

ΑΝΑΓΝΩΡΙΣΗ ΑΚΟΛΟΥΘΙΩΝ ΧΕΙΡΟΝΟΜΙΩΝ

Διπλωματική εργασία

Ιωαννίδης Χρήστος

Επιβλέποντες καθηγητές: Ποταμιάνος Γεράσιμος, Αργυρίου
Αντώνιος

7 Οκτωβρίου 2015

Πανεπιστήμιο Θεσσαλίας

Τμήμα Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών

ΠΕΡΙΕΧΟΜΕΝΑ

Εισαγωγή

Chalearn Gesture Challenge 2013

Εξαγωγή Χαρακτηριστικών

Μοντέλα Αναγνώρισης Πληροφορίας

Πειράματα και Αποτελέσματα

Isolated Testing

Embedded Testing

Επίλογος

ΕΙΣΑΓΩΓΗ

- Αλγόριθμοι και μέθοδοι που επιτρέπουν στους υπολογιστές να "μαθαίνουν" και να προβλέπουν φαινόμενα.
- Αυτοματοποιημένη, Δυναμική Ανάλυση Δεδομένων.
- Προγράμματα που παίρνουν "αποφάσεις"
- Τομείς ενδιαφέροντος :
 - Δυναμική Ανάλυση δεδομένων. (Χρηματηστήριο)
 - Μηχανές Αναζήτησης (Google, Amazon)
 - Όραση ΗΥ (Αναγνώριση Γραμματοσειράς, Χειρονομιών)

ΑΝΑΓΝΩΡΙΣΗ ΑΝΘΡΩΠΙΝΗΣ ΔΡΑΣΗΣ

- Αποτελεί μέρος των προβλημάτων Όρασης Υπολογιστών.
- Στοχεύει στην καταγραφή και εκμάθηση του τρόπου με τον οποίο ο άνθρωπος αντιλαμβάνεται το περιβάλλον μέσω της όρασης (Εικόνες, Πρόσωπα, Συναισθήματα, Καταστάσεις)
- Ανθρώπινες Χειρονομίες μπορούν να δηλώνουν προσταγή, εντολή, συναίσθημα, παράκληση κ.α.

ΑΝΑΓΝΩΡΙΣΗ ΑΝΘΡΩΠΙΝΗΣ ΔΡΑΣΗΣ

- Οφέλη για μηχανική εκμάθηση Ανθρώπινης Δράσης:
 - Αυτοματοποίηση Παραγωγής.
 - Βοηθητικές Υπηρεσίες (Εφαρμογές για άτομα με Ειδικές ανάγκες, Ηλικιωμένους)
 - Έξυπνα Σπίτια (Εκτέλεση εντολών μέσω χειρονομιών)
 - Εξυπνες Υποδομές Πόλεων. (Αναγνώριση συνθηκών επικυνδυνότητας, ληστείες, πυρκαγιές κ.α.)

CHALEARN GESTURE CHALLENGE 2013

CHALEARN GESTURE CHALLENGE 2013

- Chalearn: Οργάνωση με στόχο την προώθηση της έρευνας στον τομέα της Μηχανικής Εκμάθησης.
- Συνεργασία με δεκάδες Πανεπιστήμια και Οργανισμούς ανα τον κόσμο
- Διοργάνωση Διαγωνισμών (Challenges) γύρω από σχετικά προβλήματα.
- Chalearn Challenge 2013:
 - Αναγνώριση Ακολουθιών Χειρονομιών

CHALEARN GESTURE CHALLENGE 2013

- Τα δεδομένα:
 - Βάση Δεδομένων με 20 διακριτές χειρονομίες
 - Καθημερινές χειρονομίες που δείχνουν προσταγή ή συναίσθημα.
 - Χρήση του σένσορα Kinect για καταγραφή ήχου, εικόνας και "βάθους" για κάθε βίντεο
- Το πρόβλημα:
 - Κατασκευή ταξινομητή για την αναγνώρισή τους.

Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ



(a) Basta!



(b) Buonisimo



(c) Che due palle



(d) Che vuoi



(e) Cos'hai combinato?



(f) Cosa ti farei!



(g) Vanno d'accordo



(h) Ho fame



(i) Non me ne niente



(j) E' un furbo!



(k) Si sono messi d'accordo!

Η ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ



(l) Perfetto!



(m) Le vuoi
prendere?



(n) Sei
Pazzo?



(o) Tanto
tempo fa



(p) Sono
stufo



(q) Vattene



(r) Vieni qui



(s) Non ce
n'è più

Figure 1: Vocabulary

Ο ΣΕΝΣΟΡΑΣ KINECT

- Καλύτερη Αντίληψη του 3-D χώρου.
- Αισθητήρας Κίνησης για την ανίχνευση κινούμενων αντικειμένων στο χώρο
- Αισθητήρας Βάθους για ανίχνευση της απόστασης ενός αντικειμένου από την κάμερα
- Κυκλοφόρησε από την Microsoft το 2010



ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ II - ΤΥΠΟΙ ΠΛΗΡΟΦΟΡΙΑΣ

- RGB Video καταγεγραμένο στα 20FPS.
- Depth Mask and User body tracking.
- Εύρεση 20 σκελετικών αρθρώσεων.



Figure 2: RGB, Depth, User mask and Skeletal Information.

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ I: ΥΠΟΒΑΘΡΟ

- Κατασκευή ενός συνόλου από χρήσιμες τιμές που "χαρακτηρίζουν" την πληροφορία.
- Χειρονομία : Ένα σύνολο από διαδοχικές πόζες (frames).
- Εξαγωγή χαρακτηριστικών για κάθε frame.
- Κατασκευή συνόλων πληροφορίας για κατηγοριοποίηση χειρονομιών.
- Στόχοι:
 - Σωστή επιλογή χαρακτηριστικών.
 - "Καλή" διάσταση εξαγώμενων χαρακτηριστικών.

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ II: ΑΛΓΟΡΙΜΟΣ

- Χρήση μόνο των σημείων Σκελετικών Αρθρώσεων.
- Εξαιρετικά Χρήσιμη στην εξαγωγή στάσεων του σώματος.
- Εξαγωγή των **Pose Descriptors**.
- Ακολουθούμε τα βήματα που περιγράφονται στην εργασία των Neverova, Wolf et al. :"A multi-scale approach to gesture detection and recognition"

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ II: ΑΛΓΟΡΙΜΟΣ

Pose descriptor:

- Υπολογισμός γωνιών μεταξύ των "εικονικών" οστών που σχηματίζονται με τη χρήση των σημείων σκελετικών αρθρώσεων.
- Υπολογισμός αποστάσεων των αρθρώσεων των χεριών μεταξύ αυτών και από το κέντρο βάρους του ανθρώπινου σώματος.
- Υπολογισμός 3 σετ των 9 γωνιών (για κάθε διάσταση στον 3-D χώρο)
- 61 αποστάσεις μεταξύ αρθρώσεων.
- Σύνολο 88 χαρακτηριστικά/ frame.

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ II: ΑΛΓΟΡΙΜΟΣ

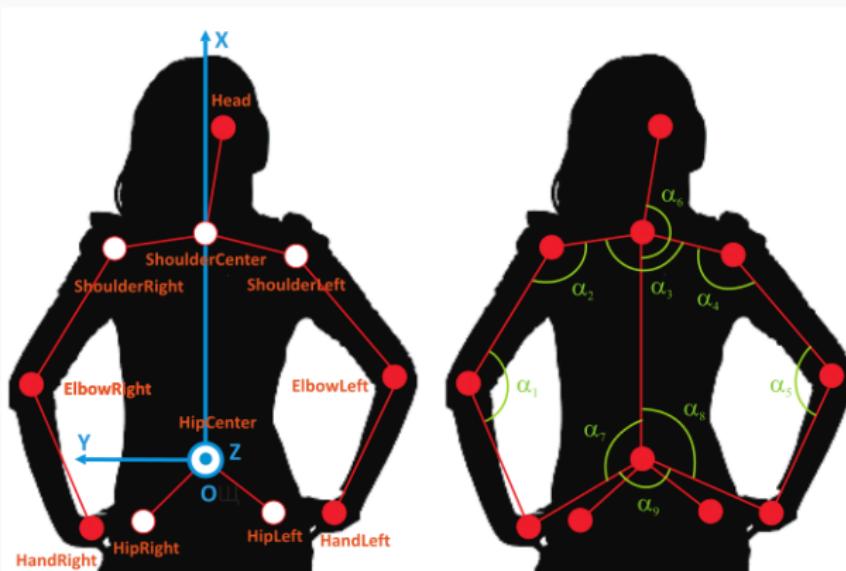


Figure 3: Angle Based Pose Descriptor.

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ III: ΒΗΜΑΤΑ

Βήμα 1^ο : Μετατροπή συστήματος αξόνων.

- Από το σύστημα αξόνων με κέντρο την κάμερα σε σύστημα αξόνων με κέντρο το κέντρο σώματος.
- Κανονικοποίηση των σημείων αρθρώσεων με βάση την απόσταση του κέντρου σώματος από τους ώμους.
- Εξάλειψη διαφορών ύψους, θέσης ατόμων στον πραγματικό χώρο.

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ III: ΒΗΜΑΤΑ

Βήμα 2^ο : Εφαρμογή Vector Median Filtering.

- "Παράθυρο" που ομαλοποιεί τιμές μεταξύ γειτωνικών frame παρατηρήσεων.
- Εξάλειψη ακραίων τιμών από πιθανό "θόρυβο" στα δεδομένα. (Αρθρώσεις που καταγράφηκαν εσφαλμένα)

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ III: ΒΗΜΑΤΑ

Βήμα 3^ο : PCA στις αρθρώσεις που απαρτίζουν το κορμό του σώματος.

- Μετατροπή του συστήματος αξόνων ξανά.
- Όλα τα δεδομένα προβάλλονται στον τελικό αξονα που χαρακτηρίζει την θέση και κλίση του κορμού του σώματος (rotation invariance).

Βήμα 4^ο : Εξαγωγή των Pose Descriptors με χρήση των σκελετικών σημείων στον νέο χώρο.

ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ IV: ΑΛΓΟΡΙΘΜΟΣ

Algorithm 1 Extract Pose descriptor

```
1: procedure SKELETON READER
2: 1. Extract Skeletal info:
3:   for  $i = startges$  to  $endges$  do
4:      $S \leftarrow$  load videoFrame-i.mat
5:     if  $S.Skeleton$  is empty then continue;
6:      $eucldist \leftarrow norm( S.Skeleton.ShoulderCenter(1,i) - S.Skeleton.HipCenter(1,i) )$ 
7:      $sjoint \leftarrow zeros(20,3)$ 
8:     for  $j = 1$  to  $20$  do
9:        $sjoint(j,:) \leftarrow (S.Skeleton.joint(j,:)/eucldist) - (S.Skeleton.joint(1,:)/eucldist)$ 
10:      for  $z = 1$  to  $20$  do
11:         $P(z).Timemtx(framenum) \leftarrow sjoint(:,z).$ 
12: 2. VMF:
13:   for  $z = 1$  to  $20$  do
14:      $P(z).Timemtx \leftarrow medfilt2(sjoints, [1, 3]).$ 
15: 3. Angle Based Pose Descriptor:
16:   for  $j = 1$  to  $20$  do
17:      $sjoint(j,:) \leftarrow P(j).Timemtx(:,j)^T$ .
18:    $coeffs = princomp(sjoints.torso).$ 
19:    $projJnt = coeffs'^*sjoint^T.$ 
20:    $jointSets = [\dots].$ 
21:
22:   # Angles a-c
23:   for  $i = 1$  to  $len(jointSet)$  do
24:      $angle-a(i) \leftarrow angleCalc( projJnt(jointSets(i,1), 1:2),$ 
 $projJnt(jointSets(i,2), 1:2), projJnt(jointSets(iter,3), 1:2)).$ 
25:      $angle-b(i) \leftarrow angleCalc( projJnt(jointSets(i,1), 2:3),$ 
 $projJnt(jointSets(i,2), 2:3), projJnt(jointSets(iter,3), 2:3)).$ 
26:      $angle-c(i) \leftarrow angleCalc( projJnt(jointSets(i,1), 1:2:3),$ 
 $projJnt(jointSets(i,2), 1:2:3), projJnt(jointSets(iter,3), 1:2:3)).$ 
27:      $mdistance - d \leftarrow mahalanobisdist(handSets, projJnt).$ 
```

ΜΟΝΤΕΛΑ ΑΝΑΓΝΩΡΙΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΤΕΧΝΙΚΕΣ ΜΑΘΗΣΗΣ

- Εποπτευόμενη Μάθηση (Supervised Learning)
- Εκάθηση χωρίς επίβλεψη (Unsupervised Learning)
- Ενισχυτική Μάθηση (Reinforcement Learning)

ΤΕΧΝΙΚΕΣ ΜΑΘΗΣΗΣ

- Εποπτευόμενη Μάθηση (Supervised Learning)
 - Παροχή των χαρακτηριστικών από τον χρήστη
 - Μηχανή προσαρμόζει τα μοντέλα της ώστε να χαρακτηρίζει τα δοσμένα δεδομένα.
 - Αναγνώριση “παρόμοιων Χαρακτηριστικών”
 - Επιλογή χαρακτηριστικών ;
 - Overtraining ;

HIDDEN MARKOV MODELS

- Στοχαστικό, πιθανοτικό μοντέλο
- Κάθε κόμβος ενός HMM σηματοδοτεί μια κατάσταση στο σύστημα.
- Ακμές στο γράφημα εκφράζουν τις εξαρτήσεις μεταξύ των καταστάσεων.

HIDDEN MARKOV MODELS

- Μοντελοποίηση δεδομένων με χρονική υπόσταση.
- “Κρυφές” καταστάσεις.
- Ορατή μόνο η συμπεριφορά του συστήματος, δηλαδή η έξοδος (emission probabilities) κάθε κατάστασης
- Μία παρατήρηση τη χρονική στιγμή t παρήχθη στην κατάσταση S_t που είναι κρυφή από τον παρατηρητή.
- Η έξοδος μιας HMM κατάστασης στον χρόνο t (S_t) εξαρτάται άμεσα από την αμέσως προηγούμενη κατάσταση ($S_t \in S_{t-1}$)

HIDDEN MARKOV MODELS

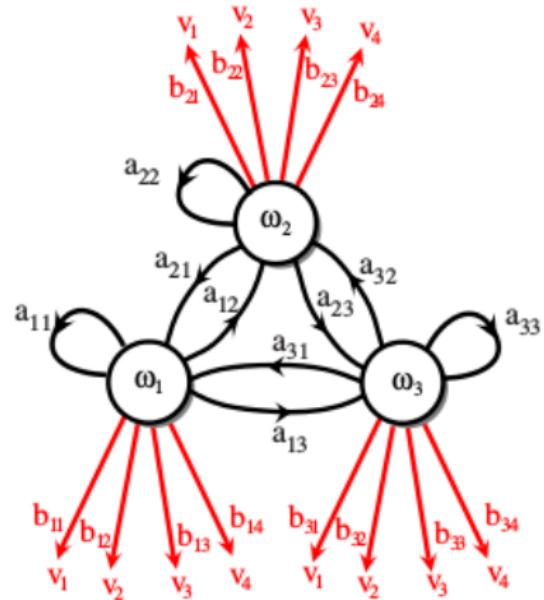


Figure 4: A first order HMM

ΓΚΑΟΥΣΙΑΝΑ ΜΟΝΤΕΛΑ ΜΙΞΗΣ

- Μία μίξη Γκαουσιανών ορίζεται ως ένας συνδυασμός κατανομών πυκνότητας πιθανότητας.
- Η κατανομή του d -διάστατου χώρου με μέση τιμή m και πίνακα συναδιασποράς C διάστασης $d \times d$, ορίζεται ως:

$$\varphi(x; \theta) = (2\pi)^{-\frac{d}{2}} \det(C)^{-\frac{1}{2}} \exp\left(\frac{-(x - m)^T C^{-1} (x - m)}{2}\right) \quad (1)$$

όπου το θ υποδηλώνει τις παραμέτρους m και C .

- Η εκπαίδευση τέτοιων μοντέλων επιτυγχάνεται μέσω του αλγορίθμου Αναμενόμενης Τιμής-Μεγιστοποίησης.

ΒΑΘΙΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

- Τεχνητά Νευρωνικά Δίκτυα:
 - Στατιστικές μέθοδοι μηχανικής μάθησης που προσπαθούν να προσομοιώσουν τη λειτουργικότητα των βιολογικών νευρωνικών δικτύων
 - Νευρώνες, κρυφά επίπεδα
 - πχ. perceptron μια μορφή νευρωνικού δικτύου
- Βαθιά Νευρωνικά Δίκτυα :
 - ANNs με πολλά κρυφά επίπεδα
 - Μοντελοποίηση μη γραμμικών σχεσιακών μοντέλων
 - Χρήση σε προβλήματα Αναγνώρισης ήχου, Όρασης HY

ΒΑΘΙΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

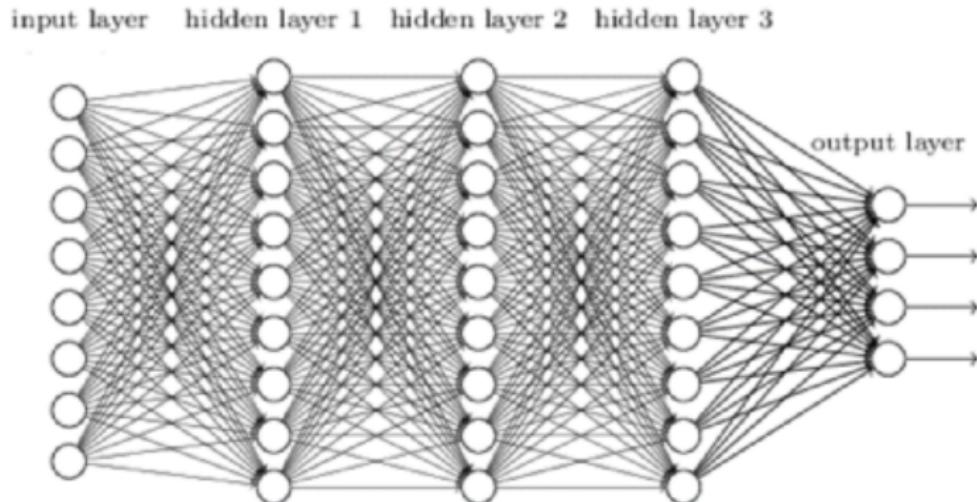


Figure 5: An Artificial Neural Network.

ΒΑΘΙΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ : ΕΚΠΑΙΔΕΥΣΗ

- Χρήση του αλγορίθμου οπισθοδιάδοσης (backpropagation)
- Ανανέωση των βαρών με χρήση του τύπου:

$$w(t+1) = w(t) + \eta \Delta w \quad (2)$$

- Χρήση συνάρτησης softmax σαν συνάρτηση κόστους:

$$\hat{y}_k = \frac{\exp(v_k^L)}{\sum_k \exp(v_{k'}^L)} \quad (3)$$

όπου y δηλώνει την εξόδο του νευρωνικού δικτύου, L είναι το πλήθος των κρυφών επιπέδων και v η συνάρτηση ενεργοποίησης.

ΠΕΙΡΑΜΑΤΙΚΑ ΜΟΝΤΕΛΑ

- GMM-HMM
 - Εκπαίδευση χαρακτηριστικών πάνω από HMM - μοντελοποίηση χρονικότητας γεγονότων.
 - GMM μοντελοποιούν τις emission πιθανότητες (έξοδοι των HMM καταστάσων)
- DNN-HMM
 - Χρήση Βαθιών νευρωνικών δικτύων στην μοντελοποίηση των εξόδων των HMM καταστάσεων.

ΠΕΙΡΑΜΑΤΑ ΚΑΙ ΑΠΟΤΕΛΕΣΜΑΤΑ

ΠΕΙΡΑΜΑΤΙΚΑ ΕΡΓΑΛΕΙΑ

- HTK: Hidden Markov Model Toolkit
 - Αναπτύχθηκε στο Machine Intelligence Laboratory, University of Cambridge
 - Εργαλεία για εξαγωγή, επεξεργασία ακουστικών αρχείων, HMMs
- Kaldi Speech Recognition Toolkit
 - Ξεκίνησε στο πλαίσιο workshop project για Subspace Gaussian Mixture Models, John Hopkins University 2009
 - Αναπτύθηκε “ανοιχτά” από την κοινότητα.
 - Ανάπτυξη σε C++
 - Πλήθος παρεχόμενων εργαλείων είτε με χρήση C++ συναρτήσεων είτε με αυτοματοποιημένα εργαλεία (shell scripting).
 - Υλοποίηση HMM & DNN μοντέλων

- Isolated Training, Testing
 - Σπάμε τις ακολουθίες του training και test σε διακριτές χειρονομίες.
 - Εκπαίδευση και Αναγνώριση μεμονωμένων χειρονομιών.
- Isolated Training, Embedded Testing
 - Εκπαίδευση με βάση μεμονωμένες χειρονομίες.
 - Εκπαίδευση αναγνώρισης παύσεων μεταξύ χειρονομιών.
 - Αναγνώριση Ακολουθιών Χειρονομιών.

- Πληροφορία, Μέγεθος Χαρακτηριστηκών
 - 88 χαρακτηριστικά / frame
 - Μεγάλο πλήθος χαρακτηριστικων.
 - Μπορούμε να πετύχουμε παρόμοια/ καλύτερα αποτελέσματα με λιγότερα features; (πχ. LDA)
- Αριθμός Γκαυσιανών, Νευρώνων, κρυφών επιπέδων ;
- Overtraining ;

- Isolated vs Embedded Testing
 - Αναγνώριση μεμονωμένων χειρονομιών:
 - Απλή Γραμματική, ζητάμε την παραγωγή 1 αποτελέσματος (χειρονομίας) για κάθε video sample.
 - Κρίσιμοι τομείς: Καλά χαρακτηριστικά, ορισμένες κλάσεις
 - Αναγνώριση ακολουθιών χειρονομιών:
 - Δεν γνωρίζουμε τον αριθμό των χειρονομιών στην ακολουθία.
 - Επιτυχής αναγνώριση παύσεων μεταξύ χειρονομιών.
 - Κρίσιμοι τομείς: “Φύση” εξαγόμενων χαρακτηριστικών, μοντέλα classification, αναγνώριση παύσεων.

1. 1. GMM-HMM

- 5 state HMM, 4 - 32 Gaussian mixtures in each state

2. 2. DNN-HMM

- 5 state HMM, 4 - 32 Gaussian mixtures in each state
- 5-20 training iterations (epochs)
- 1-3 hidden layers

WER: Word Error Rate

- Το WER ορίζεται σαν το άθροισμα των Insertion, Substitution και Deletion Errors διαιρεμένο με το πλήθος των χειρονομιών.

$$WER = \frac{S + I + D}{N} \cdot 100\% \quad (4)$$

ISOLATED TESTING

ΑΠΟΤΕΛΕΣΜΑΤΑ KALDI I: ORIGINAL FEATURE SET VS LDA

Data	WER (%)	Substitution Errors	Total not found
Original Feature set 88feat./frame	31.38	166	166/529
LDA→ 19feat./frame	43.10	228	228/529
LDA→ 10feat./frame	33.46	177	177/529

Figure 6: Original Feature Set vs LDA projections

ΑΠΟΤΕΛΕΣΜΑΤΑ KALDI III: ΣΥΜΠΕΡΑΣΜΑΤΑ Ι

Original Feature Set vs LDA:

- Τα αρχικά χαρακτηριστικά καλύτερα από τις προβολές αυτών με χρήση LDA.
- LDA “χάνει” πληροφορία, πιθανότητα στην απεικόνιση της χρονικότητας των χαρακτηριστικών.

Isolated testing:

- Αναγνώριση περίπου των 2/3 των χειρονομιών.

ΑΠΟΤΕΛΕΣΜΑΤΑ KALDI II: DNN-HMM CLASSIFICATION SCORES

Epochs	Additional Epochs	Hidden Layers	WER(%)
15	5	2	19.66
15	5	3	21.74
7	3	2	24.39
3	2	2	25.90
3	2	1	25.33
HMM			31.38

Figure 7: DNN-HMM scoring results

Νευρωνικά Δίκτυα, DNN-HMM model

- Εμφανής βελτίωση σε σχέση με το μοντέλο GMM-HMM.
- 20 epochs, 2 hidden layers

ΑΠΟΤΕΛΕΣΜΑΤΑ HTK I: GMM-HMM CLASSIFICATION SCORES

Model	Correct (%)	WER	Substitution Errors	Total found
GMM1/HMM9	48.02	51.98	275	254/529
GMM2/HMM18	57.28	42.72	226	303/529
GMM4/HMM18	62.57	37.43	198	331/529
GMM4/HMM27	56.52	43.48	230	299/529
GMM8/HMM27	60.49	39.31	209	320/529
GMM8/HMM36	56.71	43.29	229	300/529
GMM16/HMM36	59.36	40.64	215	314/529

Figure 8: GMM-HMM scoring results in HTK

EMBEDDED TESTING

ΑΠΟΤΕΛΕΣΜΑΤΑ KALDI : GMM-HMM CLASSIFICATION SCORES

Model	WER (%)	Insertion Errors	Deletion Errors	Substitution Errors	Total not found
Original Feature set (88feat./frame)	93.77	1787	110	979	2876 / 3067

Figure 9: GMM-HMM scoring results in KALDI

ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

- Κακή απόδοση υπάρχοντος μοντέλου
- Ποσοστό λάθους στο 93.77%
- Μεγάλος αριθμός insertion errors. Αναγνώριση λιγότερων ακολουθιών από τις πραγματικές

ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Τί πήγε στραβά:

1. Υπάρχουσα Γραμματική και λεξιλόγιο

- 1 gesture = 1 word
- 1 gesture = x poses
- Το υπάρχον μοντέλο δεν μας καλύπτει επαρκώς.

2. HMM μοντέλο

- Η αύξηση των HMM states επιβεβαιώνει το πρόβλημα σε επίπεδο γραμματικής και λεξιλογίου.
- Ελάχιστη μείωση του λάθος με την αύξηση των HMM καταστάσεων.

3. Παύσεις μεταξύ χειρονομιών. (pauses, short pauses)

ΑΠΟΤΕΛΕΣΜΑΤΑ KALDI : DNN-HMM CLASSIFICATION SCORES

Epochs	Additional Epochs	Hidden Layers	WER(%)	IE	DE	SE	Total Not Found
15	5	2	57.68	698	336	735	1769 / 3097
15	5	3	59.24	784	285	748	1817 / 3067
7	3	2	59.90	719	341	771	1837 / 3067
3	2	2	60.81	655	364	846	1865 / 3067
3	2	1	60.25	656	375	817	1848 / 3097
GMM-HMM		93.77					

Figure 10: GMM-HMM scoring results in KALDI

ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

- Αισθητά καλύτερη απόδοση συστήματος με χρήση DNNs.
- Περίπου 1/2 των ακολουθιών αναγνωρίζεται.
- DNNs vs GMMs

ΕΠΙΛΟΓΟΣ

- Μελέτη της επιστήμης πίσω από την αναγνώιση χειρονομιών
- Χρήση πληροφορίας από depth sensors για εξαγωγή χαρακτηριστικών.
- Κατασκευή δυο μοντέλων ταξινομητών βασισμένων σε HMMs.
- Εξοικίωση και χρήση των εργαλείων HTK και KALDI

ΜΕΛΛΟΝΤΙΚΕΣ ΚΑΤΕΥΘΥΝΣΕΙΣ

- Χρήση πολλαπλών μοντέλων στην ταξινόμηση (RGB, depth, audio) και σύνθεση αυτών.
- Κατασκευή νέου λεξιλογίου και γραμματικής, βασισμένο σε poses.
- Νέες τεχνικές μάθησης.

Ευχαριστώ πολύ για τον χρόνο
σας

Ερωτήσεις; Απορίες;