

Building a Simulation Platform for Chinese Calligraphy Characters

WANG Chung-Shing^a, CHANG Teng-Ruey^b, LIN Man-Ching^c, and
WANG Ya-Hui^a

^aTung-Hai University

^bNan-Kai University of Technology

^cFeng-Chia University

Abstract

This research looked at an existing Chinese calligraphy font, the Kai-font, and carried out geometric modeling and analysis to set up a stroke feature model database. This allowed creation of an independent calligraphy simulation platform for extracting and analyzing fonts. Users can then input through the use of a mouse, digit pad or other systems to draw out the strokes of a Chinese character. The system will then automatically recognize the stroke starting coordinates, angles, area ratio and other parameters for stroke feature analysis and extraction. These data can then be compared with stroke feature information in the database to produce a simulated calligraphy font. This platform also allows the user to output the font simulation result to 3D drawing software such as SolidWorks for further design and applications. The advantage of this research hoped to develop a calligraphy font simulation system with a user friendly interface to allow traditional Chinese beginners having the opportunity to write out a Chinese calligraphy font on the computer. This will in turn serve to stimulate interests in Chinese calligraphy learning. Furthermore, this research enables the user to output the results to a computer aided design field for further application while promoting and spreading the art of Chinese calligraphy.

Keywords: *computer geometric modeling, Chinese calligraphy, stroke recognition, feature extraction, stroke database, calligraphy characters platform*

1. Introduction

The Chinese calligraphy is a very unique and important asset in Chinese culture and its high artistic value is praised by people all over the world. Although currently there are many different kinds of Chinese fonts available for selection for the computer system, the ability for the user to provide modification to the font itself is very limited. For example: we cannot offer a calligraphy font style that caters to personal handwriting habits to create new or composite words for processing and application. This research uses computer geometry modeling to analyze and extract characteristic strokes to produce a handwritten Chinese calligraphy font in order to build a calligraphy character platform. This offers the users an opportunity to further study and understand the writing structure of Chinese calligraphy.

In this study, we analyzed the Kai-font in Microsoft True Type fonts. Through analysis of the characteristic strokes of the Kai-font and carrying out simultaneous processing of strokes from non-traditional writings, we can remove, process and summarize duplicated strokes to create a stroke characteristic database [Chung & Tseng, 1995]. After that, by using the interface for the system platform, users can input handwritten font strokes with input devices such as the mouse or digit pad. The system will then automatically retrieve the corresponding strokes to carry out “Thinning” and “Stroke feature extraction” process for comparison with the characteristic stroke database to find the corresponding stroke to complete font output. The simulated results are then outputted as a .jpg file. These files are automatically exported into 3D computer aided design software, SolidWorks for further processing and application.

2. Paper Survey

Stroke characteristics are an important feature unique to Chinese calligraphy. Each individual stroke is composed of points, lines and curves. The point corresponds to the black spot; line refers to the linear movement of the tip of the pen and the curve is characterized by curve motion of the pen [Hornby, 1972; Stallings, 1976, 1977]. Strokes can be defined using operation points and parameters. Operation points may include the origin, end point, length, direction... etc. of the stroke. Parameters are then used to define the shape, width, position... etc. for the stroke [Wong & Hsu, 1995]. However, Chinese

calligraphy font strokes have a huge amount of variations which result in an extremely complex way for describing strokes. Even for similar a stroke structure, a small change in the shape, angle, position length or proportion can result in a completely different character.

Calligraphy font recognition, also known as reverse recognition, aims to achieve an effective way to obtain the point cloud data for the fonts. [Lin et al., \(2001\)](#) used an automatic threshold value setting to identify fonts and create a common words database. The search function of the database was then used to identify the fonts in the document. Detection and identification of common characters allows recognition of handwritten or printed fonts. [Romero et al., \(1997\)](#) used artificial neural network algorithm to enhance the identification rate for similar words along with the use of different search directions to modify the neural parameters to improve recognition speed. [Lin et al., \(1996\)](#) used a “Trend-followed” translation technology to find the contour segment and overall characteristic information for the font. [Chuang et al., \(1995\)](#) used heuristic algorithm to identify printed fonts while utilizing contour features to increase recognition rate.

Previous research by the authors used image processing and reverse engineering (RE) in Chinese calligraphy [[Wang, et al., 2006](#)]. This research combined processes of RE, grey prediction theory in pattern processing, geometric modeling for constructing Chinese calligraphy characters and rapid prototyping (RP) for model making. First of all, the written Chinese calligraphy was scanned by scanner. Next, the contours of the Chinese characters were detected with image processing. These contours were then converted to point data, which can be easily processed in any CAD software by using B-Spline curves to fit the points. An example was illustrated below using a Chinese compound word “the fortune and treasure are coming” to show the steps of the process. Finally, a RP model was constructed to show various applications for the products in Chinese calligraphy pattern. With this research, a collaboration between the Chinese calligraphy in handcraft and digital virtual design can be realised; but more importantly, the aesthetic aspects in the characters can be preserved ([Figure 1](#)).



(a) Image file of compound character



(b) Grey prediction contour detection



(c) Compound character in a paper weight



(d) A hanging decoration

Figure 1 Reverse design process for a Chinese compound character

3. Methodologies

In this research, we developed a simulation platform for Chinese calligraphy so the user can quickly create simulated calligraphy fonts for further application. The research framework consisted of the following three components:

- (1) Stroke database construction
- (2) Stroke feature recognition and coding.
- (3) Creating a simulation platform and user interface for writing stroke input.

3.1 Stroke database construction

In order to create a client based calligraphy simulation system, we first need to generate a corresponding stroke database. We used the existing built-in Kai-font within Microsoft Windows operating system as basis to obtain stroke information. Kai-fonts are TrueType vector fonts with characters constructed from individual strokes and we can extract its font structure data using Microsoft Development Network (MSDN). With TrueType vector fonts, font data structures are arranged as such that one character can be composed of many stroke contours and every contour can be composed of many curves. Statistics show that the Kai-font has 23,230 characters and 178,196 contour data.

Using the character 「永」 as example: The creation of the character with Kai-font is separated into six stroke contours, as shown in Figure 2. The meanings of each original stroke contour data are represented below. The mathematical relationship between ①②③④⑤ is:

$$\textcircled{1} = 16 + (\textcircled{2} + \textcircled{3}) * 4 + (\textcircled{4} + \textcircled{5}) * 8$$

① Size of the contour in the font file in bytes; ② Number of straight lines; ③ Number of curves; ④ Number of points used by the straight line; ⑤ Number of points used by the curve; ⑥ Angle of xy; ⑦ The actual Chinese character ; ⑧ What number stroke it is in the Chinese character. Consider the first stroke in the character 「永」, the “point” stroke, the contour has 0 straight lines, 8 curves, 16 points used by the curves (2 points for 1 curve) and a 39.2 degree angle. The total size is $16 + (0+8)*4 + (0+16)*8 = 176$ (Bytes).

①	②	③	④	⑤	⑥	⑦	⑧
176	0	8	0	16	39.2	永	1

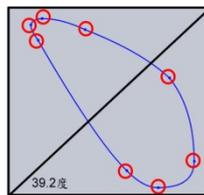


Figure 2 First stroke contour information of the character 「永」

If we input all the contours of the Kai-font into the contour database, a great amount of time is required for the contour recognition search process, which becomes highly inefficient. Therefore, duplicate contours were removed to reduce number of comparisons required in order to increase operational efficiency. The simplest method was to analyse common contours structures between different characters and removing those contour structures that were the same between them.

This research used the GetGlyphOutline function in Win32 API to access contour data in selected characters. The following stroke contour data were then analyzed to determine which contours need to be deleted.

1. Size occupied

Because Kai-font characters were created from individual strokes, the strokes with same structures take up the same amount of data space. For example, the first stroke in 「永」 and 「汁」 and the second stroke in 「汀」 both have the 「、」 structure and occupied 176 bytes. However, apart from the size, we also need to consider other attributes such as the angle between xy, to decide whether to remove the selected stroke contour as illustrated in Figure 3.



Figure 3 Stroke contours with same data size

2. Number of lines, curves and points

Each individual stroke may contain many lines or curves and we can distinguish between them by the number of lines, curves and points. For example: the first stroke of 「刀」、 「刁」、 「力」 all contain the same number of lines, curves and points; Therefore, they all have the same structure (Figure 4).



Figure 4 Strokes with same of lines, curves and points

3. Angle of xy

We can determine the structure of the stroke contour by looking at the diagonal angle formed by a box containing the stroke. As shown in Figure 5, the first stroke of the character 「九」 has an angle of 35.2 degrees and the second stroke 62.3 degrees. Calligraphy fonts with same strokes but with different stroke contour height and widths can be distinguished this way.

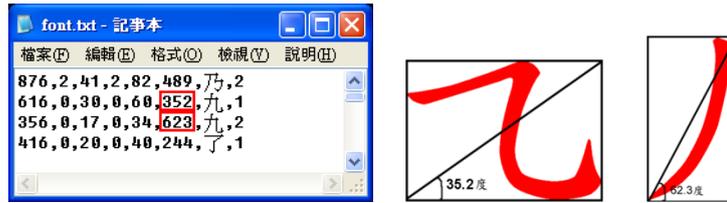


Figure 5 xy angle of the character 「九」

Based on the three criteria mentioned above, same stroke contours can be deleted from the extracted font file. If the three values were the same between different strokes, the one with the largest dimensions was kept because smaller stroke contours were more likely to have errors when carrying out stroke feature analysis. After analysis and deletion, a total of 10,218 stroke contour data remains. This contour characteristic database was then stored in Microsoft Access as shown in Table 1.

Table 1 Contour coding of contour database

stroke order	lx	ly	lx2	ly2	lx3	ly3	lx4	ly4	lx5	ly5	lx6	ly6	lx7	ly7	lx8	ly8	lx9	ly9	lx10	ly10	lx11	ly11	lx12	ly12	lx13	ly13	lx14	ly14	lx15	ly15	lx16	ly16	lx17	ly17	lx18	ly18	lx19	ly19	lx20	ly20	lx21	ly21	lx22	ly22	lx23	ly23	lx24	ly24	lx25	ly25	lx26	ly26	lx27	ly27	lx28	ly28	lx29	ly29	lx30	ly30	lx31	ly31	lx32	ly32	lx33	ly33	lx34	ly34	lx35	ly35	lx36	ly36	lx37	ly37	lx38	ly38	lx39	ly39	lx40	ly40	lx41	ly41	lx42	ly42	lx43	ly43	lx44	ly44	lx45	ly45	lx46	ly46	lx47	ly47	lx48	ly48	lx49	ly49	lx50	ly50	lx51	ly51	lx52	ly52	lx53	ly53	lx54	ly54	lx55	ly55	lx56	ly56	lx57	ly57	lx58	ly58	lx59	ly59	lx60	ly60	lx61	ly61	lx62	ly62	lx63	ly63	lx64	ly64	lx65	ly65	lx66	ly66	lx67	ly67	lx68	ly68	lx69	ly69	lx70	ly70	lx71	ly71	lx72	ly72	lx73	ly73	lx74	ly74	lx75	ly75	lx76	ly76	lx77	ly77	lx78	ly78	lx79	ly79	lx80	ly80	lx81	ly81	lx82	ly82	lx83	ly83	lx84	ly84	lx85	ly85	lx86	ly86	lx87	ly87	lx88	ly88	lx89	ly89	lx90	ly90	lx91	ly91	lx92	ly92	lx93	ly93	lx94	ly94	lx95	ly95	lx96	ly96	lx97	ly97	lx98	ly98	lx99	ly99	lx100	ly100	lx101	ly101	lx102	ly102	lx103	ly103	lx104	ly104	lx105	ly105	lx106	ly106	lx107	ly107	lx108	ly108	lx109	ly109	lx110	ly110	lx111	ly111	lx112	ly112	lx113	ly113	lx114	ly114	lx115	ly115	lx116	ly116	lx117	ly117	lx118	ly118	lx119	ly119	lx120	ly120	lx121	ly121	lx122	ly122	lx123	ly123	lx124	ly124	lx125	ly125	lx126	ly126	lx127	ly127	lx128	ly128	lx129	ly129	lx130	ly130	lx131	ly131	lx132	ly132	lx133	ly133	lx134	ly134	lx135	ly135	lx136	ly136	lx137	ly137	lx138	ly138	lx139	ly139	lx140	ly140	lx141	ly141	lx142	ly142	lx143	ly143	lx144	ly144	lx145	ly145	lx146	ly146	lx147	ly147	lx148	ly148	lx149	ly149	lx150	ly150	lx151	ly151	lx152	ly152	lx153	ly153	lx154	ly154	lx155	ly155	lx156	ly156	lx157	ly157	lx158	ly158	lx159	ly159	lx160	ly160	lx161	ly161	lx162	ly162	lx163	ly163	lx164	ly164	lx165	ly165	lx166	ly166	lx167	ly167	lx168	ly168	lx169	ly169	lx170	ly170	lx171	ly171	lx172	ly172	lx173	ly173	lx174	ly174	lx175	ly175	lx176	ly176	lx177	ly177	lx178	ly178	lx179	ly179	lx180	ly180	lx181	ly181	lx182	ly182	lx183	ly183	lx184	ly184	lx185	ly185	lx186	ly186	lx187	ly187	lx188	ly188	lx189	ly189	lx190	ly190	lx191	ly191	lx192	ly192	lx193	ly193	lx194	ly194	lx195	ly195	lx196	ly196	lx197	ly197	lx198	ly198	lx199	ly199	lx200	ly200	lx201	ly201	lx202	ly202	lx203	ly203	lx204	ly204	lx205	ly205	lx206	ly206	lx207	ly207	lx208	ly208	lx209	ly209	lx210	ly210	lx211	ly211	lx212	ly212	lx213	ly213	lx214	ly214	lx215	ly215	lx216	ly216	lx217	ly217	lx218	ly218	lx219	ly219	lx220	ly220	lx221	ly221	lx222	ly222	lx223	ly223	lx224	ly224	lx225	ly225	lx226	ly226	lx227	ly227	lx228	ly228	lx229	ly229	lx230	ly230	lx231	ly231	lx232	ly232	lx233	ly233	lx234	ly234	lx235	ly235	lx236	ly236	lx237	ly237	lx238	ly238	lx239	ly239	lx240	ly240	lx241	ly241	lx242	ly242	lx243	ly243	lx244	ly244	lx245	ly245	lx246	ly246	lx247	ly247	lx248	ly248	lx249	ly249	lx250	ly250	lx251	ly251	lx252	ly252	lx253	ly253	lx254	ly254	lx255	ly255	lx256	ly256	lx257	ly257	lx258	ly258	lx259	ly259	lx260	ly260	lx261	ly261	lx262	ly262	lx263	ly263	lx264	ly264	lx265	ly265	lx266	ly266	lx267	ly267	lx268	ly268	lx269	ly269	lx270	ly270	lx271	ly271	lx272	ly272	lx273	ly273	lx274	ly274	lx275	ly275	lx276	ly276	lx277	ly277	lx278	ly278	lx279	ly279	lx280	ly280	lx281	ly281	lx282	ly282	lx283	ly283	lx284	ly284	lx285	ly285	lx286	ly286	lx287	ly287	lx288	ly288	lx289	ly289	lx290	ly290	lx291	ly291	lx292	ly292	lx293	ly293	lx294	ly294	lx295	ly295	lx296	ly296	lx297	ly297	lx298	ly298	lx299	ly299	lx300	ly300	lx301	ly301	lx302	ly302	lx303	ly303	lx304	ly304	lx305	ly305	lx306	ly306	lx307	ly307	lx308	ly308	lx309	ly309	lx310	ly310	lx311	ly311	lx312	ly312	lx313	ly313	lx314	ly314	lx315	ly315	lx316	ly316	lx317	ly317	lx318	ly318	lx319	ly319	lx320	ly320	lx321	ly321	lx322	ly322	lx323	ly323	lx324	ly324	lx325	ly325	lx326	ly326	lx327	ly327	lx328	ly328	lx329	ly329	lx330	ly330	lx331	ly331	lx332	ly332	lx333	ly333	lx334	ly334	lx335	ly335	lx336	ly336	lx337	ly337	lx338	ly338	lx339	ly339	lx340	ly340	lx341	ly341	lx342	ly342	lx343	ly343	lx344	ly344	lx345	ly345	lx346	ly346	lx347	ly347	lx348	ly348	lx349	ly349	lx350	ly350	lx351	ly351	lx352	ly352	lx353	ly353	lx354	ly354	lx355	ly355	lx356	ly356	lx357	ly357	lx358	ly358	lx359	ly359	lx360	ly360	lx361	ly361	lx362	ly362	lx363	ly363	lx364	ly364	lx365	ly365	lx366	ly366	lx367	ly367	lx368	ly368	lx369	ly369	lx370	ly370	lx371	ly371	lx372	ly372	lx373	ly373	lx374	ly374	lx375	ly375	lx376	ly376	lx377	ly377	lx378	ly378	lx379	ly379	lx380	ly380	lx381	ly381	lx382	ly382	lx383	ly383	lx384	ly384	lx385	ly385	lx386	ly386	lx387	ly387	lx388	ly388	lx389	ly389	lx390	ly390	lx391	ly391	lx392	ly392	lx393	ly393	lx394	ly394	lx395	ly395	lx396	ly396	lx397	ly397	lx398	ly398	lx399	ly399	lx400	ly400	lx401	ly401	lx402	ly402	lx403	ly403	lx404	ly404	lx405	ly405	lx406	ly406	lx407	ly407	lx408	ly408	lx409	ly409	lx410	ly410	lx411	ly411	lx412	ly412	lx413	ly413	lx414	ly414	lx415	ly415	lx416	ly416	lx417	ly417	lx418	ly418	lx419	ly419	lx420	ly420	lx421	ly421	lx422	ly422	lx423	ly423	lx424	ly424	lx425	ly425	lx426	ly426	lx427	ly427	lx428	ly428	lx429	ly429	lx430	ly430	lx431	ly431	lx432	ly432	lx433	ly433	lx434	ly434	lx435	ly435	lx436	ly436	lx437	ly437	lx438	ly438	lx439	ly439	lx440	ly440	lx441	ly441	lx442	ly442	lx443	ly443	lx444	ly444	lx445	ly445	lx446	ly446	lx447	ly447	lx448	ly448	lx449	ly449	lx450	ly450	lx451	ly451	lx452	ly452	lx453	ly453	lx454	ly454	lx455	ly455	lx456	ly456	lx457	ly457	lx458	ly458	lx459	ly459	lx460	ly460	lx461	ly461	lx462	ly462	lx463	ly463	lx464	ly464	lx465	ly465	lx466	ly466	lx467	ly467	lx468	ly468	lx469	ly469	lx470	ly470	lx471	ly471	lx472	ly472	lx473	ly473	lx474	ly474	lx475	ly475	lx476	ly476	lx477	ly477	lx478	ly478	lx479	ly479	lx480	ly480	lx481	ly481	lx482	ly482	lx483	ly483	lx484	ly484	lx485	ly485	lx486	ly486	lx487	ly487	lx488	ly488	lx489	ly489	lx490	ly490	lx491	ly491	lx492	ly492	lx493	ly493	lx494	ly494	lx495	ly495	lx496	ly496	lx497	ly497	lx498	ly498	lx499	ly499	lx500	ly500	lx501	ly501	lx502	ly502	lx503	ly503	lx504	ly504	lx505	ly505	lx506	ly506	lx507	ly507	lx508	ly508	lx509	ly509	lx510	ly510	lx511	ly511	lx512	ly512	lx513	ly513	lx514	ly514	lx515	ly515	lx516	ly516	lx517	ly517	lx518	ly518	lx519	ly519	lx520	ly520	lx521	ly521	lx522	ly522	lx523	ly523	lx524	ly524	lx525	ly525	lx526	ly526	lx527	ly527	lx528	ly528	lx529	ly529	lx530	ly530	lx531	ly531	lx532	ly532	lx533	ly533	lx534	ly534	lx535	ly535	lx536	ly536	lx537	ly537	lx538	ly538	lx539	ly539	lx540	ly540	lx541	ly541	lx542	ly542	lx543	ly543	lx544	ly544	lx545	ly545	lx546	ly546	lx547	ly547	lx548	ly548	lx549	ly549	lx550	ly550	lx551	ly551	lx552	ly552	lx553	ly553	lx554	ly554	lx555	ly555	lx556	ly556	lx557	ly557	lx558	ly558	lx559	ly559	lx560	ly560	lx561	ly561	lx562	ly562	lx563	ly563	lx564	ly564	lx565	ly565	lx566	ly566	lx567	ly567	lx568	ly568	lx569	ly569	lx570	ly570	lx571	ly571	lx572	ly572	lx573	ly573	lx574	ly574	lx575	ly575	lx576	ly576	lx577	ly577	lx578	ly578	lx579	ly579	lx580	ly580	lx581	ly581	lx582	ly582	lx583	ly583	lx584	ly584	lx585	ly585	lx586	ly586	lx587	ly587	lx588	ly588	lx589	ly589	lx590	ly590	lx591	ly591	lx592	ly592	lx593	ly593	lx594	ly594	lx595	ly595	lx596	ly596	lx597	ly597	lx598	ly598	lx599	ly599	lx600	ly600	lx601	ly601	lx602	ly602
--------------	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------

reduce width from both sides of the line until the width is only one pixel but still retain the connection of the line.

This research proposed a stroke feature and coding method as basis for stroke feature recognition. Each stroke was encoded as a 15 bit unit to generate the stroke characteristic value. Coding in this research was divided into four parts as shown in [Table 2](#). Font stroke distribution characteristic values, which were mainly used as a point of comparison with written stroke characteristic values, can be separated into four categories ([Table 3](#)).

Table 2 Stroke characteristic coding index

items		bits
1. Stroke aspect ratio (length v.s. width)		2
2. Stroke vertex position		2
3. Angle between vertices		1
4. Stroke distribution	(1) The segments in y-axis when scanning in x direction	5
	(2) The segments in x-axis when scanning in y direction	5

Table 3 Stroke distribution characteristic value

Items	Index name
1. The segments area percentage in x-axis when scanning in y direction	bnx0 ~ bnx4
2. The segments area percentage in y-axis when scanning in x direction	bny0 ~ bny4
3. Fitting the partition percentage in x-axis when scanning in y direction	cx0 ~ cx4
4. Fitting the partition percentage in y-axis when scanning in x direction	cy0 ~ cy4

4. A Simulation Platform for Chinese Calligraphy Characters

Current research aims to develop a simulation platform for Chinese calligraphy character recognition. Softwares used for development are shown in [Table 4](#).

Table 4 Software for a simulation platform

Items	Software
Operation system	Windows XP
Programming language	Borland Delphi 7.0
Strokes database	Microsoft Access 2003
3D software	SolidWorks 2010 SP0

The system platform used the “Kai-font” as output for simulated fonts. Figure 6 illustrated the writing simulation process. In addition to outputting as solid fonts, we can also choose to output as font outlines or thin fonts. In addition to providing an output mode on screen, the font simulation platform can utilize the API function in SolidWorks to import simulated fonts as 3D models into software so designers can easily carry out follow-up applications. After importing to SolidWorks, each stroke became an individual unit and can be modified as shown in Figure 7. In Chinese calligraphy, certain auspicious phrases can be written as a compound word. For example: 「招財進寶」、「囍」 ...etc. Figure 8 and Figure 9 showed the compound word, 「招財進寶」 and the output screen in SolidWorks.



Figure 6 A process for Chinese character 「永」

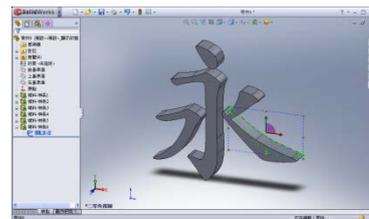


Figure 7 Output 「永」 to SolidWorks



Figure 8 Compound character writing

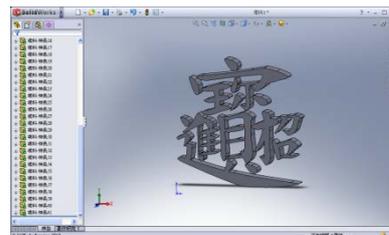


Figure 9 Output to SolidWorks

5. Conclusion

This research aims to develop a platform for Chinese calligraphy font simulation. The results and advantages are as follows:

- A calligraphy font simulation platform was constructed for the Chinese Kai-font. Beginners can easily write out Chinese character strokes with a mouse, digital pen or other input devices and the system platform will construct a simulated calligraphy font in real-time.
- The character stroke database for the Kai-font was reconstructed. Through character stroke feature extraction and analysis we combined strokes that were not practical under normal writing habits in order to delete duplicated strokes. This allowed us to reduce the amount of character stroke data for the Kai-font from 178,196 to 10,218, which increased system processing efficiency.
- A coding system was made for the Kai-font character strokes. Each stroke was encoded as a 15bit unit for comparison with both the index value within the database and the hand-written input strokes.
- With the use of API method within SolidWorks, we can quickly and efficiently export simulated fonts to 3D drawing software such as SolidWorks. This allowed further processing of Chinese compound words, creation of rarely used words, curve fitting and other additional applications.

6. Acknowledgement

This research was partially supported by the National Science Council, Taiwan under the Project (NSC 100-2221-E-029-010).

References

- Chuang, C.T., and Tseng, L.Y. (1995). A Heuristic Algorithm for the Recognition of Printed Chinese Characters. *IEEE Transactions on System, Man and Cybernetics*, 25(4), 710-717.
- Hornby, A.S. (1972). *The Advanced Learner's Dictionary of Current English*, 2nd Ed.,

Oxford Press.

- Lin, J.R., Chen, C.F. (1996). Stroke Extraction for Chinese Characters Using a Trend-Followed Transcribing Technique. *Pattern Recognition*, 29(11), 1789-1805.
- Lin, C.F., Fang, Y.F., and Juang, Y.T. (2001). Chinese Text Distinction and Font Identification by Recognizing Most Frequently Used Characters. *Image and Vision Computing*, 19, 329-338.
- Romero, R.D., Touretzky, D.S., and Thibadeau, R.H. (1997). Optical Chinese Character Recognition Using Probabilistic Neural Networks. *Pattern Recognition*, 30(8), 1279-1292.
- Stallings, W.W. (1976). *Approaches to Chinese Character Recognition*, *Pattern Recognition*, Pergamon Press, Great Britain, 8, 87-98.
- Stallings, W.W. (1977). Chinese Character Recognition, in Fu, K.S., 1977, *Syntactic Pattern Recognition and Application*, Springer-Verlag, New York, 95-123.
- Wang, C.S., Hsiao, C.Y., Chang, T.R., and Teng, C.K. (2006). Product Development for Chinese Calligraphy Using Reverse Engineering and Rapid Prototyping, *Virtual and Physical Prototyping*, 1(4), 259-269.
- Wong, P.Y.C & Hsu, S.C. (1995). Design Chinese Typeface Using Components, *IEEE Graphics*. 0730-3157/95, 416-421.
- Zhang, S., and Fu, K.S. (1984). A Thinning Algorithm for Discrete Binary Images. *Proceedings of the International Conference on Computers and Application*, Beijing, China. 879-886.