34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

On Just Noticeable Difference for Bionic Eye

Yi Li, Chris McCarthy, and Nick Barnes

***Abstract*— We propose to use Just Noticeable Difference (JND) as the principle in visualizing results for image processing modules for prosthetic vision. Current Bionic Eye hardware implants have limited levels of separately perceivable brightness (i.e., low dynamic range in visualizing images). Therefore, it is important to ensure that the critical contrast must remain perceivable by maintaining of visual differences in downsampled images with reduced dynamic range. JND provides a mathematical framework for these psychophysics events. An increase by 1 in JND space corresponds to the smallest detectable change in visual space (i.e., just noticeable). Combining this principle and the dynamic range constraint, we cast the visualization problem to a linear optimization problem, which enables us to generate optimal visualization images. We demonstrate the usefulness of this principle on visualizing ground-plane segmentation. Experiments show that the proposed principle effectively provides critical visual information at different dynamic ranges, and generates consistent results for image sequences.**

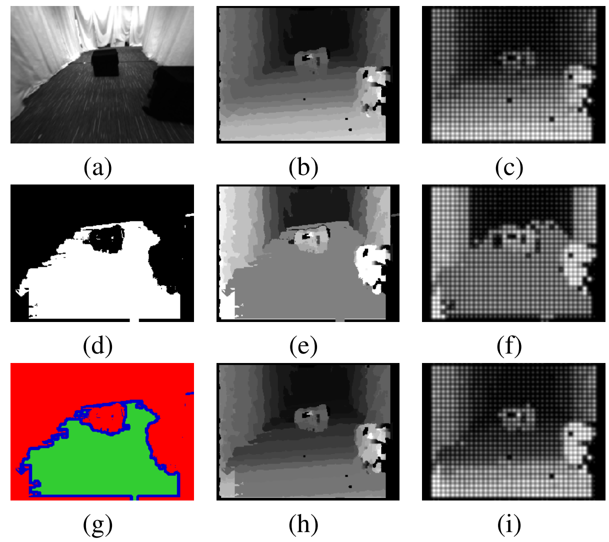


Fig. 1. (a) Intensity image; (b) Depth representation from a stereo camera; (c) Rendering (b) using a 30×35 phosphene rendering with 6 bits dynamic range; (d) Ground-plane segmentation by [6]; (e) Ground plane of (b) was manually set to a fixed value; (f) Rendering (e) using the same phosphene representation as (c); (g) red: pixels in the plane region, blue: boundary pixels of the ground region, green: pixels inside the ground region; (h) Optimal augmentation by our proposed method; (i) rendering (g) using the same phosphene rendering as (c).

I. Introduction

Visual processing for Bionic Eye aims to provide assistive modules based on computer vision techniques. A number of algorithms have been developed for this purpose. Ground plane segmentation [6], saliency [9], and face zooming [2] have demonstrated effective assistance for individuals with vision impairment.

A central question in all visual processing modules is how to augment results for various computer vision algorithms. Current hardwares for prosthetic vision have both limited dynamic range and levels of separately perceivable brightness [5]. Therefore, contrast may be lost during downsampling and dynamic range reduction. Any visualization for integrating these modules must be properly designed to ensure that critical contrast must remain perceivable by maintaining of visual differences in processing.

Many visual processing modules output so called "importance maps". The perception of these "important regions" must be preserved at different dynamic ranges. For instance, objects must be visually different from their background, given the ground-plane segmentation [6]. Therefore, it is ideal to augment these importance maps to the original image to highlight the importance contrasts while keeping other regions unchanged.

Fig. 1 illustrates such an example. Current state of the art in ground-plane segmentation effectively identifies the ground region (white pixels in Fig. 1d, computed from the depth map Fig. 1b of the scene Fig. 1a). A straightforward visualization strategy is to manually set the ground region to a fixed value (Fig. 1e). In this resolution and dynamic range, the obstacles pop up undoubtedly. However, this does not guarantee automatically the contrast is still visible in low resolution and low dynamic range phosphene rendering (Fig. 1f), and a Bionic Eye user may lose its perception of the slope of the ground as well. Fig. 1g would be more preferred, where values of pixels in the ground region are reduced accordingly while obstacles are noticeable in phosphene images (Fig. 1h).

We propose to use Just Noticeable Difference (JND) as the principle in rendering results for visual processing modules for Bionic Eye. An increase by 1 in JND space corresponds to the smallest detectable change in visual space (i.e., just noticeable). Combining this principle with the dynamic range constraint in Bionic Eye, we are able to ensure critical information surviving during visual processing.

In this paper, we use depth representation to illustrate the usefulness of the proposed principle. Recent work has reported the use of surface depth as an alternative scene representation of intensity images [1]. This is achieved by computing stereo disparities between two space-separated

Yi Li, Chris McCarthy, and Nick Barnes are with National Information and Communication Technology Australia (NICTA) and College of Engineering and Computer Science at the Australian National University, Canberra, ACT Australia 2601. Email: {yi.li, chris.mccarthy nick.barnes}@nicta.com.au

parallel cameras. In general, bright means close and darker means further away. We will demonstrate how this JND transforms the visualization problem to an optimization problem in Sec. II. We then apply this principle to augment results for ground plane segmentation (Sec. III). Experiments in Sec. IV show that the visualization guarantees visual differences for critical regions.

II. JND: THE PRINCIPLE

JND refers to methods that use psychological functions in visualization. Visual adaptation [7] and power law were exploited in early work. These effectively model the human visual system, thus JND has been adopted to many fields (e.g., tone mapping in high dynamic range [8]).

Just noticeable difference is originally from the Psychophysics community, with a strong mathematical flavour in its definition. Assuming there is a patch that has a critical contrast we want to preserve at certain dynamic range, we define the following two constraints:

1) If its contrast is larger than the JND, or
2) if its contrast is larger than 1 level of dynamic range.

We now formulate these two intuitive constraints in the solution space. In this paper, we assume the intensity range is normalized to 1. The goal is to "compress" the levels of brightness from 64 in the orignal depth map to an arbitrary k ($k < 64$) in the phosphene rendering.

Denote $\mathcal{I}$ as a gray scale depth representation patch (e.g., any 5×5 region) that has n pixels and c is its center pixel index. The center pixel is "just noticeable" if and only if:

$$\left| \frac{\mathcal{I}_c - \frac{1}{n-1} \sum_{j \neq c} \mathcal{I}_j}{\frac{1}{n-1} \sum_{j \neq c} \mathcal{I}_j} \right| \geq \delta, \tag{1}$$

where $j = 1...n$, and $\mathcal{I}_j$ is the pixel value for the j^{th} pixel. δ is the JND value, which is normally small (e.g., 0.1).

Eq. 1 can be easily extended to any linear contrast filter. For simplicity, we will stick to the simpler formulation (Eq. 1) in this paper.

The contrast of $\mathcal{I}$ can still be perceived at k levels of brightness if:

$$\mathcal{I}_c - \frac{1}{n-1} \sum_{j \neq c} \mathcal{I}_j \geq \frac{1}{k}, \tag{2}$$

Thus, we state that this patch $\mathcal{I}$ is "just noticeable" under a quantization $\frac{1}{k}$ if and only if it satisfies Eq. 1 and 2.

III. CASE STUDY: VISUALIZING GROUND-PLANE SEGMENTATION

A. Brief introduction to visualizing ground-plane segmentation

In mobile robot navigation, ground-plane modelling is commonly employed to determine the traversability of the immediate space. Recently, McCarthy and Barnes [6] proposed a surface detection and segmentation scheme based on the examination of iso-disparity contours. Using iso-disparity analysis for ground surface segmentation, the method infers all traversable space in the image, from which all non-traversable surfaces (e.g., walls, obstacles) are also obtained.

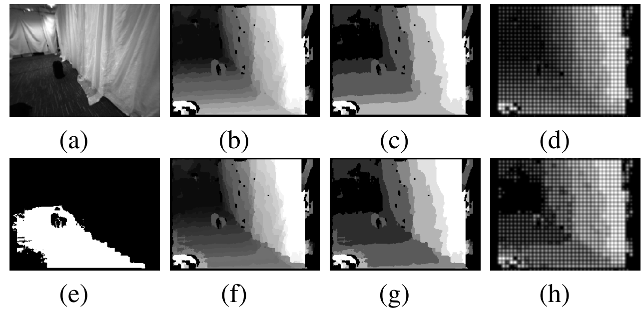


Fig. 2. An example. (a) Scene; (b)-(d): Depth map, direct dynamic range reduction, and its phosphene rendering, respectively. The phosphene is rendered using a 30×35 regular grid at $k = 8$ levels of dynamic range; (e) Ground region of (a); (f)-(h): Our augmented depth map, its dynamic range reduction, and phosphene rendering, respectively.

The goal of the visualization is to augment depth-based phosphene scene rendering to provide both a cleaner visualisation of the ground surface, and to enhance the distinction between traversable and non-traversable space; in particular, small ground surface obstructions. This goal can be interpreted as filling in new values to the traversable regions such that the critical difference between traversable and non-traversable space is noticeable in low resolution low dynamic range phosphene images.

B. Visualizing ground-plane segmentation using the JND principle

Define $\mathcal{X}$ as the solution of visualizing an image $\mathcal{P}$ at k levels of dynamic range.

In general, there are three sets variables of $\mathcal{X}$: 1) pixels of plane regions (red pixels in Fig. 1g), 2) boundary pixels of plane regions (blue pixels in Fig. 1g), and 3) pixels inside ground regions (green pixels in Fig. 1g).

The solution of visualization must keep the variables in Set 1 unchanged and the ones in Set 2 "just noticeable". Further, the ordering of Set 3 must be preserved as much as possible. Therefore, we formulate our visualization method as follows.

1) Applying the JND principle to Set 2: For variables x of Set 1, we take a $\sqrt{n} \times \sqrt{n}$ patch where i is its center pixel index. According to the proposed JND principle, we have:

$$\begin{cases} x_i - \frac{1+\delta}{n-1} \sum_{j \neq i} x_j - u_i^1 = 0 \\ x_i - \frac{1}{n-1} \sum_{j \neq i} x_j - u_i^2 = \frac{1}{k} \end{cases}, \tag{3}$$

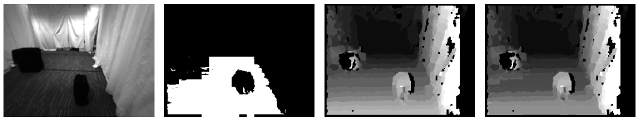
if $p_i \geq \frac{1}{n-1} \sum_{j \neq i} p_j$
or

$$\begin{cases} -x_i + \frac{1-\delta}{n-1} \sum_{j \neq c} x_j - u_i^1 = 0 \\ -x_i + \frac{1}{n-1} \sum_{j \neq i} x_j - u_i^2 = \frac{1}{k} \end{cases}, \tag{4}$$

if $p_i < \frac{1}{n-1} \sum_{j \neq i} p_j$

In both cases, u_b^1 and u_b^2 are nonnegative. This can be written as a linear equation system:

$$A_1 x - u = y_1, \tag{5}$$



(a) Scene (b) Ground (c) Depth (d) Ours

Fig. 3. Another example.

2) Applying the JND principle to Set 3: The constraints in the pixels of set 3 is simply the ordering constraints between pairs of variables (x_i, x_j).

$$x_i - x_j - v_{ij} = 0 \iff p_i > p_j, \tag{6}$$

where v_{ij} is a nonnegative variable.

Further, if the pixels have the same values in the depth map, they must have the same value in the visualization space.

$$x_i = x_j \iff p_i = p_j, \tag{7}$$

This gives us a second linear system.

$$A_2 x - v = 0, \tag{8}$$

3) Applying the JND principle to Set 1: For the pixels in the plane region, we simple set

$$x_i = p_i, \tag{9}$$

This gives us the third linear system

$$A_3 x = y_3, \tag{10}$$

Therefore, we can rewrite the whole visualization problem as a linear programing problem

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} A_1 & -I_1 & 0 \\ A_2 & 0 & -I_2 \\ A_3 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} x \\ u \\ v \end{bmatrix} \tag{11}$$

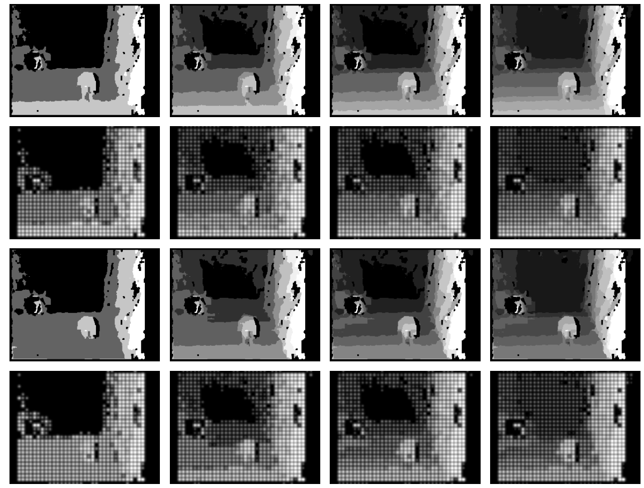
subject to $x, u, v \geq 0$, where I_1 and I_2 are two identity matrices. Once we have the solution x, we can re-normalize it to $0 - 1$ and render the solution in phosphene maps by downsampling and reuducing dynamic range.

C. Nonnegative Conjugate Gradient for solving visualization problem

Realtime response and small memory are two key features in Bionic Eye systems. Improper selection of an optimization method can easily slow down the computation. In our algorithm, we choose Conjugate Gradient as a possible choice [4].

Optimization over a large number of variables is a challenging task. Only a handful of methods are suitable for large scale problems [4]. Conjugate Gradient based methods are highly suitable for this purpose because only matrix multiplications are involved. In conjugate direction methods, the Nonnegative Conjugate Gradient [3] methods are more suitable for some computational imaging problems.

We used the Nonnegative Conjugate Gradient (NCG). NCG has similar advantages as other Conjugate Direction method where only matrix multiplications are required, with a small overhead to satisfy the boundary condition. The key idea of NCG is to go backward if any variable is out of bounds, and re-start the algorithm if necessary. Please refer to [3] for details.



4 Levels 8 Levels 12 Levels 16 Levels

Fig. 4. Comparison between direct rendering (First and Second row) and the augmentation based on our principle (Third and Forth row) of Fig. 3 at different dynamic ranges.

IV. EXPERIMENTS

We first show some results of our algorithm. Then, we demonstrate the efficacy of our algorithm at different dynamic range. Finally we show the results for an image sequence. These experiments suggest our principle is effective in preserving critical information in phosphetic vision.

A. Demonstration

Fig. 2 shows our result in an indoor scene. The input of our method is a depth map where each pixel is represented by 6 bits (i.e. 64 levels) (Fig. 2b). One can see that the obstacle in the middle of Fig. 2a cannot be effectively identified in Fig. 2b because the depth of the obstacle is very similar to the depth of the ground. If we quantize Fig. 2b directly to 8 levels of brightness, the contrast of the object is lost completely (Fig. 2c). Therefore, the contrast of the obstacle is not perceivable in the phosphene representation (Fig. 2d).

The ground plane segmentation gives us the ground region (Fig. 2e) Therefore, we can apply our method in Sec. III to ensure the visual difference between obstacles and ground region is noticeable (Fig. 2f). Therefore, if we quantize Fig. 2c to 8 levels of brightness, the obstacle in the center is well preserved (Fig. 2g). We can also see that the slope of the ground is preserved in the phosphene rendering because we impose an ordering constraint (Fig. 2h).

Fig. 2h clearly provides a better space perception compared to Fig. 2d. This demonstrate the usefulness of our augmentation based on the JND principle.

B. Results for different dynamic ranges

As we pointed out at Sec. I, quantization of the dynamic range is an important constraint in the visualization. Here we show the results for another scene (Fig. 3) at a number

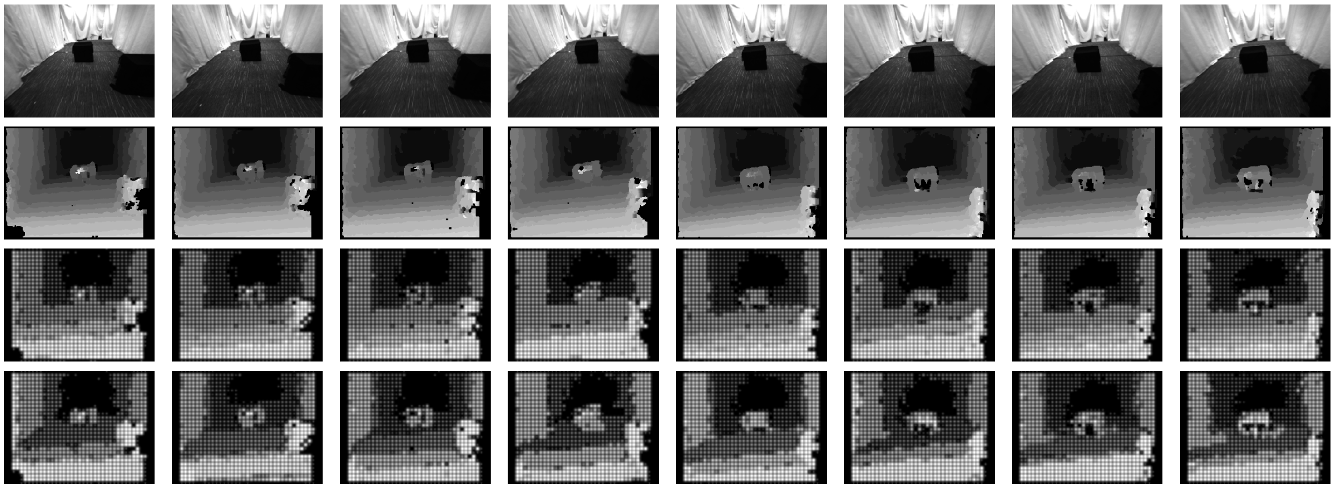


Fig. 5. A video example. The first row, intensity image; Second row, depth representations; third row, phosphene rendering of direct dynamic range reduction ($k = 8$); Fourth row, phosphene rendering of our augmentations ($k = 8$).

of quantizations. Similar to the previous experiment, we show the original depth representation (Fig. 3c) and the augmentated depth maps by our proposed method (Fig. 3d).

Fig. 4 shows the results for Fig. 3c and 3d at 4, 8, 12, and 16 levels of dynamic ranges, respectively. At each quantization, one can see that the obstacles are visible in the phosphene maps generated by our method, but not in the original depth representations.

This experiment shows our algorithm is adaptive to various levels of brightness, and provides effective results for the Bionic Eye.

C. Results for image sequences

We finally show results for images in a video. An ideal phosphene augmentation highlights obstacles consistently. This suggests that obstacles must appear unambiguously, and their shapes/distance must also be consistent.

Fig. 5 shows results for 8 frames in a video. In our method (4th row), one can see the obstacles in the middle of the scene has a clear difference compared to the background, and the size/shape of the obstacle can be identified easily. These are two advantages over the direct visualization (3rd row). Further, the obstacle in the lower right corner gradually disappear when the ground-plane method stops reporting this object as a non-traversable area.

This experiment shows our method generates consistent visualization across frames in a video. Obstacles in videos can be perceived because critical contrast is preserved.

V. Conclusion

We present a JND-based principle for augmenting results for visual processing modules for Bionic Eye. This principle ensures that critical contrast must remain perceivable by maintaining of visual differences in visual processing. Experiments show that the proposed method is effective in visualizing results for a state of the art ground-plane segmentation. Objects on the ground are guaranteed to have visual differences at different dynamic ranges.

VI. Acknowledgement

NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications, and the Digital Economy, and the Australian Research Council (ARC) through the ICT Centre of Excellence Program. This research was also supported in part by ARC through its Special Research Initiative (SRI) in Bionic Vision Science and Technology grant to Bionic Vision Australia (BVA).

References

[1] N. Barnes, P. Lieby, H. Dennet, J. G. Walker, C. McCarthy, N. Liu, and Y. Li. Investigating the role of single-viewpoint depth data in visually-guided mobility. In *Vision Sciences Society (VSS)*, 2011.

[2] Xuming He, Chunhua Shen, and Nick Barnes. Face detection and tracking in video to facilitate face recognition in a visual prothesis. In *ARVO*, May 2011.

[3] M.R. Hestenes. *Conjugate direction methods in optimization*. Applications of mathematics. Springer-Verlag, 1980.

[4] Quoc V. Le, Jiquan Ngiam, Adam Coates, Ahbik Lahiri, Bobby Prochnow, and Andrew Y. Ng. On optimization methods for deep learning. In *ICML*, pages 265–272, 2011.

[5] P. Lieby, N. Barnes, C. McCarthy, N. Liu, H. Dennet, J. Walker, V. Botea, and A Scott. Substituting depth for intensity and real-time phosphene rendering: Visual navigation under low vision conditions. In *EMBC*, 2011.

[6] C. McCarthy, N. Barnes, and P. Lieby. Ground surface segmentation for navigation with a low resolution visual prosthesis. In *Proceedings of IEEE EMBC 2011*, pages 4457–4460. IEEE, 2011.

[7] Sumanta N. Pattanaik, Jack Tumblin, Hector Yee, and Donald P. Greenberg. Time-dependent visual adaptation for fast realistic image display. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, 2000.

[8] Erik Reinhard and Kate Devlin. Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Vis. and Comp. Graphics*, 11:13–24, January 2005.

[9] A. Stacey, Y. Li, and N. Barnes. A salient information processing system for bionic eye. In *Proceedings of IEEE EMBC 2011*. IEEE, 2011.