Design Engineering

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238

A Novel Image Enhancement Method based on Visual Saliency Detection

Xiaoguang Xu[1], Hongda Shen[2], Hai Li[1,*]
[1]College of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, P. R. China
[2]Dept. of Electrical Engineering, University of Alabama in Huntsville, Huntsville, 35816, USA
*Corresponding Author: Hai Li

Abstract:

We present a novel image enhancement method based on the visual saliency detection model. This is a two-stage method including visual saliency detection and salient detail enhancement. In general, visual salient regions of the image will be detected first. Then, unsharp masking technique with different enhancement weights is applied to salient and non-salient regions respectively. Both of the visual experimental results and objective measurement reveal that this proposed method can enhance the image with much less noise-amplification and over-enhancement problems.

Keywords*: Image enhancement, Adaptive unsharp mask, Visual saliency detection, Noise-resilient.*

I. INTRODUCTION

Image enhancement technique has attracted significant attention from the researchers and engineers for past decades because of the increasingly growing need for high-quality images in many computer vision applications: medical image analysis [1], video surveillance, face recognition [2] etc. Many image enhancement methods have been proposed to improve the image quality and reveal the hidden details so far. Histogram-based methods focus on the image statistics: pixel intensity distribution and then certain transfer functions can be applied to map the original distribution to a desired one, which actually changes the original values of each pixel. In return, contrast or detail enhancement can be expected. For instance, Histogram equalization (HE) [3] is the most well-known histogram-based contrast enhancement due to its effectiveness and simplicity. However, HE “enhances” the image noise as well as image detail and also creates certain artifacts as a result of over-enhancement, In contrast to HE, a local histogram-based method named, contrast limited adaptive histogram equalization (CLAHE) [4], was proposed to address noise amplification problem by using the local histogram in predefined non-overlapping image blocks whereas HE is normally considered a global transform. Though CLAHE alleviates the noise amplification, it still suffers from some unpleasant halos or

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238

blocking effect. In addition to histogram-based image enhancement methods, direct enhancement approaches use the direct expression of contrast to reconstruct a contrast-enhanced image from the enhanced contrast expression [5]. This type of algorithms succeeds in some applications. However, as we all know, there is no perfectly suitable contrast model for the images so far. Given this fact and the dependency on the contrast expression used in this type of algorithms, direct enhancement methods also have those noise amplification and over enhancement issues as same as histogram-based methods due to the inaccurate contrast model. Also, the computational complexity of this type of methods increases with the complexity of contrast definition.

Multiscale decomposition method has also been investigated recently and applied to enhance the medical images [6]. Generally, the image is decomposed into different frequency bands first. Then different enhancement schema will be applied to those frequency bands accordingly. The enhanced image can be reconstructed from those enhanced bands. Since image noise and fine details are taken care of respectively in a different frequency band, noise amplification and over enhancement is reduced after the reconstruction. But this type of methods has to adjust some parameter values to achieve the optimal performance. Furthermore, its complexity is relatively higher than the other methods since signal decomposition and reconstruction are introduced. Unsharp Masking (UM) [3] only decomposes the image into low and high-frequency subimages, which is considered as the simplest multiscale decomposition method. In this paper, we use an adaptive UM algorithm as the enhancement operator which will be discussed in the next section.

Human visual system (HVS) indicates that the image quality is highly correlated with image structural content. Most of existing image enhancement methods do not fully exploit human visual features. Recently, scientists have developed some quantitative models to detect the human visual salient regions of images. Salient regions mainly refer to those discriminative regions of the image, which are very distinguishable from their neighborhoods. Limited by this subjective nature of visual saliency, it was not very applicable in the computer vision area at its early stage. However, the object model makes it possible for visual saliency to be used in certain applications. Itti et al. [7] use low-level vision features to form a bottom-up model. Following their original work, many visual saliency models have been proposed in the recent decade from different angles: image entropy [8], local and global contrast [9, 10]. Furthermore, high-level features involving face detection [11] and image segmentation [12] are reported to build a saliency model. Besides, learning based models compute the saliency map using features learned from large-scale image databases. With the development of the eye tracker, eye tracking data has been investigated to build a saliency model [13]. Saliency model can help us distinguish the important regions with relevant image information from the irrelevant background regions. Originated from the natural correlation between the human visual system and image quality, we propose a novel image enhancement method based on the HVS. More specifically, we intend to introduce human visual saliency detection model to divide the image

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238

into salient regions against non-salient regions and enhance them using UM with different parameter settings.

The rest of the paper is organized as follows: In Section 2, we present our image enhancement algorithm. Experimental results and objective measurements are reported in Section 3. Finally, we make conclusions and give direction for the future research in Section 4.

II. THE PROPOSED ALGORITHM

As a well-known fact, there is always certain noise existing in one image. Traditional image Enhancement methods do not take existing image noise effect into account particularly. Therefore, they enhance the noise along with the enhancement of other important images structural content. In addition, some fine details of the image are enhanced to create unpleasant perceptual distortion or distraction in certain applications. Visual saliency (VS) detection technique models the HVS and detects the salient regions of the image objectively. Hence, the properties of VS make itself a good candidate for image enhancement. In this paper, our proposed image enhancement method will apply one VS detection method to find salient regions of the image. Once the salient regions are identified, they can be enhanced by UM with high gain weight while non-salient regions will be enhanced by UM with small enhancement weight. Fig.1 illustrates the basic procedure of our proposed algorithm. Note that the T in the figure refers to the empirical threshold which is predefined to determine whether the current pixel belongs to the salient region.

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238

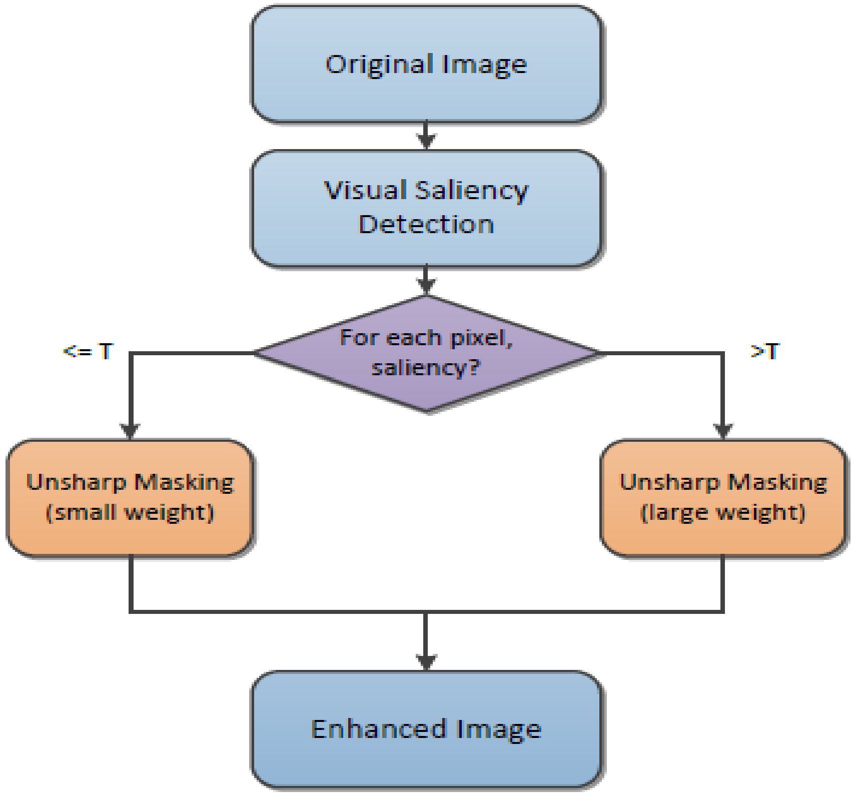


Fig 1: The proposed algorithm

2.1 Visual Saliency Detection

T Tremendous attention has been given to visual saliency detection for the past two decades. In order to model the HVS and measure the VS objectively and quantitatively, much effort has been made. In the paper [14], the author proposed a novel VS detection algorithm based on textural contrast. Our image enhancement method utilizes this detection because of its simplicity and its nature of high correlation with visual contrast. This VS detect on model has two major parts: luminance contrast and local structure descriptor. Luminance contrast can be computed as follows:

$$LC(i) = | I_{\mu} - \frac{1}{N} \sum_{j \in W_i} I(j) |^2 \qquad (1)$$

where I_{μ} denote the global luminance mean value for the entire image; W_i and N refer to local image patch centered at pixel *I(i)* and the patch size respectively. Eq. (1) actually computes the squared difference between global mean value and local patch mean value.

In HVS, it is widely accepted that human eyes are more sensitive to well-structured local image content i.e. edges, corners, lines and contours etc. Hence, local gradient tensor which models the dominant orientation and the energy along this direction is applied here. So the gradient tensor is given as:

$$GT(i)=\begin{bmatrix}\sum_{j\in W_i}G_x^2(j) & \sum_{j\in W_i}G_x(j)G_y(j)\\ \sum_{j\in W_i}G_x(j)G_y(j) & \sum_{j\in W_i}G_y^2(j)\end{bmatrix} \quad (2)$$

where G_x and G_y denote the horizontal and vertical gradient respectively. In order to find the dominant direction of the gradient tensor and the energy associated with it,we compute the eigenvalues of this gradient tensor matrix λ_1 and λ_2 .The eigenvectors, v_1 and v_2 , corresponding to the directions of maximum and minimum variation. This gradient tensor provides a good estimation of the underlying image structure. The local region containing the dominant structure is more likely to be salient given human eyes can perceive this structured region even under distortion. The concept of directional consistency is proposed to measure how intensively the local image patch pixels are distributed along the dominant direction. This directional consistency is computed as:

$$\phi(i)=(\lambda_1-\lambda_2)^2 \quad (3)$$

Hence, the higher the directional consistency is, the more the local patch is well-structured and salient. However, the exact symmetrical local patch may cancel out those oppositely oriented gradients to generate unreliable measurements. Hence, following the center-surrounding directional pattern, the difference of directional consistency between the center and surrounding pixels is defined as follows in [14]:

$$D(i)=\sum_{j\in W_i}|\phi(j)-\phi(i)| \quad (4)$$

In this fashion, this local structure descript can model the image structure of the central pixel against its neighborhood more effectively. One salient region should have high luminance contrast and discriminative directional structured pattern, so the final saliency measurement is determined by the product of these two terms.

$$S(i)=LC(i)\times D(i) \quad (5)$$

where *S(i)* denote the saliency measure at the current pixel *I(i)* while *LC(i)* and *D(i)* refer to luminance contrast and directional consistency at this pixel respectively. Gaussian filtering will be applied to smoothe the saliency map. The final saliency map will be normalized to [0 255] in the end for the further use.

2.2 Adaptive Unsharp Masking

We propose an adaptive unsharp masking (AUM) method to further enhance the image according to the visual saliency map generated from the previous stage. The traditional UM method applies low-pass filtering to the image I and subtract the filtered image I_{low} from the original image low to obtain the difference image I_{high}. Difference image will mainly contain majority of image high structural content including edges, lines and object boundaries etc. along with certain level of noise. Then, the difference image I_{high} will be multiplied by a scalar and added back to the high low-pass filtered image I_{low} to generate an enhanced image, which is

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238

described in the Eq.(6). However, as we discussed above, this method also enhances the perceptual noise since this type of noise is very likely contained in the high-frequency component of the image signal because of uniform enhancement.

$$E(i) = I_{low} + \alpha I_{high}, \alpha \in \Re^{+} \tag{6}$$

With the help of saliency map, we propose this method to enhance the salient regions with higher enhancement weight and enhance non-salient regions with much lower weight. α in the Eq.(6) denotes the enhancement weight. In our AUM, this weight is redefined as:

$$\alpha_i = \begin{cases} C_{max} & if \quad S(i) > T \\ C_{min} & if \quad S(i) \le T \end{cases} \tag{7}$$

where C_{max} and C_{min} refer to two constants with $C_{max} > C_{min}$. Threshold T is defined by the user to decide if this pixel belongs to the salient region or not. Salient regions with important structural details will be fully enhanced while non-salient regions will not be enhanced as much so that unpleasant artefacts and noise-amplification can be avoided as hoped.

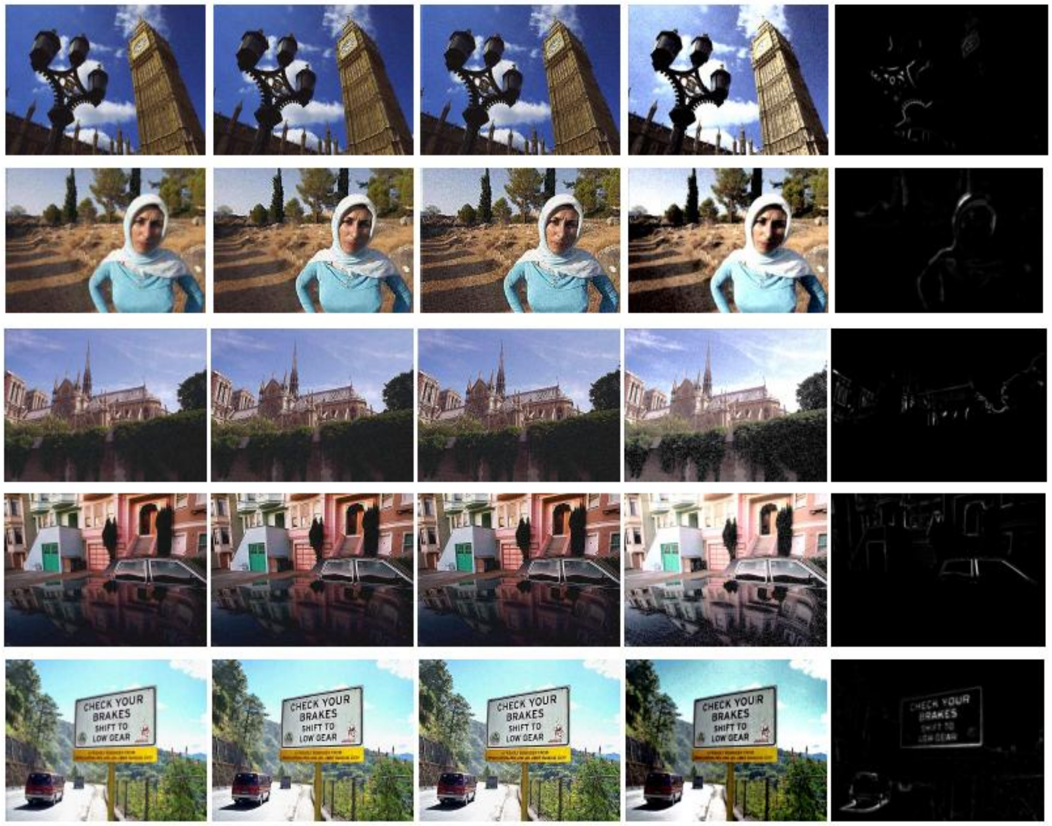
For this AUM, it is worth noting that this method is independent from the selection of low-pass filers. In our implementation, we choose the simple gaussian filtering for its simplicity and effectiveness. We believe other low-pass filters will produce similar results. Threshold T is a control parameter which gives the users the freedom to adjust the algorithm performance according to the specific application.

III. EXPERIMENTAL RESULTS

In this section, we present our experimental results of testing this proposed image enhancement algorithm on the online collected natural image database. Due to the limited space and relatively large image size, we only randomly select some images from this database to present here. In order to simulate the pictures of low quality in the real world, blurring effect and additive white Gaussian noise are added to those images manually. Our visual saliency-based image enhancement algorithm will be compared with UM and HE methods. Following the visual experiment, BRISQUE [15], a newly developed referenceless image quality measurement metric are applied to evaluate these three algorithms objectively.

Prior to the analysis of the experimental results, parameter settings are given here: T, saliency threshold, is defined as the global mean value of the saliency map, i.e. pixels with saliency values larger than this threshold T are enhanced as salient pixels while others are enhanced as non-salient pixels. Enhancement weights are $C_{max} = 5$ and $C_{min} = 2$ for salient and other regions respectively. Once we discriminate the salient pixels from non-salient pixels, AUM with two enhancement weights is applied separately. Note that the size of the local image patch in our implementation is fixed at 5 × 5.

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238



Original Proposed UM HE Saliency Map

Fig 2: Comparative Visual Experimental Results

To fully present the superiority of our proposed algorithm, Gaussian blur and *w*hite Gaussian noise effects (variance = 5) are added manually to those test images as mentioned earlier. Our proposed method, UM and HE methods are applied to those test images. Fig 2 shows the visual experiment results. Note that these five test images are of different size so we adjust them only for display purpose. In order to compare the details of the images, a zoom-in operation is highly recommended. The first column in the Fig 2 shows those test images with different distortions while Column 2 - 4 show the enhanced images from our method, UM and HE respectively. Column 5 shows the VS maps generated in our method to indicate where we enhance the most. It is very obvious that HE method creates unnatural looking for all these five test images in spite of the expected contrast enhancement. Also, the additive noise has been amplified with a certain level of over-enhancement, especially in the non-salient regions. UM method enhances the contrast without creating any obvious unnatural artefacts in general. However, UM method suffers the noise amplification and over-enhancement to some extent as well as HE. Moreover, similar to HE, UM creates some unnatural artefacts in the non-salient regions in almost all the test images. Compared to HE and UM, salient regions are enhanced more than non-salient regions in our proposed method so that existing additive noise is much

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238

less amplified. Besides, over-enhancement has been avoided to the most extent due to the signal adaptive weights applied in the proposed method. It is worth pointing out that our method depends on accurate saliency detection. In other words, if saliency detection does not work well, our method may not produce enough enhancement as we expect, for instance, the face region in the test image “Women” and also the words on the yellow board of test image “Sign”. In general, this saliency detection model applied in our proposed method produce very dependable results to produce good enhancement. The saliency maps shown in the Fig. 2 provide a good understanding of the image structure and capture almost all the major salient objects, which leads to an effective adaptive technique to enhance the image and alleviate the issues of those traditional methods.

In addition to visual experiment, we have conducted a quantitative test to verify what we have observed in the visual experiment. Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [15] is chosen in this work to measure the image quality. BRISQUE is newly proposed no-reference image quality assessment model which uses scene statistics of locally normalized luminance coefficients to quantify potential losses of ‘naturalness’ in the image due to the presence of various distortions [15]. Therefore, we believe BRISQUE can produce a holistic measure of natural image quality due to its inherent correlation with human visual perception. Note that BRSIQUE score has been normalized between 0 and 100: the less BRISQUE score is, the better image quality it indicates. All these properties of BRISQUE make it a good tool for this quantitative experiment.

TABLE I. BRISQUE score

	Big Ben	**Women**	**Notredam**	**Car**	**Sign**
Original	11.00	7.68	7.64	14.15	16.58
UM	21.47	5.61	17.48	8.02	14.30
HE	12.88	6.59	14.69	12.61	15.66
Proposed	6.73	0.32	1.13	5.22	9.87

TABLE I shows the BRISQUE scores of five test images processed by our method, UM and HE respectively and the best scores are highlighted. In most of the cases, the proposed method shows its superiority against UM and HE has given its overall BRISQUE score is lower than others. However, given the fact that there is no a perfect no-reference image quality estimator so far, this BRISQUE score can not objectively quantify how better/worse one algorithm is compared with others but can only indicate whether it is better or worse in general. Overall, this BRISQUE test result matches what we have noticed in the subjective visual experiment and verifies the superiority of our proposed method on image enhancement especially for those images with various type of distortions.

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238

IV. CONCLUSION AND FUTURE WORK

The proposed visual saliency-based image enhancement method utilizes the visual saliency detection model to detect the salient regions of the image and applies UM with adaptive weights accordingly. The experimentations provided in this paper show considerable improvements in the image quality and enhanced salient details over traditional methods. Structural information of the image is significantly enhanced while irrelevant details and noise are suppressed to avoid unpleasant artefacts. In addition, over-enhancement has also been reduced to its minimum, which produces better quality in the background regions of the image. In the future, multiscale adaptive UM can be applied to assist the enhancement from different frequency bands. Also, enhancement weights used in the AUM will be designed according to the actual image structural or content to further improve the image quality. Moreover, more tests using our proposed algorithm will be conducted on the images taken in the severe environment, i.e., higher noise level, unbalanced lighting and possible shaking effects.

ACKNOWLEDGEMENT

This work is supported by Wuhu Science and Technology Plan Foundation Projects (2020cg12) and Anhui Polytechnic University Scientific Research Project Funding (xjky02201909);

REFERENCES

[1] P. C. Tay, H. Shen, R. D. Adams, J. Z. Zhang. (2013) A microcalcification enhancement method for mammogram images, in: Proc. IEEE Intl. Conf. on Image Process. 2013: 1232-1236.
[2] S. Du, R. Ward. (2006) Adaptive region-based image enhancement method for face recognition under varying illumination conditions, in: Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Process. 2: 2
[3] R. C. Gonzalez., R. E. (2006) Woods, Digital Image Processing (3rd Edition), Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
[4] A. Boschetti, N. Adami, R. Leonardi, M. Okuda. (2010) High dynamic range image tone mapping based on local histogram equalization, in Proc. IEEE Intl. Conf. on Multimedia and Expo (ICME) 2010:1130-1135.
[5] S. C. Nercessian, K. A. Panetta, S. S. Agaian. (2013) Non-linear direct multi-scale image enhancement based on the luminance and contrast masking characteristics of the human visual system, IEEE Trans. Image Process., 22(9): 3549-3561.
[6] X. Liu, J. Tang, X. Zhang. (2009) A multiscale image enhancement method for calcification detection in screening mammograms, in: Proc. IEEE Intl. Conf. on Image Process. 2009: 677-680.
[7] L. Itti, C. Koch, E. Niebur. (1998) A model of saliency-based visual attention for rapid scene analysis, IEEE Trans. Pattern Anal. Mach. Intell., 20(11): 1254-1259.
[8] T. Kadir, M. Brady, Saliency. (2001) Scale and image description, Intl. J. Computer Vision 45(2):83-105, 2001.
[9] M. Cheng, N. J. Mitra, X. Huang, P. H. S. (2015) Torr, and S. Hu, Global contrast based salient region detection, IEEE Trans. Pattern Anal. Mach. Intell. 37(3):569-582.

ISSN: 0011-9342
Issue: 11 | Pages: 229 - 238

[10] F. Perazzi, P. Krahenbuhl, Y. Pritch, A. Hornung. (2012) Saliency filters: Contrast based filtering for salient region detection, in Proc. IEEE Intl. Conf. on Computer Vision and Pattern Recognition 2012: 733-740.

[11] S. Goferman, L. Zelnik-Manor, A. Tal. (2010) Context-aware saliency detection, in Proc. IEEE Intl. Conf. on Computer Vision and Pattern Recognition 2010: 2376-2383.

[12] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H. Shum. (2011) Learning to detect a salient object, IEEE Trans. Pattern Anal. Mach. Intell. 33(2): 353-367.

[13] T. Judd, K. Ehinger, F. Durand, A. Torralba. (2009) Learning to predict where humans look, in Proc. IEEE Intl. Conf. on Computer Vision 2009: 2106-2113.