



## Perekayasaan Histogram Orientasi Kecerunan Mengeskan Erotan dan Pencongan Manuskrip Merong Mahawangsa

MOHD SANUSI BIN AZMI, KHAIRUDDIN OMAR & AZIZI ABDULLAH

### ABSTRAK

*Pengesanan Teks Optik (PTO) adalah popular dan telah diperkembangkan kepada teks jawi. Pengkajian ditumpukan kepada penentuan sudut erotan dan pencongan pada teks jawi. Terdapat pelbagai kaedah dalam menentukan sudut erotan dan pencongan teks jawi di antaranya Kuasa Dua Terkecil (KDT) dan Histogram Orientasi Cerunan (HOC). Kaedah yang direkayasa ialah HOC dipelopori oleh Sun dan Si untuk mengesan erotan dan pencongan tulisan rumi. Teknik ini turut popular di kalangan penyelidik-penyelidik tulisan Jawi untuk mengesan erotan dan pencongan tulisan tangan Jawi. Walau bagaimanapun, kecekapan pengesanan sudut ini perlu diselidiki dan diperbaiki dengan algoritma rekayasa agar sesuai dengan kecekapan dan ketepatan sudut pengesanan. Perekayasaan algoritma HOC sedia ada mungkin dapat menyumbangkan kepada ketepatan dalam mengesan erotan dan pencongan tulisan Jawi. Pengujian dilaksanakan untuk membandingkan HOC asal dan yang direkayasa serta algoritma KDT.*

### ABSTRACT

*Optic Text Recognition (PTO) is popular and has been extended to Jawi (Arabic character) text. The research is focused on the determination of the degree of skew and slant in Jawi text. There are a variety of methods to determine the degree of skew and slant in Jawi text like the Smallest Square (KDT) and Gradient Orientation Histogram (HOC). The reengineered method is the HOC pioneered by Sun and Si to detect the skew and slant of Latin alphabet. This technique is also popular among Jawi researchers to detect the skew and slant of Jawi writing. However, the efficiency of the degree determination need to be investigated and enhanced with a reengineered algorithm so that it is suitable with the efficiency and accuracy of the determination degree. The reengineering of the present HOC algorithm may contribute to the accuracy in the skew and slant detection of Jawi writing. Tests are conducted to compare the original HOC and the enhanced HOC, as well as the KDT algorithm.*



## PENDAHULUAN

Pengecaman tulisan tangan atau lebih dikenali sebagai Pengecaman Teks Optik Aksara (PTOA) telah lama dibangunkan dalam menterjemahkan tulisan tangan yang diimbas ke dalam bentuk teks. Bidang ini termasuk dalam bidang pengecaman corak iaitu sub bahagian kecerdasan buatan. Pengecaman tulisan tangan Rumi telah banyak dilakukan dan diikuti oleh tulisan-tulisan lain contohnya tulisan Arab, Jepun, Cina dan juga tulisan Jawi. Menurut Scattolin (1995), kajian pengecaman tulisan tangan berlaku sejak lebih 30 tahun lalu dan hanya sedikit aplikasi yang berjaya menggunakan pengecaman tulisan tangan. Sistem Pengecaman Teks Optik komersial yang pertama telah diperkenalkan pada awal tahun 1950an menggunakan logik pengecaman berdasarkan kepada perkakasan (Khairuddin 2000). Walau bagaimanapun, bidang ini telah dibangunkan oleh penyelidik-penyelidik tempatan untuk digunakan pada tulisan Jawi. Sistem tulisan Jawi mempunyai aksara yang sama dengan aksara-aksara Arab (Khairuddin 2000). Oleh itu, permasalahan tulisan tangan Jawi adalah serupa dengan permasalahan tulisan Arab.

Kajian ini menumpukan terhadap imej Jawi dari manuskrip Merong Mahawangsa. Pengkajian adalah untuk menormalkan semula tulisan tangan manuskrip ini jika terdapat pencongan imej yang disebabkan oleh posisi imbasan dan juga faktor semula jadi tulisan. Selain itu, aksara-aksara Jawi juga terdapat erotan kerana faktor semula jadi tulisan tangan dan juga faktor perkaitannya dengan khat al-Qalam al-Sittah (Mohd Sanusi dan Khairuddin 2002a).

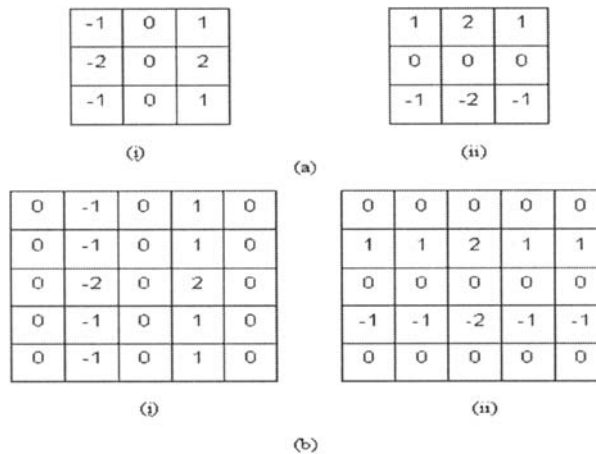
Khairuddin (2000) menggunakan pendekatan HOC model warna skala perduaan. Model warna ini hanya sesuai dalam persekitaran DOS. Manakala Sun dan Si (1997) menumpukan kajiannya kepada tulisan rumi menggunakan pendekatan HOC model warna skala kelabu. Walau bagaimanapun kesamaan antara Sun dan Si (1997), Khairuddin (2000) serta Juzlinda (2002) ialah mereka menggunakan HOC yang berasaskan Topeng Penjejak Sisi Perluasan Sobel (TPSPS) untuk mengesan dan membetulkan erotan dan pencongan.

### SUMBER PEREKAYASAAN HOC

HOC adalah teknik menentukan erotan dan pencongan menggunakan pendekatan kecerunan yang diperoleh dengan menyelusur piksel-piksel imej kajian menggunakan TPSPS dan berbeza dari kebanyakan kaedah lain yang menggunakan persamaan garis lurus (Mohd Sanusi 2003). Perekayasaan HOC dilakukan terhadap penggunaan topeng penjejak sisi yang digunakan dan model warna yang diwakilkan. Topeng penjejak sisi yang digunakan oleh Khairuddin (2000) dan Juzlinda (2002) ialah Topeng Penjejak Sisi Sobel (TPSS)  $5 \times 5$ . Ciri topeng  $5 \times 5$  ini merupakan perluasan dari TPSS asal iaitu  $3 \times 3$ . Oleh yang demikian, adalah lebih praktikal topeng  $5 \times 5$  ini dikenali sebagai Topeng Penjejak Sisi Perluasan Sobel (TPSPS)  $5 \times 5$  kerana bentuk ini hanya



dikemukakan oleh Sun dan Si (1997). Rajah TPSS dan TPSPS adalah seperti dalam Rajah 1.



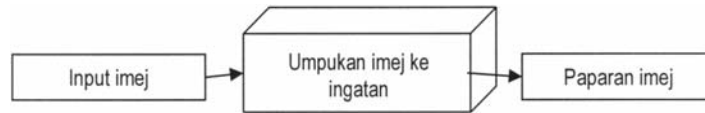
RAJAH 1. (a) Topeng Penjejak Sisi Sobel (b) Topeng Penjejak Sisi Perluasan Sobel

#### SENI BINA PROSES

Perekayasaan terhadap HOC melalui beberapa langkah awal yang tidak dapat dikecualikan kerana pengimplementasian algoritma ini menggunakan persekitaran yang berbeza. Sebelum ini, Khairuddin (2000) menggunakan persekitaran DOS dan diklasifikasikan kepada generasi ketiga bahasa pengaturcaraan. Kali ini persekitaran yang digunakan ialah generasi keempat. Oleh yang demikian, langkah-langkah pengecaman dan pembedaan erotan dan pencongan diterangkan dari pemerolehan imej hingga kepada imej dibetulkan. Oleh itu, peringkat yang terlibat ialah peringkat prapemprosesan. Pemerolehan imej perlu diterangkan kerana perbezaan dari bahasa pengaturcaraan yang digunakan oleh (Khairuddin 2000). Proses penggunaan piksel-piksel imej menggunakan penunjuk. Penunjuk ditakrifkan sebagai piksel yang mewakili imej.

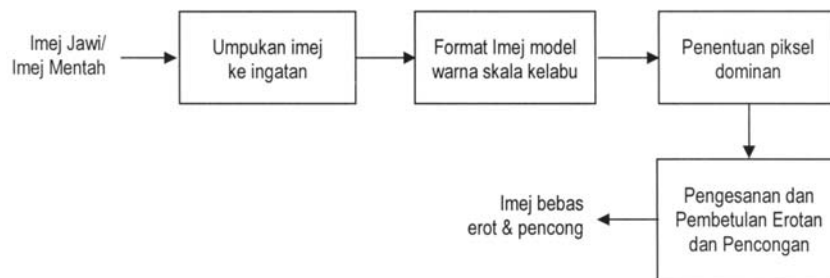
Pemprosesan imej oleh komputer memerlukan penapisan semula piksel dan penentuan piksel-piksel dominan yang digunakan sebagai penunjuk. Penunjuk piksel dalam kajian ini ialah piksel hitam tulisan jawi. Manipulasi imej mentah memerlukan penapis piksel yang menggunakan TPSS. Dalam kajian ini, operasi imej dibahagi kepada 3 peringkat asas iaitu i) input imej, ii) umpukan imej ke ingatan, dan iii) paparan imej seperti yang ditunjukkan di dalam Rajah 2 berikut:

Walaupun secara umum gambaran pemprosesan imej adalah seperti di dalam Rajah 2, namun ia tidak dapat menunjukkan dengan jelas bahagian



RAJAH 2. Peringkat umum operasi imej

pemprosesan imej terutamanya pemprosesan untuk mengesan dan membetulkan erotan dan pencongan tulisan Jawi kajian. Rajah 3 merupakan perincian pemprosesan imej kajian dan ianya menggambarkan proses-proses yang perlu dilalui untuk mengesan dan membetulkan erotan dan pencongan.



RAJAH 3. Rajah blok bagi modul prapemprosesan imej Jawi terpencong

#### IMEJ JAWI/ IMEJ MENTAH

Imej Jawi manuskrip Merong Mahawangsa adalah input kepada kajian ini dan berbeza dengan pendekatan oleh Khairuddin (2000). Pendekatan input imej Jawi oleh Khairuddin (2000) menggunakan sampel imej daripada 28 orang. Imej-imej Jawi asal manuskrip Merong Mahawangsa diperoleh dari kumpulan penyelidik Abdul Razak et al. (2002). Manuskrip ini mengandungi 116 halaman dan hanya 10 halaman sahaja yang diambil secara rawak iaitu halaman 1, 2, 3, 4, 5, 16, 27, 38, 49, dan 100. Kemudiannya sebanyak 10 petikan imej secara rawak yang lebih kecil diambil dari setiap halaman. Tujuannya ialah untuk mengetahui pola erotan dan pencongan imej. Menurut Muhammad Khatim (2003), 30 hingga 50 input imej sudah memadai untuk pengujian, dan semakin banyak input imej pengujian maka semakin kurang ralat. Sampel imej jawi digunakan berdasarkan kepada Mohd Sanusi (2003).

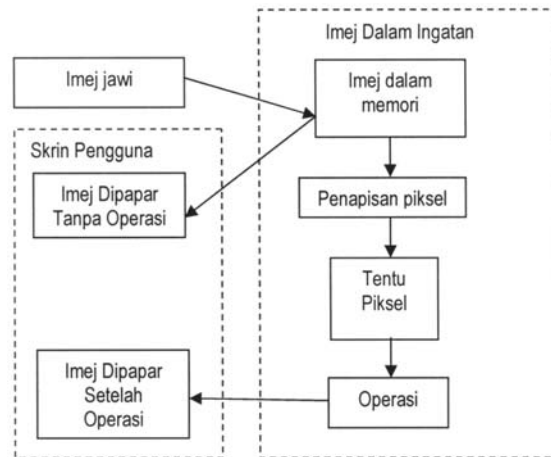
#### UMPUKAN IMEJ KE INGATAN

Fasa seterusnya ialah melaksanakan proses mengumpukkan imej ke dalam ingatan sama ada untuk dilakukan operasi seterusnya atau memaparkan imej. Pada peringkat ini, imej boleh dipaparkan terus tanpa melakukan operasi pengecaman piksel untuk tujuan manipulasi imej jika imej itu hendak





dipaparkan sahaja. Rujuk Rajah 2 dan Rajah 4. Apabila imej hendak dilakukan manipulasi maka perlu dilaksanakan penapisan piksel dan penentuan piksel-piksel dominan yang dijadikan sebagai penunjuk. Rajah 4 menunjukkan gambaran proses keseluruhan imej dalam bentuk yang berbeza. Paparan imej boleh dilakukan tanpa operasi dan juga dengan operasi.



RAJAH 4. Perincian operasi imej kajian

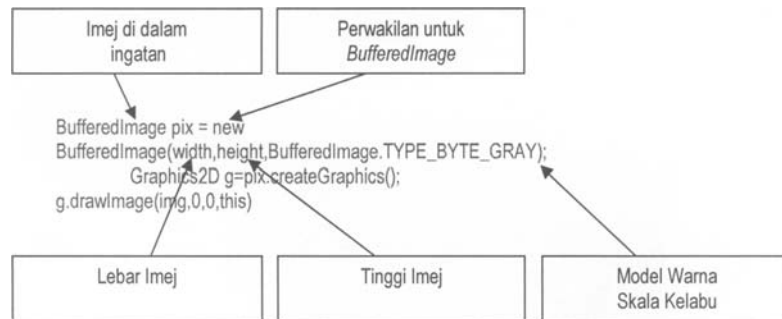
#### FORMAT IMEJ RGB KE MODEL WARNA SKALA KELABU

Sebelum imej dibuat penapisan menggunakan TPSS maka perlu pengumpulan imej ke dalam ingatan yang dikenali sebagai *BufferedImage* di dalam pakej *java.awt.image*. Pada peringkat ini imej akan disimpan ke dalam ingatan dalam bentuk format RGB, skala kelabu, perduaan, atau format-format lain yang disokong oleh aplikasi Java. Pendekatan skala kelabu adalah kerana penggunaan saiz ingatannya yang kecil berbanding dengan skala RGB dan ARGB (Efford 2000). Walau bagaimanapun, saiz ingatannya lebih besar dari skala perduaan. Mohd Sanusi (2003) telah membuktikan bahawa penggunaan saiz skala kelabu tidak sesuai untuk melaksanakan HOC kerana sudut-sudut berlaku pencongan yang dicerap begitu kecil. Selain itu, menurut Sun dan Si (1997), penggunaan skala kelabu adalah ideal untuk pemprosesan imej. Oleh itu, penggunaan skala kelabu merupakan kembangan skala imej dari (Khairuddin 2000) yang menggunakan skala perduaan. Imej diumpukkan ke dalam ingatan menggunakan model warna skala kelabu seperti yang ditunjukkan pada Rajah 5.

Imej yang disimpan ke dalam ingatan *BufferedImage* adalah ditukarkan kepada format skala kelabu menggunakan pernyataan berikut:

```
BufferedImage(width,height,BufferedImage.TYPE_BYTE_GRAY);
```





RAJAH 5. Imej diumpukkan ke dalam skala kelabu

Pembolehkan imej di dalam ingatan disingkatkan kepada *pix* sahaja. Perwakilan *pix* ini memainkan peranan penting untuk kelas-kelas di bawah kelas utama. Imej yang hendak dilakukan manipulasi format perdua seperti yang dilaksanakan oleh Khairuddin et al. (1999) boleh mengubah `TYPE_BYTE_GRAY` kepada `TYPE_BYTE_BINARY`.

#### PENENTUAN PIKSEL DOMINAN

Penentuan piksel imej kajian penting kerana untuk melakukan proses penjejakan TPSS. Khairuddin (2000) telah menggunakan piksel hitam dalam model warna perdua untuk melakukan penjejakan pinggir. Kajian ini menggunakan model warna skala kelabu sebagai penunjuk menggunakan format RGB. Piksel putih diwakilkan dengan nilai 255,255,255 manakala hitam diwakilkan 0,0,0 dalam format RGB.

#### PENGESANAN DAN PEMBETULAN EROTAN DAN PENCONGAN

Pengesanan dan pembaikan erotan dan pencongan menggunakan pendekatan Khairuddin (2000) yang menggunakan model warna skala perdua dan juga pendekatan yang digunakan oleh Sun dan Si (1997) yang menggunakan model warna skala kelabu dan TPSS 5x5 diperbincangkan oleh Mohd Sanusi dan Khairuddin (2002b). Persamaan di antara kaedah HOC Khairuddin (2000) dengan Sun dan Si (1997) ialah mereka menggunakan TPSS 5x5 walaupun berbeza model warna. Perakayaan yang telah dilakukan ialah menggunakan TPSS 3x3 dan juga menggunakan model warna skala kelabu seperti yang dipaparkan dalam Algoritma 1 berikut:

#### Algoritma 1:

Mula

Langkah 1: Operasi Kecerunan Dilaksanakan menggunakan TPSS 3x3.

Langkah 2: Hasilkan histogram orientasi dan dihaluskan lagi dengan penapis median.

68



- Langkah 3: Cari maksimum dari Histogram untuk sudut pencongan awal.
- Langkah 4: Diperhalusi lagi dengan membuat persamaan kubik polinomial dan dikira maksimumnya.
- Langkah 5: Imej diputarkan dari sudut yang diperoleh.
- Tamat

#### UJI KAJI

Uji kaji adalah proses untuk menentukan bahawa algoritma yang direkayasa dan disyorkan memberikan keputusan yang lebih baik atau pun tidak. Imej-imej yang diambil dari manuskrip Merong Mahawangsa dibahagikan kepada 4 kategori ujian. Tiga kategori pertama adalah pendekatan yang digunakan dan diperbaiki. Kategori-kategori ujian tersebut adalah:

1. HOC yang digunakan oleh Khairuddin (2000) pada imej tulisan tangan Jawi menggunakan TPSPS 5x5 pada model warna skala perduaan (Mohd Sanusi 2003)
2. HOC yang dipelopori oleh Sun dan Si (1997) untuk teks imej Rumi menggunakan TPSPS 5x5 dan model warna skala kelabu (lihat bahagian '*Hasil Uji Kaji HOC Menggunakan TPSPS 5x5 Skala Kelabu*')
3. HOC yang diperbaiki menggunakan TPSS 3x3 (HOC TPSS 3x3) dan menggunakan model warna skala kelabu (lihat bahagian '*Hasil Uji Kaji HOC Menggunakan TPSS 3x3 Skala Kelabu*')
4. Uji kaji kaedah KDT untuk perbandingan dengan HOC TPSS 3x3 (lihat bahagian '*Perbandingan Antara HOC TPSS 3x3 Dan KDT*')

Setiap uji kaji menggunakan 100 sampel imej mentah manuskrip Merong Mahawangsa. Hampir kesemua uji kaji menggunakan sudut  $45^\circ$  dan  $63^\circ$  kecuali uji kaji (2) dan (3). Alasan mengapa tidak mengambil sudut tersebut dibincangkan pada bahagian berikutnya. Pengujian adalah untuk memperoleh sudut-sudut berlaku pencongan dan erotan supaya dapat diperbaiki algoritmanya. Setiap sudut yang diperoleh dicatatkan dan diplotkan graf untuk menunjukkan kekerapan maksimum. Imej yang sama akan diuji pada atur cara yang berlainan supaya data yang diperoleh dapat dibuat perbandingan.

#### HASIL UJI KAJI HOC MENGGUNAKAN TPSPS 5X5 SKALA KELABU

Hasil uji kaji menunjukkan terdapat lebih banyak sudut berlaku pencongan iaitu pada  $45^\circ$ ,  $63^\circ$ ,  $71^\circ$ ,  $75^\circ$ ,  $78^\circ$ ,  $80^\circ$ ,  $81^\circ$ ,  $82^\circ$ ,  $83^\circ$ ,  $84^\circ$ ,  $85^\circ$ ,  $86^\circ$ ,  $87^\circ$ ,  $88^\circ$ , dan  $89^\circ$  (Mohd Sanusi 2003). Sudut  $45^\circ$  ditolak menurut Sun dan Si (1997) kerana sudut ini sentiasa maksimum nilainya pada skala kelabu. Hasil uji kaji penyesuaian lengkung persamaan polinomial kubik, menunjukkan tanpa sudut  $45^\circ$  nilai-nilai residual setiap sudut kekerapan pencongan lebih kecil. Menurut

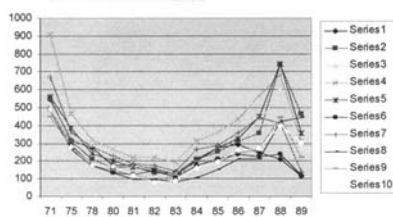


Mohd Sanusi (2003), graf tanpa 45° lebih baik bentuknya berbanding terdapat 45° kerana sudut 45° kekerapan maksimumnya adalah begitu tinggi seperti hasil uji kaji yang ditunjukkan pada Jadual 5 dan Jadual 6 dalam bahagian *Hasil Uji Kaji Penyesuaian Lengkuk Persamaan Polinomial*. Jadual 1 menunjukkan keputusan uji kaji dari halaman 5, dan 49. Sudut berlaku pencongan adalah tidak seragam antara imej-imej yang diuji kaji. Untuk hasil uji kaji pada halaman lain, sila rujuk Mohd Sanusi (2003).

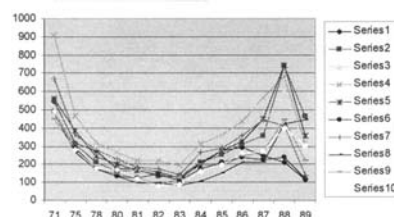
JADUAL 1. Sudut berlaku kekerapan dan sudut pencongnya

Imej Halaman 5	Sudut Kekerapan °darjah	Sudut Pencong °darjah	Imej Halaman 49	Sudut Kekerapan °darjah	Sudut Pencong
1	71	19	1	71	19
2	88	2	2	89	1
3	71	19	3	89	1
4	71	19	4	89	1
5	88	2	5	71	19
6	71	19	6	88	2
7	71	19	7	71	19
8	71	19	8	89	1
9	71	19	9	89	1
10	71	19	10	89	1

Rajah 6 hingga Rajah 7 menunjukkan graf kekerapan melawan darjah diplotkan berdasarkan uji kaji (Mohd Sanusi 2003). Pada garis tertinggi menunjukkan pada sudut terbahit terdapat banyak kekerapan berlaku pencongan imej.



RAJAH 6. Graf kekerapan lawan darjah untuk imej halaman 5



RAJAH 7. Graf kekerapan lawan darjah untuk imej halaman 49

#### HASIL UJI KAJI HOC MENGGUNAKAN TPSS 3X3 SKALA KELABU

Jadual 2 adalah keputusan uji kaji menggunakan HOC TPSS 3x3 dari halaman 5 dan 49. Sudut-sudut berlaku pencongan pada halaman 5 dan halaman 49

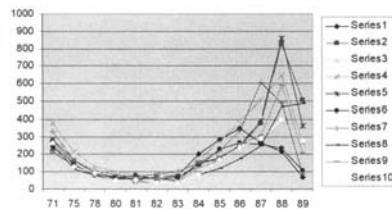




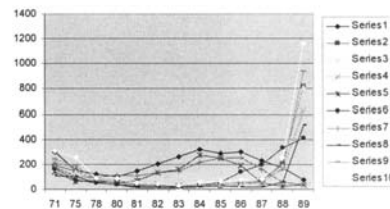
ditunjukkan dalam Jadual 2. Pencongan imej Jawi berlaku pada sampel tersebut dan sudutnya tidak sama di antara sampel walaupun sampel yang diambil adalah pada halaman yang sama. Graf dalam Rajah 8 dan Rajah 9 di bawah, menunjukkan pola sampel dari ujian di atas. Daripada graf-graf dalam Rajah 8 dan Rajah 9, setiap garis mewakili sampel. Garis yang tertinggi pada paksi-y menunjukkan kekerapan berlaku paling tinggi dan di situ diperoleh sudut pada paksi-x.

JADUAL 2. Sudut berlaku kekerapan dan sudut pencongnya

Imej Halaman	Sudut Kekerapan °darjah	Sudut Pencong °darjah	Imej Halaman	Sudut Kekerapan °darjah	Sudut Pencong
5			49		Pencong
1	86	4	1	84	6
2	88	2	2	89	1
3	88	2	3	89	1
4	88	2	4	89	1
5	88	2	5	84	6
6	71	19	6	89	1
7	87	3	7	86	4
8	89	1	8	89	1
9	88	2	9	89	1
10	88	2	10	89	1



RAJAH 8. Graf kekerapan lawan darjah untuk imej Halaman 5



RAJAH 9. Graf kekerapan lawan darjah untuk imej Halaman 49

#### PERBANDINGAN HOC 3X3 DAN HOC 5X5 SKALA KELABU

Kajian terhadap HOC TPSS 3x3 adalah sebanyak tiga kajian. Perbandingan TPSPS berbeza pendekatan model warna tidak dijadikan perbandingan untuk model warna skala perduaan yang hanya memberikan beberapa sudut berlaku pencongan (Mohd Sanusi 2003). Selain itu, pembaikan yang dilaksanakan ditumpukan kepada TPSPS. Jadual 3 menunjukkan perbandingan sudut hasil uji kaji berasaskan kepada HOC TPSPS dan HOC TPSS.





Jadual 3 menunjukkan perbandingan antara dua pendekatan. Terdapat berbeza nilai-nilai sudut pada beberapa sudut yang diambil secara rawak dari halaman 5. Pada imej 1, HOC TPSS 3x3 memberi nilai sebanyak 4<sup>o</sup> manakala HOC TPSPS 5x5 ialah 19<sup>o</sup>. Imej 2 nilai sudut HOC TPSS 3x3 ialah 2<sup>o</sup> dan begitu juga pada HOC TPSPS 5x5. Imej 3, dan 4 mempunyai sudut 2<sup>o</sup> pada HOC TPSS 3x3 dan 19<sup>o</sup> pada HOC TPSPS 5x5. Imej 5 dan 6 nilai sudut adalah sama iaitu sudut untuk kedua-duanya ialah 2<sup>o</sup>, dan imej 6 sudut untuk kedua-duanya ialah 19<sup>o</sup>. Imej 7, 8, 9, dan 10 sudutnya sama iaitu 19<sup>o</sup> untuk HOC TPSPS 5x5. Manakala HOC TPSS 3x3 ialah 3<sup>o</sup>, 1<sup>o</sup>, 2<sup>o</sup>, dan 2<sup>o</sup>. Selain itu, uji kaji untuk halaman 49 dirujuk pada Jadual 4. Sampel-sampel imej yang diperoleh dari halaman 49 manuskrip Merong Mahawangsa dibandingkan dengan hasil uji kaji pada Jadual 4. Terdapat beberapa sampel yang mempunyai sudut yang berbeza iaitu imej 1, 5, 6, dan 7. Imej-imej terbabit ditunjukkan pada Rajah 10.

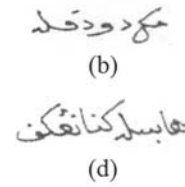
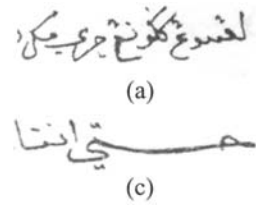
JADUAL 3. Perbandingan kekerapan dan sudut pencong

Imej	HOC TPSS 3x3		HOC TPSPS 5x5	
	Kekerapan	Sudut	Kekerapan	Sudut
1	350	4	546	19
2	824	2	739	2
3	784	2	719	19
4	647	2	908	19
5	859	2	742	2
6	277	19	489	19
7	606	3	669	19
8	485	1	484	19
9	587	2	451	19
10	397	2	492	19

JADUAL 4. Perbandingan kekerapan dan sudut pencong

Imej	HOC TPSS 3x3		HOC TPSPS 5x5	
	Kekerapan	Sudut	Kekerapan	Sudut
1	322	6	618	19
2	815	1	763	1
3	812	1	830	1
4	631	1	593	1
5	278	6	303	19
6	414	1	420	2
7	255	4	411	19
8	514	1	533	1
9	938	1	910	1
10	1162	1	1232	1





RAJAH 10. Sampel sudut berbeza: (a) imej 1, (b) imej 5, (c) imej 6, dan (d) imej 7

HASIL UJI KAJI PENYESUAIAN LENGKOK PERSAMAAN POLINOMIAL

Data-data dari pengujian terhadap HOC TPSS 3x3 dan HOC TPSPS 5x5 dilakukan penyesuaian lengkok menggunakan perisian Curve Expert 1.3. Ujian ini mencari lengkok terbaik untuk persamaan polinomial kubik. Penyesuaian lengkok dilakukan untuk mendapatkan nilai residual. Nilai residual ialah jarak antara koordinat piksel yang disesuaikan dengan koordinat titik asal (Muhammad Khatim 2003). Semakin kecil nilai residual maka semakin tepat graf yang terhasil. Jadual 5 dan Jadual 6 menunjukkan nilai-nilai residual untuk setiap sudut imej 2 yang diambil dari halaman 4. Taburan nilai residual yang semakin kecil menunjukkan semakin baik graf yang diplotkan.

Graf persamaan polinomial kubik untuk HOC TPSS 3x3 dirujuk Rajah 11 yang terdapat nilai 45° dan 63°. Rajah 11 menunjukkan graf diplotkan dengan baik dan nilai-nilai residualnya rujuk Jadual 5 lebih kecil berbanding graf yang menggunakan sudut 45° dan sudut 63°. Rajah 12 pula ialah graf persamaan polinomial kubik untuk HOC TPSPS 5x5. Graf diplot dengan

JADUAL 5. Residual HOC TPSS 3x3

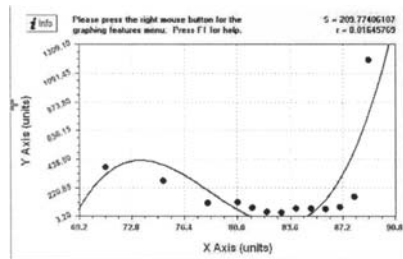
Sudut	HOC TPSS 3x3	HOC TPSS 3x3	HOC TPSS 3x3
45°	-9.96314	Tanpa 45°	Tanpa 45°
63°	159.1087	63.55728	Tanpa 63°
71°	-418.115	-304.314	55.38144
75°	39.3032	110.195	-120.817
78°	111.7386	130.1343	-92.7005
80°	178.0821	164.4512	54.70287
81°	139.5545	113.3305	70.97067
82°	92.64291	57.22979	79.95738
83°	46.90884	6.553529	84.91912
84°	17.91389	-22.2937	95.11203
85°	-62.7804	-96.9074	35.79225
86°	-174.612	-195.883	-78.7841
87°	-282.02	-282.816	-219.361
88°	-351.443	-323.302	-358.682
89°	513.6815	580.0641	393.5088



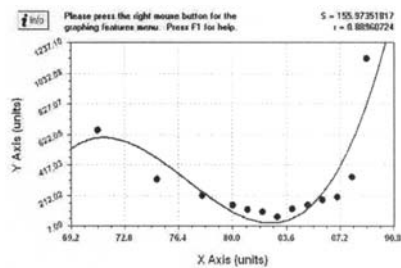
JADUAL 6. Residual HOC TPSPS 5x5

Sudut	HOC TPSPS 5x5	HOC TPSPS 5x5	HOC TPSPS 5x5
45 <sup>0</sup>	25.57685346	Tanpa 45 <sup>0</sup>	Tanpa 45 <sup>0</sup>
63 <sup>0</sup>	-216.1713226	29.12312454	Tanpa 63 <sup>0</sup>
71 <sup>0</sup>	179.8165442	-112.3279019	52.49116634
75 <sup>0</sup>	149.9726488	-32.01693135	-137.870994
78 <sup>0</sup>	128.3999305	81.1755269	-20.93151455
80 <sup>0</sup>	58.51693029	93.50928076	43.22055353
81 <sup>0</sup>	11.03338829	78.35419327	58.94411898
82 <sup>0</sup>	-30.89670854	60.01397442	70.42817511
83 <sup>0</sup>	-100.8372554	2.760757889	38.66932419
84 <sup>0</sup>	-97.35214743	5.866677379	59.66416853
85 <sup>0</sup>	-126.0052798	-38.39613343	22.4093104
86 <sup>0</sup>	-158.3605478	-103.7555408	-50.09864789
87 <sup>0</sup>	-210.9818464	-208.9394112	-179.8631041
88 <sup>0</sup>	-159.4330709	-231.6756107	-247.8874558
89 <sup>0</sup>	546.7218835	376.3079942	290.8248992

penyesuaian terbaik. Perbandingan graf yang diplotkan HOC TPSS 3x3 lebih baik pemplotannya dari HOC TPSPS 5x5. Rujuk Rajah 11 dan Rajah 12 untuk perbandingan graf tanpa sudut 45<sup>0</sup> dan 63<sup>0</sup> yang dikaji.



RAJAH 11. Graf polinomial kubik HOC TPSS 3x3 tanpa sudut 45<sup>0</sup> dan 63<sup>0</sup>



RAJAH 12. Graf polinomial kubik HOC TPSPS 5x5 tanpa sudut 45<sup>0</sup> dan 63<sup>0</sup>

#### HASIL UJI KAJI KUASA DUA TERKECIL

Kuasa Dua Terkecil (KDT) merupakan pendekatan yang digunakan oleh Mazani (2002). Pengujian turut dilakukan ke atas kaedah KDT bertujuan untuk membezakan algoritma yang diperbaiki iaitu HOC menggunakan TPSS 3x3. Uji kaji KDT dilaksanakan menggunakan 100 sampel, sama seperti yang digunakan pada uji kaji HOC TPSPS dan HOC TPSPS.

Uji kaji yang dilaksanakan untuk KDT mempunyai sudut yang baik terhadap imej yang terpencong. Nilai sudut positif imej pada Jadual 7



menunjukkan imej pencong di atas paksi asalan. Manakala nilai negatif menunjukkan imej terpencong di bawah paksi asalan. Kaedah KDT diketahui menggunakan pendekatan garis lurus. Maka, pendekatan cara untuk mencari sudut pencong adalah berbeza dari kaedah HOC.

JADUAL 7. Uji kaji kuasa dua terkecil

Imej	M/S 1	M/S 2	M/S 5	M/S 49
1	0.1829	0.4188	2.3759	1.2206
2	0.5703	-0.889	0.5085	4.3892
3	-0.069	-1.755	0.1053	1.2943
4	0.4435	-1.309	-0.142	0.7251
5	0.5621	0.9543	-0.872	-0.129
6	0.0276	0.8371	-5.565	0.5788
7	1.9255	0.8854	-0.503	0.3512
8	-0.067	-1.171	3.4596	2.9261
9	0.6295	0.9943	0.8902	2.1318
10	-0.138	-2.361	-0.505	0.9506

PERBANDINGAN ANTARA HOC TPSS 3x3 DAN KDT

Penggunaan HOC merupakan pendekatan purata piksel pada imej. Setiap piksel akan ditelusuri. Penggunaan kaedah ini menapis piksel berdasarkan kepada jenis TPSS yang digunakan sama ada 3x3 atau 5x5. Penggunaan 3x3 mempunyai kernal pada baris 3x1 yang akan dilaksanakan sebanyak tiga kali untuk setiap satu topeng. TPSS 5x5 menggunakan kernal 5x1 dan pelaksanaan penapisannya sebanyak 5 kali untuk satu topeng Sobel. Kawasan penyaringan TPSS lebih luas berbanding dengan TPSS. Selepas penapisan setiap piksel imej akan dijelmakan ke bentuk imej yang telah dilakukan penjejakan sisi. Tujuan dilakukan penapisan TPSS atau TPSS untuk mendapatkan  $\theta$  yang digunakan mencari sudut berdasarkan piksel yang telah ditetapkan.

KDT tidak menapis piksel imej keseluruhannya dan hanya menggunakan piksel tertentu sebagai penunjuk. Pendekatan yang digunakan ialah berdasarkan kepada garis lurus  $y = mx + c$  yang merentasi paksi-y. Garis lurus ini diperoleh sudut yang berlandaskan kepada garis datar.

HOC pula menggunakan pendekatan purata terhadap setiap piksel imej Jawi. Koordinat-koordinat piksel dalam kaedah HOC akan dicari kecerunan  $dy/dx$ . Dari kecerunan setiap piksel imej ini dicari kekerapan berlaku pencongan setiap sudut dalam selang  $0^\circ \leq \theta \leq 90^\circ$ . Kekerapan yang paling tinggi menunjukkan pencongan berlaku pada sudut tersebut.

Kelebihan HOC ialah pengekstrakan setiap piksel imej dan diterjemahkan kepada bentuk TPSS dan dapat menyumbangkan kepada proses seterusnya dalam pemprosesan imej Jawi seperti penyingkiran aksara sekunder dan



penipisan. Manakala KDT tidak menyumbang kepada proses seterusnya kerana piksel-piksel tidak dibuat penapisan dan diperoleh sisi setiap imej. Kaedah ini memulangkan nilai sudut yang berlaku pencongan sahaja. Jadual 8 menunjukkan perbezaan sudut antara HOC TPSS 3x3 dengan KDT.

Jadual 8 juga menunjukkan KDT memberi nilai pencongan sudut yang kecil secara puratanya. Faktor ini disebabkan oleh dua pendekatan mencari sudut yang berbeza. HOC terekayasa menggunakan purata setiap piksel imej manakala KDT tidak menggunakan purata dan sebaliknya menggunakan persamaan garis lurus.

Faktor masa pemprosesan imej pengesanan erotan dan pencongan menggunakan kaedah HOC adalah agak lama jika dibandingkan dengan kaedah KDT. Pemprosesan akan menjadi lebih efektif jika menggunakan perkakasan berkuasa tinggi, misalnya unit pemprosesan pusat yang terkini, dan ingatan capaian rawak yang besar.

JADUAL 8. Perbandingan HOC TPSS 3x3 dengan KDT

HOC TPSS 3x3 Halaman 3	KDT Halaman 3	HOC TPSS 3x3 Halaman 4	KDT Halaman 4	HOC TPSS 3x3 Halaman 5	KDT Halaman 5
3	0.1677	1	-0.596	4	2.3759
2	0.8838	1	-0.355	2	0.5085
2	1.1586	1	-1.041	2	0.1053
3	1.4977	1	1.1272	2	-0.142
2	2.1607	1	-1.249	2	-0.872
3	1.4424	1	0.8427	19	-5.565
2	2.4133	1	-0.394	3	-0.503
2	0.459	1	-0.211	1	3.4596
2	0.6651	1	-0.352	2	0.8902
4	-0.243	1	1.1659	2	-0.505




#### PEMBETULAN IMEJ TERPENCONG

Pembetulan imej dilakukan dengan memutarakan imej pada koordinat baru berdasarkan kepada sudut pencong yang diperoleh menggunakan kaedah HOC TPSS 3x3. Jadual 9 menunjukkan imej yang diproses, dan imej yang dibetulkan. Pembetulan imej-imej pada Jadual 9 menunjukkan imej-imej yang berjaya dikesan sudut pencongan dan diputarakan ke koordinat baru. Maklumat lanjut tentang sudut-sudut pencongan untuk imej-imej terbabit boleh dirujuk daripada Mohd Sanusi (2003).





JADUAL 9. Imej yang diproses dan pembetulan pencongan

Halaman 5	Proses Imej	Hasil Imej
Imej 1		این مکات بودی
Imej 2		ایت پتاله ادان کفوت
Imej 5		دف نبی الکریمان کارن

#### PERBINCANGAN DAN KESIMPULAN

Uji kaji yang dilaksanakan ke atas kaedah HOC yang dikaji dan juga perekayasaan kaedah HOC menggunakan 100 sampel imej memperoleh sudut yang ada kalanya sama. Namun, terdapat banyak perbezaan sudut terhasil dari HOC perekayasaan yang mempunyai nilai sudut yang lebih tepat jika dibandingkan kaedah Sun dan Si (1997) dengan Khairuddin (2000). Hasil uji kaji terperinci terhadap 100 sampel imej ini boleh dirujuk daripada Mohd Sanusi (2003) dan dibandingkan seperti dalam Jadual 10. Hasil uji kaji yang diperoleh dicari berapa banyak sudut di bawah  $5^\circ$ . Hasil yang diperoleh menunjukkan HOC TPSS  $3 \times 3$  mempunyai ketepatan di bawah  $5^\circ$  sebanyak 86 peratus manakala HOC TPSS  $5 \times 5$  sebanyak 69 peratus.

JADUAL 10. Peratus sudut imej di bawah  $5^\circ$

Halaman	HOC TPSS $3 \times 3$	HOC TPSS $5 \times 5$
1	0	0
2	10	10
3	10	2
4	10	10
5	9	2
16	10	10
27	9	8
38	10	10
49	8	7
100	10	10
Peratus	86	69



Model warna perduaan yang digunakan oleh Khairuddin (2000) di dapati jurang masa proses tidak begitu ketara di antara TPSS 3x3 dan TPSPS 5x5. Faktor ini disebabkan muatan ingatan diperlukan untuk memproses imej adalah kurang kerana skala perduaan mempunyai nilai 1 bit untuk setiap piksel. Model warna skala kelabu nilainya 8 bit untuk 1 piksel. Oleh yang demikian, penggunaan model warna skala kelabu menggunakan masa yang agak lama. Penggunaan masa dapat dikurangkan dengan penggunaan TPSS 3x3.

Penjimatan masa pemprosesan penting kerana terdapat banyak proses yang bakal dilalui dalam bidang ilmu Pengecaman Optik Jawi sebelum imej tersebut dapat diubah kepada bentuk *teks Jawi boleh edit*. Selain itu, pengguna akhir menitik beratkan masa yang digunakan untuk memproses dan menghasilkan output selain dari ketepatannya.

Pendekatan HOC TPSPS 5x5 adalah merupakan perancangan berdasarkan kepada uji kaji dan perbandingan dengan kaedah lain.

#### RUJUKAN

- Abdul Razak bin Hamdan, Wan Hilmi bin Mohd Nor, Mohd Faidzul bin Nasruddin, Roslim bin Mohamad, Khairuddin bin Omar, Mazani bin Manaf, Abdul Razak bin Salleh, Marini binti Abu Bakar & Miswan bin Surip. 2002. Manuskrip Merong Mahawangsa. Kumpulan Projek IRPA. Fakulti Teknologi Sains Maklumat, Universiti Kebangsaan Malaysia, Bangi.
- Juzlinda Mohd Ghazali. 2002. Penganalisa tulisan tangan berasaskan aspek pencongan dan kecondongan. Tesis Sarjana. Fakulti Teknologi Sains Maklumat, Universiti Kebangsaan Malaysia, Bangi.
- Khairuddin bin Omar. 2000. Pengecaman tulisan tangan teks Jawi menggunakan pengkelas multiaras. Tesis Ph.D. Fakulti Sains Komputer, Universiti Putra Malaysia, Serdang.
- Khairuddin bin Omar, Ramlan bin Mahmud, Md Nasir bin Sulaiman, & Abdul Rahman bin Ramli. 1999. Pembetulan pencongan dan erotan tulisan Jawi. Siri laporan teknikal FTSM/Julai 1999/LT74. Fakulti Teknologi Sains Maklumat, Universiti Kebangsaan Malaysia, Bangi.
- Mazani bin Manaf. 2002. Pengecaman tulisan tangan perkataan Jawi dengan menggunakan rangkaian Neural Rekuren Bama. Tesis Dr. Fal. Fakulti Teknologi Sains Maklumat, Universiti Kebangsaan Malaysia, Bangi.
- Mohd Sanusi bin Azmi. 2003. Perancangan erotan dan pencongan histogram orientasi kecerunan manuskrip Merong Mahawangsa. Tesis Sarjana. Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia, Bangi.
- Mohd Sanusi bin Azmi dan Khairuddin bin Omar. 2002a. Ulasan ciri-ciri tulisan Jawi dalam manuskrip Merong Mahawangsa. Laporan Teknikal FTSM/Jun 2002/LT9. Fakulti Teknologi Sains Maklumat, Universiti Kebangsaan Malaysia, Bangi.
- Mohd Sanusi bin Azmi dan Khairuddin bin Omar. 2002b. Ulasan jenis-jenis tulisan khat Jawi secara umum serta kaitannya dengan manuskrip Merong Mahawangsa. Laporan Teknikal FTSM/Julai 2002/LT12. Fakulti Teknologi Sains Maklumat, Universiti Kebangsaan Malaysia, Bangi.





- Muhammad Khatim Bin Hasan. 2003. Komunikasi persendirian, ruang legar Fakulti Teknologi Sains Maklumat, Universiti Kebangsaan Malaysia, 20 Februari.
- Efford, N. 2000. *Digital image processing: a practical introduction using Java*. Leeds: Addison-Wesley.
- Scattolin, P. 1995. Recognition of handwritten numerals using elastics matching. Tesis Sarjana. University of Montreal, Quebec, Canada.
- Sun, C. & Si, D. 1997. Skew and slant correction for document images using gradient direction. *Proceedings of the 4<sup>th</sup> International Conference on Document Analysis and Recognition* 1(18-20): 142-146.

Mohd Sanusi Bin Azmi & Khairuddin Omar  
Jabatan Sains dan Pengurusan Sistem  
Fakulti Teknologi dan Sains Maklumat  
Universiti Kebangsaan Malaysia  
43600 UKM Bangi  
Selangor Darul Ehsan  
e-mail: sanusi2u@yahoo.com, ko@ftsm.ukm.my

Azizi Abdullah  
Jabatan Sains Komputer  
Fakulti Teknologi dan Sains Maklumat  
Universiti Kebangsaan Malaysia  
43600 UKM Bangi  
Selangor Darul Ehsan  
e-mail: azizi@ftsm.ukm.my