

SEMINAR TEKNOLOGI MAKLUMAT

Inovasi Pengurusan Teknologi Maklumat: Paradigma dan Aplikasi

UCAPTAMA :

Prof. Dr. Khairuddin Omar

27 OKTOBER 2010

HOTEL AVILLION LEGACY

MELAKA

PENGANJUR :

Sekolah Teknologi Maklumat

dan

KYM-STES



JAWATANKUASA PENGANJUR

Penaung : Datuk Saroni bin Judi
Ketua Eksekutif KYM INTERNATIONAL

Pengarah Bersama : Dr. Abdul Karim bin Mohamad
Puan Salina Mohd Said

Timb. Pengarah : Puan Norsheila Ismail

Setiausaha Bersama : Cik Noraini Amin
Cik Noor Diana Mohamed Hassim

Bendahari bersama : Cik Farah Farhana Md Lani
Puan Haslinda Desa

Buku Aturcara : Puan Norazlin Jaliludin
Puan Suhana Samat
Cik Siti Amirah Tarmizi

Protokol, Keraian : Puan Ruziah Puasa
& Penaja Puan Juliah Mohamat
Cik Norlizah Kassim
Puan Sharidzah Anilizah

Teknikal : En.Nik Man Nik Soh
En Dzulfitri Othman

Pendaftaran & : Puan Noraniza Md Jani
Lokasi Puan Julia Bujang
Cik Nurul Ain Hashim

Hebahan & : Cik Hafeeza Husain
Maklumbalas Cik Suzi Iryanti Fadilah
Puan Ahamalina Munandar

Editor & Penerbitan : Puan Faridah Johar
Puan Rosniza A. Rahim
Cik Nor Hidayah Mohd Jamil

KANDUNGAN

Inovasi Pengurusan Teknologi Maklumat: Paradigma dan Aplikasi	Halaman
Kerangka Paleografi Jawi Digital : Satu Cadangan Awal Prof Dr. Khairuddin Omar	1-14
<i>Recognizing Botnet through DNS Traffic</i> Dr. Mohd Faizal Abdollah	15-19
<i>The Impact of Packet Retransmission on QoS of Audio-Video Transmission Over IEEE 802.11e HCCA WLANs</i> Dr. Zul Azri Muhammad Noh	20-23
<i>Web Based Avatar Game Maker : Adapting the concept of IMalaysia</i> Mohamad Lutfi Dolhalit	24-27
Kesediaan Guru Bahasa Melayu menggunakan Aplikasi Multimedia dalam Pengajaran Erman Hamid	28-32
<i>The Anatomy of DNS Spoofing using ARP Poisoning: Root Cause and Prevention Method</i> Mohd Zaki Mas'ud	33-37
<i>Support in IEEE 802.11g DCF Ad-hoc Networks through Variation of DIFS values</i> Mohd Najwan Md Khambari	38-42
<i>Investigating the Effect of Network Entry Phase Optimization in WiMAX Network</i> Mohd Firdaus Ghazali	43-45
<i>Web-Based Device Control and Communication through Parallel Port</i> Nor Azman bin Mat Ariff	46-48

Kerangka Paleografi Jawi Digital : Satu Cadangan Awal

Khairuddin O.^a, Mohd Sanusi A.^b, Siti Norul Huda S.A.^c, Azizi A.^d,
Mohammad Faidzul N.^e

^{a,c,d,e}Pusat Teknologi Kecerdasan Buatan
Fakulti Teknologi dan Sains Maklumat
Universiti Kebangsaan Malaysia
Bangi, Selangor

^bJabatan Kejuruteraan Perisian
Fakulti Teknologi Maklumat dan Komunikasi
Universiti Teknikal Malaysia Melaka (UTeM)
Melaka

Email: ko@ftsm.ukm.my, sanusi@utem.edu.my, mimi@ftsm.ukm.my, azizi@ftsm.ukm.my,
faidzul@ftsm.ukm.my

Abstrak

Kajian paleografi adalah kajian untuk mengetahui tarikh dan juga tempat di mana manuskrip lama ditulis. Pada masa ini, bilangan manuskrip jawi yang terkumpul di perpustakaan utama di Malaysia sahaja dianggarkan berjumlah 7789 buah. Malah terdapat manuskrip yang tidak dapat diketahui tarikh dan tempat asalnya. Selain itu, faktor khat arab yang mempengaruhi tulisan jawi juga memainkan peranan ke atas kajian paleografi. Di dalam manuskrip jawi, wujudnya variasi jenis tulisan khat arab yang menggambarkan pengaruh penulis dan juga membuktikan bahawa manuskrip tersebut ditulis oleh penulis yang ramai. Keperluan kepada paleografi jawi digital adalah amat perlu memandangkan manuskrip yang banyak dan juga kini terdapat salinan manuskrip yang dibuat di dalam bentuk digital. Dalam kertas ini, kajian menjurus kepada perspektif ahli sains komputer. Kaedah penyelidikan-penyelidikan untuk tulisan Latin, India, Ibrani dikaji, difahami dan diperbandingkan. Satu kerangka kajian untuk paleografi jawi dicadangkan untuk mengetahui jenis pengaruh khat arab ke atas manuskrip-manuskrip jawi. Fitur-fitur global dan setempat pada imej jawi akan dikaji dan digunakan untuk tujuan pengelasan. Pengelasan fitur berasaskan jenis-jenis tulisan khat di dalam manuskrip akan menyumbangkan kepada pengenalan pastian manuskrip. Selanjutnya ianya akan dapat menyumbangkan kepada asas penentuan jenis-jenis tulisan di dalam manuskrip melayu.

Kata kunci: paleografi, pengekstrakan fitur, pengelasan, pengelasan fitur.

1.0 Pengenalan

Paleografi adalah kajian manuskrip tulisan tangan kuno (Yosef et al. 2004). Paleografi juga boleh ditakrifkan sebagai seni melihat dan memahami (Bischoff 1990) kajian tentang sejarah skrip termasuk huruf singkatan dan tanda bacanya serta penafsirannya (Brown 1994). Mendapatkan tarikh dan tempat asal bagi manuskrip yang tidak diketahui tarikh dan tempat asalnya (Derolez 2008). Malah, menurut Sijpesteijn (2008) dipetik dari (Witkam 2009), paleografi adalah sains untuk mengenal pasti tarikh dokumen-dokumen lama, manakala paleografi arab adalah kajian kepada

pembangunan skrip-skrip arab melalui masa dan tempat. Menurut Witkam (2009), maksud barat di dalam kajian paleografi dunia barat boleh diaplikasikan kepada kajian paleografi arab.

Perspektif sains komputer dan yang lebih khususnya dalam bidang pemrosesan imej, kajian paleografi adalah subset kepada kajian ke atas tulisan. Cuma kajian ini menumpukan kepada manuskrip lama yang tidak diketahui tarikh dan tempat asalnya manuskrip tersebut. Kajian tulisan tangan pula telah banyak dilakukan di dalam pelbagai domain seperti mengenal pasti penulis,

pengelasan jenis tulisan, forensik, mahu pun kajian paleografi yang menyumbangkan kepada mendapatkan gambaran atau maklumat sejarah yang tersurat di sebalik tulisan.

Kajian ke atas tulisan tangan secara amnya ke atas tulisan roman dilakukan oleh Bulacu & Schomaker (2007), Izadi & Suen (2008), Khaled Mohammed Bin Abdl (2010), Schomaker et al. (2007), Srihari et al. (2001), tulisan india secara umumnya dilakukan oleh penyelidik Joshi et al. (2007), Padma & Vijaya (2009), tulisan hebrew dilakukan oleh Yosef et al. (2005), tulisan tangan farsi dilakukan oleh Shahabi & Rahmati (2006), Shahabi & Rahmati (2009). tulisan tangan arab dilakukan oleh Abandah & Khedher (2009), Abdi & Khemakhem (2010), Abuhaiba et al. (2002), Al-Habian & Assaleh (2008), Al-Ma'adeed et al. (2008), Bulacu et al. (2007), Gazzah & Amara (2008). Kebanyakan dari kajian ini bertujuan mengenal pasti identiti penulis, mengenal pasti jenis tulisan, gaya tulisan dan juga kajian paleografi. Umumnya kajian mengenal pasti identiti penulis lebih mendapat sambutan penyelidik.

Walau pun kajian ke atas identiti penulis berasaskan kepada dokumen tulisan tangan telah banyak dilaksanakan, namun kepentingannya kepada kajian paleografi tidak begitu signifikan (Moalla et al. 2006). Menurutnya, berdasarkan kepada kajian tulisan latin yang dilakukannya, jenis tulisan sentiasa berubah bergantung kepada penggunaan. Menurutnya, pada abad VIII hingga ke XI, tulisan jenis Caroline digunakan secara meluas di barat. Seterusnya tulisan jenis Gothic lahir dan berkembang di utara eropah. Perubahan kepada jenis tulisan bagi pengkaji paleografi berlaku dengan secara perlahan-lahan yang menyebabkan kepada kerumitan mengenal pasti

kategori jenis tulisan. Moalla et al. (2006) turut memberikan hasil dapatan dari pemerhatian beliau ke atas tulisan latin berjenis Caroline yang mempunyai unsur-unsur tulisan berjenis Gothic. Menurutnya lagi, evolusi jenis tulisan Caroline kepada *Cursive Gothic* kemudiannya kepada *Batarde Gothic* dan akhirnya kepada *Textualis Gothic* telah menambahkan lagi kerumitan bagi pengkaji Paleografi. Kerumitan ini bukan hanya pada tulisan latin malah turut dikongsi bersama oleh tulisan yang berasaskan kepada aksara arab.

Dalam dunia ketamadunan Islam, bentuk tulisan tangan juga mempunyai kategori yang tersendiri. Terdapat jenis-jenis khat yang popular seperti Kufi, Riq'ah, Nas'akh, Diwani, Diwani Jali dan Raihani (Sakkal 2003), (Mohd Sanusi, 2003) dan (Mohd Sanusi dan Khairuddin, 2002). Setiap jenis khat ini diperkenalkan pada masa dan tempat yang berbeza. Khat kufi dicipta di Kuffah dan mula dijadikan piawai penulisan Al-Quran pada abad ke 8 selama hampir 300 tahun, manakala khat nasakh pula dibangunkan pada abad ke 10 dan telah diperkemarkan di Turki pada abad ke-16. Kemudiannya, tulisan ini telah dijadikan sebagai tulisan piawai untuk menulis Al-Quran. Khat jenis Nastaliq ataupun farsi ditemui pada abad ke 14 dan 15 di Iran. Uniknyanya bagi setiap khat, ia lahir pada zaman yang berbeza dan juga pada tempat yang berbeza.

Di Tanah Melayu, kajian yang dilakukan oleh Mohd Sanusi (2003) ke atas manuskrip Merong Mahawangsa telah membawa kepada penemuan bahawa erotan dan pencongan tulisan jawi mempunyai kaitan kepada jenis tulisan yang digunakan. Malah menurutnya, di dalam manuskrip tulisan jawi tersebut terdapat beberapa jenis tulisan iaitu nasakh, thuluth dan riq'ah. Hasil dapatan

beliau mampu memberikan input berguna kepada kajian paleografi tulisan jawi.

Di dalam kitab Undang-undang Melayu, menurut Jajuli (1994), kitab ini ditulis oleh 3 orang penulis yang berlainan bersandarkan jenis khat yang berbeza pada halaman 1-10, 11-16 dan 17-32. Menurutnya, pada bahagian pertama iaitu halaman 1-10, tulisan khat yang digunakan berbeza dengan khat nasakh atau thuluth yang piawai. Tampaknya, bentuk khatnya lebih bulat dan pendek seolah-olah ianya tulisan jawa atau pun sanskrit. Pada bahagian kedua, ianya mempunyai elemen khat thuluth dan pada bahagian ketiga, bentuknya berubah kepada lebih berbentuk empat segi.

Dapatan yang diperoleh dari tarikh dan tempat khat dibangunkan, diguna secara meluas dan dikembangkan (Sakkal 2003) dan pengaruh tulisan khat di dalam manuskrip (Sanusi 2003) dan (Mohd Sanusi Azmi 2005) boleh digunakan sebagai petunjuk kepada kajian Paleografi. Malah ciri-ciri teknikal (Kodikologikal) khat tersebut boleh dijadikan indikasi penentuan jenis khat dan usia khat terbabit.

Berbalik kepada bidang pemprosesan imej, selama ini kajian paleografi jawi dilakukan oleh penyelidik seperti (Gallop 2005) dan (Witkam 2009) tidak dilakukan secara digital. Penyelidik tempatan yang mengkaji paleografi jawi ialah (Wan Ali Wan 2006) yang telah menerbitkan kertas kerja di Jurnal Filologi Melayu, jilid 14, tahun 2006. Beliau juga tidak menggunakan kaedah digital di dalam kajian paleografi. Penyelidik-penyelidik ini menentukan tarikh dan asal dokumen ditentukan oleh mereka berdasarkan kepada ciri-ciri kodikologikal yang wujud di dalam tulisan jawi.

Oleh itu, penyelidikan untuk menghasilkan aplikasi paleografi jawi secara digital adalah amat bersesuaian memandangkan kajian ini akan dapat membantu mengenal pasti tarikh, jenis tulisan dan tempat manuskrip ditulis tanpa perlu kehadiran pakar paleografi.

Faktor ini disokong oleh Jajuli (1994), menurutnya, terdapat manuskrip yang tidak diketahui bila ditulis. Malah penulis menganggap bahawa manuskrip kajian beliau iaitu Undang-undang Melayu ditulis pada pertengahan abad ke 18. Buat masa ini sahaja, berdasarkan kajian pada tahun 2007 manuskrip terkumpul di Malaysia dianggarkan berjumlah 7789 buah, yakni Perpustakaan Negara Malaysia sebanyak 3699 buah, Perpustakaan Universiti Malaya sebanyak 307 buah, Perpustakaan Tun Sri Lanang sebanyak 57 buah, Dewan Bahasa dan Pustaka sebanyak 226 buah dan Muzium Seni Islam Malaysia pula sebanyak 3500 buah (Mohd Hilmi Md Rifin 2007). Dengan bilangan yang dinyatakan di atas, keperluan kepada paleografi digital adalah penting di dalam menentukan jenis tulisan, pentarikhan, dan tempat asal manuskrip.

Untuk menjayakan paleografi jawi secara digital, teknik yang digunakan oleh penyelidik dalam pemprosesan imej untuk mengenal pasti penulis dokumen dapat digunakan memandangkan teknik ini dapat mengekstrak fitur imej tulisan yang menjadi input kepada proses klasifikasi dan akhirnya kepada mengenal pasti penulis. Kajian juga dilakukan ke atas teknik paleografi yang telah dibangunkan di dalam tulisan latin, Ibrani, India secara umum and tulisan arab. Beberapa contoh akan dinyatakan dan dikaji prosesnya.

2.0 Kepentingan Kajian Paleografi

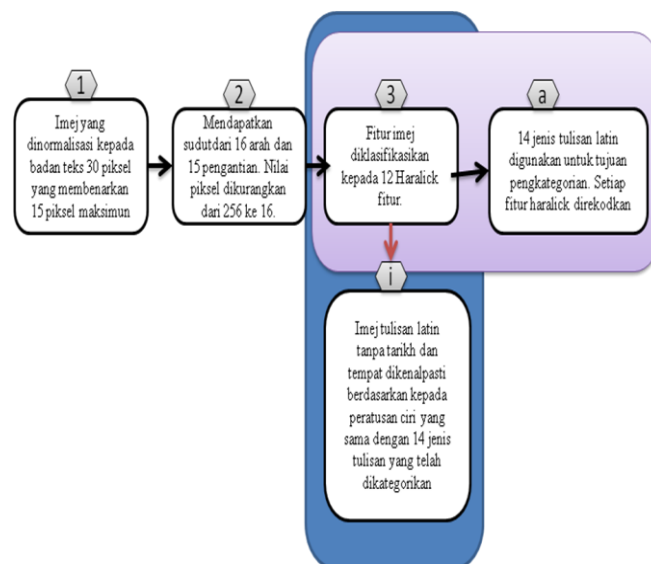
Kajian Paleografi tulisan tangan Jawi dapat memberikan maklumat berharga kepada generasi sekarang. Melalui kajian ini, maklumat seperti berikut dapat dikenal pasti:

- 1) ketulenan manuskrip dapat ditentukan,
- 2) mengesan sama ada terdapat lebih dari seorang penulis di dalam manuskrip,
- 3) mengenal pasti jenis tulisan yang digunakan (Joshi et al. 2007) dan (Witkam 2009),
- 4) mengenal pasti tempat dimana manuskrip itu ditulis berdasarkan kepada kodikologikal (ciri-ciri teknikal) dan juga latar belakang atau pun dengan lebih mudah dikenali sebagai rasam (Joshi et al. 2007) dan (Witkam 2009), dan akhir sekali
- 5) mengenal pasti tarikh ataupun julat tahun bila tulisan tangan ditulis berdasarkan kepada kodikologikal atribut dan gaya tulisan (Moalla et al. 2006).

Kelima-lima maklumat yang dinyatakan diperolehi dari Yosef et al. (2005) dan dilakukan penyesuaian kepada manuskrip tulisan jawi.

3.0 Kajian Paleografi secara digital oleh penyelidik-penyelidik di luar Gugusan Tanah Melayu

Kajian paleografi tulisan roman dilakukan oleh penyelidik Moalla et al. (2006). Beliau mengelaskan manuskrip kepada 14 jenis tulisan dan setiap tulisan menurutnya mewakili selang masa kegunaannya. Proses yang digunakan oleh beliau dapat digambarkan dalam Rajah 1 di bawah.



Rajah 1: Ringkasan Model (Moalla et al. 2006)

Moalla et al. (2006) telah menetapkan ciri imej input yang khusus. Beliau mengambil bahagian teks sahaja tanpa melibatkan keseluruhan muka surat. Beliau menggunakan Pergantungan Ruang Paras-Kelabu (PRPK). Dari PRPK Moalla et al. (2006) telah mendapatkan 12 fitur haralick yang digunakan untuk mengenalpasti jenis tulisan berdasarkan kepada jadual jenis tulisan melawan fitur. Hit bagi setiap fitur yang menyamai jenis tulisan direkod peratusannya. Fitur Haralick yang digunakan adalah ditunjukkan pada Jadual 1 di bawah:

Jadual 1: Fitur Haralick

	Jenis-jenis fitur Haralick
<i>f1</i>	Sudut momen kedua atau pun kehomogenan
<i>f2</i>	Elemen momen berbeza
<i>f3</i>	Korelasi
<i>f4</i>	varians (VAR) (Jumlah empat segi)
<i>f5</i>	Momen perbezaan songsang (IDM)
<i>f6</i>	Jumlah purata
<i>f7</i>	Jumlah varians
<i>f8</i>	Jumlah entropi
<i>f9</i>	Entropi
<i>f10</i>	Perbezaan varians
<i>f11</i>	Perbezaan entropi
<i>f12</i>	Pengiraan Entropi

Yosef et al. (2005) telah mengkaji tulisan hebrew (Ibrani) dan mengelaskan tulisan tersebut

kepada jenis kaligrafi yang wujud di dalam Ibrani. Beliau telah melakukan proses seperti digambarkan pada Rajah 2 di bawah:



Rajah 2: Methodologi klasifikasi jenis gaya ibrani oleh Yosef et al. (2005)

Penjelasan pada Rajah 1, Yosef et al. (2005) mengatakan bahawa kualiti tulisan manuskrip lama menyusut mengikut masa. Pernyataan ini turut disahkan oleh penyelidik Jawi yang mengkaji pembaikan kualiti imej jawi dengan kerosakan pada latar belakang (Sitti Rachmawati et al. 2010). Pada fasa pra pemrosesan, pengkaji (Yosef et al. 2005) telah menggunakan kaedah pemperduaan. Beliau telah menggunakan kaedah Multi-paras pengambangan. Di dalam kaedah ini, beliau melakukan pengambangan global, diikuti menyisihkan objek tidak berkaitan, diikuti oleh pemprosesan komponen tempatan dan langkah terakhir di dalam kaedah pemperduaan yang dicadangkan beliau ialah pasca-pemrosesan. Tujuan untuk setiap langkah di dalam pra pemrosesan dirumuskan di dalam Jadual 2 di bawah:

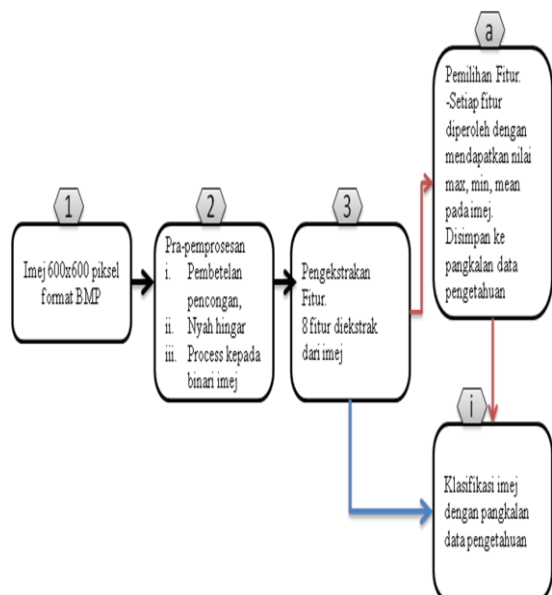
Jadual 2: Multi-Paras Pengambangan

Langkah	Tujuan
Pengambangan Global	1) mengecilkan ruang pencarian kepada latar depan calon 2) menghasilkan maklumat spatial.
Menyisihkan objek tidak berkaitan	1) Menyisihkan objek-objek yang tidak berkaitan seperti piksel asing yang bukan kepunyaan garis. 2) Melakukan proses Pelebelan Komponen Terkait dan pisahkan

		kelebaran komponen yang mungkin wujud.
	Pemrosesan imej tempatan	1) Mendapatkan threshold berdasarkan kepada paras kelabu latar depan imej. 2) Mendapat calon-calon piksel dari latar belakang yang menyambung kepada latar depan imej.
	Pasca-pemrosesan	1) Menutupi rongga-rongga kecil

Fitur yang digunakan oleh penyelidik ini menggunakan daerah latar belakang yang dominan yang diperoleh daripada komponen-komponen yang terhubung.

Kajian paleografi di dalam bidang pemrosesan imej juga dikaji oleh penyelidik india seperti Padma & Vijaya (2009), Beliau telah menggunakan pendekatan global di dalam mengenal pasti jenis tulisan yang terdapat pada dokumen. Walau pun, kajian yang digunakan tidak menyebut mengenai kajian paleografi, tetapi kajian yang dilakukan mengenal pasti jenis tulisan adalah sebahagian daripada kajian paleografi. Pada Rajah 3 di bawah menunjukkan model yang digunakan oleh Padma & Vijaya (2009):



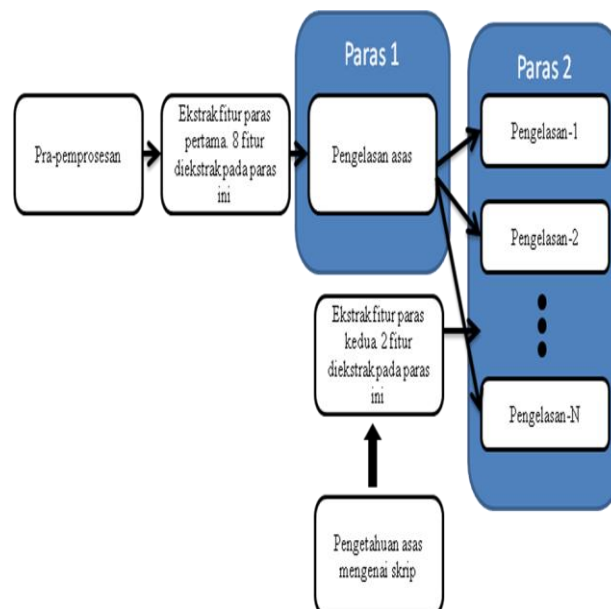
Rajah 3: Model Padma & Vijaya (2009).

Fitur yang digunakan oleh (Padma & Vijaya 2009) seperti ditunjukkan pada Jadual 3 di bawah.

Jadual 3: Fitur oleh Padma & Vijaya (2009)

	Fitur
1	Bil-baris-max-bawah
2	Garis-mendatar-atas
3	Calit-puncak
4	komponen-bawah
5	Saiz-paip-puncak
6	Saiz-paip-bawah
7	Ketumpatan-paip-atas
8	Ketumpatan-paip bawah

Selain dari Padma & Vijaya (2009), penulis ingin membincangkan kerangka kerja umum yang diperkenalkan oleh Joshi et al. (2007). Pengelasan oleh Joshi et al. (2007) telah dibahagikan kepada 3 jenis berasaskan kepada Penapis Log-Gabor Bank. Pengelasan pertama dinamakan sebagai fitur aras pertama, pengelasan kedua pula ialah fitur aras kedua dan yang terakhir ialah profil mengufuk. Gopal et al (2007) telah memperkenalkan model pada Rajah 4 di bawah:



Rajah 4: Model (Joshi et al. 2007).

Jadual 4 di bawah menunjukkan fitur aras pertama dan juga fitur aras kedua yang digunakan oleh penyelidik ini.

Jadual 4: Fitur dari Penapis Log-Gabor Bank

	Fitur paras pertama	Fitur paras kedua
1	Profil tenaga $E(\theta_i); i=1, \dots, 8$ untuk semua 10 skrip tulisan india.	Kadar tenaga dinormalkan $R = (i,j) = E(\theta_i)/E(\theta_j)$
2	relatif kekuatan dalam orientasi bersebelahan. $\Delta E_i; 1, \dots, 8$	

4.0 Perbandingan dan kesimpulan kaedah-kaedah penyelidik paleografi di dalam pemrosesan imej

Jadual 5 menunjukkan ringkasan kaedah yang digunakan oleh penyelidik di dalam kajian paleografi. Keempat-empat kajian pada domain tulisan yang berbeza dikaji. Perbandingan dapat dilihat pada jadual ini. Hanya satu kajian menggunakan pendekatan setempat manakala selebihnya menggunakan kaedah global.

Bahagian yang paling penting di dalam menentukan jenis tulisan dan juga pentarikan ialah fitur-fitur yang diekstrak dari manuskrip. Setiap fitur menyumbang kepada proses pengelasan dan juga mengenal pasti manuskrip berdasarkan kelas yang telah dibina berdasarkan fitur-fitur ini.

Jadual 5: Perbandingan kaedah penyelidikan Paleografi

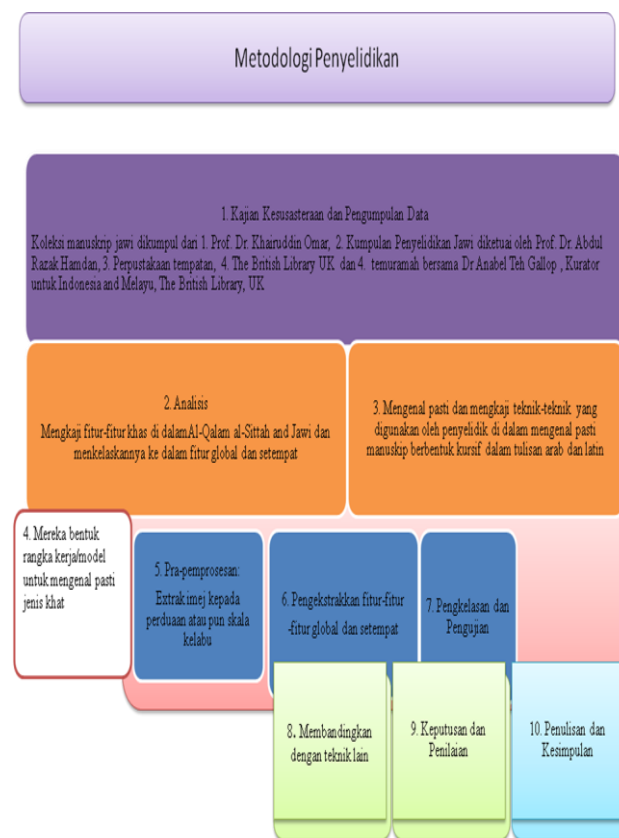
	Paleografi Latin (Moalla et al. 2006)	Paleografi Ibranni (Yosef et al. 2004)	Paleografi India (Padma & Vijaya 2009)	Paleografi India (Joshi et al. 2007)
Pendekatan	Global	Setempat	Global	Global
Fitur	12 fitur Haralick	Daerah latar belakang yang diperoleh dari komponen terhubung	8 fitur diekstrak dari baris tulisan	Filter diperoleh dari penapis Log-Gabor. Fitur dibahagikan kepada 2 paras. Paras pertama ialah profil kekuatan dan kedua ialah relatif kekuatan
Algoritma digunakan	Spatial Gray-Level Dependence (SGLD) Dan Hiralick	Multi-paras pengambangan	Tidak dinyatakan dengan jelas	Penapis Log-Gabor
Keputusan	Kategori imej berdasarkan kepada peratusan ciri-ciri 14 jenis tulisan latin terdapat di dalam imej manuskrip	Mengenalpasti huruf ibrani lamed dan aleph diperoleh secara manual.	Peratusan keberadaan 8 fitur di dalam imej manuskrip.	Plot graf tenaga dan diperoleh kadar kesalahan.

Jenis tulisan ditentukan berdasarkan kepada peratusan fitur-fitur dikenal pasti berada pada pengelasan yang telah dibina. Melalui jenis tulisan ini, maka dapat diketahui tarikh ataupun

sele masa manuskrip tersebut ditulis (Moalla et al. 2006)

5.0 Cadangan Kerangka Metodologi Kajian Paleografi Jawi

Kerangka Paleografi Jawi pada Rajah 5 di bawah dicadangkan oleh Mohd Sanusi Azmi (2010).



Rajah 5: Kerangka Kajian Paleografi Jawi Digital

6.0 Hala Tuju Kajian

Hala tuju kajian paleografi jawi di dalam pemrosesan imej dapat diperkembangkan kepada kajian yang lebih komprehensif memandangkan kepada beberapa faktor berikut.

- i. Penglibatan organisasi antarabangsa dan juga tempatan. Pada peringkat antarabangsa, terdapat organisasi yang secara aktif menjalankan

seminar, latihan, geran penyelidikan, biasiswa, penerbitan, dan laman web yang aktif mengemaskini aktiviti manuskrip islam. Organisasi terbabit ialah “*The Islamic Manuscript Association (TIMA)*”. Organisasi ini berpangkalan di Britain. Segala maklumat dan persidangan boleh diperoleh dari laman web TIMA iaitu <http://www.islamicmanuscript.org/>.

Selain dari TIMA, British Library UK dalam jabatan *South and Southeast Asia Section*, turut terlibat di dalam pengkajian manuskrip melayu. Terdapat banyak buku yang diterbitkan dan juga manuskrip yang disimpan di dalam British Library. Maklumat mengenai buku-buku yang berkaitan dengan manuskrip melayu dan juga Paleografi melayu bolehlah berkunjung ke laman sesawang ini

http://www.bl.uk/researchregister/1.9/?app_cd=RR&page_cd=PUBLICATION&researcher_id=20

Selain di Britain, terdapat sebuah universiti di Holland yang menjalankan penyelidikan dan penyimpanan manuskrip melayu. Terdapat juga kajian paleografi yang dijalankan oleh penyelidik di dalam universiti ini. Maklumat lanjut disertakan pada laman sesawang ini <http://www.library.leiden.edu/special-collections/oriental-collections/intro-se-asia.html>

Manakala pada peringkat tempatan, Perpustakaan Negara Malaysia (PNM) turut memainkan peranan dalam kajian Paleografi dengan jayanya

menerbitkan *Jurnal Filologi Melayu*. Sehingga kini PNM telah menerbitkan 16 penerbitan *Jurnal Filologi Melayu*. Maklumat lanjut bolehlah rujuk pada alamat web ini <http://www.pnm.my/index.php?id=42>

ii. Pakar-pakar paleografi jawi antarabangsa dan tempatan.

Dewasa ini, pakar paleografi jawi yang terkenal di peringkat antarabangsa dan tempatan adalah seperti pada jadual berikut. Penerbitan yang direkodkan di Jadual 6 di bawah hanyalah yang berkaitan khusus pada kajian paleografi.

Jadual 6: Pakar-pakar paleografi jawi antarabangsa dan tempatan

Nama Penyelidik	Lokasi	Penerbitan	Tahun	Penerbit/ ISBN
Anabel Teh Gallop	Britain	1.Beautifyin g Jawi: between calligraphy and palaeography , penerbit: Universiti Pendidikan Sultan Idris	2005	Universiti Pendidikan Sultan Idris, ISBN: 9834143125
		2. Ottoman influences in the seal of Sultan Alauddin Riayat Syah of Aceh (r.1589-1604)	2004	Indonesia and the Malay world, Vol. 32, No. 93, pp. 176-190
Jan Just Witkam	Universiti Leiden, Holland	The didactics of Palaeography	2009	TIMA’s Workshop on Codicology, Cambridge, UK
Wan Ali Wan Mat	Malaysia	Paleografi Jawi: Satu Pengenalan	2006	Jurnal Filologi Melayu, Jilid 14, Perpustakaan Negara Malaysia

iii. Paleografi dan pemprosesan imej.

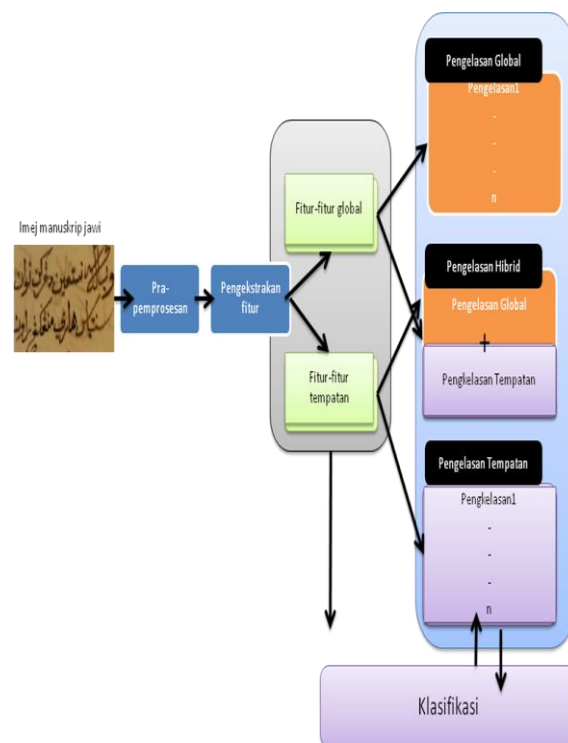
Pada bidang pemprosesan imej, input dari pakar paleografi adalah penting dalam menentukan fitur yang berguna di dalam proses pengekstrak fitur, pengelasan dan juga menentukan ketepatan pengecaman yang dilakukan secara digital. Jadi, penerbitan-penerbitan akademik di dalam pemprosesan imej juga penting di dalam membangunkan kerangka dan algoritma paleografi jawi. Jadual 7 bawah menunjukkan beberapa sumber penerbitan yang boleh dijadikan rujukan dan juga pengembangan kajian paleografi di dalam pemprosesan imej. Selain dari pada itu, penerbitan-penerbitan lain yang mungkin berguna bolehlah dirujuk pada rujukan.

Jadual 7: Penerbitan Rujukan untuk Paleografi digital

	Nama Penerbitan
i.	Using codebooks of fragmented connected-component contours in forensic and historic writer identification (Schomaker et al. 2007)
ii.	Text-independent writer identification and verification using textural and allographic features(Bulacu & Schomaker 2007)
iii.	Combining global and local features for writer identification(Siddiqi & Vincent 2008)
iv.	Establishing handwriting individuality using pattern recognition techniques (Srihari et al. 2001)
v.	Discriminatory power of handwritten words for writer recognition(Tomai et al. 2004)
vi.	Neural Networks and Support Vector Machines Classifiers for Writer Identification Using Arabic Script(Gazzah & Amara 2008)

iv. Kerangka kerja/ Model Mengenal pasti jenis khat di dalam manuskrip jawi.

Kerangka kerja jawi ini adalah berdasarkan analisis dan cerapan penyelidik yang telah dikaji. Penyelidik di dalam mengenal pasti penulis dan juga jenis tulisan ada yang menggunakan fitur-fitur global, tempatan dan ada yang menggabungkan global dan tempatan yang dinamakan sebagai pengelasan hibrid seperti yang ditunjukkan pada di bawah. Rajah menunjukkan kerangka umum yang boleh diguna pakai oleh penyelidik paleografi jawi di dalam membuat klasifikasi manuskrip. Terpujanglah pada mereka sama ada menggunakan fitur global, tempatan mahupun hibrid.



Rajah 6: Kerangka kerja umum Jawi digital

7.0 Kesimpulan

Kajian paleografi adalah satu bidang kajian yang menarik. Dari kajian ini, maklumat secara tersurat yang tidak terkandung pada teks dapat dirungkai. Jenis tulisan, cara menulis, hiasan pada manuskrip

dapat menentukan asal usul manuskrip, masa ditulis, bilangan penulis dan malah keaslian manuskrip dapat ditentukan.

Penyelidik di dalam bidang kajian paleografi pada domain tulisan latin dan arab, menggunakan fitur-fitur yang diekstrak dari manuskrip lama. Fitur-fitur boleh jadi setempat, global dan hibrid. Namun, pengekstrakan fitur-fitur pengkaji-pengkaji tersebut belum pasti kesesuaiannya ke atas tulisan jawi yang walaupun serupa dengan tulisan arab namun terdapat perbezaan ketara dalam penulisan, penyebutan malah tulisan jawi mempunyai tambahan aksara yang disesuaikan dengan sebutan bahasa Melayu.

Oleh itu, kajian secara komprehensif ke atas potensi fitur tulisan jawi sama ada global, setempat dan hibrid dapat menyumbang pada

kajian paleografi secara digital. Pada fasa pengelasan, fitur-fitur tersebut akan dikelaskan kepada kategori al-qalam al-sittah. Terdapat 6 kategori utama dan mungkin melalui kajian ke atas manuskrip lama, hanya beberapa jenis di dalam al-qalam as-sittah yang digunakan di dalam manuskrip melayu. Beberapa sampel imej-imej manuskrip jawi disertakan pada lampiran.

Kerangka yang dicadangkan adalah kerangka umum yang boleh digunakan oleh penyelidik paleografi secara digital.

Rujukan


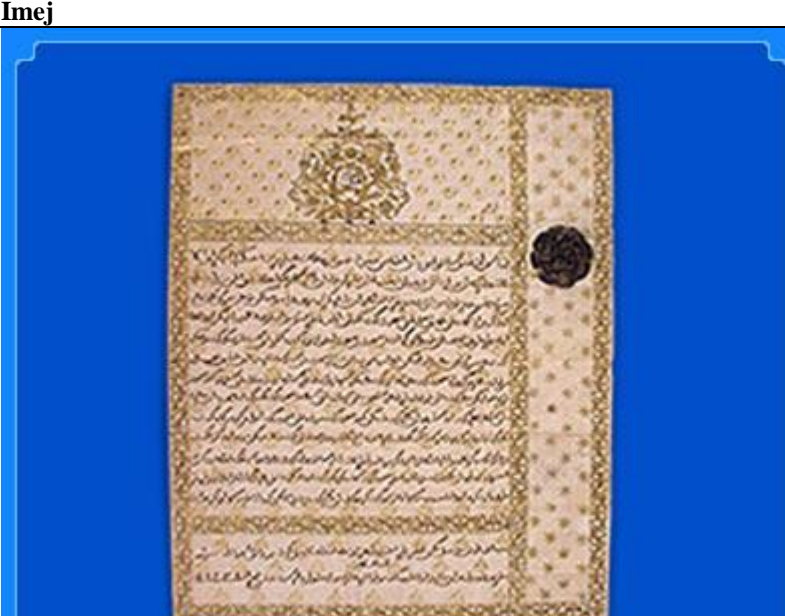
- Abandah, G. A. & Khedher, M. Z. 2009. Analysis of handwritten arabic letters using selected feature extraction techniques. *World*. 22 (1): 49–73.
- Abdi, M. N. & Khemakhem, M. 2010. Off-line text-independent arabic writer identification using contour-based features. *International Journal of Signal and Image Processing*. 1: pp 4-11.
- Abuhaiba, I. S. I., Mahmoud, S. A. & Green, R. J. 2002. Recognition of handwritten cursive arabic characters. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. 16 (6): 664–672.
- Al-Habian, G. & Assaleh, K. 2008. Online arabic handwriting recognition using continuous gaussian mixture hmms. *Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on*, pp. 1183–1186.
- Al-Ma'adeed, S., Al-Kurbi, A. A., Al-Muslih, A., Al-Qahtani, R. & Al Kubisi, H. 2008. Writer identification of arabic handwriting documents using grapheme features. *Computer Systems and Applications, 2008. AICCSA 2008. IEEE/ACS International Conference on*, pp. 923–924.
- Bischoff, B. 1990. Latin palaeography. Antiquity and the middle ages. Translated by dáibhí ó cróinín & david ganz. *Cambridge :University Press*.
- Bulacu, M. & Schomaker, L. 2007. Text-independent writer identification and verification using textural and allographic features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. pp: 701–717.
- Bulacu, M., Schomaker, L. & Brink, A. 2007. Text-independent writer identification and verification on offline arabic handwriting. *Document Analysis and Recognition, 2007. ICDAR 2007. Ninth International Conference on*, pp. 769–773.
- Derolez, A. 2008. The palaeography of gothic manuscript books. From the twelfth to the early sixteenth century. *Cambridge :University Press*.


- Gallop, A. T. 2005. *Beautifying jawi: Between calligraphy and palaeography*. Tanjung Malim: Universiti Pendidikan Sultan Idris.
- Gazzah, S. & Amara, N. B. 2008. Neural networks and support vector machines classifiers for writer identification using arabic script. *The International Arab Journal of Information Technology*. 5 (1): pp 93–102.
- Izadi, S. & Suen, C. Y. 2008. Online writer-independent character recognition using a novel relational context representation. *2008 Seventh International Conference on Machine Learning and Applications*, pp. 867-870
- Jajuli, M. 1994. *Teks undang-undang melayu pertengahan abad kelapan belas*. 1. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Joshi, G. D., Garg, S. & Sivaswamy, J. 2007. A generalised framework for script identification. *Int. J. Doc. Anal. Recognit.* 10 (2): 55-68.
- Khaled Mohammed Bin Abdl, S. Z. M. H. 2010. Swarm-based feature selection for handwriting identification. *Journal of Computer Science*. 6 ((1)): 80-86.
- Michelle P. Brown. 1994. Understanding illuminated manuscripts. A guide to technical terms. *Los Angeles: Getty Publications*.
- Mohd Hilmi Md Rifin, A. N. Z. 2007. Creating a digital library to handle malay manuscripts using greenstone. *International Conference on Libraries, Information and Society*, pp. 223-231.
- Mohd Sanusi Azmi, K. O., Azizi Abdullah. 2005. Perekayasaan histogram orientasi kecerunan mengesan erotan dan pencongan manuskrip merong mahawangsa. *Jurnal Teknologi Maklumat & Multimedia*. 2: 63-79.
- Padma, M. C. & Vijaya, P. A. 2009. Identification of Teḷugu, devanagari and english scripts using discriminating. *International Journal of Computer science & Information Technology (IJCSIT)*. Vol 1 (No 2): 64-78.
- Sanusi, M. 2003. Perekayasaan erotan dan pencongan histogram orientasi kecerunan manuskrip merong mahawangsa. Tesis Sarjana, Fakulti Teknologi dan Sains Maklumat Tesis, Universiti Kebangsaan Malaysia.
- Schomaker, L., Franke, K. & Bulacu, M. 2007. Using codebooks of fragmented connected-component contours in forensic and historic writer identification. *Pattern Recognition Letters*. 28 (6): 719–727.
- Shahabi, F. & Rahmati, M. 2006. Comparison of gabor-based features for writer identification of farsi/arabic handwriting. *Tenth International Workshop on Frontiers in Handwriting Recognition*.
<http://hal.inria.fr/docs/00/10/44/66/PDF/cr1086130064985.pdf> (18 Oktober 2010)
- Shahabi, F. & Rahmati, M. 2009. A new method for writer identification of handwritten farsi documents. *2009 10th International Conference on Document Analysis and Recognition*, pp. 426-430.
- Siddiqi, I. & Vincent, N. 2008. Combining global and local features for writer identification. *Proceedings of the 11. Int. Conference on Frontiers in Handwriting Recognition, Montreal*, 19-21 August 2008. http://opus.bsz-bw.de/fhhv/volltexte/2009/286/pdf/writer_recogniti_on_combining_local_global_methods.pdf (18 Oktober 2010)
- Sijpesteijn, P. 2008. ‘palaeography’, in kees versteegh (ed.). *Encyclopedia of Arabic language and Linguistics*. vol. 3 ((Leiden, E.J. Brill, 2008)): pp. 513-524.
- Sitti Rachmawati, S.N.H. Sheikh Abdullah, K. Omar, M.S. Zakaria, & C.Y. Liong, 2010. Review on Image Enhancement Methods of Old Manuscript with the Damaged Background. *International Journal on Electrical Engineering and Informatics*. 2(1): 1-14. ISSN 2085-6830.
- Srihari, S. N., Cha, S. H. & Lee, S. 2001. Establishing handwriting individuality using pattern recognition techniques. *Sixth International Conference on Document Analysis and Recognition (ICDAR'01)*, pp. 1195-1204.
- Tomai, C. I., Zhang, B. & Srihari, S. N. 2004. Discriminatory power of handwritten words for writer recognition. *Pattern Recognition*. 2: 638–641.
- Wan Ali Wan, M. 2006. Paleografi jawi : Satu pengenalan. *Jurnal Filologi Melayu*. Jilid 14: 65-69.


Yosef, I. B., Kedem, K., Dinstein, I. h., Beit-Arie, M. & Engel, E. 2004. Classification of hebrew calligraphic handwriting styles: Preliminary results.


Proceedings of the First International Workshop on Document Image Analysis for Libraries (DIAL'04), pp. 299 - 305.

Lampiran

Imej	Maklumat
 <p>MSS 2138 - Surat</p>	<p>Contoh surat diperoleh dari Perpustakaan Negara Malaysia http://www.pnm.my/manuskrip/melayu/02spenulisan/303bent_surat_03.htm</p>
 <p>MSS 2673 - Surat</p>	<p>Diperoleh dari Perpustakaan Negara Malaysia http://www.pnm.my/manuskrip/melayu/02spenulisan/303bent_surat_02.htm</p>

Imej	Maklumat
	<p>“Ini Syair Raja Tedung dengan Raja Katak” Tahun 1863. Manuskrip ini disimpan di University of Leiden, Holland. Pengkaji Manuskrip: Jan Just Witkam</p>

Imej	Maklumat
 <p data-bbox="427 1704 836 1733">MSS 427 - Hikayat Amir Hamzah</p>	<p>Manuskrip ini diperoleh dari Perpustakaan Negara Malaysia http://www.pnm.my/manuskrip/melayu/03koleksi/mss427.htm “Menceritakan kegagahan dan keberanian Amir Hamzah”</p>

Imej	Maklumat
 <p data-bbox="411 779 863 810">MSS 1714 - Hikayat Bayan Budiman</p>	<p data-bbox="1070 219 1390 309">Manuskrip ini diperoleh dari Perpustakaan Negara Malaysia</p> <p data-bbox="1070 315 1390 398">http://www.pnm.my/manuskrip/melayu/03koleksi/mss1714.htm</p> <p data-bbox="1070 495 1390 920">“Disalin pada 17 Syawal 1336H. [26 Julai 1918M.] Hikayat ini merupakan satu hikayat yang popular yang juga dikenali dengan <i>Hikayat Khoja Maimun</i>. Ia mengisahkan kebijaksanaan seekor burung bayan yang telah menyampaikan berbagai cerita dengan tujuan menghindar Bibi Zainab, isteri Khoja Maimun, daripada berlaku curang semasa ketiadaan suaminya.”</p>