## 2 Проектът

Проектът, който ще Ви представя е разработен в университетът WestBohemia university в Пилзен, Чехия, където имах удоволствието да замина на обмен по програма Еразъм+. Още когато проучвах възможностите за обмен, интересът ми беше привлечен от техните изследвания в сферата на Motion Capture или така нареченото проследяване на движение и 3D моделиране. Още повече бях впечатлена от социалната насоченост на проекта. Аз самата имах период, в който бях лишена от гласа си и въпреки че беше временно, щом видях че имат проект свързан с това веднага изявих желание да се включа. Бях приета да участвам в този проект. Той е под егидата на министерство на образованието на Чехия. Негова цел е да се създаде система за улесняване комуникацията на хора със слухови и гласови увреждания и достъпът им до информация. Проектът е мултидисциплинарен. Разработва се от чешки специалисти в различни области – лингвисти и специалисти в ИТ. Проектът все още не е завършен.

## 3 Езикът на глухонемите

По света съществуват над 300 различни езика на глухонемите. Също както говоримите езици се различават от региона, в който се развиват така съществуват различни форми и диалекти на този тип езици. Съществена разлика между говоримия език и езика на знаците е че докато говоримият език е линеен, думите се подреждат в изречения, изреченията в текстове, една дума носи едно значение, при езика на знаците една информационна единица носи повече от едно значение и това се дължи на факта че значението се предава в пространството и то се предава едновременно от движението, формата и позицията на жестовете, както и от мимиките на лицето. Тази разлика в предаването на информация води и до усложняването на представянето на знаците и автоматичната им обработка. Друга разлика между говоримият език и езикът на знаците е че докато при добре познатия ни говорим език има ясно обособена система за писменото му представяне (азбуката) тук няма общоприета единна система с ясно изразени писмени и граматични правила.

Има различни школи, които развиват различни системи. Българският жестомимичен език е вариация на руския, който пък е по австро-унгарска линия, а той от своя страна – по френска. Към същата група принадлежат чешки, словашки, полски, унгарски и румънски жестомимичен език.

Близостта в говоримия език не гарантира сходство в жестомимичния – например американски – британски английски, испански – каталунски.

Друга особеност е начина на мислена на хората родени с подобни увреждания. Докато говорещите хора мислят на конкретен език, хората които никога не са чували говорима реч мислят на жестомимичен език. Поради което много по-трудно биха се научили да пишат и четат на езика на страна в която живеят.

Съобразявайки се с всички тези разлики между двата езици и трудностите пред хората с увреждания в университетът в Хмабург развиват система за означаване на знаците.

Системата цели да представи как точно се извършва знака, а не какво означава. Например отбелязва се дали знакът симетричен – огледално/ паралелно, описание на мимиките, начална форма, ориентация и позиция на ръката, и движение ако е динамичен знак. Предимство на системата е универсалност спрямо различните езици. Тя е възприета от работещите по проекта. И моето изследване също е базирано на тази система за обозначаване на знаци.

## 4 Съществуващи решения

Sign 3D и SignCom са проекти разработени за френския език на знаците, ViSiCAST и eSign editor са международни проекти, спонсорирани от ЕС. Всички тези съществуващи и работещи системи са базирани на знание. За създаването на базата от знания за тези системи се използват средствата за MoCap (проследяване на движение). Първите версии на тези системи или първите опити в тази област са базирани на БД от знаци записани със стандартно видео. С времето се установява, че този подход е тромав. В същото време оптичните системи за проследяване и запис на движение стават по-достъпни. Тяхно предимство в тази област е че дават гъвкавост за съставяне на нови знаци, без да има нужда от презаписване на ново видео., освен това преносът на данни е много по-малък от колкото при стандартно 2Д видео. Проблем/ недостатък на тези системи е, че съставянето на базата от данни е съпътствано от дълъг процес на последваща обработка, костващ както времеви така и човешки ресурс.

При тези системи различните компоненти на знаците (жестове и мимики) са записани отделно, докато при базата на този проект те са записани едновременно.

## 5 Цел

Целта на моят проект е да създаде инструмент за автоматичен анализ и обработка на база от записани знаци, по този начин да се съкрати и улесни процеса за създаване на база от знания. За целта трябва дълга последователност от записани жестове да се разбие на по-къси, които после да се анализират и опишат по системата за обозначаване на Хамбургския университет. След това на базата на описанието на свойствата на знаците и с помощта на методи за надзиравано машинно самообучение да се класифицират.

## 6 Архитектура на системата

Системата се състои от три модула:

Модул за входни данни – прочита се информацията за записаната сесия : кадрова честота, списък от етикетите на използваните маркери, общ брой кадри/ дължина на записа; начално преработване на данните и подготовката им за по-нататъшна обработка

Модул за сегментация – извършва се груба сегментация на базата на кинематични свойства на движението

Модул за анализ на знак – уточняват се моментите на начало и край на същинската част на знака, извличане на свойствата на знака съобразени с хамбургската система за означаване

### 7-8 Модул за входни данни

Описание на входния файл, процеса на заснемане и 3д визуализация на входните данни

- Входен файл: речников файл представлява запис от една сесия, винаги започва и завършва в Т-поза (изискване на с-ма за запис на движение), съдържа последователност от морфологични единици (които може да са един знак или композиция от знаци) разделени от статична поза наречена „rest pose”.

* Глобален Центъра на координатната система зададен по време на калибриране на системата
* Създаден метод за промяна на координатната система. По времена сегментация, се задава средата между двата предни маркера на талията. Методът има възможност за връщане както в абсолютна координатна система, така и за динамично задаване на друг център, базиран на група маркери.
* Метод за определяне на активната ръка, защото сегментацията се извършва на базата на движението на по-активната ръка

### 9 Модул за сегментация

Под сегментация се разбира разбиването на дълга последователност от движения/ жестове на по-кратки, значими части. В смисъла на този проект два вида сегментация се извършват. Движения записани по време на една сесия да се разбият на речникови единици – от излизането от начална поза до влизането в начална поза.

Всяка една такава единица се състои от преход ( transition), същинска част, …… (retraction).

След това за всяка такава единица да се определят началният и крайният момент на съществената част.

Изследване на кинематичните свойства на движението

### 10 Модул за сегментация – метод

Преди да обясня методът по който е съм реализирала сегментацията, ще направя едно уточнение. Когато говоря за движение на ръката имам предвид движение на доминантната ръка, защото ако знакът се изпълнява с една ръка, то това е доминантната, а ако е с двете ръце и е симетричен, винаги се обозначава само доминантната, ако знакът не е симетричен се обозначават двете ръце като първо е доминантната. При сегментация от водещо значение е движението на доминантната ръка.

Друго което бих искала да отбележа е, че за този процес е от значение цялостното движение на цялата ръка. Затова съм реализирала прост метод, който на базата на 4 маркера на китката, изчислява положението на един маркер, позициониран между тях.

Методът за идентифициране на моментите в които ръката напуска изходна позиция и навлиза в такава, който съм реализирала се базира на изследване на кинематичните свойства на движението на ръката. Идеята е, че промяна в траекторията на движение на ръката, може да бъде индикирано от промяна в скоростта на движение.

Това което виждаме на графиката на слайда е графика на скоростта на движение на дясна ръка по време на цял запис. И погледната така изглежда доста очевидна връзката между скоростта и промяната в траекторията.

Сините точки на графиката, показват локални екстремуми.

11

Реално погледнато за да пресметнем скоростта на движение на ръката е достатъчно да пресметнем 1ва производна на функцията на траекторията. А за да определим моментите на локалните екстремуми е достатъчно да се пресметне 2ра производна (или иначе казано ускорението) и да се определят моментите, в които функцията променя знака си.

Както вече казах, за целта на сегментация се интересувам от цялостното движение на ръката, затова е възможно свеждането на функцията до едно измерение. За целта върху данните за скоростта на движение прилагам нормализация чрез уравнение на Фробениус. На фигурата е показан графика на функцията на скоростта по 3те оси ( червена, зелена, синя линия) и резултатната лилава линия – след приложена нормализация.

12

На тази графика е показана графика на функцията на ускорение на движението на ръката. Както се вижда на тази графика започва да се наблюдава шум. Само за сравнение ще покажа графика траекторията. При нея почти не се наблюдава подобен шум, Съществуват различни причини за шум в данните.

13

За да избегна грешно намерени точки, прилагам ниско-честотен филтър в/у данните от ускорението на ръката. Същественото при прилагане на такъв филтър е избирането на подходяща отрязъчна честота. За целта проведох експеримент, след направено проучване. В изследване за американския език на знаците МкДоналд изследва причините за възникване на шум. Според неговото изследване, шум с честота над 12Hz се дължи на записваща техник и може да бъде елиминиран. Друг резултат от това изследване е че представя таблица с части от тялото и честоти на движение които те могат да предизвикат. На базата на праговите честоти представени в тази таблица – за китка това е 4-5 Hz направих сравнение с резултатните графики и след прилагане на филтъра. Експериментът извърших в/ъ различни сегменти от 4 тестови входни файла.

### 14 Модул за сегментация – етикиране

Вече след като обясних как се стига до тази графика и какво представлява тя, може да се върна на нея и да обясня т.нар labeling процес или иначе казано етикиране на моментите, които представляват интерес за сегментацията. За този процес съм реализирала прагова функция, която цели да филтрира моментите в които ръката е в изходна позиция – RP. За разлика от говоримия изказ, разделянето на думи става с пауза- тишина, тук скоростта на ръката никога не е нула. Затова и е необходимо реализирането на прагова функция. Ако стойността на скоростта в този момент е под прага се отбелязва с 0, иначе с 1. Така последователността от 0 и 1 се разглежда последователно и за старт се избират кадри които са отбелязани с 0 и са предходвани от поне една 0 и последвани от поне две 1, а з край – кадри, отбелязани с 1 и са предходвани от поне една 1 и последвани от две 0. Тъй като се цели определяне на моменти, в които ръката напуска/ навлиза изходна позиция, се прави и проверка, дали позицията и е в областта на ханша. Крайният резултат от тази функция е списък от двойки номера на начален и краен кадър. Знаейки тези номера, ги използвам за да анализирам статистически, стойностите на скоростта по време на така наречената изходна позиция, които в идеалния случай биха били нула, но в реалния никога не са. За да съм сигурна в коректността на крайния метод, повтарям процедурата с нова стойност на параметъра за праговата функция, избран от стойностите на скоростта по времена изходната поза.

### 15 Модул за анализ на знак

Разполагайки с моментите, в които ръката напуска изходна позиция и навлиза в такава, мога да анализирам всеки сегмент индивидуално.

1ва стъпка в този процес е да определя началото и краят на съществената част от знака. Този процес отново се базира на скоростта на движение на ръката, и за началото имаме - 1я минимум след 1я максимум и съответно последния минимум след последния максимум. Това се обяснява по чисто физически причини- за да достигне до позицията в която ръката ще извърши знака или когато приключи знака и за да се върне в изходна позиция, има момент на подготовка. За съжаление в някои случаи, крайният момент не винаги е ясно изразен. Просто защото когато човек извършва движението, в началото е много по концентриран в това какво трябва да направи от колкото в края, понякога ръката просто плавно продължава движението си, намалявайки скоростта си докато достигне изходна позиция.

### 16 Модул за анализ на знак

Вали сняг - sněžit

Ето пример за един знак и изход на системата. Определено е дали знакът се изпълнява с една или две ръце, коя е доминантната ръка, местоположението на ръцете по-време на знака. За определяне на доминантна ръка, съм реализирала метод който изчислява общото изместване на всяка от двете ръце и сравняване на двете стойности спрямо прагова стойност ( =9/10 от средно аритметично отместване на 2те ръце и длани). За да се определи местоположението на ръката, пространството за извършване на знаците пред човека съм разделила на 15 области.

### 17 Модул за анализ на знак – местоположение на ръката

Ето пример за един знак и изход на системата. Определено е дали знакът се изпълнява с една или две ръце, коя е доминантната ръка, местоположението на ръцете по-време на знака. За да се определи местоположението на ръката, пространството за извършване на знаците пред човека съм разделила на 15 области. Разделението е базирано на разделението на пространството на Хмабургската система за обозначаване на знаци. Границите на тези 15 зони се определят от позицията на маркерите за всеки кадър, също така има зададени допустими граници (margin), които отново се изчисляват динамично за всеки кадър. (вертикална – 1/10 от дължината на торса, хориз – 2/10 от ширината на раменете)

Статистическата информация, която се извежда – бр промени на областите, времето прекарано в дадена област целят да бъдат използвани за понататъшно анализиране на знака, при описание на движението и определяне дали знакът е симетричен – огледален или паралелен.

## 18 Оценяване на метода

Както вече обясних сегментирането на знаците не е лесен процес за реализиране от система, дори когато този процес се извършва от експерти лингвисти или от глухо-неми събужда противоречия. За да мога да оценя коректността на данните на системата, извърших ръчно сегментиране на 4 тестови файла. Моите резултати бяха съпоставени с тези на чешки студент, който работи по последващата обработка на данните след запис.

На таблицата са показани резултати от сравняването на резултатите от автоматична и ръчно извършена сегментация.

При груба сегментация – за да се определи началния кадър на текущия знак за правилен, той трябва да е между крайния кадър на предходния знак и началния на текущия от ръчната сегментация. Съответно крайния – м/у крайния на текущия и началния на следващия.

Както се вижда определянето на началото е много по-точно от определянето на края, друго важно да се отбележи е средната грешка е 13 кадъра, което е по-малко от 0,1 сек. В проект за френския език на знаците, които използват този метод за сегментация, за да прецизират ръчно сегментирани данни отклонение от 0,15сек се смята за незначително.

При 2то ниво на сегментация, за да определя началото и края за коректни, трябва грешката да е по-малко от 6 кадъра, тъй като това е средната грешка.

## 19 Обогатяване на списъка със свойства на знака

* Реализиране на методи за определяне на ориентацията на ръката
* Реализиране на методи за определяне първоначалната форма на ръка както
* Реализиране на методи за машинно самообучение и класифициране на знаците