

Metode Perbaikan Citra Naskah Kuno Menggunakan Histogram Terekualisasi Lokal dan Global

Baihaqi^{1*}, Yeni Yanti², Malahayati³

¹Program Studi Pendidikan Teknik Elektro, Fakultas Tarbiyah dan Keguruan Universitas Islam Negeri Ar-Raniry, Banda Aceh, Indonesia

²Program Studi Teknik Komputer, Fakultas Teknik, Universitas Serambi Mekkah, Banda Aceh, Indonesia

³Program Studi Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Islam Negeri Ar-Raniry, Banda Aceh, Indonesia

*Koresponden email: baihaqi.bukhari@ar-raniry.ac.id

Diterima: 3 Juli 2024

Disetujui: 20 Juli 2024

Abstract

Ancient manuscripts are precious relics of the past. These manuscripts are very susceptible to disturbances, including changes in the colour of the paper. For example, the paper may become dark and black (known as foxing), and there may be black spots due to exposure to water. These changes in paper colour can affect the comfort and legibility of the document. In order to preserve and protect these ancient manuscripts, any manuscript that shows signs of damage must be repaired. One effort to restore scanned ancient manuscripts involves the use of local and global image restoration methods based on histogram equalisation. In this research project, histogram equalisation is used to improve image quality by eliminating foxing and improving the contrast between text and background. Based on the Mean Opinion Score (MOS) values obtained from the tests, the histogram equalisation method is effective in eliminating foxing compared to the Otsu method.

Keywords: *threshold, histogram equalization, foxing, and ancient manuscript images, MOS (Mean Opinion Score)*

Abstrak

Manuskrip kuno merupakan barang peninggalan pada masa lalu yang sangat berharga nilainya. Manuskrip kuno sangat rentan terhadap derau termasuk perubahan warna kertas. Misalnya kertas menjadi gelap dan hitam (yang jika terjadi pada satu tempat disebut dengan fox) serta terdapat bercak-bercak hitam akibat terkena air. Perubahan warna kertas tersebut dapat mengganggu kenyamanan dan tingkat keterbacaan dokumen. Untuk melestarikan dan menjaga manuskrip kuno tersebut diperlukan perbaikan untuk tiap-tiap manuskrip yang mengandung derau. Salah satu usaha untuk memperbaiki manuskrip kuno hasil pemindaian adalah dengan menggunakan metode restorasi citra berbasis histogram terekualisasi secara lokal dan global. Pada penelitian ini, metode histogram terekualisasi digunakan untuk meningkatkan kualitas citra dengan menghilangkan fox dan meningkatkan kekontrasan antara teks dengan latar belakangnya. Berdasarkan nilai uji MOS yang diperoleh, metode histogram terekualisasi efektif menghilangkan fox jika dibandingkan dengan metode Otsu.

Kata Kunci: *threshold, ekualisasi histogram, fox, dan citra manuskrip kuno, MOS (Mean Opinion Score)*

1. Pendahuluan

Pengolahan citra (image processing) adalah sebuah sistem yang mengolah citra sebagai masukan dan menghasilkan citra sebagai keluaran [1]. Awalnya, pengolahan citra dilakukan untuk meningkatkan kualitas citra. Namun, seiring dengan kemajuan di bidang komputasi, yang ditandai dengan peningkatan kapasitas dan kecepatan pemrosesan komputer [2]. Munculnya berbagai ilmu komputasi, pengolahan citra kini memungkinkan setiap orang untuk mengekstraksi informasi dari suatu citra [3]. Maka dilakukan usaha untuk merestorasi dokumen-dokumen kuno secara digital, sehingga memudahkan setiap orang untuk dapat mengakses dan mengambil informasi serta penelusuran teks juga lebih mudah [4]. Restorasi citra dapat dilakukan juga dengan menggunakan pendekatan virtual untuk memfasilitasi pendekatan manual. Penelitian ini menghasilkan citra yang lebih baik dengan pendekatan secara otomatis [5].

Salah satu teknik pengolahan citra dapat digunakan untuk merestorasi citra dokumen kuno [6]. Citra dokumen kuno telah tersimpan dalam kurun waktu yang sangat lama, sehingga dapat menyebabkan banyak gangguan [7]. Misalnya kertas dokumen tersebut terkena air dan lainnya, warnanya berubah dan tulisannya menjadi agak gelap dan kusam [8]. Dokumen yang mengalami perubahan warna yang hanya pada satu

bagian saja dari keseluruhan dokumen disebut dengan fox. Dokumen yang mengandung fox tersebut harus diolah dengan proses restorasi digital, karena restorasi biasa tidak mampu memperbaiki dokumen kuno menjadi lebih bagus [9]. Penelitian terkait dengan dokumen kuno dilakukan oleh [10] yang menggunakan metode scalling dan Generative Adversarial Network menghasilkan citra keluaran yang bersih dan dapat terbaca kembali.

Restorasi digital dapat dilakukan dengan berbagai macam teknik pengolahan citra yang bertujuan untuk memperbaiki citra dokumen kuno tersebut [11]. Banyak teknik telah digunakan untuk merestorasi manuskrip kuno seperti dengan menggunakan rational filtering untuk menghilangkan 'fox' dan menggunakan manuskrip kuno dengan tulisan latin[12]. Dengan menggunakan estimasi berbasis maximum likelihood dan k-means clustering untuk memisahkan teks dan latar dengan manuskrip uji bertuliskan huruf Arab [13]. Ada juga dengan memanfaatkan histogram sebagai penentu threshold antara teks dan latar. Threshold ditentukan melalui proses iterasi [14]. Metode histogram terekualisasi digunakan untuk menentukan nilai ambang untuk bagian dokumen yang mengandung fox saja, sedangkan bagian yang lain belum dievaluasi. Teknik ini lebih sederhana karena tidak memerlukan proses iterasi ataupun perhitungan statistik seperti maximum likelihood yang memerlukan banyak perhitungan [15].

Penelitian ini membahas penggunaan metode restorasi citra berbasis histogram terekualisasi untuk pengolahan citra manuskrip kuno. Teknik ini digunakan untuk keseluruhan dokumen. Berdasarkan nilai uji MOS yang diperoleh, metode histogram terekualisasi efektif menghilangkan fox jika dibandingkan dengan metode Otsu.

2. Metode Penelitian

Tempat dan Waktu Penelitian

Penelitian ini dimulai dengan tahapan pertama yang dilakukan pada penelitian ini adalah mencari sumber referensi yang berkaitan dengan pengolahan citra digital dan mempelajarinya, baik yang terdapat dalam buku-buku dan yang terdapat di Internet untuk menunjang proses dalam penyelesaian penelitian ini. Penelitian ini dilakukan di laboratorium Grup Penelitian Multimedia dan Pengolahan Sinyal, bagian dari jurusan Teknik Elektro dan Komputer di Universitas Syiah Kuala, Banda Aceh. Objek penelitian adalah dokumen jawi kuno seperti yang terlihat pada **Gambar 1** berikut. Dalam simulasi ini pengujian citra dilakukan dengan dua cara yaitu secara utuh (global) dan secara per bagian (lokal).



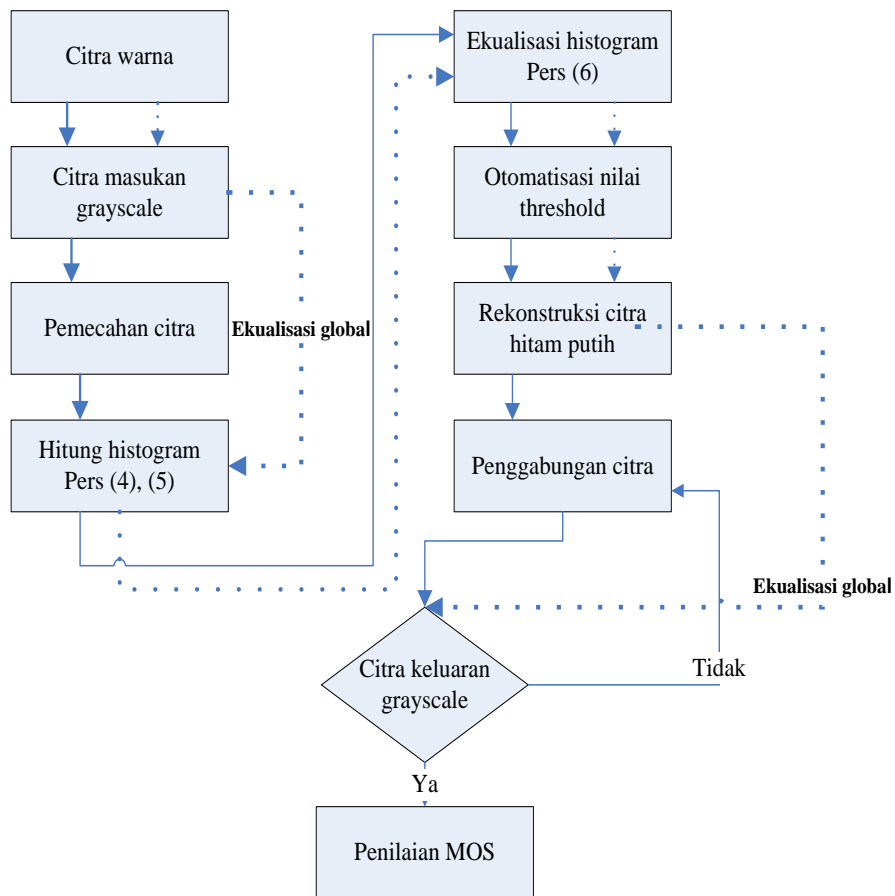
Gambar 1. Naskah dokumen kuno digital



Gambar 2. Naskah dokumen kuno grayscale

3. Hasil dan Pembahasan Rancangan Penelitian

Penelitian ini diawali dengan tahap pengumpulan citra dokumen Jawi kuno. Ada beberapa tahapan dalam melakukan proses pengolahan citra seperti ditunjukkan dalam diagram alir pada Gambar 3:



Gambar 3. Diagram Alir Pengolahan Citra Dokumen Kuno

Dalam simulasi ini pengujian citra dilakukan dengan dua cara yaitu secara utuh (global) dan secara per bagian (lokal). Tanda panah putus-putus dalam diagram alir pada Gambar 3 menunjukkan proses

pengujian secara global, sedangkan tanda panah yang tidak putus-putus menunjukkan proses pengujian secara lokal.



Gambar 4. Citra Dokumen Yang Telah Dilakukan Pemecahan



Gambar 5. Citra Hasil Ekualisasi Lokal

Pengujian Citra Secara Utuh (Ekualisasi Global)

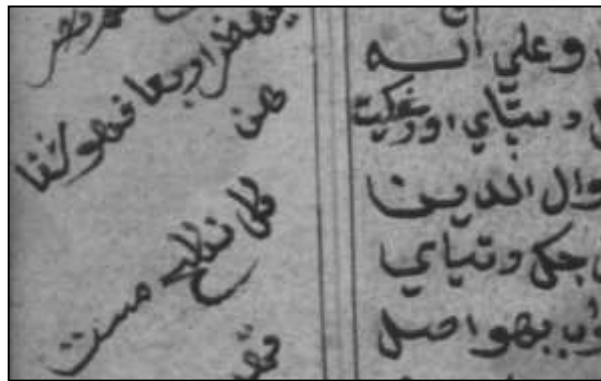
Ekualisasi global merupakan proses pengujian citra secara utuh tanpa melakukan pemecahan terhadap citra terlebih dahulu. Sehingga citra yang dihasilkan juga secara utuh tanpa perlu melakukan penggabungan.

Pengujian Citra Per Bagian (Ekualisasi Lokal)

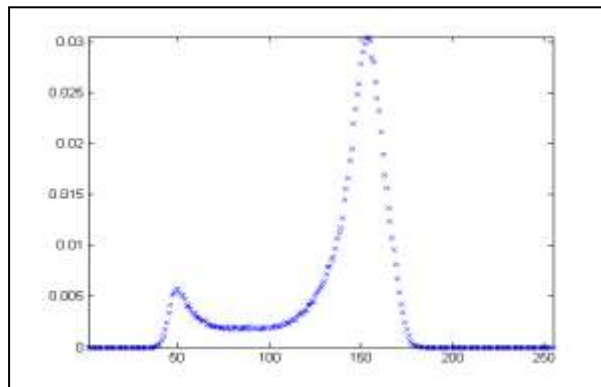
Ekualisasi lokal merupakan proses pengujian citra secara satu per satu bagian citra yang telah dipotong-potong seperti ditunjukkan pada **Gambar 5**. Untuk mendapatkan citra keluaran secara utuh maka perlu dilakukan penggabungan terhadap citra yang dihasilkan. Gambar 5 merupakan hasil dari pemecahan citra grayscale. Selanjutnya nilai ambang diotomatisasi dan citra hitam putih direkonstruksi. Hasil keluaran citra tersebut digabungkan kembali sehingga menjadi utuh.

Citra Hasil Ekualisasi Lokal

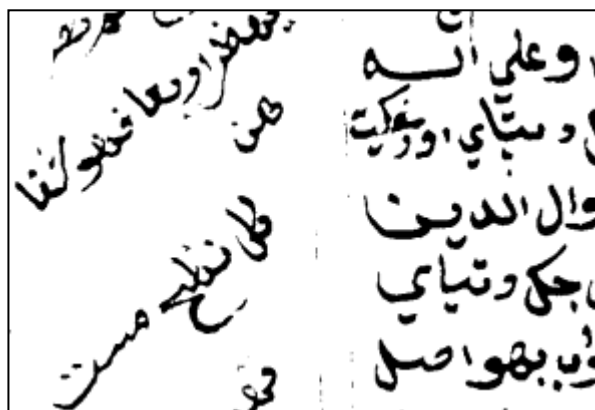
Dari semua potongan citra yang telah diuji terdapat beberapa potongan citra yang memiliki keadaan atau kondisi yang berbeda-beda antara lain:



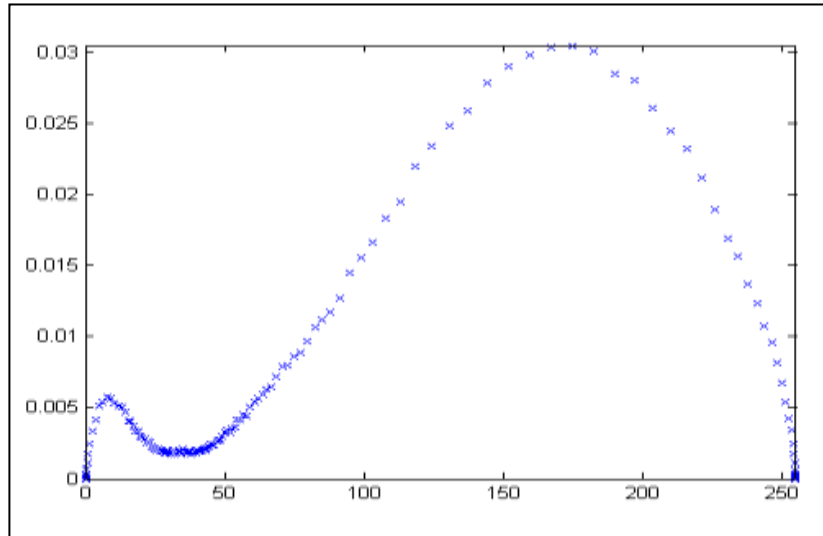
(a)



(b)



(c)

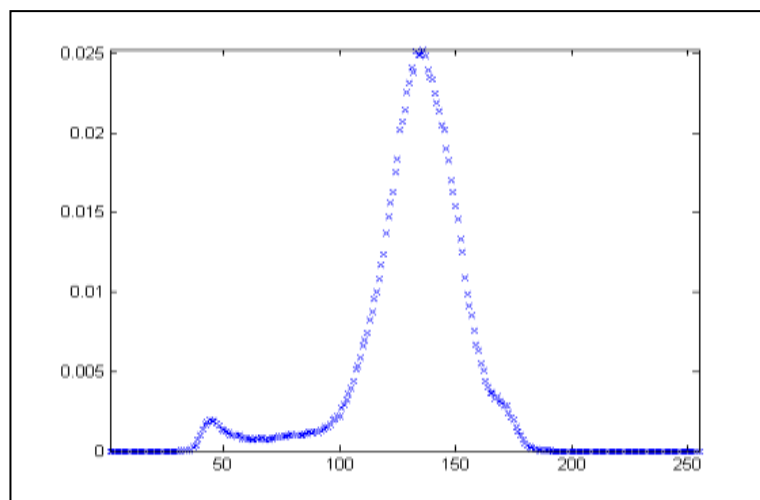


(d)

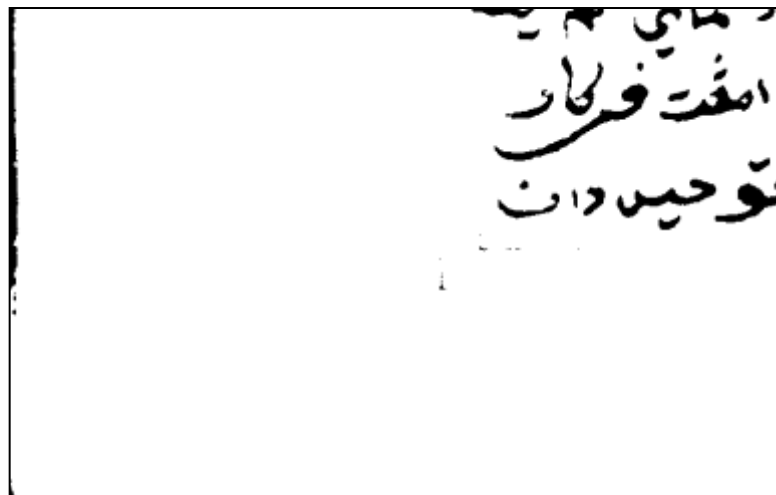
Gambar 6. (a) Citra *Grayscale*, (b) Citra Setelah Ekualisasi, (c) Histogram Ternormalisasi, dan (d) Histogram Ekualisasi



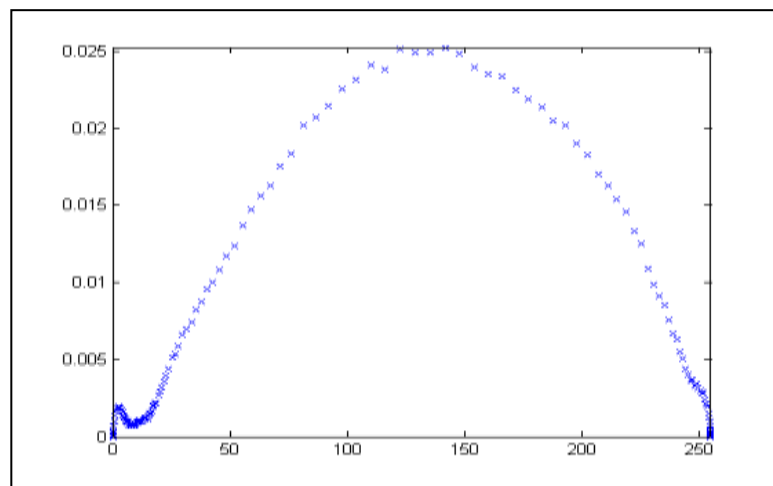
(a)



(b)



(c)



(d)

Gambar 7. (a) Citra *Grayscale*, (b) Citra Setelah Ekualisasi, (c) Histogram Ternormalisasi, dan (d) Histogram Ekualisasi

Penjelasan Gambar

m3_c2 (potongan citra pada baris ke 2 dan kolom 1)

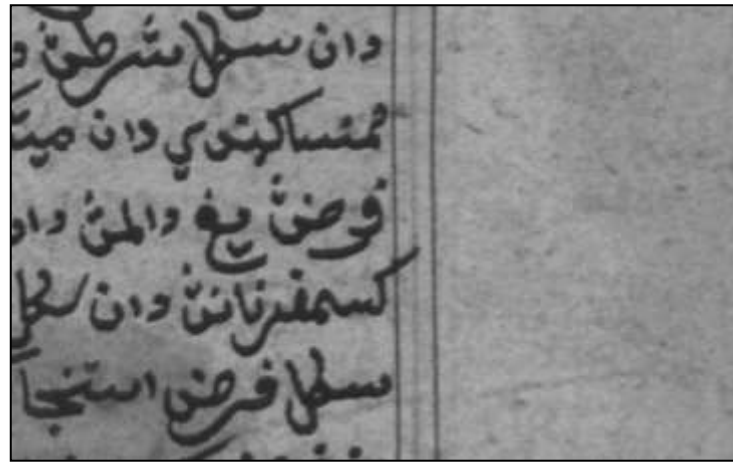
Gambar 6(a) merupakan salah satu potongan citra yang tidak mengandung fox. **Gambar 6(b)** merupakan citra setelah proses ekualisasi, sedangkan gambar 6(c) dan (d) masing-masing menunjukkan histogram ternormalisasi dan histogram ekualisasi. Nilai ambang minimum yang didapat setelah ekualisasi antara range (0-255) adalah 28.

m3_c4 (potongan citra pada baris ke 4 dan kolom 1)

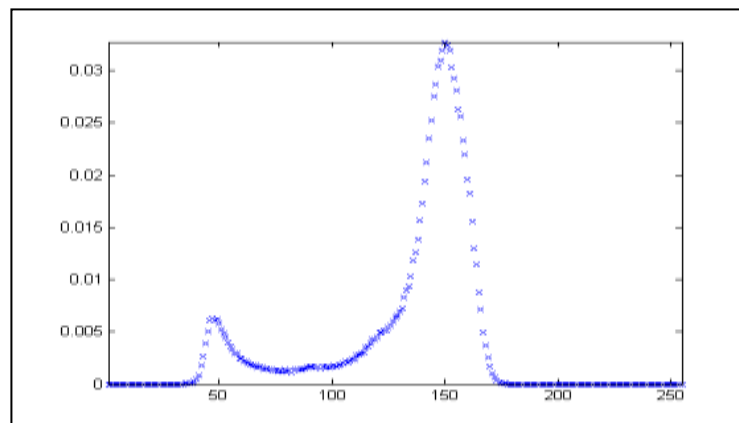
Gambar 7(a) merupakan salah satu potongan citra yang mengandung noise pada sudut dokumen terlihat berwarna putih. **Gambar 7(b)** merupakan citra setelah proses ekualisasi, sedangkan **Gambar 7(c)** dan (d) masing-masing menunjukkan histogram ternormalisasi dan histogram ekualisasi. Nilai ambang minimum yang didapat setelah diekualisasi antara range (0-255) adalah 7. **Gambar 7(a)** menunjukkan bahwa derau yang ada pada sudut dokumen telah berhasil dihilangkan setelah proses ekualisasi.

m3_c14 (potongan citra pada baris ke 2 dan kolom 4)

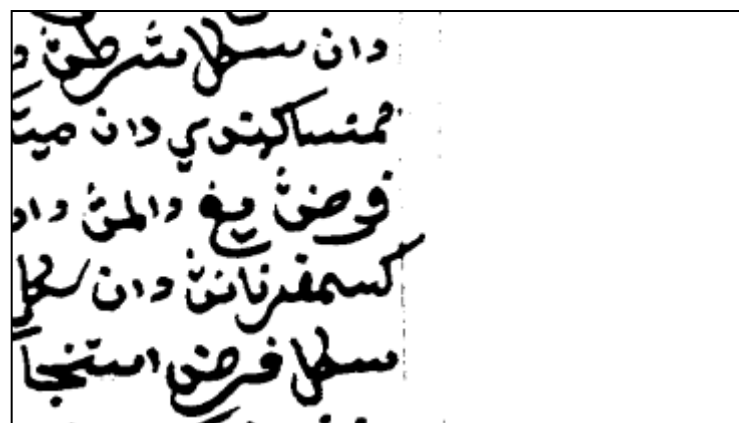
Gambar 8(a) merupakan salah satu potongan citra yang mengandung derau yang berupa fox. **Gambar 8(b)** adalah citra setelah mengalami proses ekualisasi sedangkan **Gambar 8(c)** dan (d) masing-masing menunjukkan histogram ternormalisasi dan histogram terekualisasi. Nilai ambang minimum yang didapat setelah ekualisasi antara range (0-255) adalah 29. **Gambar 8(a)** menunjukkan bahwa fox yang ada pada dokumen telah berhasil dihilangkan setelah proses ekualisasi. Selanjutnya citra digabungkan menjadi utuh kembali seperti terlihat pada **Gambar 11**.



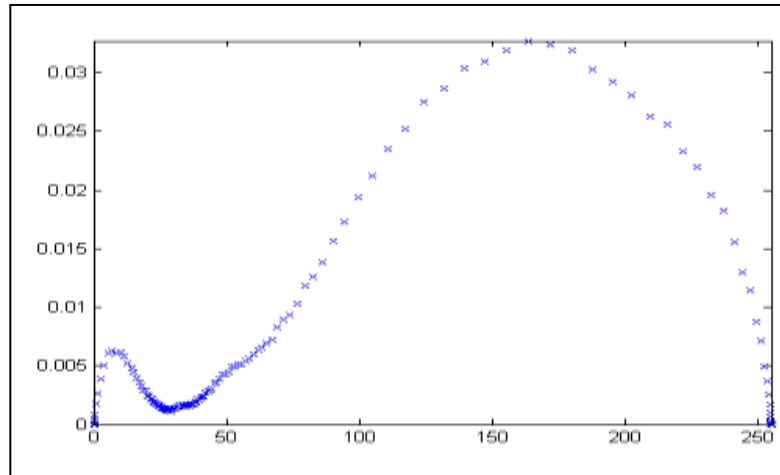
(a)



(b)



(c)



(d)

Gambar 8. (a) Citra Grayscale, (b) Citra Setelah Ekualisasi, (c) Histogram Ternormalisasi, dan (d) Histogram Ekualisasi.



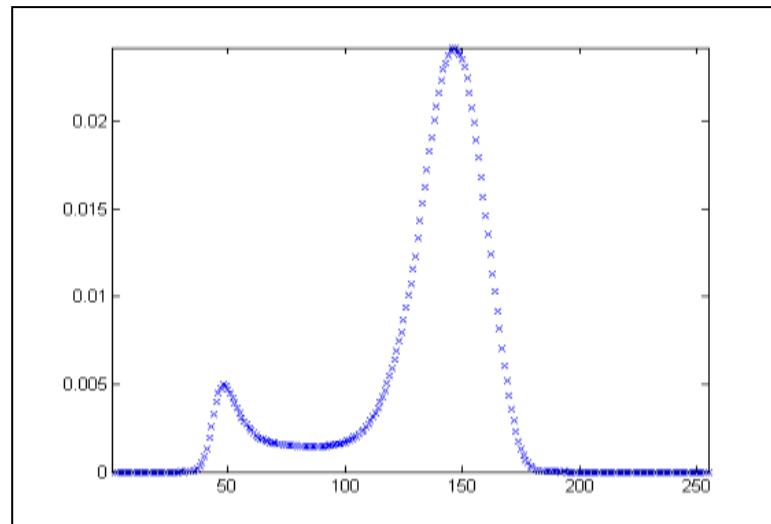
Gambar 9. Citra hasil gabungan ekualisasi lokal

Citra Ekualisasi Global

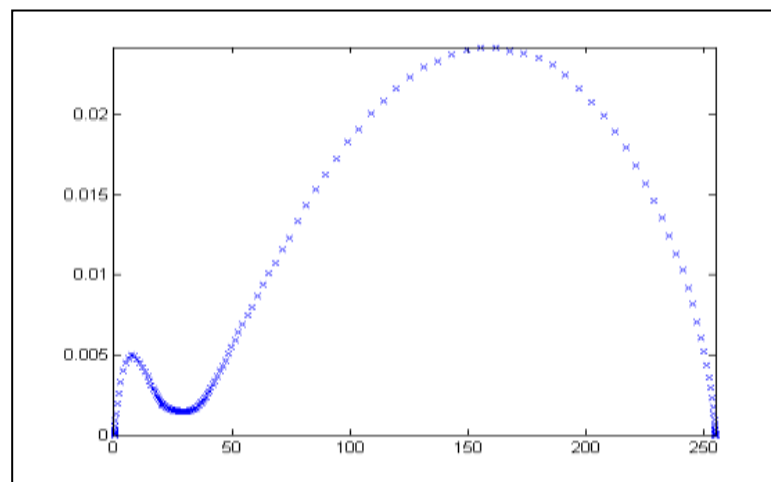
Proses pengolahan citra dengan histogram ekualisasi global merupakan perbaikan citra secara langsung untuk keseluruhan citra. Dengan proses ini citra yang dihasilkan memiliki kualitas yang hampir sama dengan proses ekualisasi lokal. Seperti terlihat pada Gambar 10(a).



(a)



(b)



(c)

Gambar 10. (a) Citra Setelah Ekualisasi, (b) Normalisasi Histogram, dan (c) Histogram Terekualisasi

Gambar 10(a) menunjukkan bahwa citra keluaran menjadi lebih bersih. Namun masih terdapat bintik-bintik hitam pada citra. Nilai ambang minimum antara range (0-255) adalah 27.



Gambar 11. Citra Hasil Otsu Lokal



Gambar 12. Citra Hasil Gabungan Dengan Otsu Lokal



Gambar 13. Citra Hasil Dengan Menggunakan Otsu Global

Citra Hasil Otsu Lokal

Gambar 11 merupakan potongan-potongan citra yang diperbaiki dengan menggunakan otsu. Selanjutnya citra tersebut akan digabungkan kembali menjadi utuh seperti terlihat pada Gambar 12.

Citra Hasil Otsu Global

Dengan menggunakan Otsu untuk memperbaiki keseluruhan citra maka didapatkan hasil yang memiliki kualitas yang hampir sama dengan menggunakan Otsu lokal seperti terlihat pada Gambar 13. Kualitas citra yang dihasilkan tidak begitu bagus karena terdapat bintik-bintik hitam dan belum bisa menghilangkan fox secara keseluruhan. Hurufnya tidak begitu halus namun masih bisa terbaca.

Pembahasan

Perbandingan Hasil Antara Ekualisasi Lokal dengan Ekualisasi Global

Dengan menggunakan ekualisasi lokal citra yang dihasilkan memiliki kualitas yang sedikit lebih bagus jika dibandingkan dengan citra dengan menggunakan ekualisasi global. Citra hasil ekualisasi lokal memiliki tingkat keterbacaan yang tinggi karena tulisannya menjadi lebih jelas dan lebih halus dibandingkan dengan citra aslinya. Seperti terlihat pada Gambar 9 citra yang dihasilkan memiliki foreground (latar depan) yang bagus dan background yang jelas.

Sedangkan citra hasil ekualisasi global (secara keseluruhan) memiliki kualitas yang hampir sama dengan hasil ekualisasi lokal. Seperti terlihat pada Gambar 10(a) masih terdapat bintik-bintik pada citra yang dihasilkan. Citra hasil ekualisasi lokal memiliki kualitas yang lebih bagus jika dibandingkan dengan citra hasil ekualisasi global.

Perbandingan Hasil Antara Otsu Lokal dengan Otsu Global

Citra hasil otsu lokal memiliki kualitas yang hampir sama dengan hasil otsu global, citra hasil otsu lokal hurufnya terlihat lebih hitam dan terdapat bintik-bintik hitam yang dapat mengganggu tingkat keterbacaan terhadap dokumen tersebut. Hal ini disebabkan oleh nilai batas ambang yang besar, sehingga banyak dari jumlah piksel yang dirubah menjadi hitam. Seperti terlihat pada Gambar 13 latar belakang pada sudut dokumen menjadi hitam dan terdapat bintik-bintik hitam diantara huruf-hurufnya.

Operasi otsu global memiliki hasil yang lebih bagus jika dibandingkan dengan otsu lokal karena tidak banyak terdapat bintik-bintik hitam. Seperti pada Gambar 13, latar belakang pada sudut dokumen tidak terlalu hitam. Namun masih terdapat bintik-bintik hitam yang menjadikan dokumen tersebut tidak terlihat indah walaupun tidak mengganggu tingkat keterbacaan dokumen. Jadi, citra hasil otsu global memiliki kualitas yang lebih baik jika dibandingkan dengan citra hasil otsu lokal.

Perbandingan Untuk Keempat Citra Yang Telah Dihadirkan

Dari keempat citra yang telah dihasilkan pada penelitian ini, yang memiliki kualitas yang bagus diantara keempat citra yang diperoleh adalah citra hasil ekualisasi lokal (**Gambar 11**) dan citra hasil ekualisasi *global* (**Gambar 10(a)**). Kemudian citra hasil Otsu *global* (**Gambar 13**) lebih bagus jika dibandingkan dengan Otsu lokal (**Gambar 11**).

Pada keempat gambar di atas terlihat perbedaan hasil karena citra yang dihasilkan dengan menggunakan otsu tulisannya lebih gelap dan terdapat bintik-bintik hitam, sedangkan dengan menggunakan histogram terekualisasi citra yang dihasilkan memiliki kualitas yang lebih baik. Sehingga tujuan dari penelitian ini berhasil memperbaiki kualitas citra dokumen kuno dengan metode histogram terekualisasi, serta penggabungan tiap-tiap potongan citra juga berhasil dilakukan.

Penilaian Mean Opinion Score (MOS)

Mean Opinion Score merupakan penilaian yang sifatnya subjektif didasarkan pada pengamatan mata manusia sebagai responden. Penilaiannya dengan menggunakan standar angka 1 sampai dengan 5. Angka 5 menunjukkan penilaian yang sangat baik sedangkan angka 1 menunjukkan kualitas yang sangat buruk. Untuk lebih jelasnya dapat dilihat pada lampiran MOS **Tabel 1** berikut sampel rata-rata yang dilakukan dengan MOS.

Tabel 1. Hasil Penilaian MOS

No	Lembar citra	Ekualisasi lokal	Ekualisasi global	Otsu lokal	Otsu global
1.	m1_way	4,10	2,65	2,90	1,90
2.	m2_way	4,05	3,95	2,90	2,50
3.	m3_way	4,10	4,05	3,05	2,95
4.	m4_way	3,30	3,45	3,10	3,00
	Rata-rata	3,8875	3,5250	2,9875	2,5875

Berdasarkan hasil MOS pada **Tabel 1** untuk masing-masing lembaran citra yaitu m1_way, m2_way, m3_way, dan lembaran yang keempat m4_way didapat hasil penilaian yang berbeda-beda untuk masing-masing citra tersebut. Nilai MOS yang paling tinggi diperoleh dengan menggunakan metode histogram terekualisasi secara lokal yaitu 4,10. Sedangkan dengan menggunakan histogram terekualisasi secara global adalah 4,05 untuk nilai MOS tertinggi. Dengan menggunakan metode otsu secara lokal didapat nilai MOS tertinggi adalah 3,10.

Sedangkan secara global didapat nilai tertinggi 3,00. Untuk keseluruhan citra didapat nilai MOS tertinggi adalah dengan menggunakan metode histogram terekualisasi secara lokal yaitu 4,10 sedikit berbeda dengan menggunakan metode histogram terekualisasi secara global yang memiliki nilai 4,05. Sedangkan nilai paling rendah didapat dengan menggunakan metode otsu secara global yaitu 1,90. Hal ini disebabkan oleh kualitas citra yang dihasilkan terlalu gelap sehingga dapat mengurangi tingkat keterbacaan terhadap citra dokumen tersebut.

Dari hasil MOS pada **Tabel 1** dapat disimpulkan juga bahwa metode histogram ekualisasi secara lokal memiliki nilai rata-rata yang lebih tinggi yaitu 3,8875 jika dibandingkan dengan ekualisasi secara global yaitu 3,5250 dan metode otsu masing-masing secara lokal dan global adalah 2,9875 dan 2,5875.

4. Kesimpulan

Setelah melakukan penelitian dengan menggunakan metode histogram terekualisasi untuk restorasi citra manuskrip kuno, maka dapat disimpulkan bahwa nilai ambang yang digunakan untuk meningkatkan kualitas citra dokumen kuno adalah nilai ambang minimum. Metode histogram terekualisasi lebih efektif untuk menghilangkan *fox* pada citra jika dibandingkan dengan metode otsu. Berdasarkan hasil uji MOS metode histogram terekualisasi secara lokal efektif menghilangkan *fox* jika dibandingkan dengan metode Otsu. Dari keempat citra yang dihasilkan, yang memiliki kualitas paling bagus adalah citra hasil ekualisasi lokal. Kualitas paling rendah diperoleh pada citra hasil Otsu secara *global*.

5. Referensi

- [1] I. Pratikakis, K. Zagoris, G. Barlas, and B. Gatos, "ICFHR 2016 handwritten document image binarization contest (H-DIBCO 2016)," *Proc. Int. Conf. Front. Handwrit. Recognition, ICFHR*, vol. 0, pp. 619–623, 2016, doi: 10.1109/ICFHR.2016.0118.
- [2] B. Su, S. Lu, and C. L. Tan, "Robust document image binarization technique for degraded document

- images,” *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1408–1417, 2013, doi: 10.1109/TIP.2012.2231089.
- [3] H. Lombok *et al.*, “Denoising Dan Binarization Untuk Pengolahan Citra,” pp. 28–29, 2016.
- [4] A. T. Gallop, “a Jawi Sourcebook for the Study of Malay Palaeography and Orthography,” *Indones. Malay World*, vol. 43, no. 125, pp. 13–171, 2015, doi: 10.1080/13639811.2015.1008253.
- [5] A. Tonazzini, P. Savino, E. Salerno, M. Hanif, and F. Debole, “Virtual restoration and content analysis of ancient degraded manuscripts,” *iJIST Int. J. Inf. Sci. Technol.*, vol. 3, no. 5, pp. 16–25, 2019.
- [6] O. Sihombing, E. Buulolo, H. K. Siburian, G. Batak, and M. O. Morfologis, “Hasil Segmentasi Citra Digital Gorga Batak,” *KOMIK (Konferensi Nas. Teknol. Inf. dan Komputer)*, vol. 2, pp. 40–48, 2018.
- [7] A. Baldominos, Y. Saez, and P. Isasi, “A survey of handwritten character recognition with MNIST and EMNIST,” *Appl. Sci.*, vol. 9, no. 15, 2019, doi: 10.3390/app9153169.
- [8] S. N. H. S. Abdullah, S. M. Ismail, M. K. Hasan, and P. Shivakumara, “Novel Adaptive Binarization Method for Degraded Document Images,” *Comput. Mater. Contin.*, vol. 67, no. 3, pp. 3815–3832, 2021, doi: 10.32604/cmc.2021.014610.
- [9] G. Lazzara and T. Géraud, “Efficient multiscale Sauvola’s binarization,” *Int. J. Doc. Anal. Recognit.*, vol. 17, no. 2, pp. 105–123, 2014, doi: 10.1007/s10032-013-0209-0.
- [10] S. U. Khan, I. Ullah, F. Khan, Y. Lee, and S. Ullah, “Historical Text Image Enhancement Using Image Scaling and Generative Adversarial Networks,” *Sensors*, vol. 23, no. 8, 2023, doi: 10.3390/s23084003.
- [11] F. Jia, C. Shi, K. He, C. Wang, and B. Xiao, “Degraded document image binarization using structural symmetry of strokes,” *Pattern Recognit.*, vol. 74, pp. 225–240, 2018, doi: 10.1016/j.patcog.2017.09.032.
- [12] U. Rani, A. Kaur, and G. Josan, “A new binarization method for degraded document images,” *Int. J. Inf. Technol.*, vol. 15, no. 2, pp. 1035–1053, 2023, doi: 10.1007/s41870-019-00361-3.
- [13] W. Boussellaa, A. Zahour, H. Elabed, A. Benabdelhafid, and A. Alimi, “Unsupervised block covering analysis for text-line segmentation of Arabic ancient handwritten document images,” *Proc. - Int. Conf. Pattern Recognit.*, no. August, pp. 1929–1932, 2010, doi: 10.1109/ICPR.2010.475.
- [14] Y. Chang, “Improving the Otsu method for MRA image vessel extraction via resampling and ensemble learning,” *Healthc. Technol. Lett.*, vol. 6, no. 4, pp. 115–120, 2019, doi: 10.1049/htl.2018.5031.
- [15] O. Patel, Y. P. S. Maravi, and S. Sharma, “A Comparative Study of Histogram Equalization Based Image Enhancement Techniques for Brightness Preservation and Contrast Enhancement,” *Signal Image Process. An Int. J.*, vol. 4, no. 5, pp. 11–25, 2013, doi: 10.5121/sipij.2013.4502.