

VILNIAUS UNIVERSITETAS  
MATEMATIKOS IR INFORMATIKOS FAKULTETAS  
PROGRAMŲ SISTEMŲ KATEDRA

# **Gestų kalbos atpažinimas naudojant įprastą web kamerą**

**Sign language recognition using web camera**

Kursinis darbas

Atliko: 3 kurso I grupės studentas  
Pranciškus Ambrazas (parašas)

Darbo vadovas: asist. Linas Petkevičius (parašas)

Vilnius, 2017

## TURINYS

IVADAS .....	2
1. APSIMOKANČIOS SISTEMOS IR GESTŲ KALBA .....	3
1.1. Apsimokančios sistemos .....	3
1.2. Gestų kalba .....	3
1.2.1. Gestų kalbos skirstymas .....	3
2. GESTŲ KALBOS ATPAŽININIMAS .....	4
2.1. Problematika .....	4
3. REKURENTINIAI NEURONINIAI TINKLAI .....	5
3.1. Principas .....	5
4. MEDŽIAGOS DARBO TEMA DĖSTYMO SKYRIAI .....	6
4.1. Poskyris.....	6
4.1.1. Skirsnis .....	6
4.1.1.1. Straipsnis .....	6
4.1.2. Skirsnis .....	6
5. SKYRIUS .....	7
5.1. Poskyris.....	7
5.2. Poskyris.....	7
REZULTATAI IR IŠVADOS .....	8
SANTRUMPOS .....	9
PRIEDAI .....	9
1 priedas. Niauroninio tinklo struktūra .....	10
2 priedas. Eksperimentinio palyginimo rezultatai .....	11

## Įvadas

Daugiau nei 360 milijonų žmonių pasaulyje kenčia dėl klausos ir kalbos įvairių problemų, o daugiau 32 milijonai jų yra vaikai ir šis skaičius vis auga [<http://www.who.int/mediacentre/factsheets/fs300/en/>]. Gestų kalba yra pagrindinis šių žmonių bendravimo įrankis. Tačiau reiktų pastebėti ir tai, jog dalis jų moka skaityti iš lūpų. Norint žmogui be šių ydų bendrauti su gestakalbiu (*gestų kalba kalbantis žmogus*) reikia vertėjo, kuris išverstų gestų kalbą į įprastinę ir atvirkščiai.

Kiekviena pasaulyje esanti kalba turi ir savo gestų kalbą. Tai reiškia, skiriasi tiek gestų kalbos gramatika, tiek netgi patys gestai. Pasaulyje randama net dialektų pagal regionus, ne tik pagal šalis. Pavyzdžiui, amerikiečių anglų kalba šnekančių žmonių pasaulyje yra apie 500 tūkstančių [**remiantis United States Census Bureau's data from the 2009-2013 American Community Survey**]. Todėl netgi bendraujant dviem žmonėms, mokantiems gestų kalbą neretai iškyla vertimo problema, tad tenka ieškoti gestų vertimų. Paiešką galima atlikti atsižvelgiant į delno padėtį, rankos judesį (ar net abiejų rankų), rankos padėtį ir kiek rankų atlieka gestą. Tuomet pagal gesto išvaizdos paveikslėlius ar kartais net vaizdo įrašus, gestakalbiai gali išsiversti gestus. Tam yra skirtos tiek internetinės svetainės - žodynai, tiek įvairūs rašytiniai žodynai.

Šio tyrimo tikslas - padėti ne tik gestakalbiams tarpusavyje, bet ir žmonėms, nesuprantantiems gestų kalbos bendrauti su gestakalbiais tam pasitelkiant technologijas, taip suteikiant šiems žmonėms pilnavertį gyvenimą bendraujant su kitais. Šiuo tyrimu siekiama apžvelgti ir įvertinti ar naudojantis įprasta kompiuterine kamera įmanoma paversti gestų kalbą rašytiniu tekstu ar net garsine kalba ir lygiai taip pat versti rašytinę ar garsinę kalbą į gestų kalbą. Taip pat siekiama, kad vėliau būtų sukurtas visiems gestakalbiams prieinamas produktas ar programinė įranga, kurią kiekvienas įsidiegęs į savo įrenginį - kompiuterį, mobiliųjį telefoną ar planšetinį kompiuterį galėtų naudotis šiuo vertėju. Taip pat tai galėtų tapti ir mokomąja gestų kalbos priemone.

Šiuo metu yra gaminamas vienas iš tokių produktų, tačiau tai yra atskiras įrenginys, turintis įmontuotą kamerą, kuri be vaizdo taip pat fiksuoja ir atstumą iki tam tikrų objektų (šiuo atveju rankos), tačiau produkto kūrėjai sako, jog jų įrenginys galės versti gestų kalbą į anglų ir kitas kalbas ir lygiai taip pat versti įprastą kalbą į rašytinę kalbą. Plačiau: <http://www.motionsavvy.com/>

# 1. Apsimokančios sistemos ir gestų kalba

## 1.1. Apsimokančios sistemos

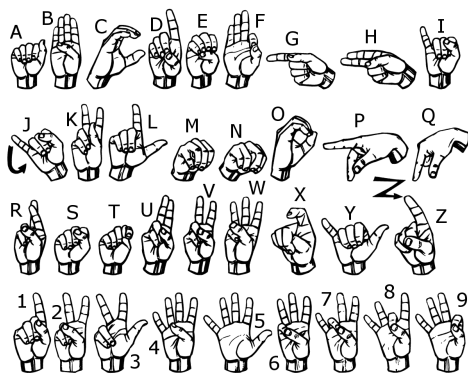
**Apsimokančios sistemos** (*angl. machine learning*) - tai mokslas apie tai, kaip kompiuterius užprogramuoti taip, jog jie patys darytų sprendimus be žmogaus įsikišimo neužprogramuojant kiekvienos galimos situacijos. Kitais žodžiais tariant leisti kompiuteriui pačiam nuspręsti kaip elgtis esant tam tikroms aplinkybėms. Dar kitaip tariant tai yra pradžia norint sukurti dirbtinį intelektą.

Savaime apsimokančių sistemų ir jų algoritmų sukūrimo dėka gatvėmis pradėjo važinėti pačios save vairuojančios (*angl. self-driving cars* arba dar kitaip vadinamos autonominės transporto priemonės, įvairūs paieškos varikliai tokie kaip „Google“ ar „Yahoo“ taikydami šiuos alogirtmus naudojami rodo kiekvienam asmeniškai sugeneruotą turinį. Taip pat reikėtų paminėti ir kalbos atpažinimo sistemas tokias kaip „Siri“ ar „Google Assistant“, kurios iš joms duotų komandų atlieka tam tikrus veiksmus.

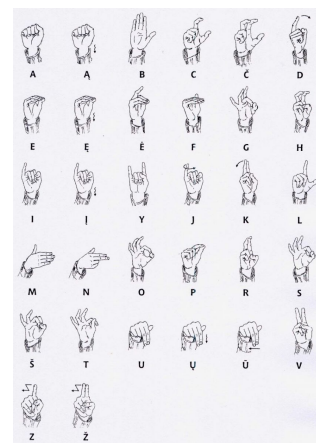
## 1.2. Gestų kalba

### 1.2.1. Gestų kalbos skirstymas

Gestų kalba susideda iš dviejų dalių - statinių ir dinaminių ženklų. Statiniais ženklais atvaizduojama didžioji abėcėlių raidžių dalis. O dinaminiais - žodžiai ir kai kurios gestų abėcėlių raidės. *Pavyzdžiui*, amerikiečių gestų kalbos abėcėlėje J ir Z raidės atvaizduojamos dinaminiais judesiais (žr. 1 pav.), o lietuvių - A, D, I, K ir kt. (žr. 2 pav.)



1 pav. Amerikiečių gestų kalbos abėcėlė



2 pav. Lietuvių gestų kalbos abėcėlė

## 2. Gestų kalbos atpažinimas

Gestų kalba ir jos atpažinimas naudojant kompiuterinę techniką yra viena iš daugelio apsimokančių sistemų sričių.

### 2.1. Problematika

Norint atpažinti gestus, paversti juos į žodžius ar sakinius susiduriama su problemomis, kurios susijusios tiek su statinių, tiek su dinaminių gestų atpažinimu.

Pagrindinės problemos išskylančios atpažįstant statinius gestų kalbos ženklus yra:

1. Kiekvienos kalbos abėcėlę sudaro skirtingas raidžių (statinių ženklų) skaičius. *Pavyzdžiui*, lietuvių kalbos abėcėlę sudaro 32 ženklai, o amerikietiška - 26;
2. Gestų panašumai. *Pavyzdžiui*, raidės A, E, N, S, T yra atvaizduojamos sugniaužtus kumštį, o net trijose iš jų (A, E ir S) skiriasi tik nykščio padėtis;
3. Kampas, kuris susidaro atpažįstant gestą. *Pavyzdžiui*, kai A raidė rodoma ne iš priekio, o iš šono;
4. Apšvietimas. *Pavyzdžiui*, gestų atpažinimas esant prieblandai ir dienos šviesai.

Pagrindinės problemos išskylančios atpažįstant dinامينius gestų kalbos ženklus yra:

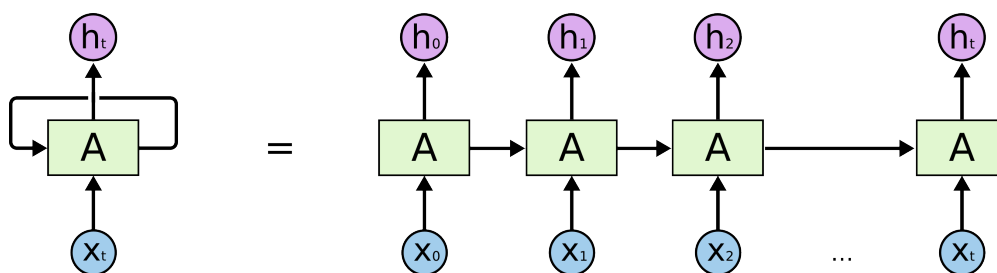
1. Nauji gestų kalbos žodžiai. *Pavyzdžiui*, kiekvienas uraganas turi savo pavadinimą, todėl tai gali reikšti naujo gesto atsiradimą;
2. Gesto kelios reikšmės. *Pavyzdžiui*, vienas gestas gali turėti kelias reikšmes, kaip kad lietuvių kalboje vienas žodis „kasa“ gali turėti net tris skirtingas reikšmes;
3. Kampas, kuris susidaro atpažįstant gestą. *Pavyzdžiui*, kai A raidė rodoma ne iš priekio, o iš šono;
4. Žodžių apjungimas į vieną sakinį. *Pavyzdžiui*, keli gestai einantys vienas po kito gali reikšti vieną žodį, tačiau tuo pačiu būti panašūs į vieną gestą, kuris jau reikš tik vieną žodį.

### 3. Rekurentiniai neuroniniai tinklai

**Dirbtinis neuroninis tinklas** (*angl. Artificial neural network*) - struktūra, skirta apdoroti dideliame kiekiui informacijos, sukurta remiantis žmogaus nervų sistemos veikimo principu. Šiam dideliame informacijos kiekiui apdoroti, yra komponentai, atliekantys tam tikrus skaičiavimus, dar vadinami neuronais. Vienas neuronas yra tarpusavyje sujungtas su dar daug kitų. Atėjus signalui, neuronai perduoda informaciją vienas kitam, kol galiausiai pasiekiamas tam tikras veiksmas, kurį reikia atlikti. Tam kad signalas keliautų teisingu keliu, reikia, kad jungtys tarp neuronų būtų užtektinai stiprios. Tam, dirbtiniuose neuronų tinkluose yra pasitelkiamos savaime apsimokančios sistemos.

**Rekurentiniai neuroniniai tinklai** (*angl. Recurrent neural network*) - tai dirbtinis neuroninis tinklas, kurie saugo informaciją apie praeituose žingsniuose (neuronuose) atliktus veiksmus ar skaičiavimus.

#### 3.1. Principas



3 pav. Rekurentinių neuroninių tinklų schema

- $x_t$  - įeiga momentu  $t$
- $g_t$  - išeiga momentu  $t$
- $A_t$  - būseną momentu  $t$

## 4. Medžiagos darbo tema dėstymo skyriai

Medžiagos darbo tema dėstymo skyriuose pateikiamos nagrinėjamos temos detalės: pradinė medžiaga, jos analizės ir apdorojimo metodai, sprendimų įgyvendinimas, gautų rezultatų apibendrinimas. Šios dalies turinys labai priklauso nuo darbo temos. Skyriai gali turėti poskyrius ir smulkesnes sudėtines dalis, kaip punktus ir papunkčius.

Medžiaga turi būti dėstoma aiškiai, pateikiant argumentus. Tekstas dėstomas trečiuoju asmeniu, t.y. rašoma ne „aš manau“, bet „autorius mano“, „autorius nuomone“. Reikėtų vengti informacijos nesuteikiančių frazių, pvz., „...kaip jau buvo minėta...“, „...kaip visiems žinoma...“ ir pan., vengti grožinės literatūros ar publicistinio stiliaus, gausių metaforų ar panašių meninės išraiškos priemonių.

### 4.1. Poskyris

Citavimo pavyzdžiai: cituojamas vienas šaltinis [**PvzStraipsnLt**]; cituojami keli šaltiniai [**PvzStraipsnEn**; **PvzKonfLt**; **PvzKonfEn**; **PvzKnygLt**; **PvzKnygEn**; **PvzElPubLt**; **PvzElPubEn**; **PvzMagistrLt**; **PvzPhdEn**].

#### 4.1.1. Skirsnis

##### 4.1.1.1. Straipsnis

#### 4.1.2. Skirsnis

## **5. Skyrius**

### **5.1. Poskyris**

### **5.2. Poskyris**



## **Rezultatai ir išvados**

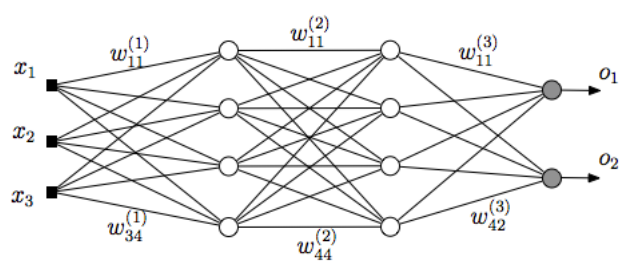
Rezultatų ir išvadų dalyje turi būti aiškiai išdėstomi pagrindiniai darbo rezultatai (kažkas išanalizuota, kažkas sukurta, kažkas įdiegta) ir pateikiamos išvados (daromi nagrinėtų problemų sprendimo metodų palyginimai, teikiamos rekomendacijos, akcentuojamos naujovės).

## **Santrumpos**

Sąvokų apibrėžimai ir santrumpų sąrašas sudaromas tada, kai darbo tekste vartojami specialūs paaiškinimo reikalaujantys terminai ir rečiau sutinkamos santrumpos.

## Priedas nr. 1

### Niauroninio tinklo struktūra



4 pav. Paveikslėlio pavyzdys

## Priedas nr. 2

### Eksperimentinio palyginimo rezultatai

1 lentelė. Lentelės pavyzdys

Algoritmas	$\bar{x}$	$\sigma^2$
Algoritmas A	1.6335	0.5584
Algoritmas B	1.7395	0.5647