

# PROPER GUIDANCE IMAGE GENERATION BASED ON SALIENCY FACTOR FOR BETTER TRANSMISSION REFINEMENT IN IMAGE DEHAZING

Libao Zhang\* and Xiaohan Wang

College of Information Science and Technology, Beijing Normal University, Beijing 100875, China

## ABSTRACT

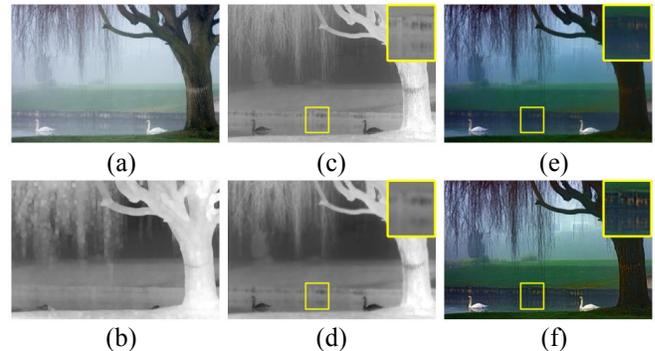
Guided image filter is one of the most commonly used ways to refine transmission maps. However, since this filter transfers the structures of the guidance image to the filtering output, when the guidance image is the input image itself, even small textures in the input image will cause the change of transmission, which is obviously contrary to the principle that transmission changes only when scene depth changes. In this paper, saliency detection, which simulates the way human eyes work, is introduced into haze removal to tackle the above issue. We first use saliency detection to capture the depth change regions, and then the saliency value is used as an adjustable factor to compute proper guidance images, in which most texture details are blurred but the depth change regions are remained clearly visible. Experimental results show that our method has great superiority in detail recovery compared with other state-of-art methods.

**Index Terms**—Image restoration, haze removal, saliency detection, transmission refinement, guidance image

## 1. INTRODUCTION

Common phenomenon such as haze and fog, usually cause outdoor images to suffer from bad visibility, which greatly limits the image application in the later stage. Therefore, haze removal has become the first and indispensable step when we deal with these degraded images.

Numerous valuable studies have been developed to remove haze component over the past decades. Early dehazing methods usually need multiple images or additional information of the same scene [1-4]. However, since this additional information is not available all the time, these dehazing methods are difficult to be widely used. Single image haze removal has drawn much attention in recent years. Fattal et al. removed the haze component by assuming that surface shading and scene transmission are locally uncorrelated [5]. Tan et al. assumed that haze-free images have higher contrast than images plagued by haze [6]. He et al. proposed dark channel prior (DCP) to estimate the



**Fig. 1.** Transmission refined by different guidance images and their corresponding dehazing results. (a) Input image. (b) Primary transmission map. (c) Refined transmission whose guidance image is the input image. (d) Refined transmission whose guidance image is computed by our proposed method. (e) Dehazing result recovered by (c). (f) Dehazing result recovered by (d).

transmission and airlight [7]. Enlightened by DCP, Tarel et al. proposed a median filtering based method to achieve fast haze removal [8]. Meng et al. used boundary constraint and contextual regularization to recover high-quality images with fine details [9]. Sulami et al. acquired more accurate results benefiting from a better estimation of the orientation and magnitude of atmospheric light [10].

Saliency detection is an attractive topic of region-of-interest extraction. Human visual system (HVS) is the basis of several saliency detection approaches, including Itti's model [11], and the context-aware saliency model [12]. Some purely computational models were also proposed, such as the spectral residual [13], and the bottom-up model based on wavelet transform [14]. Besides, some researchers tried to combine the biological and computational models [15, 16]. Recently, region-based saliency detection has been a valuable perspective. Yan et al. proposed a hierarchical saliency detection model [17]. Cheng et al. extracted the salient region by a region contrast based model [18].

A lot of dehazing methods can first roughly estimate the primary transmission maps, in which one famous example is dark channel prior [7]. For these dehazing methods, guided image filter (GIF) [19], a filter which would transfer the structure of guidance image to filtering output, is widely used in the next step. To make the refined

\*Corresponding author: libaozhang@163.com. This research was supported by National Natural Science Foundation of China (61571050 and 41771407) and Beijing Natural Science Foundation (L182029).

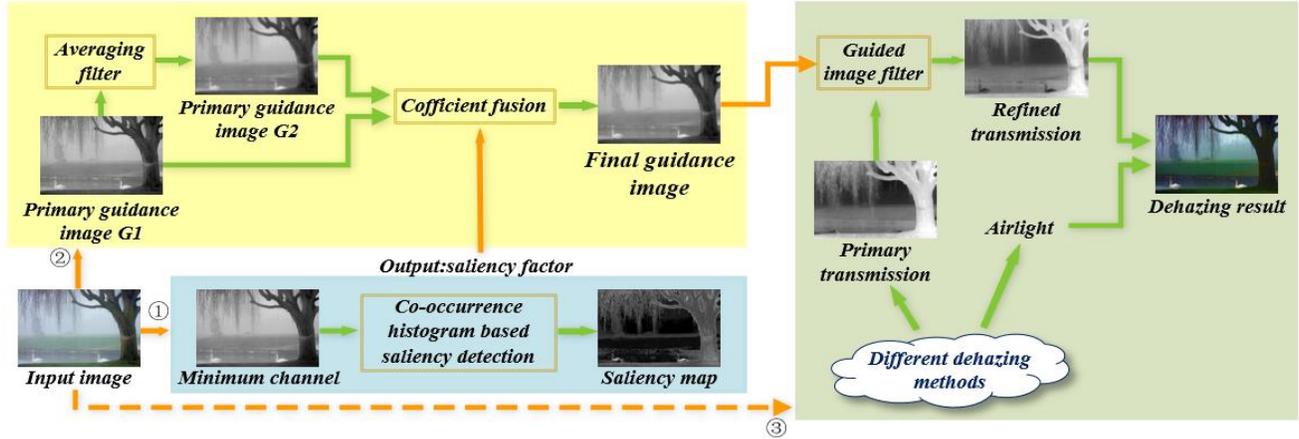


Fig. 2. The framework of the proposed method.

transmission maps match the input hazy images as exactly as possible, those input hazy images themselves tend to be chosen as the guidance images. However, as shown in Fig. 1, one problem is that these refined transmission maps have exactly the same structures and details with the input images, which means even small textures in the input image will cause the change of transmission, and this is obviously contrary to the principle that transmission changes only when scene depth changes. (Dark channel prior is used to obtain the primary transmission map in this paper.) The above problem directly leads to the loss of detail in the final dehazing results generated by this kind of methods.

In this paper, we introduce saliency detection to tackle the above issue. We first utilize saliency detection to capture the depth change regions. Then, the saliency value is used as an adjustable factor to generate proper guidance images. Figure 1 also shows our transmission and dehazing result. Compared with the results shown in Fig. 1 (c) and (e), our method suppresses the influence of texture information to the transmission map, and the final dehazing result shows more abundant details.

There are three contributions in this paper. (1) Saliency detection, which simulates the way human eyes work, is introduced into haze removal. (2) We use saliency detection to capture the depth change regions for hazy images. (3) The proposed method generates proper guidance images for better transmission refinement based on the saliency factor.

## 2. METHODOLOGY

In this section, our saliency detection method and the way to generate proper guidance images will be explained in detail. Figure 2 shows the framework of our proposed method.

### 2.1 The feature of proper guidance images

Generally, the formation of a hazy image can be described as the following model [7]:

$$I(x) = J(x) * t(x) + A(1 - t(x)) \quad (1)$$

where  $I$  is the observed intensity,  $J$  is the scene radiance,  $t$  is the medium transmission and  $A$  is the global atmospheric light. From Eq. (1), we can see that the key to recover haze-free images is to estimate correct transmission  $t$  and airlight  $A$  from the input hazy images.

Assuming that the atmosphere is homogenous, the transmission  $t$  can be further expressed as:

$$t(x) = e^{-\beta d(x)} \quad (2)$$

where  $\beta$  is the scattering coefficient of atmosphere, and  $d$  is the scene depth. This equation indicates that transmission is only dependent on scene depth when  $\beta$  remains unchanged.

As mentioned, the refined transmission map is actually contrary to the above principle when the guidance image is the input image itself. Since one direct reason is that the textures in the guidance image are transferred to the transmission map, a natural idea is to blur the textures but keep the depth change region in the guidance image. Considering that human eyes can successfully capture depth change regions while ignoring the small texture information, we introduce saliency detection, which simulates the way human eyes work, to fulfill the above task.

### 2.2 Co-occurrence histogram based saliency detection

Co-occurrence histogram is a two-dimensional histogram which can capture global and local distribution of intensity values. The main idea of using co-occurrence histogram to identify salient regions is to find the local pixel pairs that rarely appear [20, 21]. These pixel pairs either represent local discontinuity or have global uncommonness, thus they usually have high saliency. Under hazy weather, outdoor images are degraded by the turbid medium, and suffer from great loss of detailed texture. Meanwhile, since a large part of the obtained image information is determined by the thickness of haze, which is related to the scene depth, the received information of the objects with the same depth tends to have a certain similarity. Therefore, for hazy images, the pixel pairs that have low occurrence frequency usually exist in the abrupt depth change regions. Taking advantage



**Fig. 3.** Input hazy image and its saliency map. (a) Input hazy image. (b) Saliency map.

of the feature of hazy images, a co-occurrence histogram based saliency detection method is proposed, in which high saliency value represents depth change regions [22, 23].

For a single channel image  $Im$  with  $k$  possible intensities, the co-occurrence histogram  $COH$  can be expressed by [21]

$$COH = [coh(m, n)], \quad (3)$$

where  $COH$  is a symmetric square matrix of size  $k \times k$ . A  $COH$  element  $coh(m, n)$  can be computed as follows: For each pixel whose intensity is  $m$ , we count the number of pixels whose intensity is  $n$  in its local neighborhood window of size  $z \times z$  ( $z$  is set to 3 in this paper), and  $coh(m, n)$  is the sum of all these numbers. For the convenience of the follow-up calculation,  $COH$  is normalized by its maximum element:

$$NCOH = [ncoh(m, n)] = \frac{COH}{\max(coh(m, n))}, \quad (4)$$

where  $NCOH$  is the normalized co-occurrence histogram with each element between 0 and 1.

As mentioned, pixels pairs that rarely appear should have high saliency value, thus we first compute the primary saliency value of each intensity pair by:

$$s_p(m, n) = -\ln(ncoh(m, n)), \quad (5)$$

where  $s_p(m, n)$  represents the primary saliency value of the intensity pair which is composed of intensity  $m$  and intensity  $n$ . Considering that very close intensities cannot be clearly distinguished by human eyes, and in order to further suppress the saliency value of the intensity pairs that occur frequently, the final saliency value of each intensity pair  $s(m, n)$  is given by:

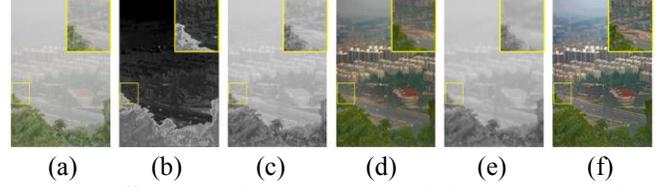
$$s(m, n) = \min_{(x, y) \in \Omega} s_p(x, y), \quad (6)$$

where  $\Omega$  is a local patch centered at  $(m, n)$  whose size is  $\omega \times \omega$  (we fix  $\omega$  to 10 in this paper). Ultimately, since there is no need to set saliency value for nonexistent intensity pairs,  $s(m, n)$  is changed to zero when  $coh(m, n)$  is equal to zero.

Finally, the saliency map of the input image can be obtained by computing the saliency value of each input pixel:

$$sa(i, j) = \sum_{i'=i-r}^{i+r} \sum_{j'=j-r}^{j+r} s(Im(i, j), Im(i', j')), \quad (7)$$

where  $sa(i, j)$  is the saliency value of the pixel whose coordinate is  $(i, j)$ , and  $r$  is the window radius which is used in computing each  $COH$  element.



**Fig. 4.** Different guidance images and their corresponding results. (a) Input image. (b) Saliency map. (c) Common guidance image. (d) Haze-free image recovered by (c). (e) Our guidance image. (f) Haze-free image recovered by (e).

Figure 3 shows an input image and its saliency map computed by our proposed method. From the result we can see that the depth change regions are marked as high saliency, while the texture part such as the leaves and trunk, has lower saliency value.

### 2.3 Saliency based guidance image generation

For each pixel, the minimum value in RGB channel is the upper limit of its thickness of haze. Hence, to better capture the depth change regions, i.e. the regions where the thickness of haze changes, a minimum channel image, which is generated by calculating the minimum value in RGB channel on each pixel, is used to compute saliency map.

For simplicity, the red channel of the input hazy image is used as one gray-scale primary guidance image  $G_1$ . In order to blur the textures, an averaging filter with a moving window is performed on  $G_1$ , which generates another primary guidance image  $G_2$ .

As mentioned, saliency value indicates depth change to some extent, thus we can use the saliency map  $sa$  as an adjustable factor to fuse  $G_1$  and  $G_2$  in order to blur most texture details while remaining depth change regions clearly visible. However, as shown in Fig. 3, the saliency map just marks the edges but not take into consideration the pixels that are close to these edges. Since these pixels are in the windows which contain depth change, they should also be less affected by the averaging filter. Thus, a refined adjustable factor  $F$  is defined as:

$$F(x) = \max_{y \in \Psi(x)} \frac{sa(y)}{d(x, y) + 1}, \quad (8)$$

where  $\Psi$  is a local patch centered at  $x$ , whose size is the same with the size of the averaging filter.  $d(x, y)$  is the normalized distance between  $x$  and  $y$ :

$$d(x, y) = \frac{E(x, y)}{\sigma}, \quad (9)$$

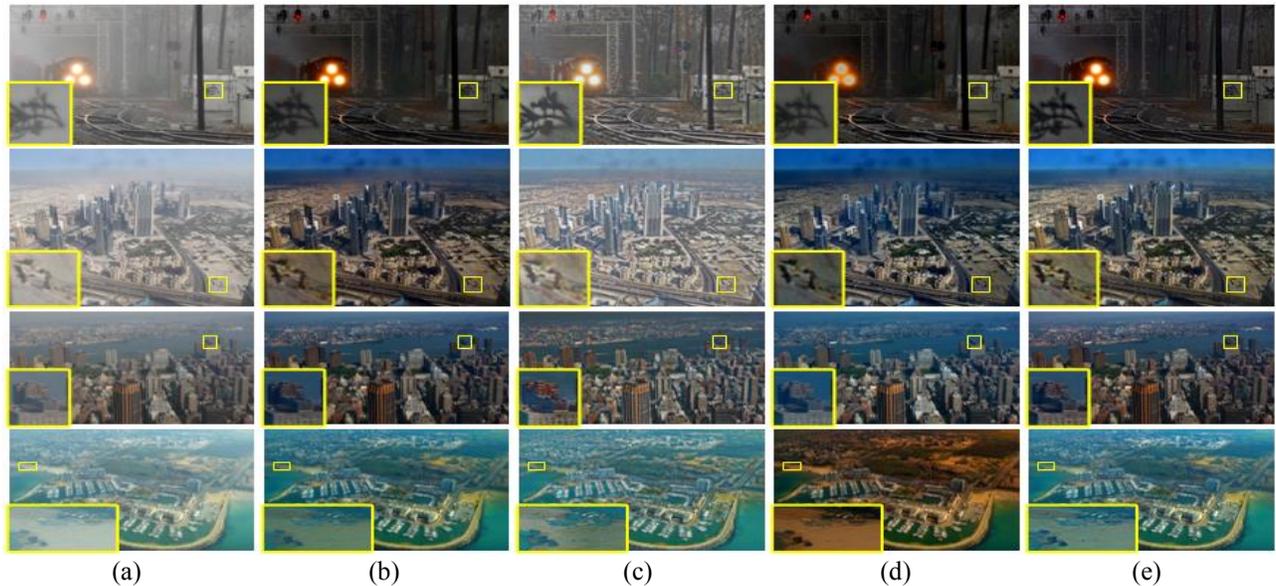
where  $E(x, y)$  is the Euclidean distance between  $x$  and  $y$ , and  $\sigma$  is the window radius of the averaging filter.

Ultimately, the final guidance image can be generated by:

$$G = F * G_1 + (1 - F) * G_2, \quad (10)$$

where  $G$  is the final guidance image.

Figure 4 gives an example of the common guidance image and our guidance image, and their corresponding dehazing results.



**Fig. 5.** Results comparison of different dehazing methods. (a) Input images. (b) He's results [7]. (c) Tarel's results [8]. (d) Sulami's results [10]. (e) Our results

### 3. EXPERIMENTAL RESULTS

In this section, we compare our results with the results of He [7], Tarel [8], and Sulami [10]. A comprehensive evaluation is made by both qualitative and quantitative analysis.

**Table 1.** Indicator  $e$  of the images in Fig. 5.

$e$	He	Tarel	Sulami	Our
First row	1.2548	1.5793	1.2682	1.5978
Second row	0.0708	0.2338	0.0685	0.0741
Third row	0.2520	0.4189	-0.2898	0.3178
Fourth row	0.4310	0.3439	0.4939	0.4890

**Table 2.** Indicator  $\bar{r}$  of the images in Fig. 5.

$\bar{r}$	He	Tarel	Sulami	Our
First row	1.0960	2.4996	1.2034	1.5588
Second row	1.3000	1.9031	1.3958	1.5084
Third row	1.2365	1.6627	2.9207	1.4736
Fourth row	1.4301	1.9807	1.5877	1.9263

**Table 3.** Indicator  $\sigma$  of the images in Fig. 5.

$\sigma$ (%)	He	Tarel	Sulami	Our
First row	0.4738	0	1.9300	0.3067
Second row	0.0023	0	0.0023	0.0113
Third row	0.0003	0	0.0424	0.0004
Fourth row	0.0285	0	0.0464	0.0064

#### 3.1 Qualitative comparison

The results of different methods are shown in Fig. 5. From the results shown in the first and second row, it can be seen

that our results have more concrete edges, while the details of other methods are relatively vague. Meanwhile, our method greatly reduces the detail loss resulting from color distortion in local region, which can be seen from the results shown in the third and last row.

#### 3.2 Quantitative comparison

We use the indicators proposed in [24] to evaluate the results of these four methods. Higher value of  $e$  in Table 1 represents higher rate of new visible edges. In Table 2,  $\bar{r}$  shows the quality of contrast restoration, and higher value denotes better contrast restoration. Table 3 shows the rate of saturated pixels  $\sigma$ . The smaller  $\sigma$  is, the better the method performs. Compared with He's method, i.e. dark channel prior with input hazy image as guided image, our proposed method has great advantage in both increasing visible edges and restoring image contrast. Compared with other state-of-art approaches, our method also gives comparable results in indicator  $e$  and  $\bar{r}$ . There is no evident superiority for our method in indicator  $\sigma$ . However, this is acceptable because we do not suppress oversaturated pixels explicitly in our method.

### 4. CONCLUSION

In this paper, we utilize a co-occurrence histogram based saliency detection method to capture the abrupt depth change regions in hazy images, then use the saliency value as an adjustable factor to generate proper guidance images for later transmission refinement step. Experimental results indicate that our method has great advantage in detail recovery compared with other state-of-art methods.

## 5. REFERENCES

- [1] Y. Y. Schechner, S.G. Narasimhan, and S.K. Nayar, "Instant dehazing of images using polarization," *2001 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 325-332, 2001.
- [2] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Polarization-based vision through haze," *Applied Optics*, vol. 42, no. 3, pp. 511-525, 2003.
- [3] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," *2000 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 598-605, 2000.
- [4] S.G. Narasimhan and S.K. Nayar, "Contrast restoration of weather degraded images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6 pp. 713-724, 2003.
- [5] R. Fattal, "Single image dehazing," *ACM Transactions on Graphics (TOG)*, ACM, vol. 27, no. 3, pp. 1-9, 2008.
- [6] R. Tan, "Visibility in bad weather from a single image," *2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-8, June 2008.
- [7] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no 12, pp. 2341-2353, Dec. 2011.
- [8] J. P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," *2009 IEEE International Conference on Computer Vision (ICCV)*, pp. 2201-2208, 2009.
- [9] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," *2013 IEEE International Conference on Computer Vision (ICCV)*, pp. 617-624, Dec. 2013.
- [10] M. Sulami, I. Glatzer, R. Fattal, and M. Werman, "Automatic recovery of the atmospheric light in hazy images," *2014 IEEE International Conference on Computer Photography (ICCP)*, IEEE, pp. 1-11, 2014.
- [11] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254-1259, 1998.
- [12] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-Aware Saliency Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, vol. 34, no. 10, pp. 1915-1926, 2012.
- [13] X. Hou and L. Zhang, "Saliency detection: A spectral residual approach," *2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp. 1-8, 2007.
- [14] N. Imamoglu, W. Lin, and Y. Fang, "A saliency detection model using low-level features based on wavelet transform," *IEEE Transactions on Multimedia*, vol. 15, no. 1, pp. 96-105, 2013.
- [15] L. Zhang, A. Li, Z. Zhang, and K. Yang, "Global and local saliency analysis for the extraction of residential areas in high-spatial-resolution remote sensing image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 7, pp. 3750-3763, 2016.
- [16] R. Achanta, S. Hemami, F. Estrada, and S. Susstrunk, "Frequency-tuned salient region detection," *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp.1597-1604, 2009.
- [17] J. Shi, Q. Yan, L. Xu, and J. Jia, "Hierarchical image saliency detection on extended CSSD," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 4, pp. 717-729, 2016.
- [18] M. Cheng, N. Mitra, X. Huang, P. Torr and S. Hu, "Global contrast based salient region detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569-582, 2015.
- [19] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, 2013.
- [20] S. Lu, C. Tan, J. Lim, "Robust and Efficient Saliency Modeling from Image Co-occurrence Histograms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 195-201, 2014.
- [21] L. Zhang, A. Li, "Region-of-Interest Extraction Based on Saliency Analysis of Co-Occurrence Histogram in High Spatial Resolution Remote Sensing Images," *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, vol. 8, no. 5, pp. 2111-2124, 2015.
- [22] L. Zhang, Q. Sun, "Saliency detection and region of interest extraction based on multi-image common saliency analysis in satellite images," *Neurocomputing*, vol. 283, pp. 150-165, 2018.
- [23] L. Zhang, J. Zhang, "A new saliency-driven fusion method based on complex wavelet transform for remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2433-2437, 2017.
- [24] N. Hautiere, J.P. Tarel, D. Aubert, and E. Dumont, "Blind contrast enhancement assessment by gradient ratioing at visible edges," *Image Analysis & Stereology Journal*, vol. 27, no. 2, pp. 87-95, 2008.