



## Summer School in Asian Language Processing

1<sup>st</sup> June – 12<sup>th</sup> August, 2006

**Course:** Script Processing  
**Course Instructor(s):** Dr. Syed Afaq Hussain (21-25), Aamir Wali (1-20, 26-30)  
**Teaching Assistant(s):** Atif Gulzar  
**Text Book(s):** Writing Systems, A Linguistic Approach by Henry Rogers (WS)  
**Grades:** Quizzes: 20%, Labs 30%, Midterm Exam 25%, Final Exam 25%  
**Online Reading Material(s):**

1. <http://developer.apple.com/FONTS/TTRefMan/index.html> (A1)
2. <http://www.microsoft.com/typography/SpecificationsOverview.mspx> (M1)
3. <http://www.microsoft.com/typography/DevArticles.mspx> (M2)
4. <http://www.microsoft.com/typography/TrueTypeHintingIntro.mspx> (M3)
5. <http://www.unicode.org/standard/WhatIsUnicode.html> (U1)
6. <http://www.unicode.org/charts/> (U2)
7. <http://www.unicode.org/reports/tr9/> (U3)
8. <http://www.unicode.org/reports/tr10/> (U4)
9. <http://www.unicode.org/reports/tr14/> (U5)

**Reference Book(s):**

1. Practical Computer Vision using C, J. Parker., John Wiley & Sons
2. Digital Image Processing by R.C. Gonzales , John Wiley
3. Digital Image Processing Using MATLAB, R.C. Gonzales , John Willey

**Grades:** Quizzes 20%, Labs and Assignments 30%, Mid-term Exam 25%, End-term Exam 25%

Schedule of Lectures		Reading	Schedule of Lab Sessions
1	Introduction /Localization		
2	Scripts/Writing Systems /Orthography	WS1,14	
3	Ancient Scripts: Deciphering and Processing, Modern Writing Systems: Types	WS2,6,11	
4	Output Methods: Bitmap& Vector Fonts, Type of Vector fonts: TTF, PCF, PCX, Type1	A1, M1	
5	True Type essentials: Contour, Glyph, Em square etc	A1, M1	
6	True Type Rasterizer	A1, M1	
7	Lab Session		Lab 1: True Type Fonts
8	Chinese Writing System and its Processing	WS3	
9	Encoding: Introduction / Different encoding schemes, Script, Codepages etc	U2	
10	Unicode	U1	
11	Brahimi Derived Scripts: Devanagari/Kannada etc	WS11	
12	Special Properties of Brahimi Derived Scripts	WS11	
13	Semitic Scripts	WS7	
14	Intelligent Fonts: TrueType Limitations, General overview	M2	

15	Midterm Exam		Midterm on Script Processing
16	OTF Architecture	M2	
17	Design of text layout engine in MS: Uniscribe&OTLS	M2	
18	Discussion of Mark Changizi and Shinsuke Shimojo. 2005. "Character Complexity and Redundancy in Writing Systems over Human History." <i>Proceedings of the Royal Society B</i> .		
19	Lab Session		Lab2: From Paper to TrueType Font
20	Writing OpenType Rules	M2	Lab3: From TrueType to OpenType
21	Introduction to OCR, Image Processing & Computer Vision		Intro to Matlab Image Processing using Matlab
22	Pre-processing (Noise Removal, Histogram manipulation, Binarization, etc)		Lab4: Preprocessing using Matlab
23	Document Layout Processing		Lab5: Segmentation, Skeletonization,
24	Skeletonization, Segmentation, Feature extraction		Lab6: Feature Extraction
25	Classifier Design, Statistical, Structural & Neural Approaches		
26	Advance Topics in Unicode (1): Line Breaking	U5	
27	Advance Topics in Unicode (2): Bi-directional Algorithm	U3	
28	Advance Topics in Unicode (3): Collation and Sorting	U4	
29	Hinting	M3	
30	Course Catch up		
			Endterm on Script Processing