quality index,
Human Visual System,
medical images

Agnieszka LISOWSKA*, Wiesław KOTARSKI**

THE UNIVERSAL QUALITY INDEX FOR MEDICAL IMAGES

The aim of this paper is to propose a new quality index which measures the distance between a reference (source) image and its corrupted copy in the way as Human Visual System (HVS) does. The new quality index called the Mean Weighted Quality Index (MW) is defined with the help of the well known easy calculated indexes. The experiments performed on a number of medical images confirmed usefulness of the new index.

1. INTRODUCTION

In many medical applications various images e.g. obtained from USG, CR, MRI are used. On the base of those images physicians put diagnosises and recommend what cure should be used. Moreover, observing small differences in a sequence of images, taken from the same patient in different time intervals, they can evaluate efficiency of medical treatment. It means that very important decisions in medicine are made upon images. Hence their quality should be highest possible. It also means that the knowledge about quality of images is one of the most important factor in proper medical diagnosis. Unfortunately, the images used there are often blurred and noised.

The notion of image quality is not obvious and quite clear. The image quality should be defined by quality indexes taking into account properties of the Human Visual System (HVS). In order to construct a good quality index one must understand how HVS works. It is known that human brain plays very important role in HVS. And, in relevance, three important features of HVS should be considered. The first one is that the background brightness of the focal area influents on the brightness of the focal area. The second one is that also surrounding contrast and texture variation of the focal area has enormous importance. And the last one is that human brain, at first look, remembers the shapes of the objects and the textures filling the shapes. In addition one should remember that human eye recognises images in logarithmic scale. Furthermore details are more visible in brightness then in darkness. Medical images are often dark and have black background. So small distortions, especially at the background, are not so visible by an observer.

* Institute of Mathematics, Silesian University, Bankowa 14, 40-007 Katowice, Poland
** Institute of Informatics, Silesian University, Będzińska 39, 41-200 Sosnowiec, Poland

More about HVS can be found in [3], [2] and the ample literature given there. Different image quality indexes are described e.g. in [4] and those basing on HVS in [3], [5].

2. MATHEMATICAL BACKGROUND

Recall the known definitions of image quality indexes. Consider a grayscale image of size $N \times M$ pixels, for which $f(x,y)$ denotes intensity value of the reference image pixel and $g(x,y)$ is intensity value of distorted image pixel; $f(x,y), g(x,y) \in \{0,\ldots,255\}$. The most popular quality index is the Peak Signal to Noise Ratio (PSNR)

$$I_{PSNR} = 10\log_{10}\frac{\sum_{x=1}^{N}\sum_{y=1}^{M}(\max f(x,y))^2}{\sum_{x=1}^{N}\sum_{y=1}^{M}(f(x,y)-g(x,y))^2},$$

where $\max f(x,y)$ often is taken as 255. This index is used frequently because of its simpicity. Similar to it is also Signal to Noise Ratio (SNR)

$$I_{SNR} = 10\log_{10}\frac{\sum_{x=1}^{N}\sum_{y=1}^{M} f^2(x,y)}{\sum_{x=1}^{N}\sum_{y=1}^{M}(f(x,y)-g(x,y))^2}.$$

But it is used rather rarely. The next indexes, Mean Difference (MD) and Mean Square Error (MSE), are similar and simply, though the MSE better reflects small distortions then MD. They are defined respectively by

$$I_{MD} = \sum_{x=1}^{N}\sum_{y=1}^{M}\frac{f(x,y)-g(z,y)}{MN}, I_{MSE} = \sum_{x=1}^{N}\sum_{y=1}^{M}\frac{(f(x,y)-g(z,y))^2}{MN}$$

In frequently use there are also Image Fidelity (IF)

$$I_{IF} = 1-\frac{\sum_{x=1}^{N}\sum_{y=1}^{M}(f(x,y)g(x,y))^2}{\sum_{x=1}^{N}\sum_{y=1}^{M} f^2(x,y)},$$

Normalized Cross Correlation (NCC)

$$I_{NCC} = \frac{\sum_{x=1}^{N} \sum_{y=1}^{M} f(x,y)g(x,y)}{\sum_{x=1}^{N} \sum_{y=1}^{M} f^2(x,y)}$$

and Structural Content (SC)

$$I_{SC} = \frac{\sum_{x=1}^{N} \sum_{y=1}^{M} f^2(x,y)}{\sum_{x=1}^{N} \sum_{y=1}^{M} g^2(x,y)}$$

The indexes mentioned above are rather typical and often used in experiments. However the results given by those indexes do not reflect the Human Visual System properly.

There are two other indexes, which are also worth to mention. The first one is the Universal Image Quality Index (Q) proposed by Zhon Wang and Alan C. Bovik in [5]. That index is defined as

$$Q = \frac{4\sigma_{fg}\bar{f}\,\bar{g}}{\left(\sigma_f^2 + \sigma_g^2\right)\left(\bar{f}^2 + \bar{g}^2\right)},$$

where

$$\bar{f} = \frac{1}{NM}\sum_{x=1}^{N}\sum_{y=1}^{M} f(x,y), \quad \bar{g} = \frac{1}{NM}\sum_{x=1}^{N}\sum_{y=1}^{M} g(x,y),$$

$$\sigma_f^2 = \frac{1}{NM-1}\sum_{x=1}^{N}\sum_{y=1}^{M}\left(f(x,y) - \bar{f}\right)^2, \quad \sigma_g^2 = \frac{1}{NM-1}\sum_{x=1}^{N}\sum_{y=1}^{M}\left(g(x,y) - \bar{g}\right)^2$$

$$\sigma_{fg} = \frac{1}{NM-1}\sum_{x=1}^{N}\sum_{y=1}^{M}\left(f(x,y) - \bar{f}\right)\left(g(x,y) - \bar{g}\right)$$

Unlike to the indexes mentioned above, this index is applied locally to an image using a sliding window approach. The authors recommend to perform the image analysis in a window size of 8×8 pixels to obtain the best results. Then the window is moved pixel by pixel from the left top corner of the image to its right bottom corner, and at the *j*–th step the Q_j is evaluated within the window. Next, after *K* steps the overall index is given by

$$I_Q = \frac{1}{K}\sum_{j=1}^{K} Q_j$$

The second quality index which reflects the HVS has been proposed by Ning Lu (see Chapter 13 in [3]). That index denoted by I_D is a very specific and cannot be expressed by a simple formula. To define I_D we need some auxiliary functions:
the focus area function

$$\Phi(x,y)=\begin{cases} 1 & \text{if } x=y=0, \\ \frac{1}{2} & \text{if } 0 \le |x|+|y| \le 2, \\ \frac{1}{4} & \text{if } |x|+|y|>2, |x|\le 3, |x|\le 3, \\ 0 & \text{otherwise,} \end{cases}$$

the intensity rescaling function

$$\Psi(x)=\begin{cases} 4x & \text{if } |x|<17, \\ 3x\pm 16 & \text{if } 17\le |x|<37, \\ 2x\pm 52 & \text{if } 37\le |x|<58, \\ x\pm 109 & \text{if } 58\le |x|<88, \\ \left[\frac{x}{2}\right]\pm 153 & \text{if } 88\le |x|<134, \\ \left[\frac{x\pm 1}{3}\right]\pm 175 & \text{if } 134\le |x|<200, \\ \left[\frac{x}{4}\right]\pm 192 & \text{if } |x|\ge 200, \end{cases}$$

and the difference function for any parameter $v\in\{0,1,...,255\}$

$$d_v(x)=\begin{cases} \frac{v}{255}\Psi\left(255-\left(\frac{x}{v}-1\right)\right) & \text{for } x\le v, \\ v+\frac{255-v}{255}\Psi\left(255\left(\frac{x-v}{255-v}\right)\right) & \text{for } x>v, \end{cases}$$

Then, at the point (x_0,y_0), we can define background brightness

$$B(x_O,y_O)=\frac{\sum_{x=1}^{N}\sum_{y=1}^{M} f(x,y)\Phi(x-x_O,y-y_O)}{\sum_{x=1}^{N}\sum_{y=1}^{M}\Phi(x-x_O,y-y_O)},$$

background variation

$$V(x_0,y_0)=\left(\frac{\sum_{x=1}^{N}\sum_{y=1}^{M}\Psi^2(f(x,y)-B(x_0,y_0))\Phi(x-x_0,y-y_0)}{\sum_{x=1}^{N}\sum_{y=1}^{M}\Phi(x-x_0,y-y_0)}\right)^{\frac{1}{2}}$$

spatial distortion (for any homeomorphism $\varepsilon=[0,N]\times[0,M]\rightarrow[0,N]\times[0,M]$)

$$E_\varepsilon(x_0,y_0)=\left(\frac{\sum_{x=1}^{N}\sum_{y=1}^{M}\left\|\varepsilon(x,y)-(x,y)\right\|\Phi(x-x_0,y-y_0)}{\sum_{x=1}^{N}\sum_{y=1}^{M}\Phi(x-x_0,y-y_0)}\right)^{\frac{1}{2}}$$

and vertical distortion of the image *g* from the image *f*

$$D_g^f(x_0,y_0)=\left(\frac{\sum_{x=1}^{N}\sum_{y=1}^{M}(d_{V(x_0,y_0)}(dst))^2\,\Phi(x-x_0,y-y_0)}{\sum_{x=1}^{N}\sum_{y=1}^{M}\Phi(x-x_0,y-y_0)}\right)^{\frac{1}{2}}$$

where

$$dst=|\Psi(g(x,y)-B(x_0,y_0))-\Psi(f(x,y)-B(x_0,y_0))|.$$

Distortion at the point (x_0,y_0) of the image *g* from the image *f* is defined by

$$Dist(x_0,y_0)=\min_{\varepsilon}\{\sqrt{(D_g^f(x_0,y_0))^2+(4E_\varepsilon(x_0,y_0))^2}\}$$

And finally, the quality index is

$$I_D=\frac{\sum_{x=1}^{N}\sum_{y=1}^{M}Dist(x,y)}{NM}$$

But the main disadvantage with the usage of that index is that computations take a long time. For example, it takes nearly 3 minutes on Athlon Thunderbird 1330 MHz processor to compute I_D for an image of size 256×256 pixels.

3. MEAN WEIGHTED QUALITY INDEX (MW)

We would like to propose a new index which measures quality of images and reflects the HVS. The main advantage of this index is that it is easy to compute and can be constructed using the other well known indexes. Our experiments showed that usage of I_{SC} and I_{NCC} indexes simulate HVS in an optimal way. The other indexes were much worse.

Consider a grayscale image of size $N \times M$ pixels and let us denote I_{SC}, I_{NCC} the quality indexes defined in the previous section. We have $I_{SC} \in [0, 255^2 MN]$ and $I_{NCC} \in [0,255]$. Then we can define the MW Index with weights w_1, w_2 as

$$I_{MW} = w_1 \left| I_{SC} - 1 \right| + w_2 \left| I_{NCC} - 1 \right| ,$$

or equivalently:

$$I_{MW} = w_1 \frac{\left| \sum_{x=1}^{N} \sum_{y=1}^{M} (f^2(x,y) - g^2(x,y)) \right|}{\sum_{x=1}^{N} \sum_{y=1}^{M} g^2(x,y)} + w_2 \frac{\left| \sum_{x=1}^{N} \sum_{y=1}^{M} (f(x,y)g(x,y) - f^2(x,y)) \right|}{\sum_{x=1}^{N} \sum_{y=1}^{M} f^2(x,y)}$$

Experimental results showed that good choise for w_1 and w_2 is w_1=0.9 and w_2=0.1. So I_{MW} equals

$$I_{MW} = 0.9 \left| I_{SC} - 1 \right| + 0.1 \left| I_{NCC} - 1 \right| .$$

Note, that for two identical images values of the indexes I_{SC} and I_{NCC} equal 1. So I_{MW}=0. Further, the larger difference between images implies the larger value of I_{MW}. The maximal value the index can take is $(255^2 MN-1)w_1+254w_2$. It is easy to see that I_{MW} depends on the size of the image and values of the weights. Despite of the upper bound is large enough, in practice the index takes values not larger than 1 (in comparison to similar images of size 256×256 pixels with weights w_1=0.9 and w_2=0.1). In addition one should remember that I_{MW} is not defined for completely black images (similarly to I_{SC} and I_{NCC}).

4. EXPERIMENTAL RESULTS

In our experiments we have investigated 11 grayscale images of size 256×256 pixels. Each image had 5 corrupted versions: 50%-gaussian noise, 10%-gaussian noise, blurring, one pixel corruption and small detail corruption (about 400 pixels). These corruptions have been added by Corel Photo-Paint program. One of the investigated original of good quality images obtained from [1] together with all its corruptions are presented in Figs. 1-6.

Our experiments have been organised in the following way. For one fixed reference original image we sorted all its corrupted images basing on our visual perception (that is the HVS). Then we compared succesively for each image values of all calculated indexes. Data gathered in Table 1 concern the sample image presented in Fig. 1 and its corrupted copies.

Analysing Table 1 it is easy to notice that our new index I_D is the closest to the HVS comparing with the other indexes. Similar results we have also obtained for all investigated images.

	50% Noise	Blurring	10% Noise	Detail	One Pixel	Oryginal
I_{MD}	0.09886	4.50848	1.86456	0.3441	0.00142	0.00000
I_{MSE}	3794.01810	42.93863	38.45776	7.63830	0.11855	0.00000
I_{PSNR}	12.33981	31.80232	32.28096	39.30084	57.39195	∞
I_{SNR}	9.03090	28.83661	29.31458	36.33670	54.42788	∞
I_{SC}	0.99719	1.00772	0.99929	1.00025	0.99999	1.00000
I_{NCC}	0.94368	0.99551	0.99460	0.99461	0.99437	1.00000
I_{IF}	0.88454	0.99869	0.99883	0.99977	0.99999	1.00000
I_Q	0.08891	0.55367	0.68890	0.99260	0.99996	1.00000
I_{ID}	10162	4665	4545	4553	4554	4554
I_{MW}	0.00816	0.00740	0.00118	0.00076	0.00057	0.00000

Tab.1. Distortions of the image "Blood"

5. CONCLUSIONS

Experiments showed that proposed quality index works quite well on medical images. So I_{MW} may be used, for example, in some control systems where there is a need to apply the quality index to quickly and automatically adjust the quality of the examined image (for example contrast and blurring) to get the best result. The I_{MW} also may be used to examine whether considered object on the image is patological or not.

Because not all medical images are grayscale ones, therefore sometimes one need to evaluate the quality index of 24-bit (or others) color images. There are some color models which one can use to evaluate the index. But, the most popular, RGB model is not good enough because the three chanels including red, green and blue are very correlated. So one should calculate the index for all three chanels. Experiments (not reported here) showed that we can apply the YIQ model and use only the luminance chanel for our computations, because the main information of the image is contained in luminance. So it is sufficent to evaluate the image quality.

So far, our experiments have been concentrated on a number of medical images. However much more images should be analysed to obtain statistical verification of the results reported here.

BIBLIOGRAPHY

[1] LEWANDOWSKI K., HELLMAN A.: *Atlas of HAEMATOLOGY*, Medical University of Gdańsk, Poland, Medyczne Wydawnictwa Multimedialne 2000, (http://www.hematologica.pl).

[2] LONCARIS S.: *A Survey of Shape Analysis Techniques*, pp. 1-46, 1998 (http://www.citeseer. nj.nec.com).

[3] LU N.: *Fractal Imaging*, Academic Press, 1997.

[4] SKARBEK W.: *Multimedia -- Algorytmy i standardy kompresji*, Akademicka Oficyna Wydawnicza PLJ, Warszawa, 1998.
[5] WANG Z., BOVIK A. C.: *A Universal Image Quality Index,* pp. 1-4, {IEEE Signal Processing Letters}, 2001.

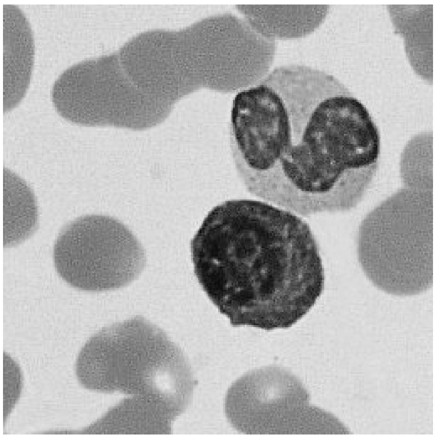

Fig.1. Original Image "Blood" (reproduced with permission from [1])

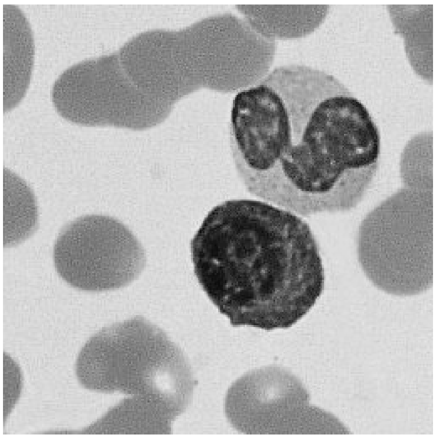

Fig.2. One Pixel Corruption

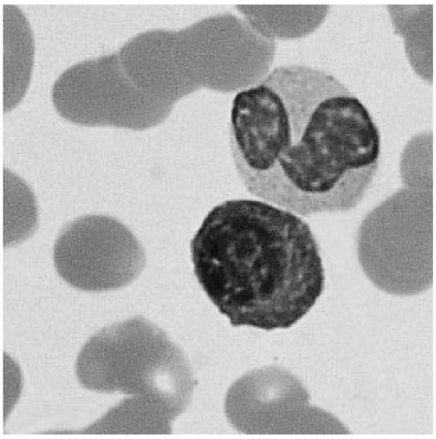

Fig.3. Patological Object Corruption

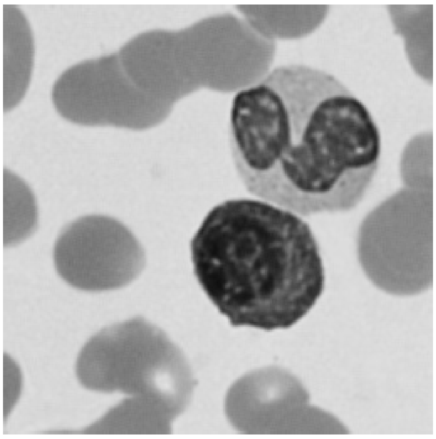

Fig.4 Blurring

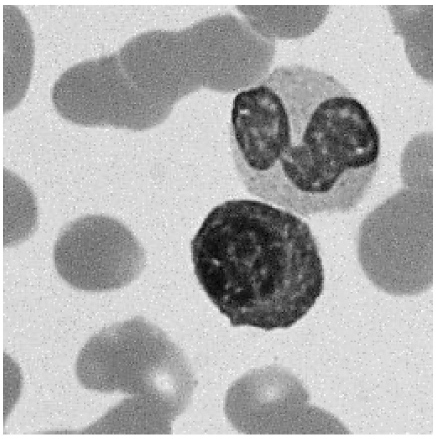

Fig.5. 10%-Gaussian Noise

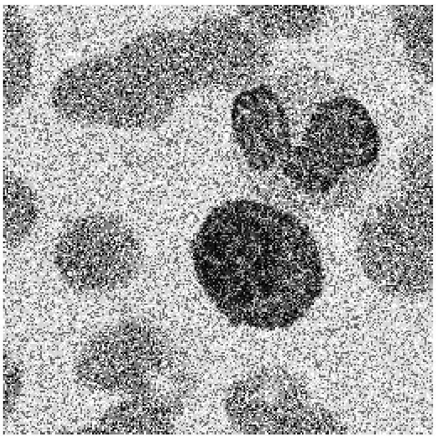

Fig.6 50%-Gaussian Noise