.02.2016).
- [12] *OpenCV [Electronic resource]* – Access mode: <http://opencv.org> (reference date обращения: 20.04.2016).
- [13] *Kinect for Windows SDK 2.0 [Electronic resource]* – Access mode: <https://www.microsoft.com/en-us/download/details.aspx?id=44561> (reference date: 20.01.2016).
- [14] <https://www.surdo.kz>