SEQM: EDGE QUALITY ASSESSMENT BASED ON STRUCTURAL PIXEL MATCHING

Won-Dong Jang and Chang-Su Kim

School of Electrical Engineering, Korea University, Seoul, Korea
E-mails: wdjang@mcl.korea.ac.kr, changsukim@korea.ac.kr

ABSTRACT

A novel quality metric for binary edge maps, called the structural edge quality metric (SEQM), is proposed in this work. First, we define the matching cost between an edge pixel in a detected edge map and its candidate matching pixel in the ground-truth edge map. The matching cost includes a structural term, as well as a positional term, to measure the discrepancy between the local structures around the two pixels. Then, we determine the optimal matching pairs of pixels using the graph-cut optimization, in which a smoothness term is employed to take into account global edge structures in the matching. Finally, we sum up the matching costs of all edge pixels to determine the quality index of the detected edge map. Simulation results demonstrate that the proposed SEQM provides more faithful and reliable quality indices than conventional metrics.

Index Terms— Image quality assessment, edge quality assessment, binary edge map, structural similarity, and pixel matching.

1. INTRODUCTION

Edge detection is essential in many image processing and computer vision applications. Different edge detectors provide different edge maps, as shown in Fig. 1. To check the suitability of an edge map in an application, it is important to assess the quality of the edge map quantitatively by comparing it with the ground-truth edge map. The peak signal-to-noise ratio (PSNR) is one of the most popular quality metrics in the field of image processing. It is easy to compute and effective for various quality assessment tasks. However, PSNR does not match well with the image quality assessment of the human visual system (HVS). Recently, several methods have been proposed to overcome this drawback. For example, Wang *et al*. [1] proposed the structural similarity (SSIM) metric, which combines the luminance, contrast, and structure terms. By considering the contrast and structural similarities, SSIM can obtain results that are more faithful to HVS. Also, Sheikh and Bovik [2] formulated the visual information fidelity (VIF), which measures the information difference between a reference image and its distorted image.

The quality assessment problem of edge maps, however, is different from that of general gray-level or color images. In edge maps, edge pixels may be displaced or disappear, whereas false edge pixels may occur. Fig. 1 illustrates the displacement, occurrence, and extinction of edge pixels between two binary edge maps. The common image metrics may assess the qualities of edge maps misleadingly.

Several specialized metrics have been developed, which can evaluate the quality of an edge map more faithfully. The figure of merit (FOM) [5] computes the distance between a detected edge pixel and the corresponding ground-truth edge pixel. It allows many-to-one mapping between detected and ground-truth edge pixels. Similar to FOM, the closest distance metric (CDM) [6] matches detected pixels to ground-truth pixels, even if they are displaced by small positional errors. However, CDM allows one-to-one mapping only and is applicable to gray-level edge maps as well as binary edge maps. The pixel correspondence metric (PCM) [7] combines the edge strength cost with the distance cost in the edge pixel matching. It uses a weighted matching algorithm for bipartite graphs and produces robust results for both binary and gray-level edge maps. FOM, CDM, and PCM are all based on the positional pixel matching, and they are conceptually simple. However, these quality metrics cannot reflect structural similarity in the pixel matching, resulting in misleading assessment results in some cases.

(a) (b)

Fig. 1. The edge maps of the "Lena" image, which are obtained by (a) the Canny edge detector [3] and (b) the Prewitt edge detector [4].

In this work, we propose a novel quality metric for binary edge maps, called the structural edge quality metric (SEQM), which employs a structural similarity term as well as a positional similarity term. For each detected edge pixel, we first compute its matching costs to candidate edge pixels in the ground-truth map. Each matching cost is the sum of the structural cost, which measures the discrepancy between the local edge structures around the two pixels, and the positional cost. Then, among the candidate pixels, we determine the best matching pixel using the graph-cut optimization [8–10]. In the optimization, we encourage the matching of neighboring detected pixels to neighboring ground-truth pixels. After the optimization, we sum up the matching costs of all detected pixels to determine the quality index of the edge map. Simulation results demonstrate that the proposed SEQM provides faithful and reliable quality indices.

It is noted that, although we develop SEQM primarily for measuring the qualities of binary edge maps, we can employ SEQM to assess various other types of binary images as well, which include document images, segment maps, and contour images. This is be-

This work was supported partly by the Global Frontier R&D Program on Human-centered Interaction for Coexistence funded by the National Research Foundation (NRF) of Korea grant funded by the Korean Government (MEST) (NRF-M1AXA003-2011-0031648), and partly by the NRF of Korea grant funded by the MEST (No. 2012-011031)

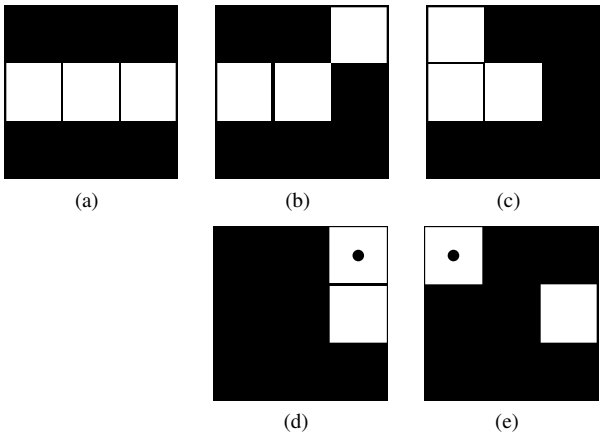


Fig. 2. (a) Source block B_S, (b) target block B_T^1, (c) target block B_T^2, (d) $B_S - B_T^1$, (e) $B_S - B_T^2$. A dotted white pixel has value -1.

cause the proposed SEQM is designed to quantify both structural and geometrical deformation in general binary images. We provide experimental results on the quality assessment of document images, and show that SEQM provides more faithful quality indices, which are more coherent to the subjective assessment of HVS, than the conventional metrics.

The rest of this paper is organized as follows. Section 2 proposes SEQM, and Section 3 provides experimental results. Finally, Section 4 concludes this work.

2. SEQM

SEQM employs the structural pixel matching, which is based on the structural similarity and the positional similarity between source and target edge pixels. Given source and target edge maps, SEQM takes three steps to compute the quality index, which measures the similarity between the source edge map and the target edge map. First, SEQM computes the pixel matching cost between each pair of a source pixel and a candidate target pixel. Second, it determines the correspondences between source pixels and target pixels using the graph-cut optimization. Third, it computes the quality index by summing up the pixel matching costs between the corresponding source and target pixels.

2.1. Structural Pixel Matching

Suppose that we determine the similarity between a source edge map I_s and a target edge map I_t. Let $\mathbf{s}$ denote an edge pixel in I_s. Also, let $\mathbf{t}$ denote an edge pixel in I_t, which is a matching candidate to $\mathbf{s}$. In this work, we search a matching pixel $\mathbf{t}$ within the 5×5 window in I_t, which is centered at the same position as $\mathbf{s}$.

First, we define the positional matching cost $C_\alpha(\mathbf{s}, \mathbf{t})$ between $\mathbf{s}$ and $\mathbf{t}$, based on the Euclidean distance, as

$$C_\alpha(\mathbf{s}, \mathbf{t}) = \frac{1}{R}\sqrt{(x(\mathbf{s}) - x(\mathbf{t}))^2 + (y(\mathbf{s}) - y(\mathbf{t}))^2}, \qquad (1)$$

where $(x(\mathbf{s}), y(\mathbf{s}))$ and $(x(\mathbf{t}), y(\mathbf{t}))$ are the coordinates of $\mathbf{s}$ and $\mathbf{t}$, respectively. Also, R is a normalizing constant to control the importance of the positional matching cost in the overall matching. We set R to 10 to make the maximum positional matching cost to 0.28.

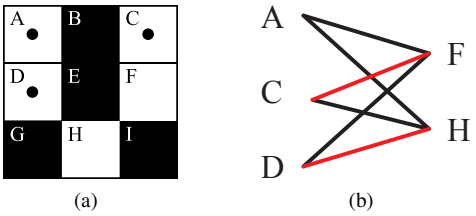


Fig. 3. (a) A difference block, and (b) the Hungarian matching result for (a). A dotted white pixel has value -1.

Next, we define the structural matching cost $C_\beta(\mathbf{s}, \mathbf{t})$, which plays a vital role in our metric SEQM. To determine $C_\beta(\mathbf{s}, \mathbf{t})$, we use the information in the blocks B_S and B_T, which are the 3×3 blocks centered at $\mathbf{s}$ and $\mathbf{t}$, respectively. We subtract the target block B_T from the source block B_S to obtain the difference block

$$B_D = B_S - B_T. \qquad (2)$$

Pixel values in B_S and B_T are binary: 0 for a non-edge pixel and 1 for an edge pixel, which are depicted by black and white pixels in this work, respectively. Thus, pixel values in B_D are ternary, *i.e.* -1, 0, or 1. Also, the center pixel of B_D should be 0, since the center pixels of B_S and B_T are $\mathbf{s}$ and $\mathbf{t}$, both of which are 1.

The absolute sum of pixel values in B_D is equal to the Hamming distance between B_S and B_T, which is a useful metric. The Hamming distance, however, is not sufficient to describe the structural similarity between B_S and B_T. Fig. 2 shows an example, in which the block in (a) is a source block B_S, and the blocks in (b) and (c) are two candidate target blocks B_T^1 and B_T^2. The difference blocks $B_S - B_T^1$ and $B_S - B_T^2$ are shown in (d) and (e). The two candidate blocks B_T^1 and B_T^2 have the same Hamming distance to the source block B_S, which is 2. But, B_T^1 has higher structural similarity to B_S than B_T^2 does. This is because we can deform B_S to B_T^1 simply by exchanging the pair of adjacent black and white pixels. The deformation from B_S to B_T^2 is more complicated.

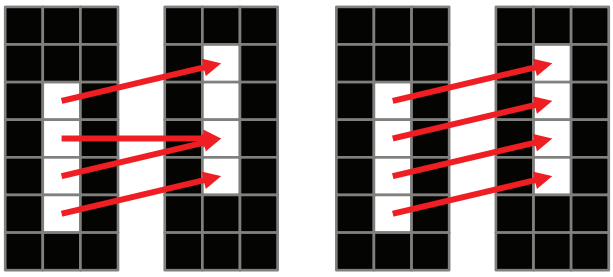
The set of pixels in the difference block B_D can be partitioned into V^-, V^0, and V^+, according to the pixel values -1, 0, and 1. Notice that '1' pixels should be near to '-1' pixels, if the deformation from B_S to B_T is not strong. Therefore, based on the graph theory, we search the optimal one-to-one matching between V^- and V^+ using the Hungarian method [11], which is the most well-known method for solving assignment problems. Specifically, we form the bipartite graph between V^- and V^+, and assign a weight $w(\mathbf{v}^-, \mathbf{v}^+)$ to the edge connecting pixel $\mathbf{v}^-$ in V^- to pixel $\mathbf{v}^+$ in V^+, which is given by

$$w(\mathbf{v}^-, \mathbf{v}^+) = H(|x(\mathbf{v}^-) - x(\mathbf{v}^+)| + |y(\mathbf{v}^-) - y(\mathbf{v}^+)|), \qquad (3)$$

where $H(d)$ is a monotonic function that returns a positive penalty according to the l_1-distance d between $\mathbf{v}^-$ and $\mathbf{v}^+$. We set $H(1) = 1$, $H(2) = 1.6$, $H(3) = 2$, $H(4) = 2$. Then, using the Hungarian method, we determine the optimal matching pairs between V^- and V^+. For instance, let us consider the difference block in Fig. 3(a). After forming the bipartite graph between $V^- = \{\mathrm{A}, \mathrm{C}, \mathrm{D}\}$ and $V^+ = \{\mathrm{F}, \mathrm{H}\}$, we obtain the optimal matching pairs (C, F) and (D, H), depicted by red edges in Fig. 3(b).

Then, we compute the structural matching cost $C_\beta(\mathbf{s}, \mathbf{t})$ between $\mathbf{s}$ and $\mathbf{t}$ as

$$C_\beta(\mathbf{s}, \mathbf{t}) = \frac{n_r + \sum_{(\mathbf{v}_i^-, \mathbf{v}_j^+) \in \mathcal{M}} w(\mathbf{v}_i^-, \mathbf{v}_j^+)}{8}, \qquad (4)$$



(a) (b)

Fig. 4. Edge pixel matching results (a) without the smoothness term and (b) with the smoothness term. The left and right blocks are the source and target images of size 7×3, respectively.

where $\mathcal{M}$ is the set of the optimal matching pairs and n_r is the number of remaining unmatched pixels. Each unmatched pixel contributes 1 in the numerator in (4). Note that $H(3) = H(4) = 2$. Therefore, when the l_1-distance between $\mathbf{v}_i^-$ and $\mathbf{v}_j^+$ is larger than 2, the two pixels share the weight $w(\mathbf{v}_i^-, \mathbf{v}_j^+) = 2$. This means that we treat two pixels as unmatched, when their l_1-distance is larger than 2. Notice that $0 \leq C_\beta \leq 1$. If $C_\beta = 0$, B_S and B_T are exactly the same.

Finally, the overall matching cost $C_{\text{pixel}}(\mathbf{s}, \mathbf{t})$ is composed of the positional cost $C_\alpha(\mathbf{s}, \mathbf{t})$ and the structural cost $C_\beta(\mathbf{s}, \mathbf{t})$, given by

$$C_{\text{pixel}}(\mathbf{s}, \mathbf{t}) = 1 - (1 - C_\alpha(\mathbf{s}, \mathbf{t})) \times (1 - C_\beta(\mathbf{s}, \mathbf{t})). \quad (5)$$

The structural similarity term $(1 - C_\beta(\mathbf{s}, \mathbf{t}))$ indicates how similar the local structure around $\mathbf{s}$ is to that around $\mathbf{t}$, whereas the positional similarity term $(1 - C_\alpha(\mathbf{s}, \mathbf{t}))$ means how close the two pixel positions $\mathbf{s}$ and $\mathbf{t}$ are. Our definition of the pixel matching cost in (5) is better than the direct multiplication of $C_\alpha(\mathbf{s}, \mathbf{t})$ and $C_\beta(\mathbf{s}, \mathbf{t})$. In case of the direct multiplication, the pixel matching cost becomes 0, if $C_\alpha(\mathbf{s}, \mathbf{t})$ or $C_\beta(\mathbf{s}, \mathbf{t})$ is 0. However, in our definition, $C_{\text{pixel}}(\mathbf{s}, \mathbf{t})$ is 0, only if both $C_\alpha(\mathbf{s}, \mathbf{t})$ and $C_\beta(\mathbf{s}, \mathbf{t})$ are 0, *i.e.* only if $\mathbf{s}$ is matched exactly to $\mathbf{t}$ both positionally and structurally.

2.2. Edge Map Matching by Energy Minimization

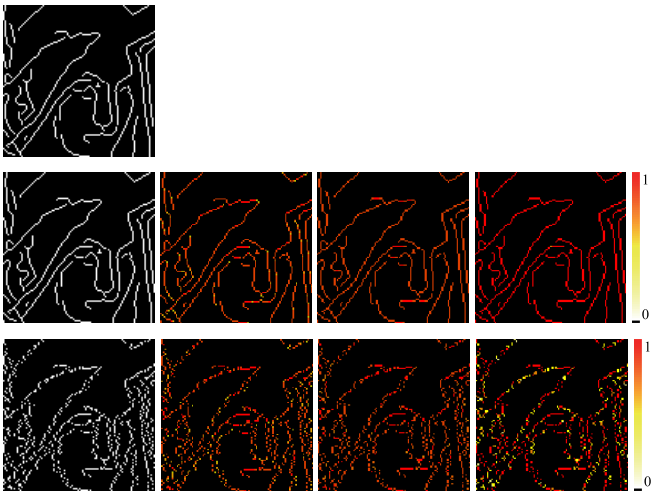
For each edge pixel $\mathbf{s}$ in the source edge map I_s, we may find the best matching edge pixel $\mathbf{t}$ in the target edge map I_t that minimizes the pixel matching cost in (5). However, the pixel matching cost considers local structures only. We propose considering global structures, as well as local structures, in order to match the edge pixels in I_s to those in I_t more reliably.

For each edge pixel $\mathbf{s}_i$ in I_s, let $\mathbf{l}_i$ denote the label, which represents the displacement vector, given by

$$\mathbf{l}_i = \mathbf{t}_i - \mathbf{s}_i, \quad (6)$$

where $\mathbf{t}_i$ is a candidate matching edge pixel in I_t. Since the matching is performed within the 5×5 window, there are 25 kinds of labels. Let L be the label map that records the labels of all edge pixels in I_s. Suppose that $\mathbf{s}_i$ and $\mathbf{s}_j$ are adjacent edge pixels in I_s. Then, if I_t is not too much degraded from I_s, $\mathbf{s}_i$ and $\mathbf{s}_j$ tend to be matched to $\mathbf{t}_i$ and $\mathbf{t}_j$, which are also adjacent. Based on the observation, we define a smoothness term

$$V(\mathbf{l}_i, \mathbf{l}_j) = \begin{cases} 0, & \text{if } \mathbf{l}_i = \mathbf{l}_j, \\ 1, & \text{if } \mathbf{l}_i \neq \mathbf{l}_j. \end{cases} \quad (7)$$



(a) (b) (c) (d)

Fig. 5. Visualization of the similarity of each edge pixel to the matching pixel in the ground-truth edge map on a part of the "Lena" image. In (a), the top is the ground-truth edge map, the middle and the bottom are distorted by '1-shift noise' and '1-swapping noise,' respectively. In (b), the similarities of edge pixels for the two types of noises are assessed by PCM. In (c) and (d), respectively, the positional similarities and the structural similarities maps are assessed by the proposed SEQM.

In other words, the smoothness term gives a penalty when adjacent pixels have different labels.

Then, we form an energy function, which consists of the data term and the smoothness term, given by

$$\begin{aligned} E(L) &= E_{\text{data}}(L) + \delta \cdot E_{\text{smooth}}(L) \\ &= \sum_{\mathbf{s}_i \in I_s} C_{\text{pixel}}(\mathbf{s}_i, \mathbf{s}_i + \mathbf{l}_i) + \delta \cdot \sum_{(\mathbf{s}_i, \mathbf{s}_j) \in \mathcal{N}} V(\mathbf{l}_i, \mathbf{l}_j), \end{aligned} \quad (8)$$

where $\mathcal{N}$ denotes the set of pairs of 8-connected edge pixels. The parameter δ controls the balance between the data term and the smoothness term. It is set to 0.1.

We obtain the optimal label map L^* by minimizing the energy function $E(L)$,

$$L^* = \arg\min_L E(L). \quad (9)$$

For the energy minimization, we use the graph-cut algorithm in [8–10]. Then, we compute the edge map matching cost $C_{\text{map}}(I_s, I_t)$ from I_s to I_t,

$$C_{\text{map}}(I_s, I_t) = E_{\text{data}}(L^*) = \sum_{\mathbf{s}_i \in I_s} C_{\text{pixel}}(\mathbf{s}_i, \mathbf{s}_i + \mathbf{l}_i^*). \quad (10)$$

Fig. 4 illustrates the impacts of the smoothness term in the optimization. Ideally, each edge pixel in the source image should be matched to the edge pixel in target image, which is shifted upward by one pixel. However, as shown in Fig. 4(a), edge pixels are poorly matched without the smoothness term, since the second pixel in the source image is matched to the second and third pixels with the identical data cost. By incorporating the smoothness term in Fig. 4(b), we resolve this ambiguity and obtain the ideal result.

Table 1. Comparison of the quality assessment results of PCM, FOM, SSIM, PSNR and the proposed SEQM. In this table, we list the quality indices of SEQM in percentage (%), by multiplying them by 100.

Image	Measure	Gaussian			Salt-and-Pepper			Speckle			Gaussian blur			1-shifting noise	1-swapping noise
		325.1	650.3	975.4	0.025	0.05	0.075	0.02	0.04	0.06	1.0	4.0	9.0		
Barbara	PCM	80.94	76.41	72.93	80.08	72.32	66.02	83.40	76.93	68.58	91.39	85.91	85.86	91.54	90.99
	FOM	87.66	83.12	79.06	87.46	80.31	72.08	90.92	84.41	73.01	94.67	88.64	86.89	91.83	91.17
	SSIM	65.81	57.78	50.76	68.67	52.84	38.17	69.23	61.15	47.39	92.02	79.44	74.61	26.26	31.89
	PSNR	12.45	11.36	10.64	12.68	10.88	9.41	13.03	11.60	10.17	19.70	14.78	13.64	8.38	8.86
	SEQM	**86.41**	**82.30**	**79.01**	**86.78**	**79.91**	**72.81**	**87.93**	**83.61**	**75.64**	**95.89**	**90.20**	**88.86**	**90.55**	**78.99**
Lena	PCM	82.64	75.60	74.63	81.96	70.88	59.29	84.06	74.43	67.38	88.88	89.58	88.72	92.22	90.71
	FOM	90.49	83.72	79.70	89.61	75.35	61.42	92.58	79.57	72.10	93.82	94.51	92.87	91.08	90.62
	SSIM	75.52	65.47	60.82	75.83	55.93	40.42	76.70	65.28	54.81	92.74	85.05	80.99	35.40	41.43
	PSNR	14.37	12.79	12.05	14.39	11.77	10.00	14.66	12.71	11.65	20.37	16.28	15.12	9.25	9.85
	SEQM	**88.02**	**83.02**	**79.84**	**88.42**	**76.98**	**66.76**	**89.27**	**80.92**	**75.72**	**95.51**	**92.66**	**91.71**	**90.19**	**78.03**
Peppers	PCM	81.98	71.67	67.14	82.03	66.88	61.91	80.34	75.14	66.97	95.61	87.30	85.14	91.46	91.65
	FOM	90.81	79.12	70.57	88.86	71.49	66.22	88.87	82.64	70.78	98.79	88.23	84.76	92.44	92.31
	SSIM	73.27	57.92	47.49	71.00	49.78	40.40	72.00	63.29	53.32	96.05	80.92	75.88	40.69	46.26
	PSNR	13.79	11.65	10.67	13.37	10.98	10.06	13.61	12.24	11.05	21.79	15.01	13.83	9.43	10.02
	SEQM	**87.16**	**79.05**	**73.59**	**87.07**	**74.88**	**68.99**	**86.47**	**81.76**	**74.65**	**97.82**	**90.42**	**87.79**	**90.76**	**80.44**
Baboon	PCM	72.57	67.92	64.43	71.98	65.99	63.13	74.63	68.52	66.20	80.91	83.54	83.55	91.25	91.47
	FOM	81.86	75.03	71.04	82.30	74.88	70.25	84.80	78.25	74.27	90.25	88.36	85.79	92.76	92.47
	SSIM	42.60	30.92	25.42	46.16	30.93	21.63	47.22	33.67	28.90	81.44	68.41	62.43	21.83	28.79
	PSNR	9.50	8.27	7.77	9.68	8.31	7.47	9.77	8.56	8.00	15.43	12.15	11.26	7.67	8.13
	SEQM	**77.51**	**72.65**	**68.52**	**78.24**	**71.28**	**66.62**	**79.57**	**73.55**	**70.67**	**92.03**	**87.57**	**86.21**	**90.93**	**80.33**

2.3. Edge Map Quality Index

We compute the edge map matching costs bidirectionally, *i.e.* from I_s to I_t and vice versa, and define the commutative metric between I_s and I_t, which is given by

$$\text{SEQM}(I_s, I_t) = 1 - \frac{C_{\text{map}}(I_s, I_t) + C_{\text{map}}(I_t, I_s)}{n_s + n_t}, \qquad (11)$$

where n_s, n_t denote the numbers of edge pixels in I_s, I_t, respectively. Notice that $0 \leq \text{SEQM}(I_s, I_t) \leq 1$.

We combine the two matching costs by

$$\frac{C(I_s, I_t) + C(I_t, I_s)}{n_s + n_t},$$

instead of the separate average

$$\frac{1}{2}\left(\frac{C(I_s, I_t)}{n_s} + \frac{C(I_t, I_s)}{n_t}\right).$$

This is because the separate average may be misleading. For example, suppose that I_s has a lot of edge pixels and I_t has few edge pixels only. Then, the matching from I_s to I_t incurs a high cost, but the matching cost from I_t to I_s is negligible. In such a case, the separate average is about 0.5, but our definition in (11) leads to the combined matching cost that is almost 1. Thus, our definition makes more sense.

3. EXPERIMENTAL RESULTS

3.1. Quality Assessment of Edge Maps

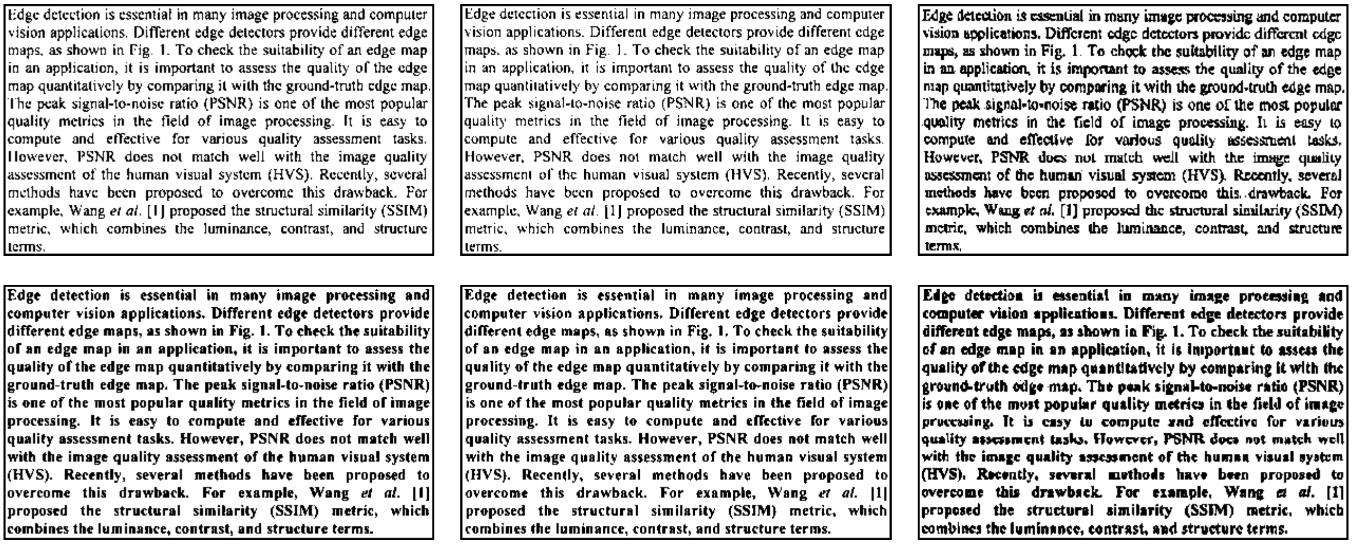
We first test the performance of SEQM on edge maps of four classical test images: "Barbara," "Lena," "Peppers," and "Baboon." We use the edge maps, obtained by the Canny edge detector [3], as the ground-truth edge maps. Then, we corrupt the test images with various noises and assess the qualities of the edge maps for the corrupted images, obtained by the same Canny edge detector. We employ four noise types: additive zero-mean Gaussian noises with variances 325.1, 650.3, and 975.4, salt-and-pepper noises with noise occurrence probabilities 0.025, 0.05, and 0.075, multiplicative speckle noises with variances 0.02, 0.04, and 0.06, and Gaussian blurs with filters of size 5×5 with variances 1.0, 4.0, and 9.0.

We also corrupt the ground-truth edge maps by two types of custom noises. The first type is '1-shifting noise' that shifts all edge pixels to the right by one pixel. The second type is '1-swapping noise' that swaps each edge pixel with its left or right pixels randomly. Note that '1-shifting noise' does not change the structural characteristics of edge maps, but '1-swapping noise' disturbs them.

Table 1 compares the edge map quality assessment results of the proposed SEQM with those of the conventional metrics PCM [7], FOM [5], SSIM [1], and PSNR. For PCM, the maximum matching distance is set to 2 as in the proposed SEQM. For FOM, the scaling parameter a is set to $1/9$, which is a typical value. We observe that, similar to the conventional metrics, SEQM quality indices decrease as the noise intensities increase for the first three noise types. For Gaussian blurs of "Lena" and "Baboon", PCM and FOM quality indices have no correlation with the noise intensities. In contrast, SEQM exhibits quality indices consistent with the noise intensities. Also, at the same noise intensity, SEQM quality indices are lower for the "Baboon" image, which has more complicated textures. Therefore, the edge detection results on the "Baboon" image are more severely affected by the same amount of noises. These results indicate that SEQM is a stable and reliable metric for edge maps.

In contrast to the conventional metrics, the proposed SEQM assesses the structural similarity of an edge map to its ground-truth one more faithfully. As shown in Fig. 5,'1-shifting noise' does not change the structure of edges, whereas '1-swapping noise' corrupts them severely. However, it can be observed in Table 1 that PCM and FOM provide similar scores for both types of noises. In case of SSIM, the scores for '1-shifting noise' are even lower than those for '1-swapping noise.' This is because the conventional metrics focus on the positional matching only, without the systematic consideration of structural similarities. In contrast, we see that SEQM provides consistently lower quality indices for '1-swapping noise' than for '1-shifting noise.' These assessment results confirm our main idea that the structural information is important in evaluating the quality of an edge map.

Fig. 5 visualizes the similarity of each edge pixel in an edge map to the matching pixel in the ground-truth edge map visually. For SEQM, we show the uni-directional similarities from detected edge



Edge detection is essential in many image processing and computer vision applications. Different edge detectors provide different edge maps, as shown in Fig. 1. To check the suitability of an edge map in an application, it is important to assess the quality of the edge map quantitatively by comparing it with the ground-truth edge map. The peak signal-to-noise ratio (PSNR) is one of the most popular quality metrics in the field of image processing. It is easy to compute and effective for various quality assessment tasks. However, PSNR does not match well with the image quality assessment of the human visual system (HVS). Recently, several methods have been proposed to overcome this drawback. For example, Wang *et al.* [1] proposed the structural similarity (SSIM) metric, which combines the luminance, contrast, and structure terms.

(a) (b) (c)

Fig. 6. Text documents. The top and the bottom contain texts in regular and boldface fonts, respectively. (a) ‘Print doc,’ (b) ‘Copy doc,’ and (c) ‘Fax doc.’

(a) (b) (c) (d)

Fig. 7. Image documents. The top and the bottom are the edge maps of “Lena” and “Baboon,” respectively. (a) The entire images of ‘Print doc’ and (b) their enlarge parts. The enlarged parts of (c) ‘Copy doc’ and (d) ‘Fax doc.’

pixels to ground-truth edge pixels. In the case of PCM in Fig. 5(b), there are small differences between the similarities for the ‘1-shifting noise’ and ‘1-swapping’ noise. On the other hand, in SEQM, we see noticeable differences between the structural similarities in Fig. 5(d), although there are negligible differences between the positional similarities in Fig. 5(c). Thus, the proposed SEQM assess the qualities of binary images reliably and faithfully, by considering the structural similarities as well as the positional similarities.

In addition, notice that both SEQM and PCM are designed such that their ideal scores for ‘1-shifting noise’ are exactly 90%. However, the actual scores in Table 1 are higher than 90% due to mismatching pairs. Specifically, the 1-pixel shift is not detected for some source pixels, which are incorrectly matched to target pixels at the same positions. But we see that SEQM provides relatively

Table 2. Comparison of the quality assessment results of PCM, FOM, SSIM, PSNR and the proposed SEQM for binary document images. In this table, we list the quality indices of SEQM in percentage(%), by multiplying them by 100.

Image		Measure	Copy doc	Fax doc
Text documents	Regular	PCM	85.63	66.90
		FOM	42.00	42.79
		SSIM	74.73	30.19
		PSNR	12.12	7.02
		SEQM	**84.38**	**72.12**
	Boldface	PCM	86.34	74.31
		FOM	40.83	42.71
		SSIM	63.02	41.32
		PSNR	9.79	7.86
		SEQM	**85.36**	**79.22**
Image documents	Lena	PCM	66.39	77.16
		FOM	69.35	95.13
		SSIM	70.63	61.37
		PSNR	12.38	9.81
		SEQM	**80.41**	**80.13**
	Baboon	PCM	73.58	75.46
		FOM	80.89	91.76
		SSIM	63.80	36.89
		PSNR	9.46	6.62
		SEQM	**82.54**	**76.54**

more accurate scores than PCM by employing the smoothness term and reducing mismatching pairs.

3.2. Quality Assessment of Document Images

Next, we examine the quality assessment performance of SEQM on binary document images. Document images can be degraded by diverse processes, such as printing, copying, and transmission. This degradation in practical applications is different from typical noise models: Gaussian, salt-and-pepper, and speckle noises. The degradation incurs displacements, occurrences, and extinction of black pixels on white background. Therefore, the proposed SEQM for edge map quality assessment, which is based on the pixel matching criterion, is also suitable for evaluating binary document images. SEQM hence can be used to measure the performance of office automation equipments, such as printers, scanners, and facsimiles.

In this test, we use two text documents (regular and boldface) and two image documents (edge maps of "Lena" and "Baboon"). We process these documents in three-ways:

1. Print $\rightarrow$ Scan
2. Print $\rightarrow$ Copy $\rightarrow$ Scan
3. Print $\rightarrow$ Facsimile Transmission $\rightarrow$ Scan

These documents are labeled as 'Print doc,' 'Copy doc,' and 'Fax doc,' respectively. Since the printing and scanning steps are common in all three processes, we regard 'Print doc' as the ground-truth binary image and measure the qualities of 'Copy doc' and 'Fax doc.'

As shown in Figs. 6 and 7, 'Fax doc' experiences heavier pixel displacements, occurrences, and extinction than 'Copy doc' in general. This is because data are more easily corrupted during the facsimile transmission. Table 2 compares the quality assessment results of the proposed SEQM with the conventional metrics. In case of the text documents, all the metrics with the exception of FOM provide higher indices for 'Copy doc' than for 'Fax doc.' However, in case of the image documents, PCM and FOM provide lower indices for 'Copy doc' than for 'Fax doc,' which is contrary to the human perception. In contrast, SEQM provides higher indices for 'Copy doc.' Note that, in the case of "Lena" in Fig. 7, edges in 'Copy doc' are thinned and the subjective quality is not very high. This is why SEQM provides similar quality indices for 'Copy doc' and 'Fax doc' of "Lena." These simulation results demonstrate that the proposed SEQM is a promising quality metric for binary document images.

4. CONCLUSIONS

We proposed a novel edge map quality metric, called SEQM, which employs not only a positional similarity term but also a structural similarity term. First, SEQM computes the matching costs of each detected edge pixel to candidate matching pixels in the ground-truth map. Then, it determines the optimal pairs of matching pixels using the graph-cut optimization. Finally, it sums up the matching costs of all detected pixels to assess the quality of the edge map. Simulation results demonstrated that the proposed SEQM provides more reliable quality indices than the conventional quality metrics by taking into account the structural differences between matching pixels. Moreover, we verified the practicality of the proposed SEQM on binary document images in real world applications.

5. REFERENCES

[1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P.Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[2] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.

[3] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[4] J. M. S. Prewitt, "Object enhancement and extraction," in *Picture Processing and Psychopictorics*, B. Lipkin and A. Rosenfeld, Eds. Academic Press, 1970.

[5] W. K. Pratt, *Digital Image Processing*, New York: Wiley Interscience, 1978.

[6] K. Bowyer, C. Kranenburg, and S. Dougherty, "Edge detector evaluation using empirical ROC curves," *Comput. Vis. Image Understand.*, vol. 84, no. 1, pp. 77–103, 2001.

[7] M. S. Prieto and A. R. Allen, "A similarity metric for edge images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1265–1273, Oct. 2003.

[8] Y. Boykov, O. Veksler, and R. Zabih, "Efficient approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 12, pp. 1222–1239, Nov. 2001.

[9] V. Kolmogorov and R.Zabih, "What energy functions can be minimized via graph cuts?," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, Feb. 2004.

[10] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, 2004.

[11] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Res. Logist. Quart.*, vol. 2, no. 1, pp. 83–97, 1955.