AN APPROACH FOR ACTIVE SEGMENTATION OF UNCONSTRAINED HANDWRITTEN KOREAN STRINGS USING RUN-LENGTH CODE

JeongSuk, J.; Kim, G.

Published in:
EPRINTS-BOOK-TITLE

An Approach for Active Segmentation of Unconstrained Handwritten Korean Strings using Run-length Code

Jason JeongSuk Yoon and Gyeonghwan Kim
Dept. of Electronic Engineering, Sogang University
CPO Box 1142, Seoul 100-611 Korea
E-mail: jasonyoon@ieee.org, gkim@ccs.sogang.ac.kr

We propose an active handwritten Hangul segmentation method. A manageable structure based on Run-length code is defined in order to apply to preprocessing and segmentation. Also three fundamental candidate estimation functions are introduced to detect the clues on touching points, and the classification of touching types is attempted depending on the structural peculiarity of Hangul. Our experiments show segmentation performance of 88.2% on touching characters with minimal over-segmentation.

1 Introduction

Conventional Hang l eg entation et o ologie gene all e a ate a a te o e ving t e ini al ve ti al o le o a t ing Ho eve t e e i ient o i ilit o i o e eg entation an t i onvin ingl lea t e eg ente atte n it e e tive a ti a t to e e ogni e in o e tl n e a le o e a a le lla le i o n in ig e lla le in o n ing o e a te t e onventional eg entation a e al o t i ila to i e ent Hang l a a te an it i ell no n t at to n e a t eg entation oint it o t e ognition i ve i lt e e o e t i a e inten to t a e o et een t e eal it an t e i eal in ing ove eg entation a e on t e t t al att i te o Hang l

e nee to e ne eve al te to avoi on ion a o t Hang l o ation in e ea Hang l e e ent a t e e i at lea t one vo el an one on onant in a Hang l lla le a Hang l lla le i a e o e e one e in a Hang l lla le a e la i e a an e en ing on t ei elative lo ation goal i to eg ent e a t Hang l lla le

Hang l one e e e o iginall e igne on gating ve ti al an o i ontal line it i le ot an to ing o t o lla le all o all o i ontal onta t a te t ong ve ti al t o e n on e en e it i e a ive to o on t ong ve ti al o onent it in n t ea o a te ati a oa e e e e ne a a e on t e n lengt o e ig a ill t ate t e t e o e tion a o ing to t e ve ti al n lengt ata t t e in i ate a a one a e t e n i ave

In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 511-516

igure omparison between conventional and ideal segmentation

igure 2 efinition of a section

igure efinition of the angle of a section

igure istribution of angles based on the sectioning scheme

o e t an t o a a ent n on it one i e i t e e ne a a a et o n et een t o an n e on l i a an n o n one i e o n t ea it i alle inall en ot i e o en it i al o e ne a an a e e nition a lie to t e o i ontal n lengt t t e ee ig

reprocessing - slant correction

e an ea il o tain t e ave age lant angle it t e angle o a e tion o n in ig a n angle o a e tion i e ne a t e angle o t e lin et een t e ente o t e egin n an t e en n in a e tion en t e e tion a o e t an t o n in i e

ig o t e i t i tion o t e angle a e on t e e tioning et o tati ti all it i lea t at Hang l a o e iagonal in lination in t i g a n a t an itten Hang l t ing gene all ave a t ong ten en to e iagonal t o e t on tantl a e t t e ave age o angle o t e t o e to e o itive ia e nli e ngli

o t e e o e tie o taine e i i all e an a e in o e to integ ate t e angle eig te i i ne l e ne in an ig in ing a o i e a ian o a ilit en it n tion tte to t e angle o en e ig

igure he results of slant correction

igure llustrations of the three fundamental functions

et an e t e eig t an t e angle o t e t e tion e e tivel an t en t e eal lengt i enote a

e e i t e n e o e tion in t e t ing t e e te lant o a t ing i gene all o taine

2

e e t e o i e a ian i an a e o e given on tant an a e an ave age o t e angle o e a le o lant o e tion a e o a e in ig

Segmentation

ivi e an e ge t ateg t e to ing o le in ngli a een e ll olve o anot e e e tive o o le an e ega e a a i ila i e o it i t e onne te o onent a e e a e o e l o eg entation eve t ele e ill o on t e ivi e o e in t i a e e a e t e e ea o e ging i till in og e

o ing o a a te all leave eve al l e on a ligat e i t o all e a o e ve an a t ange o t o e i t o i e tion at t e to ing oint en t e t o t o e a e to e ea ot e it an a te angle e on l t e tail o a a a te ten to ing o t t e e lination o t o e i t inall t e ini onta t o it a lig t to ing

e e t ee i tin tion a e on i e e a t e e l e i e e ent t e t ee to a t e a e i on ition e o ination

igure llustrations of the four touching types

igure 8 rocedures for three touching types

o t e n tion i e a an intelle t al an i ate ete to e n tion a e ill t ate in ig an t e e n tion et n e i val e o t e ollo ing et o olog

iscreteness his function measures how ad acent runs are abruptly changed. he function d , in ig. b is defined as

$$2 \qquad 2 \qquad 2$$

i erence et be the height of -th run. he di erence of the heights between a pair of runs is returned as a function value. f same heights are given for another case, the function value is same shown in ig. c , either.

Overlap he function returns the proportion of the to the overlapped portion of . When we assume that the upper-left corner is the origin of the image, the overlap function in ig. d is given as

$$2 \qquad 2$$

i i a i an a a a ave la i e to ing in igit into i t e e e ave een evi e in ing t e n a ental n tion ie a i all in o e to t t e Hang l a t e ve i ation o e e t a e ita le l e t o g eve e tion lti le to ing a e not on i e e a a e i t e e a e e a e t at t e a i n e o to ing oint it in a e tion i one e all e e to ig an ig a t e o a t o t e o e e an t e o e on ing e a le in etail e e tivel e o e o on e t o t e to ing t e a e a ollo

igure xamples of segmentation results from our system a successful segmentation b failures

estination he procedure is for detecting the -type touching points. A -type touching can be found as the run whose height is the smallest among the runs having the maximal di erence .

Wea ridge eclination of the stro e width is common on the bridge between the touched syllables. he measure of overlapping may discover the wea est points on a ligature.

onstriction When round or vertical stro es are deeply touched, we can separate them by applying the discreteness feature. he main condition is simple. he begin and the end runs of a section should be taller than the candidate run.

oncavity he concavity chec does not have to trace through a section. he candidates are given as the Hruns which have more than two upper neighbor Hruns.

perimental Results and Conclusion

te t a e e o e it t e a e i age olle te o t ent e n e o t e ting an lla le a e o nte vi all onl i t e eg ent a te eg entation ave no e e t o e ognition e ave age n e o lla le e a e t ing a e t o g t e ave age n e o onne te o onent a e i i e to t e o ation o Hang l lla le a le o t e n e o onne te o onent e o e an a te eg entation i a lie e ave age n e o onne te o onent e Hang l lla le i tantiall in ea e to neve t ele it i a e ta le en e noti e a Hang l lla le i a e o one e e e o an e o eg entation i al o e ti ate a e on Hang l lla le an t ing e te t i age ontain ingle to ing an lti le to ing o t o lla le e e l eg entation ate a e on t e n e o to ing a e o n in a le l o t o t e ingle to ing lla le a e eg ente it o t a ti a t in o e e i ent

inall e o nte t e n e o t e t ing i ave at lea t one to ing lla le o t e e i age i ave onl ingle to ing a e e ll eg ente e a e e e i ent i a lie to t e i age it lti le to ing an t e e lt i o n in a le ote t at a o t

able umber of connected components

able 2 egmentation performance for Hangul syllables

able egmentation performance for Hangul strings

o t e t ing in l e at lea t one to ing oint an t e eg en tation e o an e o t e ove all t ing i age i ite en o aging ig a an o eve al e a le o eg entation e lt o o te e an o e ve t at o t o e o o on t e to ing lla le t at a e n e ove a le

n t i a e e ave o e on e ing e o in eg entation o an itten Hang l e en ing on t e t t al e lia it t e eve al algo it an n tion e e int o e in o e to ove o e t e a a o onventional Hang l eg entation et o int o ing t i a oa eg entation a a i a ieve on Hang l lla le it to ing t e t ie o l e on t e evelo ent o e ging e e in o e to e e t e n e o eg ent a te eg entation l o t e la i ation o to ing t e nee to e e ten e to on i e on ing to ing a e

References