

# Improvements in Sub-Character HMM Model Based Arabic Text Recognition

Irfan Ahmad, Gernot A. Fink, and  
Sabri A. Mahmoud

## Outline

- Sub-character modeling
- Issues with space modeling
- Special ‘Space’ and ‘Connector’ models
- Contextual sub-character HMMs
- Multi-stream contextual sub-character HMMs
- Experiments and results
- Conclusions

## Arabic Sub-Characters

- Capture similar patterns
- Reconstruct the characters using those patterns
  - Use of 'connector' and 'space' models
- Leads to significant reduction of HMMs
- *No explicit segmentation needed*



## Sub-Character Example

شرفش الشرف سيسب الرياض الفايض  
الرصاص بورويس النصر الناضور راس شانشو

can be reconstructed ; using

ش ر ف ش ا ل ش ف س ي س ب ي ا ض  
ف ض ص ص ب و ر س ص ن ض س ن

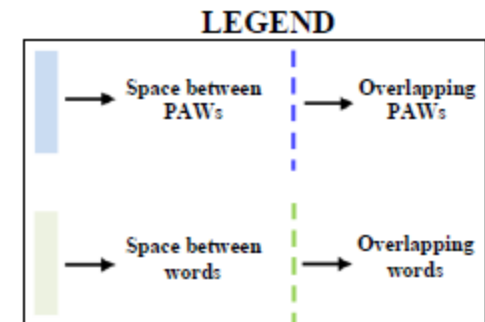
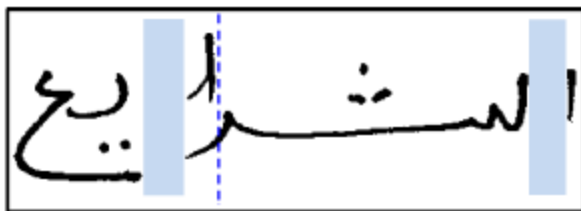
28 different character shapes models; OR

ش ر ف ن ا ل س ي ب ض ص و ن -

14 different sub-character models  
(in addition to space and connector)

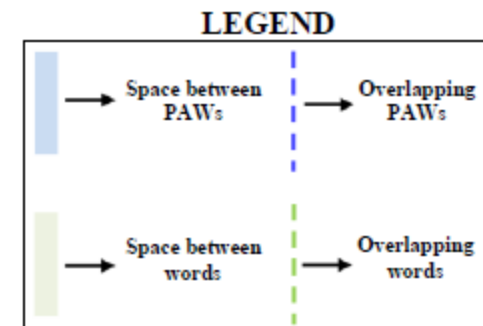
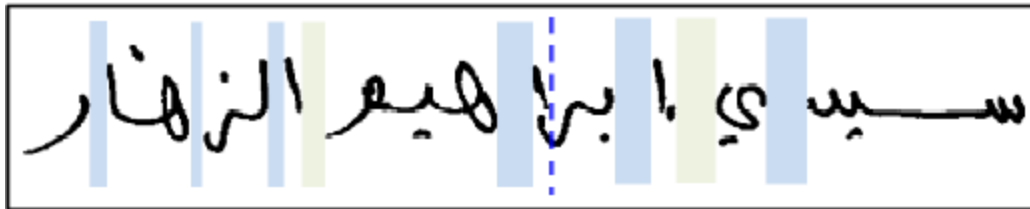
## Space Modeling Issues

- Space modeling is challenging
  - Possibility of overlapping PAWs
- Dreuw et al. presented an approach to space modeling for Arabic [Dreuw08]
- Presence of space between some PAWs but absent between other PAWs



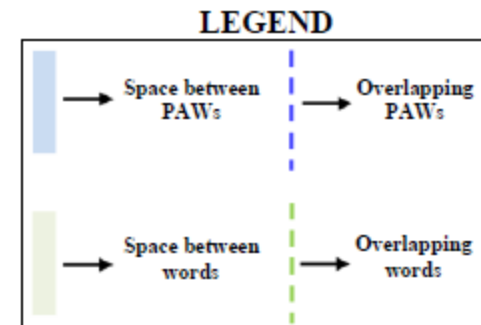
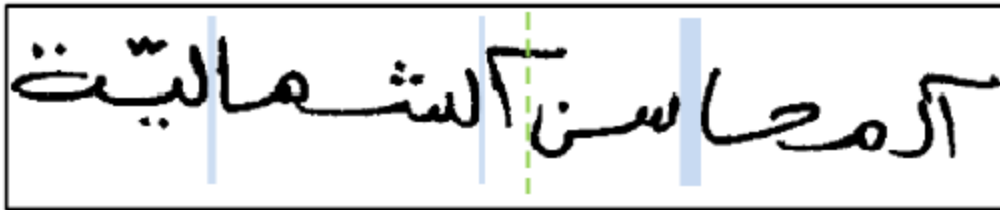
## Space Modeling Issues...

- Width of space between PAWs may be larger than the width of space between words



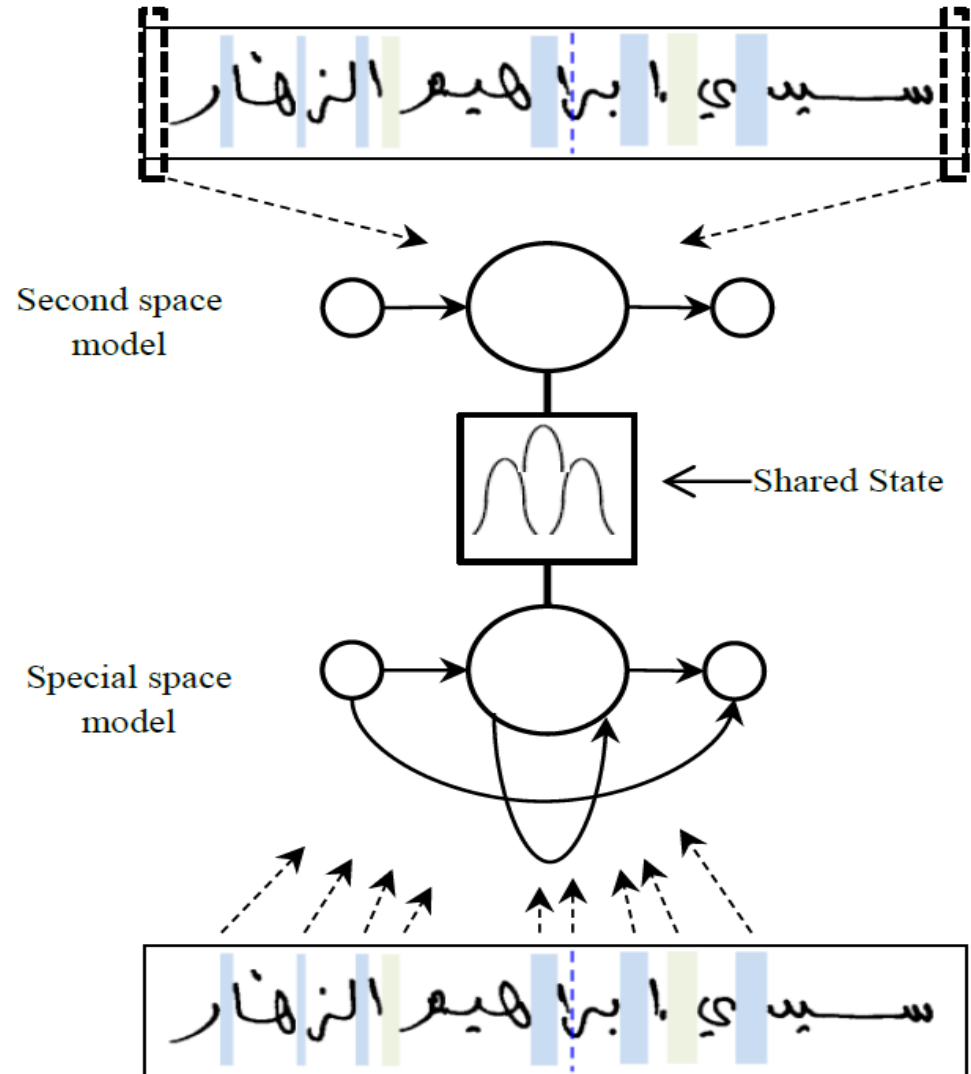
## Space Modeling Issues...

- Sometimes difficult to find space between words even though there might exist space between PAWs



# Special 'Space' and 'Connector' Models

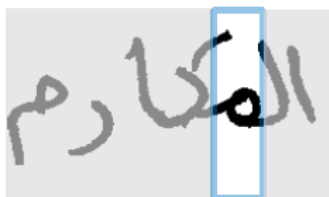
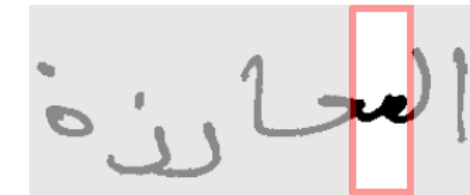
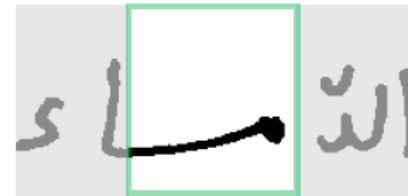
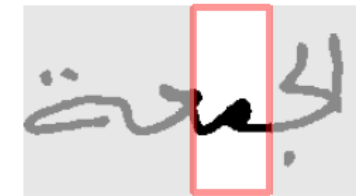
- Single state models
- Possibility of skipping the space and connector model
  - Transitions from the entry state to the exit state





## Contextual Sub-Character HMMs

- Important to model the contextual forms in handwritten Arabic text recognition



Examples show character unit (Meem) in different contexts

\* Image Source from IFN/ENIT

# Contextual Sub-Character HMMs...

## Illustrations

الدخانية

نكته

سنان

الشقراني

نكريف

المرناقي

الدخانية

تونكت

المكناشي

Examples show character unit (Noon) in different contexts

## Contextual Sub-Character HMMs...

### Approach

- Data driven approach to contextual sub-character HMMs
- Tying at different levels (data insufficiency problem)
  - Transition probabilities of all the contextual forms of a sub-character are tied
  - State tying between different contextual forms of the sub-character
    - ✓ Preserving the state sequence
- The distance  $d(x,y)$  between two states,  $x$  and  $y$  [Young02]:

$$d(x, y) = -\frac{1}{M} \sum_{m=1}^M \log [b_y(\mu_{xm})] + \log [b_x(\mu_{ym})]$$

## Multi Stream HMMs With Stream Splitting

$$b_y(o) = \prod_{s=1}^S \left[ \sum_{m=1}^M c_{ym} N(o | \mu_{ym}, \Sigma_{ym}) \right]^{\gamma_s}$$

$\gamma_s$  is the stream weight for stream 's'

- Features split into two streams ( $S = 2$ )
  - Features that were calculated from the image form one stream, and
  - Their derivative features form the second stream

## Experiments and Results

- IFN/ENIT database was used for evaluation
- Continuous HMM system using the HTK tools
- Nine statistical features + Nine derivative features  
(feature vector : 18)
- No preprocessing done on images

## Experiments and Results...

### Baseline System: Character-Shape HMMs

- 178 unique character-shapes in the IFN/ENIT dataset
  - 157 models in our system (replacing infrequent ones)
- Same number of states per character
- Initial flat start, followed by
- Alignment based training
  - Viterbi initialization for individual models
  - Baum-Welch retraining
- Viterbi decoding for word hypothesis

## Experiments and Results...

### Sub-Character System

- 97 sub-character HMMs
- Same number of states per sub-character HMM
  - Except the 'space' and the 'connector' models
- Same training procedure as the baseline system
- System extended to contextual sub-character HMMs
- Further extended to multi-stream contextual sub-character HMMs

# Experiments and Results...

## Comparison on Some Modeling Statistics

(From IFN/ENIT Database Sets a+b+c)

	<i>Baseline system</i>	<i>Sub-character system</i>
No. of HMMs	157	97
Average no. of samples per model	531	1734
Median no. of samples per model	186	492
No. of models having less than 100 samples	23 out of 157	2 out of 97
No. of models having less than 200 samples	80 out of 157	9 out of 97



# Experiments and Results...

## Evaluation Results - *Word Recognition Rate (WRR)*

<i>The Recognition System</i>	<i>Train-Test Configuration</i>			
	<i>(Statistical Significance)</i>			
	<i>abc-d</i> <i>(±0.38)</i>	<i>abcd-e</i> <i>(±0.56)</i>	<i>abcde-f</i> <i>(±0.50)</i>	<i>abcde-s</i> <i>(±1.56)</i>
Character-shape HMM system ( <i>Baseline</i> )	95.38	90.48	89.40	80.69
Sub-character HMM system	95.90	91.55	89.74	82.14
Contextual sub-character HMM system	96.67	92.91	91.57	84.49
Multi-stream contextual sub-character HMM system	97.22	93.52	92.15	85.12

# Experiments and Results...

## Comparison With Other State-of-the-Art Systems

<i>Systems</i>	<i>Train–Test Configuration</i>			
	<i>abc–d</i>	<i>abcd–e</i>	<i>abcde–f</i>	<i>abcde–s</i>
UPV-PRHLT [Margner10]	95.20	<b>93.90</b>	92.20	84.62
RWTH-OCR [Margner11, Dreuw12]	96.53	92.74	92.20	84.55
Azeem and Ahmed [Azeem13]	<b>97.70</b>	93.44	<b>93.10</b>	<b>84.80</b>
Su et al. [Su13]	96.81	93.55	-	-
<b>Present Work</b>	<b>97.22</b>	<b>93.52</b>	<b>92.15</b>	<b>85.12</b>
<b>Latest Results (unpublished)</b>				
<b>Model length adaptation + Stream weights + ...</b>	<b>97.30</b>	<b>94.76</b>	<b>93.07</b>	<b>86.33</b>

## Conclusions

- Sub-character HMMs for Arabic text recognition
  - Allow sharing of common patterns
  - Results in compact, efficient, and robust recognizer with reduced model set
- Special '*space*' and '*connector*' models
- Contextual sub-character HMMs
- Multi-stream contextual sub-character HMMs
- State-of-the-art results on the IFN/ENIT database
  - Results on set-s outperformed previously reported results