

Off-line ground-truthed Touching Character String image (*.tcs) file format

Item	Type	Length (B: Byte)	Instance
File Header			
Size of Header	Long int	4B	Number of bytes: 36+strlen(illustr)
Format code	ASCII (char*)	8B	“tcs”
Illustration	Text	Arbitrary	“#.....\0”
Code type	ASCII (char*)	20B	“ASCII” or “GB”.
Code Length (CL)	Short int	2B	1 (ASCII), 2 (GB).
Bits per pixel	Short int	2B	8 (Gray image)
String Image Records (concatenated)			
Stroke Width (SW)	Short int	2B	
Line Height (LH)	Short int	2B	
Number of Touching Points (NTP)	Short int	2B	1, 2, 3, ...
Position sequence of top and bottom terminal of touching point(s)	Short int	$2 * (2B + 2B) * NTP$	For each touching point, first top terminal, then bottom terminal; For each terminal, first row, then column.
*Number of Characters (NC)	Short int	2B	2, 3, 4, ...
Character labels (code)	Code type	$NC * CL$	
Height of image (H)	Short int	2B	
Width of image (W)	Short int	2B	
bitmap	Byte	$H * W$	Gray image

*For single-touching character string image, $NC = (NTP + 1)$; For multiple-touching character pair image, $NC \leq NTP$ and $NC = 2$.