Research Article

Digital Watermarking Method Warranting the Lower Limit of Image Quality of Watermarked Images

Motoi Iwata, Tomoo Kanaya, Akira Shiozaki, and Akio Ogihara

Graduate School of Engineering, Osaka Prefecture University, 1-1 Gakuen-cho, Sakai-shi, Osaka 599-8531, Japan

Correspondence should be addressed to Motoi Iwata, iwata@cs.osakafu-u.ac.jp

We propose a digital watermarking method warranting the lower limit of the image quality of watermarked images. The proposed method controls the degradation of a watermarked image by using a lower limit image. The lower limit image means the image of the worst quality that users can permit. The proposed method accepts any lower limit image and does not require it at extraction. Therefore lower limit images can be decided flexibly. In this paper, we introduce 2-dimensional human visual MTF model as an example of obtaining lower limit images. Also we use JPEG-compressed images of quality 75% and 50% as lower limit images. We investigate the performance of the proposed method by experiments. Moreover we compare the proposed method using three types of lower limit images with the existing method in view of the tradeoff between PSNR and the robustness against JPEG compression.

1. Introduction

Digital watermarking is a technique that embeds additional data into digital contents so that the distortion by embedding them is perceptually undetectable [1]. The distortion of watermarked images by general digital watermarking methods is fixed only after embedding. Some digital watermarking methods [2] prevent the degradation of the image quality of watermarked images by using human visual system. However the lower limit of the image quality of the watermarked images was not clear. Such obscurity of the lower limit disturbs the practical use of digital watermarking.

The method proposed by Yoshiura and Echizen [2] decided the embedding strength by introducing uniform color space so that the degradation of all regions in a image was equalized. However there is the fact that the degradation by modification in uniform color space depends on the direction of the modification described in Section 2.

In this paper, we propose a digital watermarking method warranting the lower limit of the image quality of watermarked images. The proposed method controls the degradation of a watermarked image by using a lower limit image. The lower limit image means the image of the worst quality that users can permit. The proposed method accepts any lower limit image and does not require it at extraction. Therefore lower limit images can be decided flexibly. In this paper, we introduce 2-dimensional human visual MTF model as an example of obtaining lower limit images. Also we use JPEG-compressed images of quality 75% and 50% as lower limit images, which are popular formats as degraded images.

The rest of this paper consists of five sections. We describe our approach in Section 2 and introduce the existing techniques in Section 3. Then we describe the detail of the proposed method in Section 4 and show and discuss the performance of the proposed method in Section 5. Finally we conclude this paper in Section 6.

2. Our Approach

We assume that there is a range in which the changes for pixel values are imperceptible. We call the range "perceptual capacity." Existing methods do not modify pixel values in the perceptual capacity strictly. Therefore we introduce a lower limit image which means the image of the worst quality that users can permit, that is, which provides with perceptual capacity. The contribution of the introduction of lower limit images is the separation of perceptual capacity

and watermarking procedures. The separation yields the independence of investigation.

The proposed method warrants the lower limit of the image quality of a watermarked image by approximating an original image to the corresponding lower limit image for embedding. Moreover we introduce L*a*b* color space for equalizing the degradation by embedding, where L*a*b* color space is one of the popular uniform color spaces. Then the quality of a watermarked image is between that of the original image and that of the lower limit image in L*a*b* color space. The lower limit image can be decided flexibly because the proposed method does not require it at extraction.

In general, the modification with the same quantity in a uniform color space perceptually yields the same degradation. However the direction of the modification is important, too. We found this fact by comparing the degradation of the modified images approaching human visual filtered images with that of the modified images leaving the filtered images, where the modification was done in L*a*b* color space. The human visual filter cuts off redundant component for visual sensation. Figure 1 shows the difference in quality by the direction of modification, where the human visual filter used here is mathematical 2-dimensional human visual MTF model described in Section 3.2.2. As shown in Figure 1, the degradation of the modified image approaching the filtered image is more imperceptible than that of the modified image leaving the filtered image. We utilize this feature by employing the images filtered by mathematical 2-dimensional human visual MTF model as one of the types of lower limit images. Also we use JPEG compressed images of quality 75% and 50% as lower limit images, which are popular formats as degraded images. In other words, employing the MTF model is a theoretical approach to generate lower limit images, while using JPEG-compression is a practical approach.

3. Existing Techniques

3.1. Color Spaces. In this section, we describe XYZ color space, L*a*b* color space, and opponent color space in Sections 3.1.1, 3.1.2, and 3.1.3, respectively.

3.1.1. XYZ Color Space. XYZ color space is a color space established by CIE (Commission Internationale de l'Eclairage) in 1931. The transformation of sRGB color space into XYZ color space is as follows [3].

First we obtain gamma-transformed sRGB color space by the following equations:

$$
R = \begin{cases} \dfrac{R'}{12.92}, & R' \le 0.04045, \\ \left(\dfrac{(R' + 0.055)}{1.055} \right)^{2.4}, & \text{otherwise}, \end{cases}
$$

$$
G = \begin{cases} \dfrac{G'}{12.92}, & G' \le 0.04045, \\ \left(\dfrac{(G' + 0.055)}{1.055} \right)^{2.4}, & \text{otherwise}, \end{cases}
$$

$$
B = \begin{cases} \dfrac{B'}{12.92}, & B' \le 0.04045, \\ \left(\dfrac{(B' + 0.055)}{1.055} \right)^{2.4}, & \text{otherwise}, \end{cases}
$$

$$
R' = \frac{R_s}{255},
$$

$$
G' = \frac{G_s}{255},
$$

$$
B' = \frac{B_s}{255},
$$

(1)

where R, G, and B are the values in gamma-transformed sRGB color space, and R_s, G_s, and B_s are the values in sRGB color space.

Then we obtain XYZ color space from gamma-transformed sRGB color space by the following equations:

$$
\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.412453 & 0.35758 & 0.180423 \\ 0.212671 & 0.71516 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}. \quad (2)
$$

*3.1.2. L*a*b* Color Space.* L*a*b* color space is one of uniform color spaces established by CIE in 1976 [4]. In a uniform color space, the distances between colors are fixed based on the perceptual differences between the colors [3, 5, 6].

L*a*b* color space is obtained from XYZ color space by the following equations:

$$
L^*(Y) = 116 f\left(\frac{Y}{Y_n} \right) - 16,
$$

$$
a^*(X, Y) = 500 \left\{ f\left(\frac{X}{X_n} \right) - f\left(\frac{Y}{Y_n} \right) \right\},
$$

$$
b^*(Y, Z) = 200 \left\{ f\left(\frac{Y}{Y_n} \right) - f\left(\frac{Z}{Z_n} \right) \right\},
$$

$$
f(t) = \begin{cases} t^{1/3}, & \text{if } t > 0.008856, \\ 7.787t + \dfrac{16}{116}, & \text{otherwise}, \end{cases}
$$

(3)

where X_n, Y_n, and Z_n are coefficients which depend upon the illuminant (for daylight illuminant D65, $X_n = 95.045$, $Y_n = 100$, and $Z_n = 108.892$).

3.1.3. Opponent Color Space. Opponent color space is based on input signals from L cone, M cone, and S cone in retina. Opponent color space is obtained from XYZ color space by the following equation:

$$
\begin{pmatrix} J_{w/k} \\ J_{r/g} \\ J_{y/b} \end{pmatrix} = \begin{pmatrix} 0.279 & 0.720 & -0.107 \\ -0.449 & 0.290 & -0.077 \\ 0.086 & -0.590 & 0.501 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \quad (4)
$$

where $J_{w/k}$, $J_{r/g}$, and $J_{y/b}$ represent luminance channel and the opponent channels of red-green and yellow-blue, respectively.

Original image

(a)

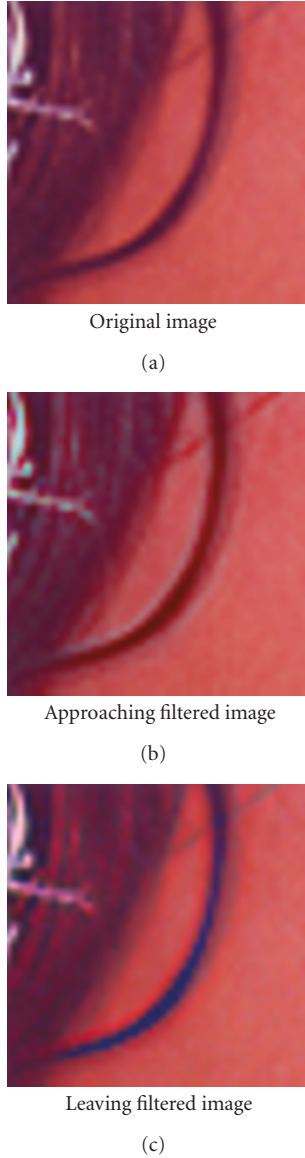Approaching filtered image

(b)

Leaving filtered image

(c)

FIGURE 1: Difference by the direction of modification.

3.2. Two-Dimensional Human Visual MTF Model

3.2.1. Modulation Transfer Function. Modulation Transfer Function (MTF) describes the relationship between spacial frequency and contrast sensitivity. Spatial frequency is a measure of how often a structure repeats per unit of distance. As shown in Figure 2, any pattern corresponds to the spacial frequency. On the other hand, contrast sensitivity is a measure of the ability to discern luminances of different levels in a static image. Contrast sensitivity depends on spatial frequency. For example, it tends to be high for medium spatial frequency, while it tends to be low for high spatial frequency.

Figure 3 shows the shape of human visual MTF for luminance. As shown in Figure 3, contrast sensitivity is numerically expressed by MTF. In human visual MTF for luminance, contrast sensitivity is high for medium spatial frequency and is suddenly low for high spatial frequency. It is known that the shape of human visual MTF for other color stimulus is similar to that for luminance.

3.2.2. Mathematical 2-Dimensional Human Visual MTF Model. Ishihara et al. [7, 8] and Miyake [9] revealed that human visual MTF depends on directivity in spatial frequency and mean of stimulus. Moreover they proposed mathematical 2-dimensional human visual MTF model about tristimulus on opponent color space.

Let u, v be horizontal and vertical spatial frequency, respectively, let $w = \sqrt{u^2 + v^2}$ be the spatial frequency on u-v plane, and let $\phi = \arctan(v/u)$ be the direction of w. Then contrast sensitivity $M(u, v)$ obtained by mathematical 2-dimensional human visual MTF model is defined as follows:

$$M(u, v) = M_0(w)[1 - \{1 - \gamma(u)\} \, |\sin 2\phi|], \qquad (5)$$

where $\gamma(u)$ represents the ratio of diagonal contrast sensitivity to horizontal contrast sensitivity when the horizontal spatial frequency is equal to u, and $M_0(w)$ is defined as follows:

$$M_0(w) = \beta(\bar{o})M_B\left(w, m_p, \sigma_p\right)$$
$$- \epsilon\beta(\bar{o})M_B\left(m_c, m_p, \sigma_p\right)M_B(w, m_c, \sigma_c), \quad (6)$$
$$M_B(w, m, \sigma) = \exp\left\{-2\pi^2\sigma^2(w - m)^2\right\},$$

where $\beta(\bar{o})$ represents the maximum value of horizontal contrast sensitivity on human visual MTF when the mean of stimulus is equal to $\bar{o}$.

We define $M_{r/g}(u, v)$ and $M_{y/b}(u, v)$ as contrast sensitivity $M(u, v)$ for red-green channel and yellow-blue channel, respectively. We can obtain $M_{r/g}(u, v)$ and $M_{y/b}(u, v)$ from (5) using the parameters shown in Table 1. The parameters $\beta_{r/g}(\overline{J_{r/g}})$, $\beta_{y/b}(\overline{J_{y/b}})$, $\gamma_{rg}(w)$, and $\gamma_{y/b}(w)$ in Table 1 are calculated by the following equations:

$$\beta_{r/g}\left(\overline{J_{r/g}}\right) = -0.07570\overline{J_{r/g}}^2 + 8.731\overline{J_{r/g}} - 1.839,$$

$$\beta_{y/b}\left(\overline{J_{y/b}}\right) = -0.054\overline{J_{y/b}}^2 + 4.851\overline{J_{y/b}} - 0.8439,$$

$$\gamma_{r/g}(u) = 0.001531u^2 - 0.06149u + 1.14,$$ (7)

$$\gamma_{y/b}(u) = 0.001919u^2 - 0.06427u + 1.09,$$

where $\overline{J_{r/g}}$ and $\overline{J_{y/b}}$ represent the means of all $J_{r/g}$ and $J_{y/b}$ in an image, respectively. In the literature [7–9], $M_0(w)$ is not calculated when $w < 1.5$. For the correction of this incident, we regard $M_0(w)$ as $M_0(1.5)$ when $w < 1.5$ so as to obey the meaning of contrast sensitivity. Figure 4 shows the shapes of $M_{r/g}(u, v)$ and $M_{y/b}(u, v)$ with or without the correction.

3.3. Filtering Based on Two-dimensional Human Visual MTF Model.

The filter of 2-dimensional human visual MTF model cuts off imperceptible components from images.

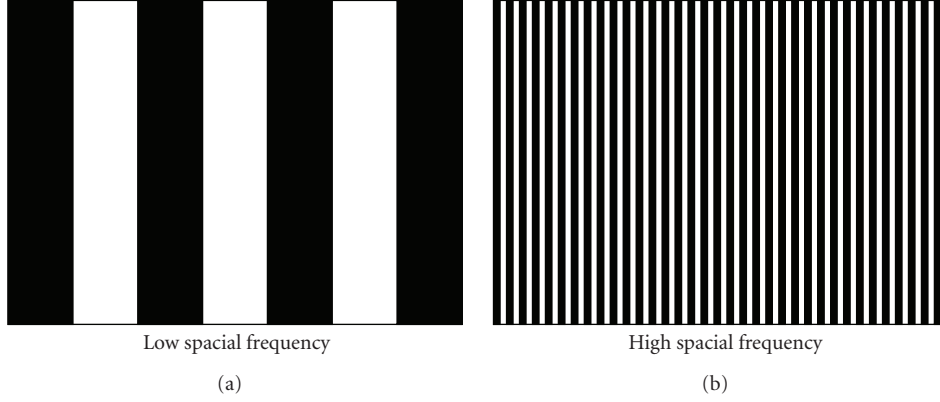| Low spacial frequency | High spacial frequency |
| (a) | (b) |

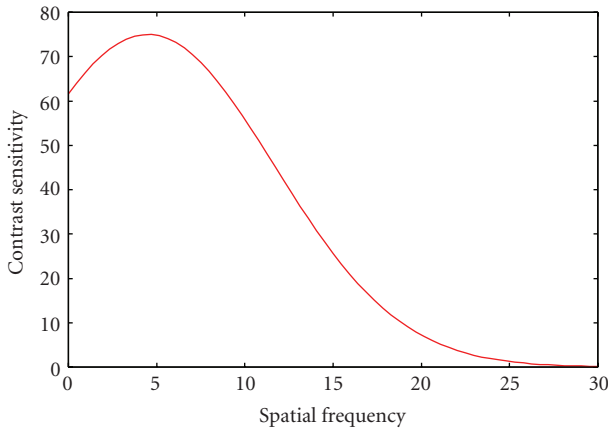FIGURE 2: Patterns with low spatial frequency and high spatial frequency.

FIGURE 3: Shape of human visual MTF for luminance.

TABLE 1: Parameters for MTF.

	$J_{r/g}$	$J_{y/b}$
$\overline{o}$	$\overline{J_{r/g}}$	$\overline{J_{y/b}}$
$\beta(\overline{o}$	$\beta_{r/g}(\overline{J_{r/g}})$	$\beta_{y/b}(\overline{J_{y/b}})$
$\gamma(u)$	$\gamma_{r/g}(u)$	$\gamma_{y/b}(u)$
m_p	1.5	1.5
σ_p	1/70	1/45
ϵ	0	1/4
m_c	—	7.5
σ_c	—	1/20

In this paper, only red-green and yellow-blue channels are filtered, which are based on the characteristic that modification in luminance is more perceptual than that in red-green or yellow-blue channel.

Step 1. An original image with $N_x \times N_y$ pixels is transformed into opponent color space. Let $J_{r/g}(x, y)$ and $J_{y/b}(x, y)$ be the values of red-green and yellow-blue channels located at the coordinate (x, y), respectively.

Step 2. $J_{r/g}(x, y)$ and $J_{y/b}(x, y)$ are transformed into $F_{r/g}(u, v)$ and $F_{y/b}(u, v)$ by discrete Fourier transform (DFT), respectively.

Step 3. The filtered discrete Fourier transform coefficients $F'_{r/g}(u, v)$ and $F'_{y/b}(u, v)$ are, respectively, obtained by the following equations:

$$F'_{r/g}(u, v) = F_{r/g}(u, v) \times M_{r/g}(u, v),$$
$$F'_{y/b}(u, v) = F_{y/b}(u, v) \times M_{y/b}(u, v). \qquad (8)$$

Step 4. The filtered pixel values in opponent color space are obtained from $F'_{r/g}(u, v)$ and $F'_{y/b}(u, v)$ by inverse DFT. Then the lower limit image is obtained by the transformation of opponent color space into sRGB color space.

4. Proposed Method

4.1. Embedding Procedure. Firstly we divide an original image with $N_x \times N_y$ pixels and the corresponding lower limit image into blocks with $L \times L$ pixels. Moreover the blocks are divided into subblocks with $L_s \times L_s$ pixels. Let $B(i, j)$ and $\hat{B}(i, j)$ be the (i, j)-th block in the original image and the lower limit image, respectively, where $0 \le i < N_x/L$, $0 \le j < N_y/L$. Let $B_s(k, l)$ and $\hat{B}_s(k, l)$ be the (k, l)-th subblock in $B(i, j)$ and $\hat{B}(i, j)$, respectively, where $0 \le k < L/L_s$, $0 \le l < L/L_s$("(i, j)" is omitted in the representation of $B_s(k, l)$ and $\hat{B}_s(k, l)$ for simplicity). The proposed method embeds one watermark bit into one block. Let $b_{ij} \in \{0, 1\}$ be the watermark bit embedded in $B(i, j)$.

The embedding procedure of b_{ij} is as follows.

Step 1. Let $g(m, n)$ be the pixel value located at (m, n) in $B(i, j)$, where $0 \le m < L$, $0 \le n < L$. Then the pixel value $g(m, n)$ is regarded as the point $P_{mn}(L^*_{mn}, a^*_{mn}, b^*_{mn})$ in L*a*b* color space. In the same manner, $\hat{g}(m, n)$ and $\hat{P}_{mn}(\hat{L}^*_{mn}, \hat{a}^*_{mn}, \hat{b}^*_{mn})$ are defined from $\hat{B}(i, j)$.

Step 2. Let $D(m, n)$ be the distance between the origin O and the point P_{mn}, and let $\hat{D}(m, n)$ be the distance between

(a)

(b)

(c)

(d)
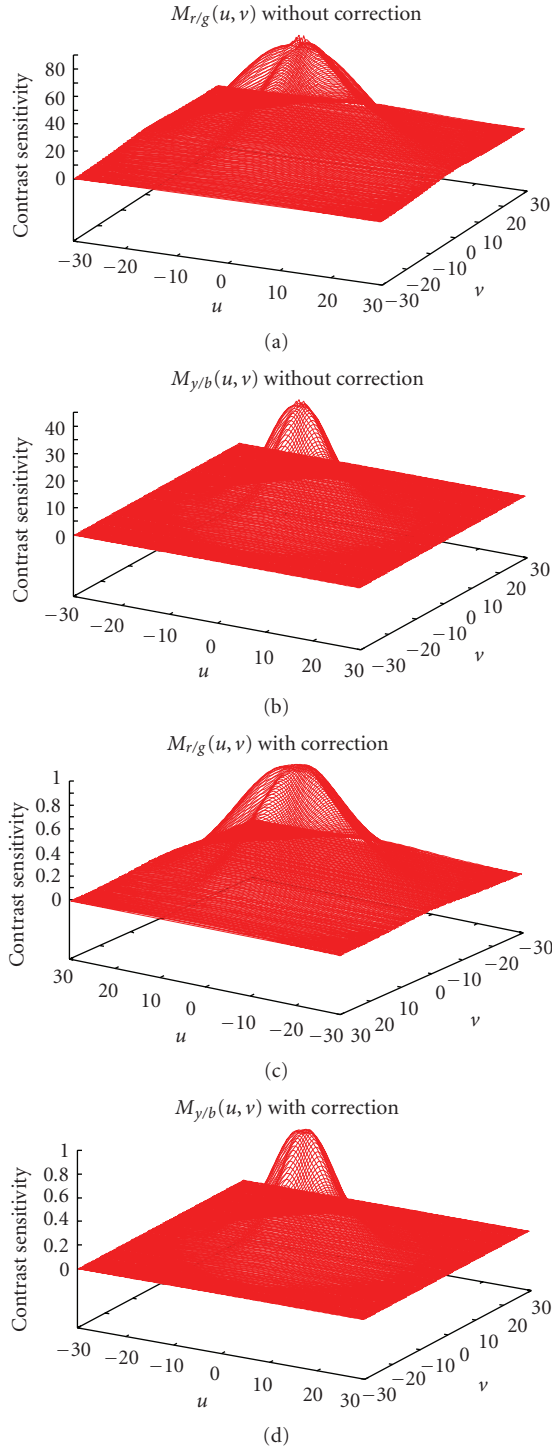
FIGURE 4: Shapes of $M_{r/g}(u, v)$ and $M_{y/b}(u, v)$.

FIGURE 5: Modification P_{mn} into P'_{mn}.

FIGURE 6: Original images.

the origin O and the point $\widehat{P}_{mn}$. $D(m, n)$ and $\widehat{D}(m, n)$ are obtained by the following equations:

$$D(m, n) = \sqrt{L_{mn}^{*2} + a_{mn}^{*2} + b_{mn}^{*2}},$$
$$\widehat{D}(m, n) = \sqrt{\widehat{L}_{mn}^{*2} + \widehat{a}_{mn}^{*2} + \widehat{b}_{mn}^{*2}}. \tag{9}$$

Aerial | Airplane

Balloon | Couple

Earth | Girl

Lena | Mandrill

Milkdrop | Parrots

Pepper | Sailboat

FIGURE 7: Watermarked images (MTF).

Aerial | Airplane

Balloon | Couple

Earth | Girl

Lena | Mandrill

Milkdrop | Parrots

Pepper | Sailboat

FIGURE 8: Watermarked images (JPEG75).

Step 3. The difference $\Delta D(m, n)$ between the norms of the pixels in the original image and the lower limit image is obtained by the following equation:

$$\Delta D(m, n) = \hat{D}(m, n) - D(m, n). \tag{10}$$

Moreover the sum Σ_+ of positive values and the sum Σ_- of negative values in $B(i, j)$ are obtained as follows:

$$\Sigma_+ = \sum_{(\forall m, n) \Delta D(m, n) \geq 0} \Delta D(m, n),$$
$$\Sigma_- = \sum_{(\forall m, n) \Delta D(m, n) < 0} \Delta D(m, n). \tag{11}$$

Step 4. The sum $W(k, l)$ of $D(m, n)$ in $B_s(k, l)$ is obtained by the following equation:

$$W(k, l) = \sum_{y=0}^{L_s - 1} \sum_{x=0}^{L_s - 1} D(L_s k + x, L_s l + y). \tag{12}$$

Step 5. The mean $\overline{W}$ of the sums $W(k, l)$ of all subblocks in $B(i, j)$ is obtained by the following equation:

$$\overline{W} = \frac{\sum_{x=0}^{L/L_s - 1} \sum_{x=0}^{L/L_s - 1} W(x, y)}{L^2 / L_s^2}. \tag{13}$$

FIGURE 9: Watermarked images (JPEG50).

Step 6. The quantized mean $\overline{W}_Q$ is obtained by the following equation:

$$\overline{W}_Q = \left\lceil \frac{\overline{W}}{Q} + 0.5 \right\rceil, \tag{14}$$

where $\lceil x \rceil$ means the maximum integer which is smaller than x. The quantizer Q acts as embedding strength.

Step 7. The quantized mean $\overline{W}_Q$ will be modified so as to be even when $b_{ij} = 0$ and be modified so as to be odd when $b_{ij} = 1$ by the following steps (Step 7~Step 9).

The watermarked value $\overline{W}'$ of the quantized mean is obtained as follows:

(i) when $W_Q \equiv b_{ij} \bmod 2$,

$$\overline{W}' = W_Q \times Q; \tag{15}$$

(ii) when $W_Q \not\equiv b_{ij} \bmod 2$,

$$\overline{W}' = \begin{cases} \left(\overline{W}_Q + 1\right) \times Q, & \text{if } |\Sigma_+| \geq |\Sigma_-|, \\ \left(\overline{W}_Q - 1\right) \times Q, & \text{if } |\Sigma_+| < |\Sigma_-|. \end{cases} \tag{16}$$

Moreover we obtain the quantity K which is added to $\overline{W}$ for embedding by the following equation:

$$K = \overline{W}' - \overline{W}. \tag{17}$$

Step 8. We obtain the quantity $K(m, n)$ which is added to each pixel value $D(m, n)$ in $B(i, j)$ for embedding as follows:

(i) when $K \geq 0$,

$$K(m,n) = \begin{cases} \dfrac{KL^2 \Delta D(m,n)}{\Sigma_+ L_s^2}, & \text{if } \Delta D(m,n) \geq 0, \\ 0, & \text{if } \Delta D(m,n) < 0; \end{cases} \tag{18}$$

(ii) when $K < 0$,

$$K(m,n) = \begin{cases} 0, & \text{if } \Delta D(m,n) \geq 0, \\ \dfrac{KL^2 \Delta D(m,n)}{\Sigma_- L_s^2}, & \text{if } \Delta D(m,n) < 0. \end{cases} \tag{19}$$

Step 9. Let $P'_{mn}(L^{*\prime}_{mn}, a^{*\prime}_{mn}, b^{*\prime}_{mn})$ be the watermarked point of P_{mn}. As shown in Figure 5, we change P_{mn} into P'_{mn} so as to satisfy $D'(m,n) - D(m,n) = K(m,n)$ by the following equation:

$$\begin{pmatrix} L^{*\prime}_{mn} \\ a^{*\prime}_{mn} \\ b^{*\prime}_{mn} \end{pmatrix} = (1-t) \begin{pmatrix} L^*_{mn} \\ a^*_{mn} \\ b^*_{mn} \end{pmatrix} + t \begin{pmatrix} L^{\hat{*}}_{mn} \\ a^{\hat{*}}_{mn} \\ b^{\hat{*}}_{mn} \end{pmatrix}, \tag{20}$$

where t is the ratio for changing of P_{mn} into P'_{mn}. The ratio t satisfies $0 \leq t \leq 1$ and the following equation:

$$\begin{aligned} &\left(\Delta L^{*}_{mn}{}^2 + \Delta a^{*}_{mn}{}^2 + \Delta b^{*}_{mn}{}^2\right) t^2 \\ &\quad + 2\left\{ L^*_{mn}\Delta L^*_{mn} + a^*_{mn}\Delta a^*_{mn} + b^*_{mn}\Delta b^*_{mn} \right\} t \\ &\quad - 2D(m,n)K(m,n) - K(m,n)^2 = 0, \end{aligned} \tag{21}$$

where $\Delta L^*_{mn}, \Delta a^*_{mn}$, and Δb^*_{mn} are obtained by the following equation:

$$\begin{aligned} \Delta L^*_{mn} &= L^{\hat{*}}_{mn} - L^*_{mn}, \\ \Delta a^*_{mn} &= a^{\hat{*}}_{mn} - a^*_{mn}, \\ \Delta b^*_{mn} &= b^{\hat{*}}_{mn} - b^*_{mn}. \end{aligned} \tag{22}$$

Step 10. The watermarked points P'_{mn} are transformed into sRGB color space, where the transformation of real numbers into integers (round-up or round-down) is decided so that the influence on P'_{mn} is minimized. Then we obtain the watermarked block $B(i, j)$.

We obtain the watermarked image after all watermark bits have been embedded.

4.2. Extracting Procedure. Firstly we obtain the blocks $B(i, j)$ and the subblocks $B_s(k, l)$ from a watermarked image in the same manner as embedding procedure.

The extracting procedure for a block $B(i, j)$ is as follows.

Step 1. The pixel values $g(m, n)$ in $B(i, j)$ are transformed into L*a*b* color space and are regarded as the points $P_{mn}(L^*_{mn}, a^*_{mn}, b^*_{mn})$ in L*a*b* color space.

Step 2. The sum $W(k, l)$ of $D(m, n)$ in $B_s(k, l)$ is obtained for each sub-block in the same manner as (12).

Step 3. The mean $\overline{W}$ of the sums $W(k, l)$ of all subblocks in $B(i, j)$ is obtained in the same manner as (13).

Step 4. The quantized mean $\overline{W}_Q$ is obtained in the same manner as (14). Then we extract b_{ij} as follows:

$$b_{ij} = \begin{cases} 0, & \text{if } \overline{W}_Q \equiv 0 \bmod 2, \\ 1, & \text{if } \overline{W}_Q \equiv 1 \bmod 2. \end{cases} \tag{23}$$

We obtain all the watermark bits after extracting for all blocks.

5. Experiments

5.1. Environments. Firstly we investigated the image quality of watermarked images and lower limit images. Then we confirmed that embedded watermark bits were perfectly extracted from watermarked images. Next we investigated the available range of the embedding strength Q because the embedding strength should be decided so that the ratio t can exist. Moreover we investigated the property of the proposed method when the embedding strength Q was variable for each block. The variable embedding strength was the maximum value for each block. Finally we investigated the robustness against JPEG compression and the comparison with an existing method in view of image quality and robustness.

As shown in Figure 6 we used twelve color images "aerial," "airplane," "balloon," "couple," "earth," "girl," "lena," "mandrill," "milkdrop," "parrots," "pepper", and "sailboat" as original images. They were standard images widely used for experiments. The size of all original images was 256×256 pixels, that is, $N_x = 256$, and $N_y = 256$. We used $L = 32$ and $L_s = 16$ as the size of blocks and subblocks, respectively. All the watermark bits b_{ij} were decided so as to satisfy $W_Q \not\equiv b_{ij} \bmod 2$. Then the watermarked images that used such watermark bits were worst degraded among those that used any watermark bit. We used $Q = M_{\min}/6$ as the embedding strength so that the ratio t in Step 9 in Section 4.1 could exist, where $M_{\min}$ represents the minimum value of the larger ones of $|\Sigma_+|$ and $|\Sigma_-|$ in each block. The lower limit images consist of three types, that is, "MTF" which is described in Section 3.2, and "JPEG75" and "JPEG50" which are JPEG-compressed images of quality 75% and 50%. The quality 75% of JPEG compression is the standard quality.

Aerial	Airplane
Balloon	Couple
Earth	Girl
Lena	Mandrill
Milkdrop	Parrots
Pepper	Sailboat

Figure 10: Lower limit images (MTF).

We used PSNR for the evaluation of image quality. PSNR was calculated by the following equation:

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\text{MSE}},$$
$$\text{MSE} = \frac{1}{3 N_x N_y} \sum_{i=1}^{3 N_x N_y}, \left(\text{img}_i - \text{oimg}_i \right)^2, \tag{24}$$

where img_i and oimg_i represent the pixels in one image and the other image, respectively. We also used mean structural similarity (MSSIM) index [10] to evaluating the similarity between watermarked images and lower limit images. MSSIM index is obtained by calculating the mean

Aerial · Airplane

Balloon · Couple

Earth · Girl

Lena · Mandrill

Milkdrop · Parrots

Pepper · Sailboat

FIGURE 11: Watermarked images (MTF, maxQ).

Aerial · Airplane

Balloon · Couple

Earth · Girl

Lena · Mandrill

Milkdrop · Parrots

Pepper · Sailboat

FIGURE 12: Watermarked images (JPEG75, maxQ).

of SSIM indices of all windows on the images. SSIM index between two window I_0 and I_1 of size 8×8 pixels was calculated by the following equation:

$$\text{SSIM}(I_0, I_1) = \frac{(2\mu_0\mu_1 + C_1)(2\sigma_{01} + C_2)}{(\mu_0^2 + \mu_1^2 + C_1)(\sigma_0^2 + \sigma_1^2 + C_2)}, \qquad (25)$$

where μ_0 and μ_1 represent the means of I_0 and I_1, respectively, and σ_0 and σ_1 represent the variances of I_0 and I_1, respectively. The constant values C_1 and C_2 are defined as default values, that is, $C_1 = (0.01 \times 255)^2$ and $C_2 = (0.03 \times 255)^2$, respectively.

5.2. Results and Discussion

5.2.1. Image Quality. Figures 7~9 show the watermarked images using "MTF," "JPEG75", and "JPEG50" as the lower limit images, respectively. As shown in Figure 7~9, the degradation of all the watermarked images was imperceptible.

Table 2 shows the PSNRs of the watermarked images and the lower limit images against the original images. As shown in Table 2, the PSNRs of the watermarked images except for "milkdrop" and "sailboat" are the lowest when the type of the lower limit images is "MTF." The PSNRs of the watermarked images "milkdrop" and "sailboat" using "MTF" are higher than those using "JPEG50," although the PSNRs of the lower limit images of type "MTF" are less than those

FIGURE 13: Watermarked images (JPEG50, maxQ).

TABLE 2: PSNRs of watermarked images and lower limit images against original images.

	Watermarked images			Lower limit images		
	MTF	JPEG75	JPEG50	MTF	JPEG75	JPEG50
aerial	34.9	37.5	35.8	26.1	28.3	26.9
airplane	39.5	45.8	43.6	25.7	30.2	28.5
balloon	43.6	48.9	46.8	32.9	34.9	33.3
couple	42.0	47.5	44.0	29.6	34.1	32.6
earth	41.0	48.0	44.6	31.4	33.7	32.0
girl	42.0	46.8	45.0	28.3	32.7	31.5
lena	42.2	45.1	44.2	26.6	32.4	30.6
mandrill	32.4	39.0	37.8	21.9	27.2	25.4
milkdrop	43.6	44.1	42.4	30.7	32.3	30.8
parrots	40.0	48.3	46.5	26.0	34.3	31.6
pepper	37.7	41.1	39.7	25.16	28.8	27.4
sailboat	43.9	44.9	43.0	27.8	31.0	29.4

TABLE 3: The minimum and maximum of Q.

	MTF		JPEG75	JPEG50
	min		max	
aerial	8	52	146	181
airplane	8	11	18	35
balloon	7	7	19	26
couple	6	66	42	89
earth	7	51	56	86
girl	6	37	44	56
lena	10	47	55	54
mandrill	6	101	136	154
milkdrop	10	44	78	89
parrots	7	7	38	43
pepper	8	81	123	162
sailboat	12	13	37	54

Such degradation tends to be imperceptible. Therefore the images filtered by 2-dimensional human visual MTF model were appropriate for lower limit images in view of the direction of modification by embedding. However the lower limit images of type "MTF" were slightly inappropriate in view of the strength of modification by embedding because some degradation was perceptible as shown in Figure 10. Therefore one of the future works is the improvement of the decision of the embedding strength.

5.2.3. Flexibility of Embedding Strength. Table 3 shows the minimum and maximum of the embedding strength Q. The minimum values of Q of "JPEG75" and "JPEG50" are similar to those of "MTF." The minimum of the embedding strength was fixed so that the embedded watermark could be perfectly extracted from the watermarked image. The maximum of the embedding strength was fixed so that the ratio t could exist (the maximum of Q is equal to $M_{min}/6$). As shown in Table 3, the range of available Q depended on images. In "balloon" and "parrots," the flexibility of Q was low because the maximum of Q is equal to the minimum of Q. It is the

of type "JPEG50." This suggests that the arbitrariness of the type of lower limit images is useful. Although the PSNRs of the watermarked images "aerial" and "mandrill" using "MTF" were less than 37 [dB] and were relatively low, the degradation of these images was imperceptible because these images mainly consisted of texture-like or noisy regions as shown in Figure 7.

5.2.2. Validity of Lower Limit Images. Figure 10 shows the lower limit images of type "MTF." As shown in Figure 10, the degradation of the lower limit images of type "MTF" appeared as emphasizing the difference of color, for example, the hair in "mandrill" or the profile of parrots in "parrots."

Original

(a)

Existing method (32.9 (dB))

(b)

Proposed method (same Q-MTF, 32.4 (dB))

(c)

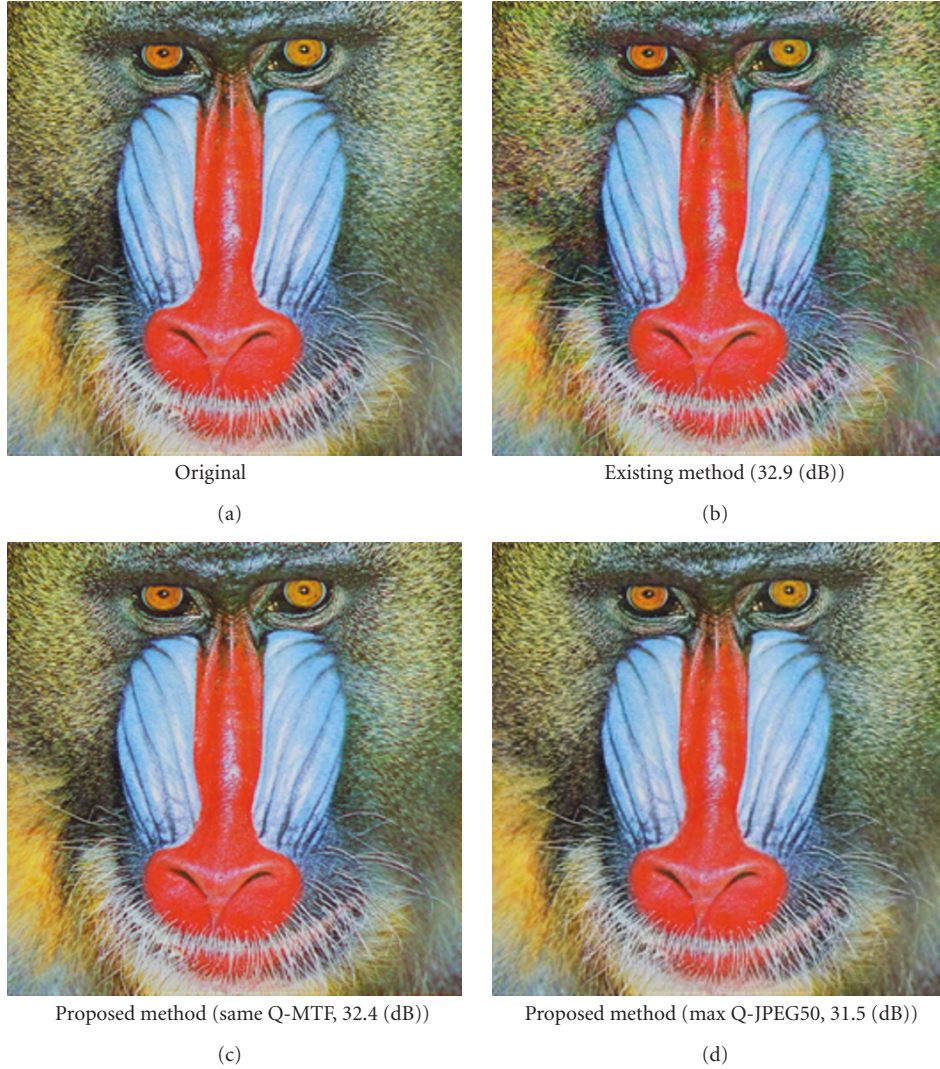Proposed method (max Q-JPEG50, 31.5 (dB))

(d)

FIGURE 14: Comparison in the image quality.

future work to investigate the relationship between the range of available embedding strengths and the robustness against attacks.

5.2.4. Performance Using the Maximum of Q of Each Block. We investigated the property of the proposed method when the maximum of Q (= $M_{min}/6$) of each block is used; that is, embedding strength Q is variable by a block. The demerit of using the maximum of Q of each block is the increase of the quantity of data saved for extracting. In the following, we call the methods using the same Q and the maximum of Q "sameQ" and "maxQ", respectively. Note that the high maximum of Q in Table 3 does not always cause the low PSNR of the watermarked image of "maxQ" with such Q because the PSNR does not depend on the maximum of Q among all blocks but on the distribution of Q for each block when the maximum of Q of each block is used.

TABLE 4: PSNRs of watermarked images using the maximum of Q of each block.

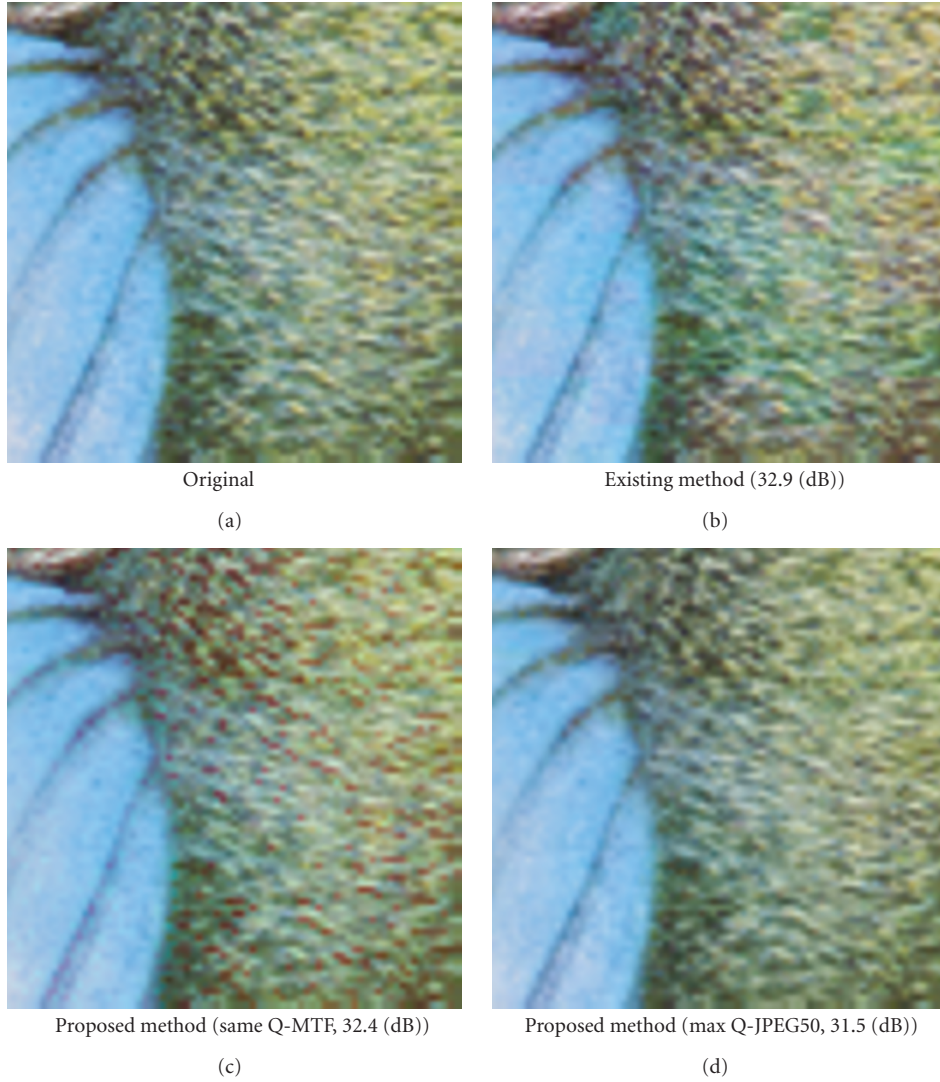|	MTF	JPEG75	JPEG50
aerial	30.0	34.4	32.8
airplane	28.6	35.1	34.9
balloon	37.6	40.1	38.7
couple	33.5	40.0	38.3
earth	34.9	40.2	38.1
girl	32.8	38.1	37.2
lena	30.1	38.2	36.3
mandrill	25.0	33.4	31.5
milkdrop	34.9	37.7	36.5
parrots	29.5	40.4	37.1
pepper	29.9	34.1	31.8
sailboat	32.6	35.9	35.2

Original	Existing method (32.9 (dB))
(a)	(b)
Proposed method (same Q-MTF, 32.4 (dB))	Proposed method (max Q-JPEG50, 31.5 (dB))
(c)	(d)

FIGURE 15: Comparison in the image quality by enlarged partial regions.

Figures 11~13 show the watermarked images using "MTF," "JPEG75", and "JPEG50" as the lower limit image, respectively. The embedding strength of all the watermarked images is "maxQ". Table 4 shows the PSNRs of watermarked images using the maximum of Q of each block. The degradation of all the watermarked images using "JPEG75" and "JPEG50" was imperceptible. The degradation of the watermarked image "mandrill" using "MTF" was slightly perceptible as scattered green dots in the hair of mandrill. Table 4 shows the PSNRs of the watermarked images using "maxQ" as the embedding strength. Although PSNR of "airplane" using "MTF" is under 30 [dB], the degradation of "airplane" was imperceptible because the degradation was chromatic. On the other hand, although the degradation of "mandrill" was mainly texture-like chromatic noise on texture-like regions, the degradation of "mandrill" was slightly perceptible because the modification by embedding was large. We confirmed that the use of "MTF" caused not only the right direction of the modification by embedding

but also too large modification by embedding. However we obtained practical results when we use "JPEG75" and "JPEG50" as the lower limit images.

5.2.5. Similarity between Watermarked Images and Lower Limit Images. Table 5 shows the MSSIMs between watermarked images and lower limit images. As shown in Table 5, we confirmed that all the watermarked images were similar to the lower limit images because all the MSSIMs were high. It is natural that the MSSIMs of "maxQ" are larger than those of "sameQ" because the use of larger Q yields the closer watermarked images to the lower limit images. It is the reason why the MSSIMs of "maxQ" are not 1.0 that there are some pixels of which $K(m, n)$ are equal to 0 in (18) or (19).

5.2.6. Robustness against JPEG Compression. We define the number of correctly extracted bits divided by the number of all embedded bits as extraction rate. Tables 6 and 7 show
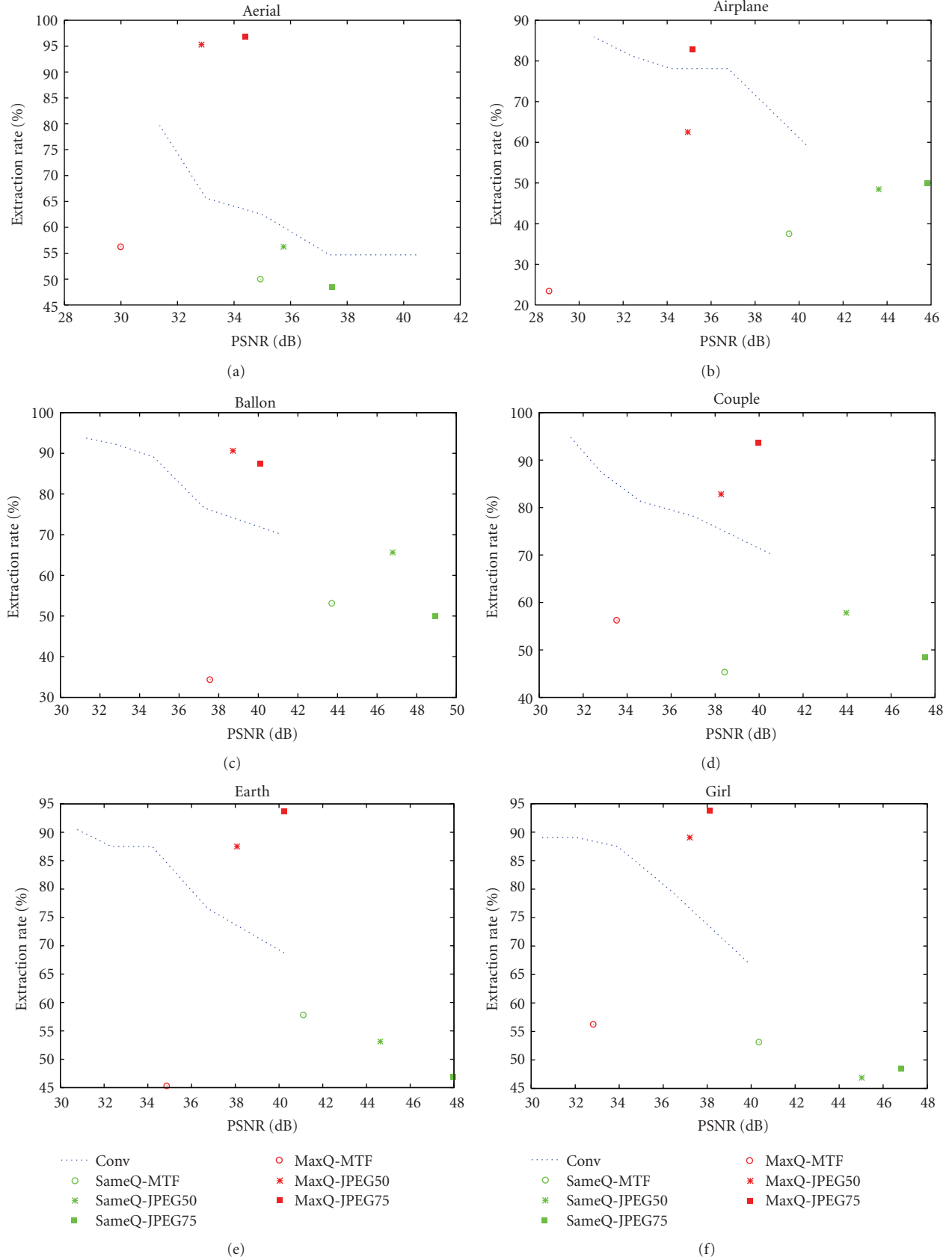
FIGURE 16: Comparison in the robustness against JPEG compression of quality 75% (1).
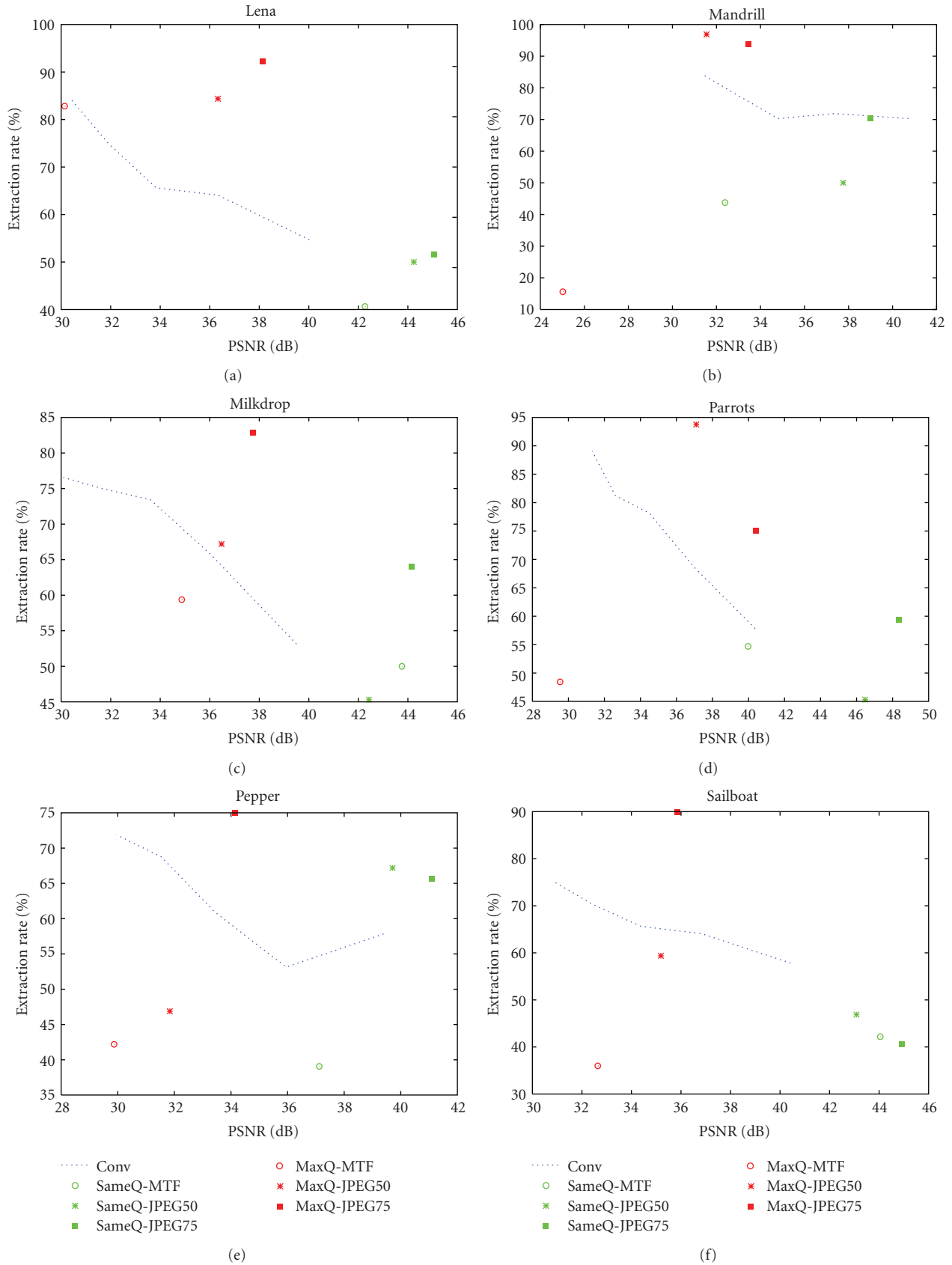
FIGURE 17: Comparison in the robustness against JPEG compression of quality 75% (2).
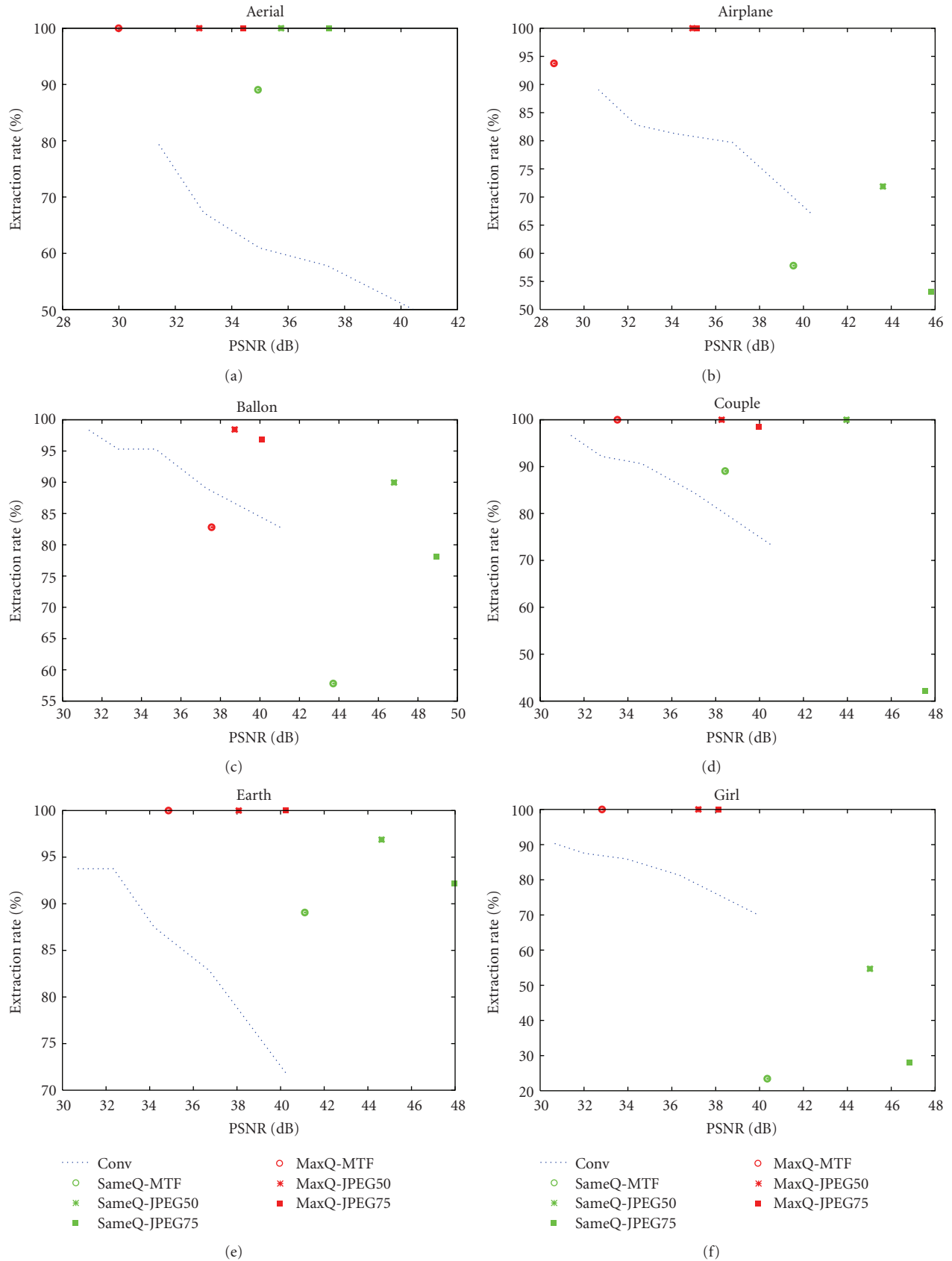
FIGURE 18: Comparison in the robustness against JPEG compression of quality 90% (1).
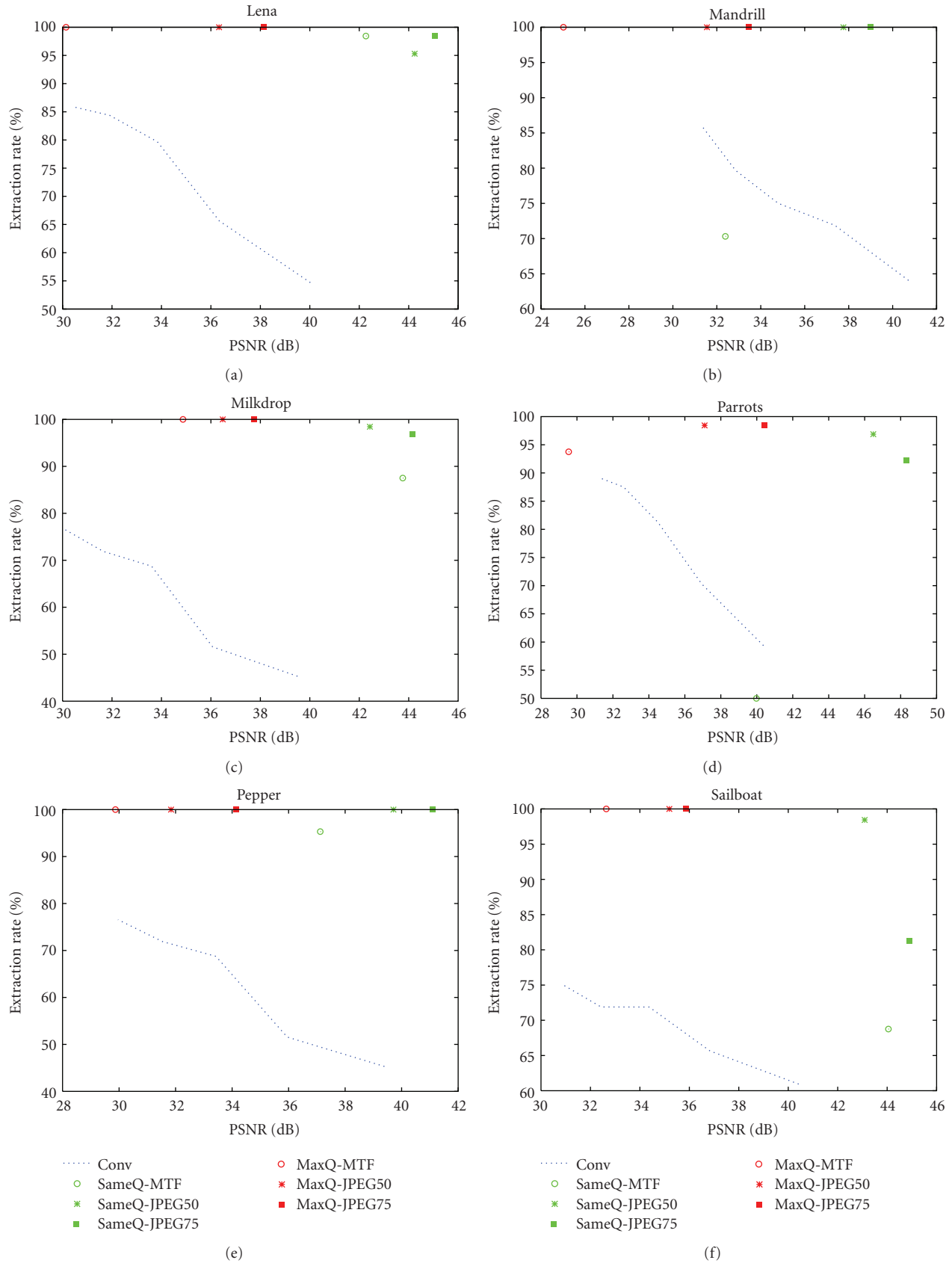
FIGURE 19: Comparison in the robustness against JPEG compression of quality 90% (2).

TABLE 5: MSSIM between watermarked images and lower limit images.

| | MTF | | JPEG50 | | JPEG75 | |
	sameQ	maxQ	sameQ	maxQ	sameQ	maxQ
aerial	0.9954	0.9972	0.9415	0.9514	0.9640	0.9702
airplane	0.9952	0.9980	0.9551	0.9710	0.9720	0.9841
balloon	0.9973	0.9986	0.9614	0.9765	0.9757	0.9855
couple	0.9881	0.9940	0.9603	0.9698	0.9713	0.9806
earth	0.9933	0.9965	0.9692	0.9776	0.9813	0.9868
girl	0.9739	0.9872	0.9550	0.9703	0.9684	0.9807
lenna	0.9750	0.9912	0.9562	0.9703	0.9738	0.9814
mandrill	0.9811	0.9911	0.9211	0.9449	0.9588	0.9696
milkdrop	0.9700	0.9840	0.9551	0.9650	0.9704	0.9768
parrots	0.9826	0.9933	0.9578	0.9733	0.9756	0.9831
pepper	0.9603	0.9782	0.9629	0.9733	0.9765	0.9819
sailboat	0.9889	0.9953	0.9662	0.9788	0.9802	0.9882

TABLE 6: Extraction rates in JPEG compression of quality 75%.

| | sameQ | | | maxQ | | |
	MTF	JPEG75	JPEG50	MTF	JPEG75	JPEG50
aerial	50.00	48.44	56.25	56.25	96.88	95.31
airplane	37.50	50.00	48.44	23.44	82.81	62.50
balloon	53.13	50.00	65.63	34.36	87.50	90.63
couple	45.31	48.44	57.81	56.25	93.75	82.81
earth	57.81	46.88	53.13	45.31	93.75	87.50
girl	53.13	48.44	46.88	56.25	93.75	89.06
lenna	40.63	51.56	50.00	82.81	92.19	84.38
mandrill	43.75	70.31	50.00	15.63	93.75	96.88
milkdrop	50.00	64.06	45.31	59.38	82.81	67.19
parrots	54.69	59.38	45.31	48.44	81.25	93.75
pepper	39.06	65.63	67.19	43.75	75.00	71.88
sailboat	42.19	40.63	46.88	35.98	89.06	59.38

TABLE 7: Extraction rates in JPEG compression of quality 90%.

| | sameQ | | | maxQ | | |
	MTF	JPEG75	JPEG50	MTF	JPEG75	JPEG50
aerial	89.06	100.00	100.00	100.00	100.00	100.00
airplane	57.81	53.13	71.88	93.75	100.00	100.00
balloon	57.81	78.13	89.06	82.81	96.88	98.44
couple	59.38	42.19	89.06	100.00	98.44	100.00
earth	89.06	92.19	96.88	100.00	100.00	100.00
girl	23.44	28.13	54.69	79.69	100.00	100.00
lenna	98.44	98.44	95.31	100.00	100.00	100.00
mandrill	70.31	100.00	100.00	100.00	100.00	100.00
milkdrop	87.50	96.88	98.44	100.00	100.00	100.00
parrots	50.00	92.19	96.88	93.75	98.44	98.44
pepper	95.31	100.00	100.00	100.00	100.00	100.00
sailboat	68.75	81.25	98.44	100.00	100.00	100.00

the extraction rates in JPEG compression of quality 75% and 90%, respectively.

As shown in Table 6, the proposed method using "sameQ" had no robustness against JPEG compression of quality 75%. Using "maxQ," some extraction rates of "JPEG75"and "JPEG50" against JPEG compression of quality 75% were larger than 90%. It was noticeable that some extraction rates of "JPEG75" were larger than those of "JPEG50" although the PSNRs of "JPEG75" were larger than those of "JPEG50." The investigation of the relationship between lower limit images and robustness is one of our future works.

As shown in Table 7, the proposed method using "sameQ" had partial robustness against JPEG compression of quality 90%. On the other hand, almost all the extraction rates using "maxQ" were equal to 100%. Therefore the proposed method using "maxQ" had the robustness against JPEG compression of quality 90%.

5.2.7. Comparison with Existing Method. We use the existing method proposed by Yoshiura and Echizen in the literature [2] for comparison. Yoshiura's method used the correlation of 2-dimensional random sequences which was one of popular watermarking procedures. Moreover Yoshiura's method took into consideration human visual system by using $L*u*v*$ color space which was one of uniform color spaces. Therefore Yoshiura's method was appropriate to the comparison.

Figure 14 shows the original image "mandrill" and the watermarked images of the existing method and the proposed methods using "sameQ-MTF" and "maxQ-JPEG50." The PSNRs of the watermarked images were approximately equalized as described in Figure 14. As shown in Figure 14, chromatic block noises were perceptible in the watermarked image of the existing method, while the degradation was imperceptible in the watermarked images of the proposed methods using "sameQ-MTF" and "maxQ-JPEG50" although the PSNRs of of the proposed methods were lower than the PSNR of the existing method. Figure 15 shows the enlarged partial regions of the images in Figure 14. As shown in Figure 15, the degradation of each watermarked image was able to be observed in detail. The degradation of the existing method was chromatic block noise. The degradation of the proposed method using "sameQ-MTF" was strong chromatic edge enhancement. The degradation of the proposed method using "maxQ-JPEG50" was imperceptible even if the partial region was enlarged. It was the reason why the degradation of the proposed method using "maxQ-JPEG50" was not block noise that the location of the pixels modified by embedding was scattered by (18) and (19).

Figures 16~19 show the comparison of Yoshiura's method and the proposed method using "MTF," "JPEG75" or "JPEG50" as the lower limit images and "sameQ", or "maxQ" as the embedding strength. The horizontal axis of the graphs in Figures 16~19 represents PSNR[dB] of watermarked images, while the vertical axis represents extraction rate[%]. In Figures 16~19, the performance of the proposed method is represented by the point for each condition, while that of

the existing method is represented by the curve. We evaluated the superiority of the proposed method by checking whether the point of the proposed method was above the curve of the existing method or not. As shown in Figures 16 and 17, only the point corresponding to "maxQ-JPEG75" was above the curve of the existing method for the results of all test images. Therefore the proposed method using "maxQ-JPEG75" was superior to the existing method for all test images in view of the robustness against JPEG compression of quality 75%. In comparison with each parameter of the proposed method, in the results of "balloon," "mandrill," and "parrots," the point corresponding to "maxQ-JPEG50" was located on the upper-left of "maxQ-JPEG75." The superiority of "maxQ-JPEG50" on the above cases would be decided depending on the importance of an extraction rate and a PSNR. As shown in Figures 18 and 19, the points corresponding to "maxQ-JPEG75" and "maxQ-JPEG50" were above the curve of the existing method for the results of all test images. Therefore the proposed method using "maxQ-JPEG75" or "maxQ-JPEG50" was superior to the existing method for all test images in view of the robustness against JPEG compression of quality 90%. Moreover the extraction rates of "maxQ-JPEG75" and "maxQ-JPEG50" for all test images were over 95%, where the errors could be recovered by using error correcting codes. In comparison with each parameter of the proposed method, the PSNRs of "maxQ-JPEG75" were higher than those of "maxQ-JPEG50" for all test images. From above discussion, the performance of "maxQ-JPEG75" was totally the best because of the imperceptibility shown in Figure 12 and the robustness against JPEG compression.

6. Conclusion

We have proposed a watermarking method warranting the lower limit of the image quality of watermarked images. The proposed method warrants the lower limit of the image quality of watermarked images by introducing lower limit images and equalizes the degradation by embedding on watermarked images by using L*a*b* color space. We have investigated the image quality of watermarked images, the validity of the lower limit images filtered by mathematical 2-dimensional human visual MTF model, the flexibility of the embedding strength, the performance using the maximum of Q of each block, the similarity between watermarked images and lower limit images, the robustness against JPEG compression, and the comparison with the existing method. Our future works should be to investigate the relationship between the robustness against general image processing and lower limit images and to improve the decision of the embedding strength for each block so as to improve the tradeoff of PSNR and an extraction rate.

References

[1] K. Matsui, *Fundamentals of Digital Watermarking*, Morikita Shuppan, 1998.

[2] H. Yoshiura and I. Echizen, "Maintaining picture quality and improving robustness of color watermarking by using human vision models," *IEICE Transactions on Information and Systems*, vol. E89-D, no. 1, pp. 256–270, 2006.

[3] "JIS handbook 61 Color 2007," Japanese Standards Association, 2007.

[4] Wikipedia, "Lab color space," November 2009, http://en.wikipedia.org/wiki/Lab_color_space.

[5] T. Oyama, *Invitation to Visual Psycology*, Saiensu-sha Co., 2000.

[6] "Colors & Dyeing Club in Nagoya," Osaka, November 2009, http://www005.upp.so-net.ne.jp/fumoto/.

[7] T. Ishihara, K. Ohishi, N. Tsumura, and Y. Miyake, "Dependence of directivity in spatial frequency responseof the human eye (1): measurement of modulation transfer function," *Journal of the Society ofPhotographic Science and Technology of Japan*, vol. 65, no. 2, pp. 121–127, 2002.

[8] T. Ishihara, K. Ohishi, N. Tsumura, and Y. Miyake, "Dependence of directivity in spatial frequency responseof the human eye (2): mathematical modeling of modulation transfer function," *Journal of the Societyof Photographic Science and Technology of Japan*, vol. 65, no. 2, pp. 128–133, 2002.

[9] Y. Miyake, T. Ishihara, K. Ohishi, and N. Tsumura, "Measurement and modeling of the two dimensionalMTF of human eye and its application for digital color reproduction," in *Proceedings of the 9th IS&T and SID Color Image Conference*, pp. 153–157, Scottsdale, Ariz, USA, 2001.

[10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.