A VISUAL SALIENCY MODULATED JUST NOTICEABLE DISTORTION PROFILE FOR IMAGE WATERMARKING

Yaqing Niu[*1,4], *Matthew Kyan*[2], *Lin Ma*[3], *Azeddine Beghdadi*[1] *and Sridhar Krishnan*[2]

[1,] L2TI lab, Institut Galilée, Université Paris 13, France
[2,] Department of Electrical and Computer Engineering, Ryerson University, Canada
[3,] Department of Electronic Engineering, The Chinese University of Hong Kong, China
[4,] Information Engineering School, Communication University of China, China
phone: + 33 1 49 40 40 57, fax: + 331 49 40 40 61, email: yaqing0930@gamil.com

ABSTRACT

Previous perceptual watermarking schemes only partially used the results from human visual system (HVS) studies. The perceptual adjustment of the watermark is mainly based on different visual sensitivity models. Numerically, visual sensitivity can be regarded as the inverse of the just noticeable distortion (JND). Another aspect affecting human perception towards visual signal is visual attention which can enhance or reduce the actual visual sensitivity and consequently the JND profile needs to be adjusted. The technique described in this paper assists image watermarking by producing a visual saliency modulated JND profile that can be used as a guide to optimize image watermarking. Experimental results with subjective test confirm the improved performance of our visual saliency modulated JND profile for image watermarking. Our saliency modulated JND profile is capable of shaping lower injected-watermark energy onto more sensitive regions and higher energy onto the less perceptually significant regions in the image, which yields better visual quality of the watermarked image.

1. INTRODUCTION

Digital image watermarking has emerged as a solution to the problem of copyright protection in the past decade. The strength of the watermarked signal is bounded by perceptual visibility. Thus, In order to maintain the image quality and at the same time increase the probability of the watermark detection, it is necessary to take the characteristic of human visual system (HVS) into consideration when engaging in watermarking research [1]. Visual sensitivity refers to the ability of human observers to detect distortion in visual field. Numerically, visual sensitivity can be regarded as the inverse of the just noticeable distortion (JND), which refers to the maximum distortion thresholds in pixels or subbands that the HVS does not perceive. Another aspect affecting human perception towards visual signal is visual attention, which can enhance or reduce the actual visual sensitivity and consequently JND profile needs to be adjusted inside and outside the salient regions in images.

Two psychophysical concepts are referred in this paper: JND and visual saliency. The former tells us how much distortion we can tolerate and the latter expresses where our visual attention is the most attracted. The technique described in this paper assists image watermarking by producing a visual saliency modulated JND profile that can be used as a guide to optimize watermarking. In this work, the saliency modulated JND profile is proposed to reflect the modulatory aftereffects of visual saliency on JND profile, which can be used as the scene-adaptive upper bounds on watermark insertion. To obtain visual saliency modulated JND profile, we need the JND value of the image. We also need the factor of visual saliency, which tells us areas of importance in the scene. From the previous work we have proposed an effective combined JND model [2] [26]. In this paper we modify the combined JND model [2][26] by the modulation of the latest efficient works on visual saliency detection [3]. Experimental results confirm the improved performance of our saliency modulated JND profile. Our saliency modulated JND profile is capable of shaping lower injected-watermark energy onto more sensitive regions and higher energy onto the less perceptually significant regions in the image, which yields better visual quality of the watermarked image. This paper is organized as follows. The next section reviews the related work on existing models for JND estimation, relevant exploration of visual saliency detection, current techniques combining visual saliency with JND profile, and JND based perceptual image watermarking. In Section 3, the proposed saliency modulated JND profile is presented. And in Section 4, a series of experiments are done to test the new saliency modulated JND profile's performances. Lastly, conclusions are drawn in Section 5.

2. RELATED RESEARCH WORK

2.1 JND Estimations

JND estimation for still images has been relatively well developed. An early perceptual threshold estimation in DCT domain was proposed by Ahumada [4]. This scheme was improved by Watson [5] after the luminance adaptation effect had been added to the base threshold, and contrast masking [6] had been calculated as the elevation factor. In [7] an additional block classification based contrast masking and luminance adaptation was considered by Zhang for digital images. A spatial JND model proposed by Zhenyu Wei [12] incorporates new spatial CSF, luminance adaptation and contrast masking. Since motion is a specific feature of videos, JND estimation for video sequences need to incorporate not only the spatial CSF, but the temporal CSF as well. A spatio-temporal CSF model was proposed by Kelly [8] from experiments on visibility thresholds under stabilized viewing conditions. Daly [9] extended Kelly's model to fit unconstrained natural viewing conditions with a consideration of eye movements. Based on Daly's model, Jia [10] estimated the JND thresholds for videos by combining other visual effects such as the luminance adaptation and contrast masking. An improved temporal modulation factor proposed by Zhenyu Wei [11][12] incorporates not only temporal CSF, but the directionality of motion is also considered. From the previous work, we have proposed an effective combined JND model [2][26] for image watermarking. We have also proposed a combined video-driven JND profile for video watermarking [13][14][27].

2.2 Visual saliency detection

Tresiman [17] proposed a theory which describes that visual attention has two stages. First, the parallel, fast, but simple pre-attentive process; and then, the serial, slow, but complex attention process [18]. The ability of human visual system to detect visual saliency is extremely fast and reliable. However, computational modeling of this basic intelligent behaviour still remains a challenge. Several computational models have been proposed to simulate human's visual attention. Itti et al. proposed a bottom-up model and built a system called Neuromorphic Vision C++ Toolkit (NVT) [19]. After that, following Rensink's theory [20], Walther extended this model, successfully applied it to object recognition tasks and created SaliencyToolBox (STB) [21]. Recently a simple and fast approach based on Fourier transform called spectral residual (SR) was proposed, which used SR of the amplitude spectrum to obtain the saliency map [5].

2.3 Combining Visual saliency with JND

More computational resource of the human brain is allocated to high attentional areas than low attentional areas, and this is the reason of visual saliency's modulation on visual sensitivity in different areas [22]. The saliency map is designed to reflect the statistical allocation of the human brain's processing resource on local visual contents. Visual saliency modulates all levels of visual perception, including visual sensitivity. Visual saliency can enhance or reduce the actual visual sensitivity. Consequently JND profile needs to be adjusted inside and outside of the salient areas in images. In [23] a computational model is proposed for incorporating a selectivity measure into the JND profile. In [22] perceptual quality significance map is proposed to reflect the modulatory aftereffects of visual attention on visual sensitivity and visual quality evaluation.

2.4 Perceptual image watermarking

Previous perceptual image watermarking researches have only partially used the results of the HVS studies [1][22][23][24][25]. An image-adaptive watermarking procedure based on Watson's spatial JND model was proposed in [23]. In [1], the DCT-based watermarking approach uses Watson's spatial JND model in which the threshold consists of spatial frequency sensitivity, luminance sensitivity and contrast masking. An Energy Modulated Watermarking Algorithm Based on Watson's spatial JND model was proposed in [22]. The main drawback of using Watson's visual models for images watermarking is that it does not satisfactorily provide the maximum strength transparent watermark. The obtained watermark is not optimal in terms of imperceptibility and robustness. We have proposed an effective combined JND model guided image watermarking scheme in DCT domain [2][26].

3. VISUAL SALIENCY MODULATED JND PROFILE

Saliency modulated JND profile is an efficient measurement incorporating visual attention's influence on visual sensitivity of the human eye to represent the additional accurate perceptual redundancies for image. Here we compute the visibility threshold of each DCT coefficient with the saliency modulated JND profile which is illustrated by Fig 1.

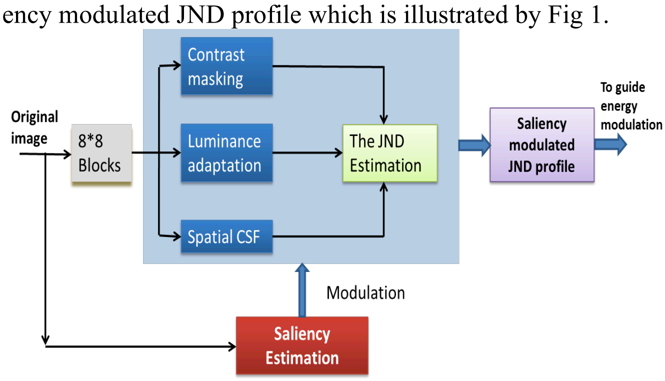


Figure 1. Diagram of saliency modulated JND profile

In the following, we proposed the effective saliency modulated JND profile in DCT domain. The JND estimation of image has been developed in our previous work [2][26]. The saliency estimation is described in [3].

Visual saliency modulates all levels of visual perception, including visual sensitivity. Visual saliency can enhance or reduce the actual visual sensitivity. Consequently JND profile needs to be adjusted inside and outside of the salient areas in images. The visual sensitivity is enhanced on salient spatial locations, due to the aftereffects of visual attention. The visibility thresholds in the salient areas are lower than the other nonattentional areas. With higher visual saliency value, the turning point of spatial contrast masking curve is pushed to higher frequencies, and the luminance adaptation tolerances are reduced. Itti *et al.* reported that visual attention could elevate the sensitivity with spatial and temporal frequencies by 30%, the sensitivity with orientations by 40%, and the sensitivity peak altitude by 5.2 dB [28].

The saliency modulated JND Profile $T_{JND}^{M}\left(n,\omega_{ij}^{M},i,j\right)$ can be expressed as (1)

$$T_{JND}^{M}\left(n,\omega_{ij}^{M},i,j\right)=T_{JND}\left(n,\omega_{ij}^{M},i,j\right)\times f_{T}^{M}\left(i,j\right) \quad (1)$$

$$\omega_{ij}^{M}\left(i,j\right)=\omega_{ij}\left(i,j\right)\times f_{\omega_{ij}}^{M}\left(i,j\right) \quad (2)$$

$$f_{T}^{M}\left(i,j\right)=\begin{cases} M_{Lum}^{\min}, & s_{block}\geq S_{\max} \\ 1-(s_{block}-0.1)\times\Phi_{Lum}, & S_{\min}\prec s_{block}\prec S_{\max} \\ M_{Lum}^{\max}, & s_{block}\leq S_{\min} \end{cases} \quad (3)$$

$$f_{\omega_{ij}}^{M}\left(i,j\right)=\begin{cases} M_{\omega}^{\max}, & s_{block}\geq S_{\max} \\ 1+(s_{block}-0.1)\times\Phi_{\omega}, & S_{\min}\prec s_{block}\prec S_{\max} \\ M_{\omega}^{\min}, & s_{block}\leq S_{\min} \end{cases} \quad (4)$$

Where $T_{JND}\left(n,\omega_{ij}^{M},i,j\right)$ is the JND estimator yielded by substituting ω_{ij} for ω_{ij}^{M} in [2][26], *n* is the index of a block in the image, *i* and *j* are the DCT coefficients' indices, ω_{ij}^{M} is modulated ω_{ij} by (2), $f_{T}^{M}\left(i,j\right)$ and $f_{\omega_{ij}}^{M}\left(i,j\right)$ are the corresponding modulation functions by (3) and (4) where s_{block} is the normalized block saliency value by [3], Φ_{Lum} and Φ_{ω} are factors to control the slope for each modulation which are set to 0.12 and 0.14, respectively. Based on our experiments, $M_{Lum}^{\max}$, $M_{Lum}^{\min}$ $M_{\omega}^{\max}$ and $M_{\omega}^{\min}$ are set to 1.0072, 0.9822, 1.0208 and 0.9916, respectively. Itti's experimental results [28] have been used in determining the actual maximum and minimum values of each modulation function.

4. EXPERIMENTAL RESULTS

We performed the experiments in Section 4.1 and 4.2 to evaluate the performance of saliency estimation and the proposed computational model for saliency modulated JND profile, respectively.

4.1 Evaluating saliency estimation

In this experiment, we follow the spectral residual method in [3] to get the saliency estimation with three main steps: (1) compute the saliency value at each pixel to form a saliency map. (2) Normalize these values in the saliency map. (3) Average values in each 8 x 8 block to attain the block based saliency estimation. Each pixel value of saliency map represents the weight for the pixel region. If a location has a larger saliency value in the saliency map (i.e., light area), it is more attention getting. We can average the saliency values of each 8 x 8 block and get a block based saliency estimation, which is used to modulate the JND profile. To evaluate the performance of saliency estimation, we also generate the saliency maps based on Itti's well-known theory [19] to compare with the spectral residual method.

Four images are chosen from salient object image database as test images for this experiment (kid, flower, couple and sandwich). The saliency estimation results are shown in Fig 2. From the result, we observe that the spectral residual method provides overall better performance than Itti's method[28]. Computationally, the cost of the spectral residual method is relatively low – this brings considerable advantage for a saliency estimator, making it easier to implement on image watermarking. Compared with Itti's method, the computational consumption of the spectral residual method is parsimonious, providing a promising solution to modulate JND profile.

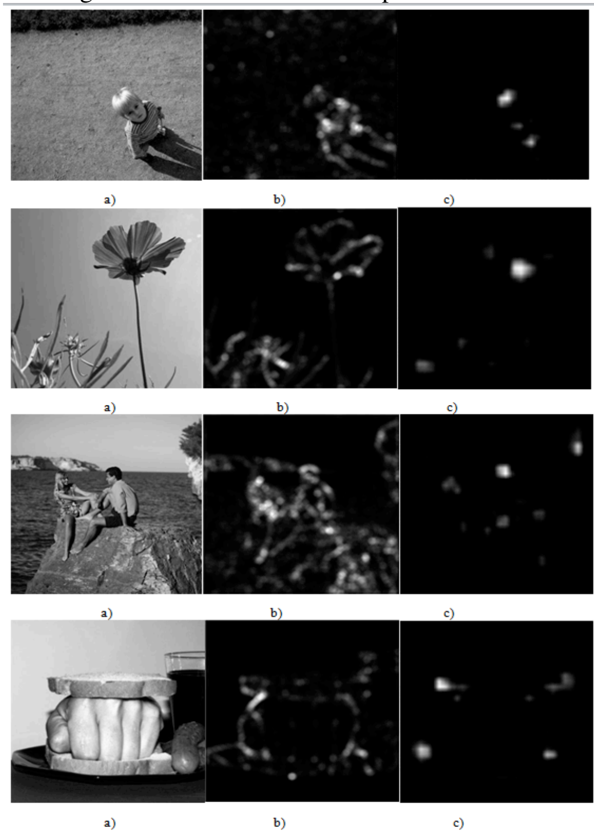


Figure 2. The saliency estimation result. a) the original image, b) saliency map generated by spectral residual, c) saliency map generated by Itti's method

4.2 Evaluating saliency modulated JND profile

To evaluate the performance of the proposed visual saliency modulated JND profile against the original JND profile [2][26], noise is injected into images according to these two JND estimations. In this experiment, the generated JND profile can be used to guide noise shaping in image to evaluate the performance of different JND profiles.

Kid image from salient object image database is used as test image for this experiment. Noise is added to each DCT coefficients of the images as (5). Where *f* takes +1 or -1 randomly; $T(n,\omega_{ij},i,j)$ represents the threshold determined by JND profile obtained via these two JND estimators; $C'(n,i,j)$ is the noise-injected DCT coefficient.

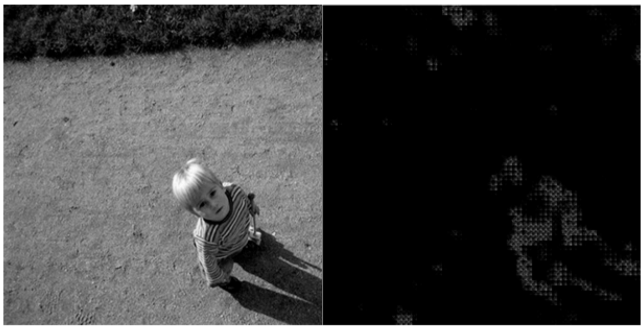
$$C'(n,i,j)=C(n,i,j)+f\times T(n,\omega_{ij},i,j) \quad (5)$$

A more accurate JND profile is supposed to derive a noise injected image with better visual quality under the same level of noise, because it is capable of shaping more noise onto the less perceptually significant regions in the image. For a convincing evaluation of the proposed visual saliency modulated JND estimator for still images, we tested it in two aspects. One is to measure how much the visual content variations are after these two JND estimation guided noise injection and the second is to assess the quality of each noise injected image. The PSNR is used here just to denote the injected noise level under different test conditions. On the other hand, the subjective viewing was used to assess the quality of the resultant visual content. A better JND profile allows higher injected-noise energy (corresponding to lower PSNR) without affecting perceptual quality.

4.2.1 Comparison of JND estimations

To compare the original unmodulated JND profile and the visual saliency modulated JND profile, the noise is injected into kid image according to these two JND estimations.

The original kid image and the difference map of these two generated JND estimation of kid image is shown in Fig 3. Fig 3(a) shows the original kid image while Fig 3(b) shows the difference map between these two JND estimations. We can see from Fig 3(b) that at the more sensitive region the visual saliency modulated JND estimation value is smaller compared to the unmodulated JND estimation value, while at the nonsensitive region the visual saliency modulated JND estimation value is larger compared to the unmodulated JND estimation value. This result confirms that the visual saliency modulated JND estimation is capable of shaping more noise onto the less perceptually significant regions and less noise onto the more perceptually significant regions in the image.

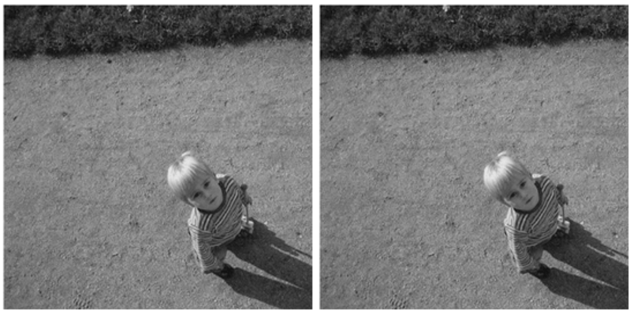


a) b)

Figure 3. The difference map between unmodulated and saliency modulated JND estimators. a) original image, b) The difference map

4.2.2 The PSNR

Fig 4 shows the result of the kid image after the JND-guided noise injection as (5) by these two JND estimations. The visual saliency modulated JND estimation yields the PSNR=27.59dB; the original unmodulated JND estimation yields the PSNR=27.61dB. The PSNR result reflects that although the visual saliency modulated JND estimation shaping more noise onto the less sensitive regions and less noise onto the more sensitive regions, the overall injected-noise energy under the saliency modulated JND estimation is a little bit higher compared to the original unmodulated JND estimation.



a) b)

Figure 4. Resultant kid image after noise injection. a) By unmodulated JND estimation, b) by saliency modulated JND estimation

4.2.3 Perceptual visual quality

From Fig 4 of the kid image after the JND-guided noise injection by the original unmodulated JND estimation and the visual saliency modulated JND estimation, it is observed that the noise is hardly noticeable in the two resultant images. Fig 5 gives a close-up view in the most sensitive region (with highest saliency values) for better visualization, and, as can be seen, saliency modulated JND estimation yields better visual quality in the noise-injected images. The perceptual visual quality result reflects that the visual saliency modulated JND estimation shaping less noise onto the more sensitive regions.

a) b)

Figure 5. A close-up view in the most sensitive region. a) by the original unmodulated JND estimation, b) by the visual saliency modulated JND estimation

From the experimental results in section 4.2.1-4.2.3, we can see saliency modulated JND estimation correlates with the HVS better with the evidence of being capable of offering higher injected-noise energy (we can see lowest PSNR in section 4.2.2) with better perceptual quality (in Figure 4 & Figure 5). And from the result in section 4.2.1, the visual saliency modulated JND estimation is useful for shaping more noise onto the less perceptually significant regions and less noise onto the more perceptually significant regions yielding better perceptual quality of more sensitive regions in image.

5. CONCLUSIONS

Previous perceptual watermarking schemes only partially used the results from HVS studies. The perceptual adjustment of the watermark is mainly based on different visual sensitivity models. Numerically, visual sensitivity can be regarded as the inverse of the just noticeable distortion (JND). The technique described in this paper assist image watermarking by producing a visual saliency modulated JND profile that can be used as a guide to optimize image watermarking. Experimental results confirm the improved performance of our visual saliency modulated JND profile for image watermarking.

6. ACKNOWLEDGMENTS

This work has been financially supported by the DGCIS France in the framework of the HD3D^2 project of the Cap Digital competitiveness cluster. We also acknowledge the funding provided by National Natural Science Foundation of China (Grant No. 60902061) and Key Construction Program of the Communication University of China ”211” Project.

REFERENCES

[1] R. B. Wolfgang, C. I. Podilchuk and E. J. Delp, “Perceptual watermarks for digital images and video”, Proc. IEEE, Special Issue on Identification and Protection of Multimedia Information, vol. 87, 1999, pp. 1108–1126.

[2] Y.Q. Niu, J.B. Liu, S. Krishnan, Q. Zhang, “Combined Just Noticeable Difference Model Guided Image Watermarking”, IEEE Int. Conf., Multimedia and Expo 2010 workshops, July 2010

[3] X. Hou and L. Zhang. Saliency Detection: A Spectral Residual Approach. Proc. CVPR, 2007.

[4] Ahumada Jr, A. J., Peterson, H. A.: Luminance-Model-Based DCT Quantization for Color Image Compression. In : Proceedings of the SPIE, vol. 1666, pp. 365-374 (1992)

[5] Watson, A. B.: DCTune: A technique for visual optimization of DCT quantization matrices for individual images. In Soc. Information Display Dig. Tech. Papers XXIV, pp. 946–949 (1993)

[6] Legge, G. E.: A power law for contrast discrimination. Vision Res, vol. 21, pp. 457–467 (1981)

[7] Zhang, X. K., Lin, W. S., Xue, P.: Improved estimation for just-noticeable visual distortion. Signal Processing, vol. 85, no. 4, pp. 795–808 (2005)

[8] D. H. Kelly,: Motion and vision. II. Stabilized spatio-temporal threshold surface. J. opt. Sot. Am. vol. 69, pp. 1340-1349 (1979)

[9] S. Daly, Engineering observations from spatio velocity and spatiotemporal visual models. In: Proceedings of SPIE, Vol. 3299, pp. 180-191(1998)

[10] Jia, Y., Lin, W., Kassim, A.A.: Estimating just noticeable distortion for video. IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no.7, pp. 820–829 (2006)

[11] Wei, Z., Ngan, K.N.: A temporal just-noticeble distortion profile for video in DCT domain. In: Proceedings of the 15th IEEE International Conference on Image Processing, pp. 1336-1339 (2008)

[12] Wei, Z., Ngan, K.N.: Spatial Just Noticeable Distortion Profile for Image in DCT Domain. In: Proceedings of IEEE International Conference on Multimedia and Expo, pp.925-928 (2008)

[13] Niu, Y.Q., Zhang, Y., Krishnan, S., Zhang, Q.: A Video-Driven Just Noticeable Distortion Profile for Watermarking. In: Proceedings of the 2009 International Conference on Engineering Management and Service Sciences (EMS 2009), (2009)

[14] Niu, Y.Q., Liu, J.B., Krishnan, S., Zhang, Q.: Spatio-Temporal Just Noticeable Distortion Model Guided Video Watermarking. In: Proceedings of the 2009 IEEE Pacific-Rim Conference on Multimedia, (2009)

[15] A. Treisman and G. Gelade. A Feature-Integration Theory of Attention. Cognitive Psychology, 12(1):97-136, 1980.

[16] J. Wolfe. Guided Search 2.0: A Revised Model of Guided Search. Psychonomic Bulletin & Review,1(2):202-238, 1994.

[17] L. Itti, C. Koch, E. Niebur, et al. A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11):1254-1259, 1998.

[18] R. Rensink. Seeing, sensing, and scrutinizing. Vision Research, 40(10-12):1469-87, 2000.

[19] D. Walther and C. Koch. Modeling attention to salient proto-objects. Neural Networks. 19, 1395-1407, 2006.

[20] Z. Lu, W. Lin, X. Yang, E. Ong and S. Yao, "Modeling Visual Attention's Modulatory Aftereffects on Visual Sensitivity and Quality Evaluation", IEEE Trans. Image Processing, vol.14(11), pp. 1928-1942, Nov. 2005.

[21] Lu, W. Lin, X. Yang, E. Ong and S. Yao, " Spatial selectivity modulated just-noticeable-distortion profile for video", IEEE Trans. Image Processing, In: Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, (2004)

[22] H.F. Ling, Z.D. Lu, F.H. Zou, R.X. Li, “An Energy Modulated Watermarking Algorithm Based on Watson Perceptual Model”, Journal of Software, Vol.17, No.5, 2006, pp.1124-1132

[23] C.I. Podilchuk and W.Zeng, “Image-adaptive watermarking using visual models”, Proc. IEEE, Vol. 16, 1998, pp. 525-539

[24] J. Huang and Y.Q. Shi, “Adaptive image watermarking scheme based on visual masking”, Electronics Letters, Vol. 34, 1998, pp. 748-750

[25] M.S. Kankanhalli and K.R. Ramakrishnan, “Content based watermarking of images”, Proc. 6th Int. ACM conf. Multimedia, 1998, pp. 61-70

[26] Niu, Y.Q., Kyan, M., Krishnan, S., Zhang, Q.: A Combined Just Noticeable Distortion Model Guided Image Watermarking. Signal, Image and Video Processing, Journal no. 11760, ISSN: 1863-1703

[27] Niu, Y.Q., Krishnan, S., Zhang, Q.: Spatio-Temporal Just Noticeable Distortion Model Guided Video Watermarking. Special Issue on Intelligent Multimedia Security and Forensics, International Journal of Digital Crime and Forensics (IJDCF), ISSN: 1941-6210 1703

[28] L. Itti, J. Braun, and C. Koch, “Modeling the modulatory effect of attention on human spatial vision,” in Advances in Neural Information Processing Systems, T. G. Dietterich, S. Becker, and Z. Ghahramani, Eds. Cambridge, MA: MIT Press, 2002, vol. 14.