International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:6, 2011

Comparing Hilditch, Rosenfeld, Zhang-Suen, and Nagendraprasad - Wang-Gupta Thinning

Anastasia Rita Widiarti

Abstract**—**This paper compares Hilditch, Rosenfeld, Zhang-Suen, dan Nagendraprasad Wang Gupta (NWG) thinning algorithms for Javanese character image recognition. Thinning is an effective process when the focus in not on the size of the pattern, but rather on the relative position of the strokes in the pattern. The research analyzes the thinning of 60 Javanese characters.

Time-wise, Zhang-Suen algorithm gives the best results with the average process time being 0.00455188 seconds. But if we look at the percentage of pixels that meet one-pixel thickness, Rosenfelt algorithm gives the best results, with a 99.98% success rate. From the number of pixels that are erased, NWG algorithm gives the best results with the average number of pixels erased being 84.12%. It can be concluded that the Hilditch algorithm performs least successfully compared to the other three algorithms.

Keywords**—**Hilditch algorithm, Nagendraprasad-Wang-Gupta algorithm, Rosenfeld algorithm, Thinning, Zhang-suen algorithm

I. Introduction

YOGYAKARTA city, as one of the centers for Javanese culture, holds a large collection of old Javanese manuscripts. The manuscripts are important for the study of Javanese culture. The general consensus agrees that preserving classical texts is an important task, as they are part of the cultural heritage. But the number of people who can read and study old Javanese texts, written in Javanese scripts, is rapidly decreasing. There is an urgent need for translation of Javanese texts into Romanized scripts so that the texts may have larger readership.

The development of document image analysis, which analyzes the visual representation of paper documents such as journals, facsimiles, office documents, spreadsheet, etc. [1], has provided valuable ways for preservation of old Javanese manuscripts in Yogyakarta. O'Gorman and Kasturi [2] delineate the stages in the process of document image analysis, which can be used and modified for the document image recognition of Javanese script. The first stage is the stage of data collection, namely obtaining data from the documents to be processed. The second stage is the pixel-level processing stage that aims to prepare the document image, and to make the features of an intermediary to help identify the image. The third phase is the phase that aims to translate a series of alphabet letters that have a variety of shapes and sizes. One of the processes that must be passed in the second stage is the process of thinning or skeletonizing.

Anastasia Rita Widiarti is a faculty member at the Department of Informatics Engineering, Sanata Dharma University, in Yogyakarta, Indonesia (phone: +6281328341628; fax: 01-0274-886529; e-mail: rita_widiarti@yahoo.com). (This research was funded by Indonesia's Directorate General of Higher Education through competitive grants: Letter of Agreement No. 304/SP2H/PP/DP2M/IV/2010 ON 12 April 2010 and Higher Education Competitive Research Grant Multi Year and No. 060/SP2H / PP/DP2M/IV/2009 DATED 6 April 2009)

Thinning is a process of reduction of the components of the image in order to obtain the most basic information about the image-forming or to obtain an image without destroying the framework of information from its original form [3]. With the framework of an image obtained, then the computer can process the data faster and more easily because the representation of the processed data becomes much simpler.

A variety of thinning algorithms is available, including Rosenfeld algorithm, Zhang Suen algorithm, Hilditch algorithm, and Nagendraprasad-Wang-Gupta (NWG) algorithm. Zhang Suen thinning algorithm [4] is known as a fast algorithm in the process of thinning when compared with Hilditch thinning algorithm [5], easy to implement [6] and can be used to attenuate various types of digital patterns. Rosenfeld algorithm [7][8] is proved to meet the requirement for eligible connectivity, and Nagendraprasad-Wang-Gupta thinning algorithms is cited to work best for Latin characters [9].

II. Literature Review

A. Thinning Algorithms

Thinning is a process of reduction of the components of the image in order to obtain the most basic information about the image-forming or to obtain an image without destroying the framework of information from its original form [3]. The result of the thinning process is called the skeleton.

The objective of thinning is to reduce the image components into an information that is essential / fundamental to further analysis and recognition can be facilitated. Such information may consist of structures of objects such as intersections (junctions), the end point (end points) and point circuit (connection points). Some of the benefits that would be obtained if the thinning process is managed properly:

1) Size becomes smaller because the data will generate important information only, and it also reduces memory usage [10][11].
2) Facilitate the structural analysis of an object [12].
3) The resulting skeleton can be used for classification in the process of pattern recognition (pattern recognition) [13].

There are several requirements that must be met so that the thinning process is generally considered good, namely:

1) Able to reduce the size of the data [10].
2) Produces a thin line image of an object, where the line is

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:6, 2011

composed of the pixels that are connected [6].

3) Able to maintain connectivity between the pixels of an object (connectivity), or in other words do not cause the shape of an object to be disconnected [14].
4) Able to maintain the important characteristics of an object [10].
5) Produces a form that resembles its original form by not introducing another new feature of an original form, nor does it eliminate the characteristics of the object [6].

There is a variety of thinning algorithms, including Zhang Suen algorithms, Rosenfeld algorithm, Hilditch algorithm, and NWG algorithm. Hilditch algorithm is relatively easy to implement; the standard procedure in the process of thinning is to do several iterations of erosion on an object, where at every iteration we check on all points in the image. A point object is then converted into a background object when these requirements are met: the number of points between the neighboring points is 2 (two) to 6 (six); there is only one (1) pattern of change from background to object; there is a background to the edge point neighbor above, or on the adjacent left side or right side; and there is a background to the neighbor above, or on the left or below it. Zhang Suen algorithm, created by Zhang and Suen [4], is included into parallel thinning algorithm, in which a new value given to a point on the current iteration depends on the value generated in the previous iteration, and the whole point representing the image are processed simultaneously, with the assumption that uses 3x3 window and each point connects with 8 point neighbors. This algorithm removes the boundary points identified from the image pattern, except the points included in the image frame. Rosenfeld algorithm is a parallel algorithm that works by successively removing a subset of the boundary/boundaries of the object or region which is also called contour pixel or contour points (pixels outside / edge) [15]. NGW algorithm is a new algorithm of serial and parallel thinning previously developed by Wang and Zhang, which then experienced improvement in processing speed [9][16].

B. The Concept of One-pixel Thickness

Thinning result or skeleton is said to have a one-pixels thickness if it does not contain the template A = {A1, A2, A3, A4} as shown in Fig. 1. A template structure is the arrangement of the pattern found in the connected components that is not a skeleton. If one or more templates A are contained in the skeleton of the thinning, there are two possibilities: either the skeleton does meet 1-pixel thickness, or the skeleton meets 1-pixel thickness when the pixels in the skeleton are critical points. Pixels in the skeleton is said to be critical points when they contain at least the structure of the template B or C as shown in Fig. 1. Pixels outside the template structure A, B, and C are pixels 0 (zero). Pixels that are circled in each template are the center of the template. Pixel "x" on each pixel of the template structure is ignored [13].

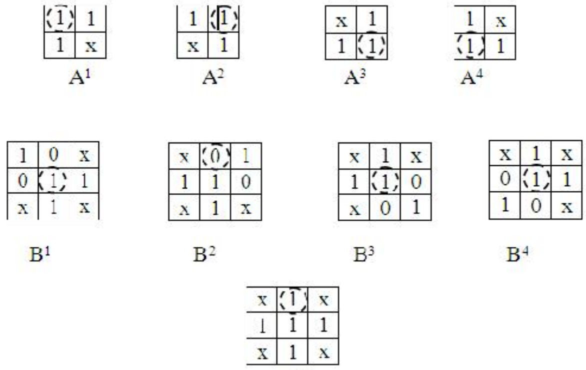


Fig. 1 The templates in the set A, B, C are patterns found in connected components that are not skeleton. Those in sets B and C are all possible configurations of critical points of a skeleton that contain one of the patterns in A [13]

III. Research Methodology

A. Preliminary Study and Analysis of Data

Before testing the algorithms, some library research, deepening of MATLAB, and selection of materials to analyze were conducted. Among different types of Javanese script, only the nglegeno Javanese letters were selected for this research. For each script, 3 data were to be made, so the total number of data used is 60 data of nglegeno Javanese characters. The Javanese characters used in this research were collected by scanning several Javanese manuscripts, both printed and handwriting manuscripts.

B. The Design of System Process

To facilitate the testing process, a computer-based thinning system was created, which will implement the four algorithms tested. See an illustration of the overview system below (Fig. 2):

Fig. 2 Overview system

Based on the overview system, a context diagram for the system as shown in Fig. 3 was created. Input for the system is an image that has not been subjected to thinning process, and output of the system is an image that has been thinned. The research then collected information related to the performance of the algorithms.

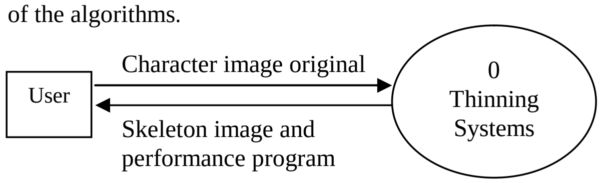


Fig. 3 Contex diagram

C. The Design of System Testing

The results of image thinning process must fulfill some requirements. The requirements used in the analysis and in the testing of the results include:

1) The size of the image before and after thinning
2) The number of pixels removed
3) The average time spent for each algorithm
4) The number of pixels that meet one-pixel thickness.

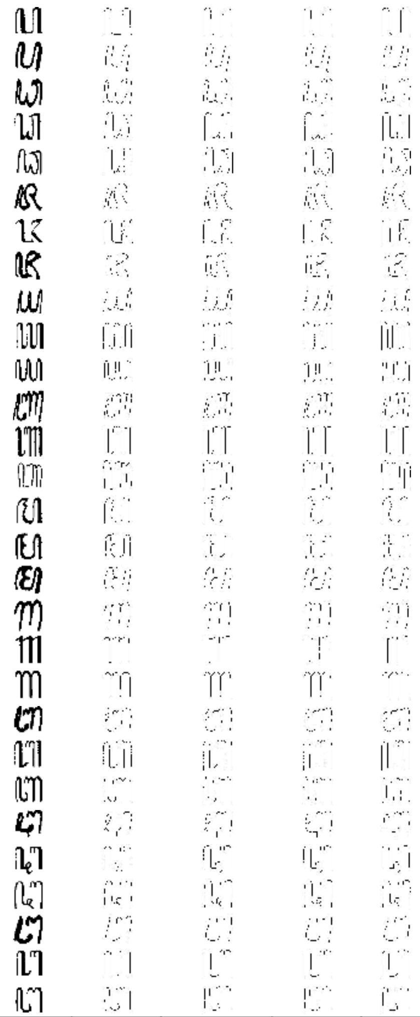
IV. Experimental Result and Analysis

A. Result of Experiment

Table I presents all original images and skeleton images after the implementation of Hilditch algorithm, Zhang-Suen algorithm, Rosenfeld algorithm, and NWG algorithm.

TABLE I
THE ORIGINAL IMAGES AND THE SKELETON IMAGE FROM THE FOUR ALGORITHMS

No	File name (JPG)	Original image	Skeleton image with Hildith	Zhang-Suen	Rosenfeld	NWG
1	HA_1					
2	HA_2					
3	HA_3					
4	NA_1					
5	NA_2					
6	NA_3					
7	CA_1					
8	CA_2					
9	CA_12					
10	RA_1					
11	RA_2					
12	RA_3					
13	KA_1					
14	KA_2					
15	KA_3					
16	DA_1					
17	DA_2					
18	DA_3					
19	TA_1					
20	TA_2					
21	TA_3					
22	SA_1					
23	SA_2					
24	SA_3					
25	WA_1					
26	WA_2					
27	WA_3					
28	LA_1					
29	LA_2					
30	LA_3					
31	PA_1					
32	PA_2					
33	PA_4					
34	DHA_1					
35	DHA_2					
36	DHA_15					
37	JA_1					
38	JA_2					
39	JA_4					
40	YA_1					
41	YA_2					
42	YA_3					
43	NYA_1					
44	NYA_2					
45	NYA_15					
46	MA_2					
47	MA_3					
48	MA_4					
49	GA_1					
50	GA_2					
51	GA_3					
52	BA_1					
53	BA_2					
54	BA_3					
55	THA_1					
56	THA_2					
57	THA_14					
58	NGA_1					
59	NGA_2					
60	NGA_3					



Based on the system testing design that has been set for this experiment, for each algorithm applied in the thinning process, we measured image size, the number of pixels removed, time spent, and the number of pixels that meet one-pixel thinness. Table II below presents a sample of the result of thinning using Hilditch algorithm, to show what is observed during the testing process. The same process was applied to the three other algorithms.

TABLE II
RESUME OF IMPLEMENTATION RESULT FOR HILDITCH ALGORITHM

Name file of original image	Image size: Original	Image size: Re-sult	Sum & % for pixels deleted: Sum	Sum & % for pixels deleted: %	Time of process (seconds)	Examination for one – pixel thickness: Sum of temp-late A	Sum of not critical points	Sum of one-pixel thick-ness	%
HA_1	1990	356	1634	82.11	0.00518	64	63	272	81.19
HA_2	1909	372	1537	80.51	0.00521	48	47	218	82.26
HA_3	1758	318	1440	81.91	0.00731	54	52	251	82.84

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:5, No:6, 2011

NA_1	1884	304	1580	83.86	0.00598	67	67	220	76.66
NA_2	1817	367	1450	79.8	0.00522	56	56	292	83.91
NA_3	2032	303	1729	85.09	0.00615	55	55	232	80.84
CA_1	1927	315	1612	83.65	0.00511	78	75	223	74.83
CA_2	1887	359	1528	80.98	0.00634	50	49	295	85.76
CA_12	1747	350	1397	79.97	0.00475	60	59	273	82.23
RA_1	1983	166	1817	91.63	0.00832	58	58	98	62.82
RA_2	1725	191	1534	88.93	0.00576	22	22	152	87.36
RA_3	1585	183	1402	88.45	0.00353	39	39	131	77.06
KA_1	1525	308	1217	79.8	0.0048	88	88	196	69.01
KA_2	1940	404	1536	79.18	0.00515	57	56	330	85.49
KA_3	1510	371	1139	75.43	0.00558	51	51	311	85.91
DA_1	1964	266	1698	86.46	0.00657	69	69	169	71.01
DA_2	2143	348	1795	83.76	0.00489	68	67	266	79.88
DA_3	2316	312	2004	86.53	0.00479	50	50	250	83.33
TA_1	2027	290	1737	85.69	0.00423	80	77	195	71.69
TA_2	1271	401	870	68.45	0.00563	63	62	319	83.73
TA_3	1846	351	1495	80.99	0.00364	74	73	257	77.88
SA_1	1720	287	1433	83.31	0.00662	71	71	191	72.9
SA_2	2096	365	1731	82.59	0.00602	68	67	280	80.69
SA_3	2086	308	1778	85.23	0.00411	61	61	232	79.18
WA_1	1670	288	1382	82.75	0.00879	61	60	215	78.18
WA_2	1962	323	1639	83.54	0.0055	42	42	268	86.45
WA_3	1506	281	1225	81.34	0.00372	36	36	236	86.76
LA_1	1673	343	1330	79.5	0.00484	89	89	225	71.66
LA_2	1781	362	1419	79.67	0.00486	32	32	324	91.01
LA_3	2270	333	1937	85.33	0.00368	49	49	271	84.69
PA_1	1868	261	1607	86.03	0.00544	58	57	195	77.38
PA_2	2083	308	1775	85.21	0.00544	41	41	248	85.81
PA_4	2171	269	1902	87.61	0.00475	66	66	184	73.6
DHA_1	1951	322	1629	83.5	0.00577	56	54	254	82.47
DHA_2	1697	343	1354	79.79	0.00605	63	63	264	80.73
DHA_15	1293	345	948	73.32	0.00357	48	48	288	85.71
JA_1	1720	305	1415	82.27	0.00633	76	76	218	74.15
JA_2	1931	292	1639	84.88	0.00392	63	62	218	77.86
JA_4	1398	292	1106	79.11	0.00553	96	96	169	63.77
YA_1	1925	291	1634	84.88	0.0043	48	47	237	83.45
YA_2	1897	420	1477	77.86	0.00524	57	57	345	85.82
YA_3	1745	319	1426	81.72	0.00333	47	47	255	84.44
NYA_1	2094	375	1719	82.09	0.00438	71	70	291	80.61
NYA_2	1916	411	1505	78.55	0.0048	33	33	372	91.85
NYA_15	1342	407	935	69.67	0.00323	41	44	352	88.89
MA_2	1973	308	1665	84.39	0.00477	62	61	230	79.04
MA_3	1964	289	1675	85.29	0.00481	64	63	200	76.05
MA_4	1290	291	999	77.44	0.00677	78	78	194	71.32
GA_1	1632	259	1373	84.13	0.00393	64	64	179	73.66
GA_2	2120	250	1870	88.21	0.00512	34	32	210	86.78
GA_3	1583	250	1333	84.21	0.00336	48	48	187	79.57
BA_1	1766	269	1497	84.77	0.00565	59	58	198	77.34
BA_2	1587	342	1245	78.45	0.00483	39	38	299	88.72
BA_3	1052	320	732	69.58	0.00478	50	50	262	83.97
THA_1	2048	236	1812	88.48	0.00569	62	60	164	73.21
THA_2	1928	291	1637	84.91	0.00417	48	48	236	83.1
THA_14	1349	300	1049	77.76	0.0024	52	51	155	75.24
NGA_1	1665	216	1449	87.03	0.00564	52	51	141	73.44
NGA_2	2121	269	1852	87.32	0.00442	21	19	242	92.72
NGA_3	1396	261	1135	81.3	0.0036	56	55	137	71.35

B. Data Analysis

Based on the results shown in Table II, we can summarize in the three tables below the following: the average amount of image sizes before and after the thinning process along with the percentage reduction in image size (Table III), the average number of pixels in the original image and the reduced number of pixels (Table IV), the average of time spent during the process of thinning (Table V), and the average number of pixels that meet one-pixel thickness (Table VI).

TABLE III
THE AVERAGE NUMBER OF PIXELS REMOVED

Algorithm	The average size of the image		Percentage reduction in image size
	Input	Output	
Zhang-Suen	14.44	1.46	89.83
NWG	14.44	1.47	89.83
Rosenfeld	14.44	1.48	89.77
Hildith	14.44	1.57	89.07

TABLE IV
THE AVERAGE NUMBER OF PIXELS OF ORIGINAL IMAGE AND SKELETON IMAGE

Algorithm	The average number of pixels			The average percentage of pixels removed
	Input image	Output image	Removed	
NWG	1800.92	280.50	1520.42	84.12
Rosenfeld	1800.92	280.52	1520.4	84.11
Zhang-Suen	1800.92	291.58	1509.33	83.38
Hildith	1800.92	310.6	1490.32	82.27

TABLE V
THE AVERAGE NUMBER OF TIME SPENT

Algorithm	The average number of time spent
Zhang-Suen	0.00455188
Hildith	0.00507102
NWG	0.00678418
Rosenfeld	0.037125383

TABLE V I
AVERAGE FOR SUM OF ONE-PIXEL THICKNESS

Algorithm	The average number of			
	Template A	Non chritical point	meet the One-pixel thickness Number	Percentage
Rosenfeld	0.97	0.05	280.47	99.98
NWG	2.05	1.02	279.48	99.64
Zhang-Suen	28.42	27.55	264.03	90.20
Hildith	56.88	56.32	235.27	79.99

V. CONCLUSION

Time-wise, Zhang-Suen algorithm gives the best results with the average process time being 0.00455188 seconds. But if we look at the percentage of pixels that meet one-pixel thickness, Rosenfelt algorithm gives the best results, with a 99.98% success rate. From the number of pixels that are erased, NWG algorithm gives the best results with the average number of pixels erased being 84.12%. It can be concluded that the Hilditch algorithm performs least successfully compared to the other three algorithms.

ACKNOWLEDGMENT

I would like to thank my students who provided me with the Javanese image characters data and helped refine the thinning system. Special thanks go to Sasongko Hadi, Kristi Krisnawati, and Margaretha Kurnianita. This research was funded by Indonesia's Directorate General of Higher Education (DIKTI) through competitive grants.

REFERENCES

[1] Srihari, S.N., Lam, S.W., Govindaraju, V., Srihari, R.K., and Hull, J.J. *Document Image Understanding*. New York: CEDAR, 1986.

[2] O'Gorman, L., and Kasturi, R. *Executive briefing: documen image analysis.* USA: IEEE Computer Society Press, 1997.

[3] M. Shimizu, H. Fukuda, and G. Nakamura. "Thinning Algorithm for Digital Figures of Characters", *Proceeding 4th IEEE Southwest Symposium on Image Analysis and Interpretation,* 2000.Pp: 83-87.

[4] Zhang, T. Y. and Suen, Ching Y., "A Fast Parallel Algorithms For Thinning Digital Patterns", *Communication of the ACM,* Vol 27, No. 3, Maret 1984, pp.236-239.

[5] Lam, L. and Suen, Ching Y., "An Evaluation Of Parallel Thinning Algorithms For Character Recognition", *IEEE Transaction On Pattern Analysis And Machine Intelligence,* Vol 17, No. 9, September 1995, pp.914-919.

[6] E. Adeline, *Enhancement of Parallel Thinning Algorithm for Handwritten Characters Using Neural Network*, Master Thesis, Department of Computer Science, Faculty of Computer Science and Information Technology, Universiti Technologi Malaysia, 2005. http://eprints.utm.my/3796/1/AdelineEngkamatMCD205ttt.pdf.

[7] Klette, Gisela. *Skeletons in Digital Image Processing.* 2002

[8] K.H. Lee, K.B. Eom, and R.L. Kashyap "Character Recognition Based on Attribute-Dependent Programmed Grammar", *IEEE Transaction On Pattern Analysis And Machine Intelligence,* VoL. 14, No. 11, November 1992, pp.1122-1128.

[9] Nagendraprasad, MV., Wang, PSP., and Gupta, A., "Algorithms for Thinning and Rethickening Binary Digital Pattern", *Digital Signal Processing,* Vol. 3, 1993, pp. 97-102. http://dspace.mit.edu/bitstream/handle/1721.1/46843/algorithmsforthi00nage.pdf?sequence=1

[10] Zhang, T. Y. dan Wang, P. S. P., "Analysis of Thinning Algorithms", *College of Computer Science Northeastern University Boston, MA 02115.* 1992, pp.763-766.

[11] L. Lam, SW Lee, and CY. Suen, "Thinning Methodologies – A Comprehensive Survey", *IEEE Transaction on Pattern Analysis and Machine Intelligence*. Vol. 14, No. 9, September 1992, pp. 869-885.

[12] Dawoud, Amer dan Kamel, Mohamed, "New Approach for the Skeletonization of Handwritten Characters in Gray-Level Images", *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003), IEEE.*

[13] Jang, BK., and Chin, RT., 'Analysis of Thinning Algorithms Using Mathematical Morphology", *IEEE Transactions on Pattern Analysis and Machine Intellegence.* Vol. 12, No. 6, 1990, pp. 541-551.

[14] Rinaldi, Munir. *Pengolahan Citra Digital dengan Pendekatan Algoritmik*. Bandung: Penerbit Informatika, 2004.

[15] Taussaint, Godfried. *Skeletons.* http://www.citr.auckland.ac.nz/

[16] Wang, PSP., and Zhang, YY. "A Fast and Flexsible Thinning Algorithm", *IEEE Transactions on Computer.* Vol. 38, No. 5, 1989, pp. 741-745.