applied sciences

MDPI

Supplementary Materials

1. Using traditional IQA methods to find clearest fecal microscopic image

In order to further prove the effectiveness of GMANet, in this part, we demonstrated the performance of 37 types of traditional image quality assessment (IQA) methods and local maximum gradient (LMG) method (proposed in Section 3.1 of our work) on finding clearest human fecal microscopic image in autofocus process. These methods contain 13 types of FR-IQA methods and 24 types of NR-IQA methods.

For each IQA method, we calculated all response values on feces dataset that contains 1036 groups of fecal microscopic images and we assumed that the image with maximum (or minimum, depending on the specific algorithm) response value is the clearest image in each image group. To simplify the comparison, we only compared the performance metric of prediction accuracy for each IQA method. The definitions of t_0, t_1 and "acc" are the same as those in Section 4.2.2 of our work. These traditional IQA methods do not need to divide dataset into training, validation and test set, so all algorithm results are for the whole feces dataset.

1.1 FR-IQA methods

FR-IQA methods require reference images as standards. Since there are no reference images for fecal microscopic images, we used the assumption in [1], that is, the clearer the image is, the greater the difference between its Gaussian blurred image and the original image is. Gaussian blur operation with kernel size 11 and sigma value 1.5 was performed on feces dataset, and then original images and corresponding Gaussian blurred images are used as reference images and distorted images, respectively. After above processing, we could use FR-IQA methods to calculate the difference between them and choose the image with largest difference as the clearest image in each image group. The performance of 13 types of FR-IQA methods is shown in Table 1. In SSIM-down-sampling [4] (SSIM-DS) algorithm, images were scaled down to 256×256 pixel size. All other parameters used the data from original literature. Among them, the algorithm with best prediction accuracy value has been marked in bold.

Table S1. The performance of 13 types of FR-IQA methods.

FR-IQA method	t_0	t_1	acc
PSNR	487	335	79.34%
WSNR[2]	497	330	79.83%
SSIM[3]	7	87	9.07%
MS-SSIM[4]	139	287	41.12%
SSIM-DS[4]	258	454	68.73%
IWSSIM[5]	37	156	18.63%
FSIM[6]	**388**	**445**	**80.41%**
GMSD[7]	53	104	15.15%
VIFp[8]	127	258	37.16%
MAD[9]	165	316	65.06%
SQMS[10]	225	338	54.34%
ADD-SSIM[11]	282	249	51.25%
PSIM[12]	387	296	65.93%

It can be seen from Table 1 that the prediction accuracy of SSIM [3], IWSSIM [5], GMSD [7], VIFp [8] and MS-SSIM [4] are unsatisfactory, we believed that the parameters of the Gaussian blur filter can affect the algorithm response values and the prediction performance. So we further discussed the influence of the standard deviation value of the Gaussian blur filter on the FR-IQA method performance. Figure 1 (a) and (b) show the relationship between the prediction accuracy of FR-IQA methods and the standard deviation value of the Gaussian blur filter. For the VIFp [8] method, the four extracted feature values VIFp1, VIFp2, VIFp3, and VIFp4 were as the response values of the four FR-IQA methods. Therefore, for the VIFp [8] method, four “acc” values were calculated.

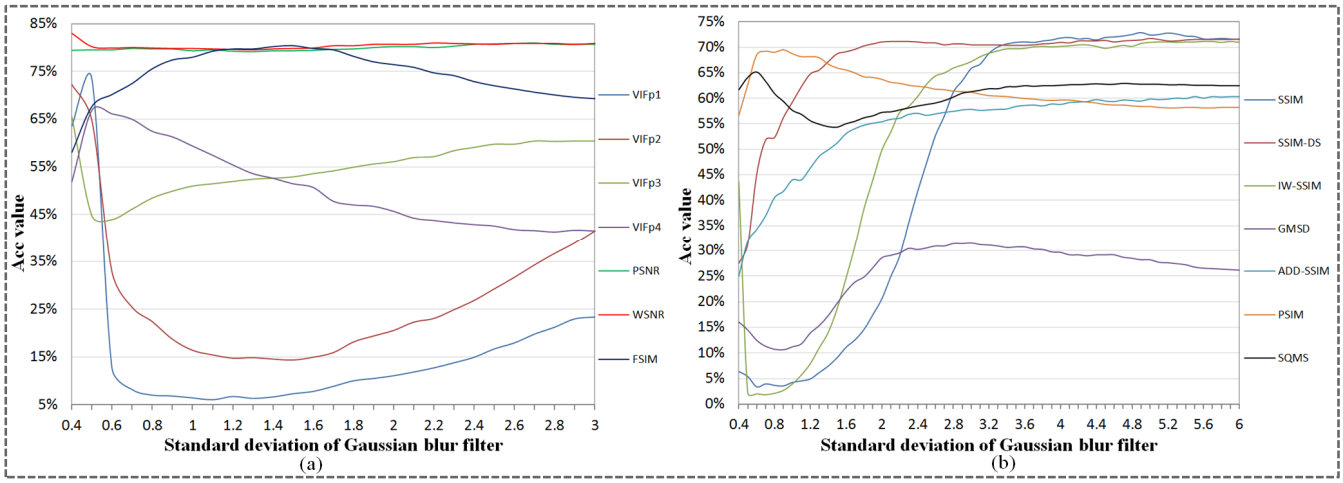

Figure S1. The relationship between the performance of the FR-IQA methods and the standard deviation value of the Gaussian blur filter. The abscissas in (a) and (b) are the standard deviation values, and the ordinates are the "acc" values. Each color curve represents an FR-IQA method.

The best "acc" values with standard deviation values for FR-IQA methods are shown in Table 2. It can be seen from Table 2 that the standard deviation value of the Gaussian blur filter has a great impact on the prediction accuracy of FR-IQA methods and the best standard deviation for each FR-IQA method is different.

Table S2. Standard deviation values of FR-IQA methods when obtaining the best "acc" values.

FR-IQA method	Standard deviation	Best acc
PSNR	2.7	80.98%
WSNR[2]	**0.4**	**83.01%**
SSIM[3]	4.9	72.88%
SSIM-DS[4]	5.0	71.72%
IWSSIM[5]	5.6	71.14%
FSIM[6]	1.5	80.41%
GMSD[7]	3.0	31.47%
VIFp1[8]	0.5	73.36%
VIFp2[8]	0.4	72.30%
VIFp3[8]	0.4	65.54%
VIFp4[8]	0.5	66.89%
SQMS[10]	0.6	65.15%
ADD-SSIM[11]	5.5	60.33%
PSIM[12]	0.7	69.21%

1.2 NR-IQA methods

The performance of gradient-based, entropy-based, and contrast-based NR-IQA methods is shown in Table 3. RDW abbreviates the "relative deviation weighting" method [13]. For LMG method, the size of φ was set to 40×40 and fecal microscopic images were divided to overlapping φ with stride 40 in horizontal and vertical directions. Due to the size of fecal microscopic image is 1600×1200, the corresponding LMG feature image is 40×30 and the response value of LMG method is the average of LMG feature image. All other parameters used the data from original literature.

Table S3. The performance of NR-IQA methods based on gradient, entropy, and contrast.

NR-IQA method	t_0	t_1	acc
Brenner[14]	411	405	78.76%
Tenengrad[15]	182	289	45.46%

EAV[16]	263	369	61.00%
SMD[17]	240	278	50.00%
Histogram	247	423	64.67%
Laplace[18]	425	327	72.59%
Energy Spectrum[19]	153	280	41.80%
RDW[13]	167	270	42.18%
RMS[20]	273	332	58.40%
JND[21]	354	399	72.68%
LMG	**149**	**673**	**79.27%**

The performance of NR-IQA methods based on edge, similarity, colour, phase, and free energy is shown in Table 4. The image block size of JNB [22], CPBD[23], Blur Effect[1], and NRSS[24] methods is 40×40 pixels, and the judgment threshold of each image block for JNB [22], and CPBD[23] is 16. All other parameters used the data from original literature.

Table S4. The performance of NR-IQA methods based on edge, similarity, colour, phase, and free energy.

NR-IQA method	t_0	t_1	**acc**
NRSS[24]	88	376	44.79%
Blur Effect[1]	139	186	31.37%
Color[25]	42	190	22.39%
Edge Width[26]	196	270	44.98%
JNB[22]	222	258	46.33%
CPBD[23]	33	213	23.75%
GPC[27]	**384**	**402**	**75.81%**
ARISM[28]	51	97	14.29%

The performance of transform-domain-based NR-IQA methods is shown in Table 5. [29] used Discrete Fourier Transform (DFT), [30] and [31] used Discrete Cosine Transform (DCT), [32] used Discrete Wavelet Transform (DWT), and [33] and [34] used Singular Value Decomposition (SVD). The image block size for [29], [30], [31], and [34] is 40×40 pixels. For [32], we used a db6 wavelet for two-layer decomposition. The neighborhood size in [33] is 41×41 pixels. All other parameters used the values from original literature.

Table S5. The performance of transform-domain-based NR-IQA methods.

NR-IQA method	t_0	t_1	**acc**
[29]	314	338	62.93%
[30]	289	321	59.89%
[31]	23	106	12.45%
[32]	**498**	**297**	**76.74%**
CW-SSIM[33]	395	322	69.21%
Q[34]	278	403	65.73%

1.3 Analysis of the low prediction accuracy of LMG method

From Table 3, we found that the prediction accuracy of LMG methods was not satisfactory, and we analyzed this problem in this section. In Figure 2, (a) and (b) are two images from the same group of fecal microscopic images. We believed that image (b) is sharper than (a) because of the clear formed elements and impurities. However, when we used Brenner [14], Tenengrad [15], EAV [16], SMD [17], Laplace entropy [18], and LMG methods, the algorithm response value of image (a) is greater than that of (b). These six IQA methods use the maximum value as the definition of clearest image, which means that these IQA methods consider image (a) to be clearer.

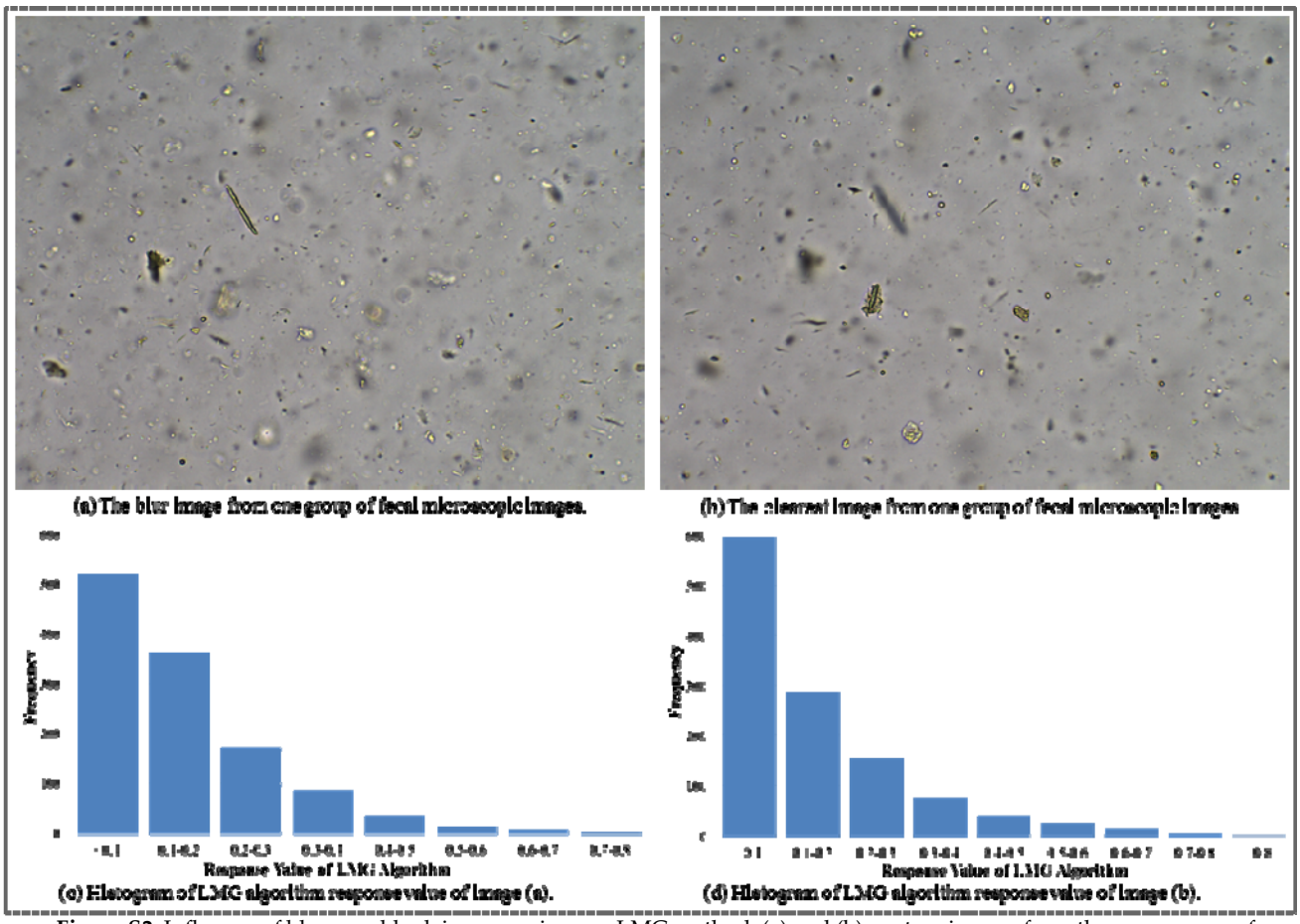

Figure S2. Influence of blurry or blank image regions on LMG method. (a) and (b) are two images from the same group of fecal microscopic images; (c) and (d) are histograms of LMG algorithm response value of (a) and (b), respectively.

There are some formed elements and impurities in image (b), and most of image regions are blank without any substance. There is some particulate matter in image (a), and most image regions are blurry (most substances are defocused). The algorithm response value in blank or blurry regions is close to 0, but the value of blurry regions is slightly higher than that of blank regions. In Figure 2, (c) and (d) are the histograms of LMG algorithm response value of (a) and (b), respectively. Due to the larger number of low response values in image (b), the mean algorithm response value of image (b) is less than that of (a), even though the maximum value in (b) is greater than that in (a). These low response values are noise to algorithm results.

To eliminate the impact of blurry or blank image regions, we added low response thresholds for the gradient-based algorithms (Brenner [14], Tenengrad [15], EAV [16], and SMD [17]). When calculating the mean value, a response value below the low response threshold is not counted. After verification, the best low response thresholds for Brenner [14], Tenengrad [15], EAV [16], and SMD [17] are 10, 19, 14 and 5, respectively. For Laplace entropy [18] method, we only counted the entropy of pixels whose Laplace convolution value was greater than 2. For LMG method, we sorted the LMG value of all block φ from large to small and only calculated the average of the top 4%. When humans judge whether images are clear, they mainly focus on sharp image regions but ignore blurry or blank image regions, and the above processing methods use a similar principle. The results are shown in Table 6, and the prediction accuracy of the six algorithms is improved.

Table S6. Performance of the six improved NR-IQA methods.

NR-IQA method	t_0	t_1	acc
Brenner[14]	500	366	83.59%
Tenengrad[15]	529	367	86.49%
EAV[16]	538	354	86.10%

SMD[17]	530	358	85.71%
Laplace[18]	489	387	84.56%
LMG	**501**	**408**	**87.74%**

The analysis in this section proves that region with low gradient value has an interference effect on judging the clearest image in autofocus process. Therefore, in our work, we introduced the gradient image of LMG algorithm into CNN as mask attention mechanism to enhance the features of high gradient area, so that CNN pays more attention to the image content of clear region during the training process.

2. Using resnet50 as GMANet backbone

We further used resnet50 as the backbone of GMANet, and the results are shown in Table 7.

Table S7. The performance of GMANet based on backbone resnet50.

		Test set				**Leucorrhea dataset**			**Blood dataset**		
	Model	t_0	t_1	acc	srocc	t_0	t_1	acc	t_0	t_1	acc
Round 1	L_0	134	68	97.115%	0.8811	359	251	87.268%	27	23	38.462%
	L_1	127	75	97.115%	0.8978	360	308	95.565%	49	65	87.692%
Round 2	L_0	123	80	97.596%	0.8817	363	274	91.130%	69	40	83.846%
	L_1	116	87	97.596%	0.8845	395	286	97.425%	31	80	85.385%

For the GMANet using resnet50 as the backbone, we used Output2 as the final output, while Output0 and Output1 were used as auxiliary outputs during the training process. For model L1, the gradient mask attention module was added on the first convolution block. It can be seen from Table 1 that GMANet based on backbone resnet50 can also obtain high prediction accuracy (equivalent to VGG16) on test set and leucorrhea dataset. However, the prediction accuracy on blood dataset is poor and unstable, so the network structure and model selection criteria need to be further adjusted for optimization. Therefore, we finally chose VGG16 with simple model structure as the backbone.

3. The performance of GMANet on assessing the clarity of objects

In order to verify the performance of GMANet on assessing the clarity of objects, we used the image fusion method to stitch clearest image patches together. We removed the Output1 of GMANet and used Score1 as the network output, which is 16 times smaller than the input image. We used fecal microscopic image with the size of 1024×1536 as network input.

For each group of microscopic images, we first input each image into GMANet, getting N score matrices with a size of 64×96 (N is the number of images in this group). Then we concatenated all the score matrices along the z direction, getting a three-dimensional matrix with a size of 64×96×10. The maximum value of the three-dimensional matrix is calculated in the z direction, and the z-axis coordinate was recorded where the maximum value is located, that is, the corresponding image number. For each small image patch (or object), the corresponding clearest image patch is cropped according to the image number. Finally, we stitched all clearest image patches together to obtain the fused image. As shown in Figure 3, (a) is the clearest image in one group of fecal microscopic images and and (b) is the corresponding fusion image. It can be seen from Figure 3 that there are defocused white blood cells in the clearest image, while most white blood cells in the fusion image are clear. The two blurred white blood cells on the left are defocused during the whole autofocus process, and there is no clear image patch for them. The fusion method based on clearest image patches is helpful to improve overall clearness of fecal microscopic image.

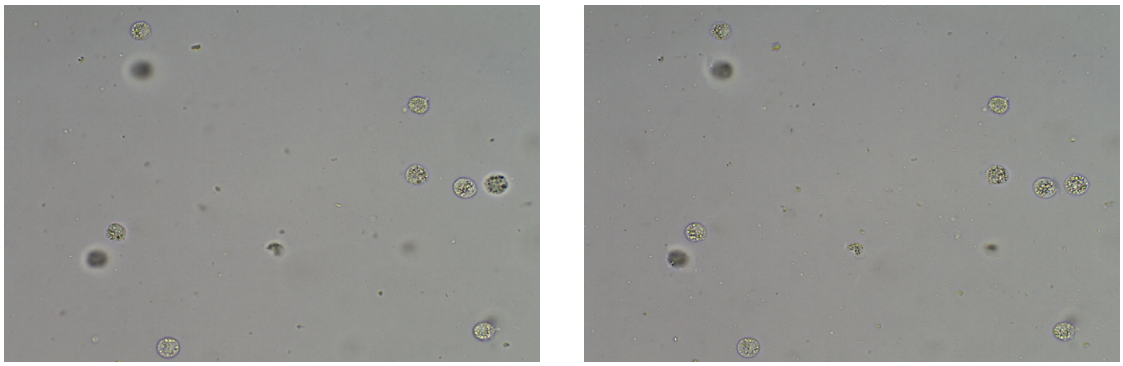

Figure S3. (a) is the clearest image of one group of fecal microscopic images. (b) is the corresponding fusion image.

However, when the sample contains lots of impurities or the material in the sample solution flows slowly, the fusion images will have obvious stitching traces, as shown in Figure 4. The image fusion effect based on the sharpness of image patches is unstable. Therefore, using deep learning method to fuse the microscopic images captured in autofocus process into one clear microscopic image is our next research direction.

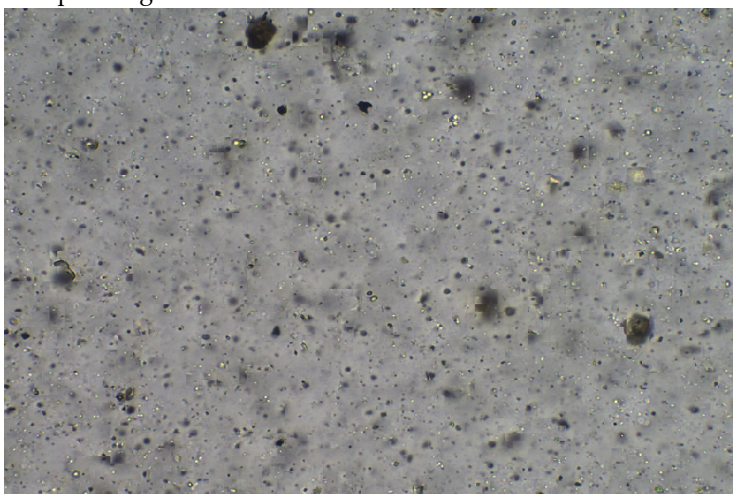

Figure S4. The fusion image of one feces sample which contains lots of impurities.

References

1. Crete, F.; Dolmiere, T.; Ladret, P.; Nicolas, M. The blur effect: Perception and estimation with a new no-reference perceptual blur metric. In Proceedings of the Conference on Human Vision and Electronic Imaging XII, San Jose, CA, Jan 29-Feb 01, 2007.
2. Damera-Venkata, N.; Kite, T.D.; Geisler, W.S.; Evans, B.L.; Bovik, A.C. Image quality assessment based on a degradation model. *Ieee Transactions on Image Processing* **2000**, *9*, 636-650, doi:10.1109/83.841940.
3. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *Ieee Transactions on Image Processing* **2004**, *13*, 600-612, doi:10.1109/tip.2003.819861.
4. Wang, Z.; Simoncelli, E.P.; Bovik, A.C. Multi-scale structural similarity for image quality assessment. In Proceedings of the 37th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, Nov 09-12, 2003; pp. 1398-1402.
5. Wang, Z.; Li, Q. Information Content Weighting for Perceptual Image Quality Assessment. *Ieee Transactions on Image Processing* **2011**, *20*, 1185-1198, doi:10.1109/tip.2010.2092435.
6. Zhang, L.; Zhang, L.; Mou, X.Q.; Zhang, D. FSIM: A Feature Similarity Index for Image Quality Assessment. *Ieee Transactions on Image Processing* **2011**, *20*, 2378-2386, doi:10.1109/tip.2011.2109730.
7. Xue, W.F.; Zhang, L.; Mou, X.Q.; Bovik, A.C. Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index. *Ieee Transactions on Image Processing* **2014**, *23*, 684-695, doi:10.1109/tip.2013.2293423.
8. Sheikh, H.R.; Bovik, A.C. Image information and visual quality. *Ieee Transactions on Image Processing* **2006**, *15*, 430-444, doi:10.1109/tip.2005.859378.
9. Larson, E.C.; Chandler, D.M. Most apparent distortion: full-reference image quality assessment and the role of strategy. *J Electron Imaging* **2010**, *19*, doi:10.1117/1.3267105.
10. Gu, K.; Wang, S.Q.; Yang, H.; Lin, W.S.; Zhai, G.T.; Yang, X.K.; Zhang, W.J. Saliency-Guided Quality Assessment of Screen Content Images. *Ieee Transactions on Multimedia* **2016**, *18*, 1098-1110, doi:10.1109/tmm.2016.2547343.

11. Gu, K.; Wang, S.Q.; Zhai, G.T.; Lin, W.S.; Yang, X.K.; Zhang, W.J. Analysis of Distortion Distribution for Pooling in Image Quality Prediction. *Ieee Transactions on Broadcasting* **2016**, *62*, 446-456, doi:10.1109/tbc.2015.2511624.
12. Gu, K.; Li, L.D.; Lu, H.; Min, X.K.; Lin, W.S. A Fast Reliable Image Quality Predictor by Fusing Micro- and Macro-Structures. *Ieee Transactions on Industrial Electronics* **2017**, *64*, 3903-3912, doi:10.1109/tie.2017.2652339.
13. Zhu, S. New Kind of Energy Spectrum Entropy Image Focusing Evaluation Function. *JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS* **1999**.
14. Brenner, J.F.; Dew, B.S.; Horton, J.B.; King, T.; Neurath, P.W.; Selles, W.D. An automated microscope for cytologic research a preliminary evaluation. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society* **1976**, *24*, 100-111, doi:10.1177/24.1.1254907.
15. Yeo, T.T.E.; Ong, S.H.; Jayasooriah; Sinniah, R. AUTOFOCUSING FOR TISSUE MICROSCOPY. *Image and Vision Computing* **1993**, *11*, 629-639, doi:10.1016/0262-8856(93)90059-p.
16. Wang, H.; Zhong, W.; Wang, J. Research of Measurement for Digital Image Definition. *Journal of Image and Graphics* **2004**, *9*, 828-831.
17. Tenenbaum, J.M. Accommodation in Computer Vision. **1970**.
18. Jarvis, R.A. Focus optimisation criteria for computer image processing. *Microscope* **1976**, *24*.
19. Muller, R.A.B., Andrew. Real-time correction of atmospherically degraded telescope images through image sharpening. *Journal of the Optical Society of America* **1974**, *64*.
20. Frazor, R.A.; Geisler, W.S. Local luminance and contrast in natural images. *Vision Research* **2006**, *46*, 1585-1598, doi:10.1016/j.visres.2005.06.038.
21. Chou, C.H.; Li, Y.C. A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile. *Ieee Transactions on Circuits and Systems for Video Technology* **1995**, *5*, 467-476, doi:10.1109/76.475889.
22. Ferzli, R.; Karam, L.J. A No-Reference Objective Image Sharpness Metric Based on the Notion of Just Noticeable Blur (JNB). *Ieee Transactions on Image Processing* **2009**, *18*, 717-728, doi:10.1109/tip.2008.2011760.
23. Narvekar, N.D.; Karam, L.J. A No-Reference Image Blur Metric Based on the Cumulative Probability of Blur Detection (CPBD). *Ieee Transactions on Image Processing* **2011**, *20*, 2678-2683, doi:10.1109/tip.2011.2131660.
24. Xie, X.F.; Zhou, J.; Qin-Zhang, W.U. No-reference quality index for image blur. *Journal of Computer Applications* **2010**, *30*, 921-924.
25. Guo, H.; Cao, J.; Zhou, Z.; Tang, L.; Ma, N. Image definition evaluation algorithm based on color relativity. *Infrared and Laser Engineering* **2013**.
26. Marziliano, P.; Dufaux, F.; Winkler, S.; Ebrahimi, T.; Ieee; Ieee; Ieee. A no-reference perceptual blur metric. In Proceedings of the IEEE International Conference on Image Processing, Rochester, Ny, Sep 22-25, 2002; pp. 57-60.
27. Blanchet, G.; Moisan, L. An explicit Sharpness Index related to Global Phase Coherence. In Proceedings of the IEEE International Conference on Acoustics, 2012.
28. Zhai, G.T.; Wu, X.L.; Yang, X.K.; Lin, W.S.; Zhang, W.J. A Psychovisual Quality Metric in Free-Energy Principle. *Ieee Transactions on Image Processing* **2012**, *21*, 41-52, doi:10.1109/tip.2011.2161092.
29. Field, D.J.; Brady, N. Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research* **1997**, *37*, 3367-3383, doi:10.1016/s0042-6989(97)00181-8.
30. Marichal, X.; Ma, W.Y.; Zhang, H.J. Blur determination in the compressed domain using DCT information. In Proceedings of the International Conference on Image Processing, 1999.
31. Kristan, M.; Pers, J.; Perse, M.; Kovacic, S. A Bayes-spectral-entropy-based measure of camera focus using a discrete cosine transform. *Pattern Recognition Letters* **2006**, *27*, 1431-1439, doi:10.1016/j.patrec.2006.01.016.
32. Wang, Y.W.; Liu, X.L.; Xie, H. A wavelet-based focus measure and 3-D autofocusing for microscope images. *Optics and Precision Engineering* **2006**, *14*, 1063-1069.
33. Sampat, M.P.; Wang, Z.; Gupta, S.; Bovik, A.C.; Markey, M.K. Complex Wavelet Structural Similarity: A New Image Similarity Index. *Ieee Transactions on Image Processing* **2009**, *18*, 2385-2401, doi:10.1109/tip.2009.2025923.
34. Shao, Y.; Sun, F.C.; Liu, Y. A No-reference Image Quality Assessment Method Using Local Structure Tensor. *Journal of Electronics & Information Technology* **2012**, *34*, 1779-1785.