BOUNDARY DEPTH INFORMATION USING HOPFIELD NEURAL NETWORK

Sheng Xu, Ruisheng Wang*

Department of Geomatics, Engineering, University of Calgary, Calgary, Alberta, Canada T2N 1N4 -
(sheng.xu2, ruiswang)@ucalgary.ca

Commission V, WG V/1

KEY WORDS: Boundary, Stereo Matching, Depth Information, Optimization, Discontinuity, Hopfield Neural Network

ABSTRACT:

Depth information is widely used for representation, reconstruction and modeling of 3D scene. Generally two kinds of methods can obtain the depth information. One is to use the distance cues from the depth camera, but the results heavily depend on the device, and the accuracy is degraded greatly when the distance from the object is increased. The other one uses the binocular cues from the matching to obtain the depth information. It is more and more mature and convenient to collect the depth information of different scenes by stereo matching methods. In the objective function, the data term is to ensure that the difference between the matched pixels is small, and the smoothness term is to smooth the neighbors with different disparities. Nonetheless, the smoothness term blurs the boundary depth information of the object which becomes the bottleneck of the stereo matching. This paper proposes a novel energy function for the boundary to keep the discontinuities and uses the Hopfield neural network to solve the optimization. We first extract the region of interest areas which are the boundary pixels in original images. Then, we develop the boundary energy function to calculate the matching cost. At last, we solve the optimization globally by the Hopfield neural network. The Middlebury stereo benchmark is used to test the proposed method, and results show that our boundary depth information is more accurate than other state-of-the-art methods and can be used to optimize the results of other stereo matching methods.

1. INTRODUCTION

Determination of correspondence between two pictures from different viewpoints of the same scene is the primary contribution of stereo matching. Various constraints have been used to find the optimal match, such as color consistency, gradient distribution, and geometric shape. Nevertheless, these constraints provide less information about discontinuities which is the bottleneck of stereo matching.

Compared with current matching methods, the proposed method is in our novel 3D space with less complexity, and has the constraint of discontinuities to find the optimal match. The main contributions of our work are in the following: 1) we obtain a new 3D space by radial basis information (RBI); 2) we form a new energy function to calculate the cost for the matching; 3) we convert the proposed objective function to a solvable Hopfield neural network (HNN) energy function.

2. RELATED WORK

Stereo matching methods can be divided into region-based and feature-based (Scharstein and Szeliski, 2002). Region-based algorithms use a sliding window to calculate the aggregation of the cost. Then various specific optimizations are followed to obtain the minimum cost for the matching, such as Graph Cut (Boykov et al., 2001) and Belief Propagation (Sun et al., 2003). Due to the propagation of errors in discontinuities, the accuracy of these methods is not satisfactory.

Feature-based algorithms pay attention to features and perform better in discontinuities. Shape based matching (Ogale and Aloimonos, 2005) has analyzed the effects of the shape of

established dense point correspondence to improve the matching, but the result in the repeated shape areas is imprecise. CrossTrees (Cheng et al., 2015) uses two priors: edge and super-pixel, which are proposed to avoid the false matching in discontinuities. However, the method fails to large planar surfaces with fewer features. Instead of using adjacent pixels for matching, LLR (Zhu et al., 2012) chooses neighborhood windows and believes there is a linear relationship between pixel values and disparities. However, it is sensitive to the size of the window and large areas devoid of any features. To eliminate improper connections on the boundaries of two objects, imprNLCA (Chen et al., 2013) uses the boundary cue of the reference image which is more reliable than the color cue in areas with similar colors. Nevertheless, imprNLCA shows less improvement in discontinuities compared to original NLCA (Yang, 2012). The Borders (Mattoccia et al., 2007) obtains the precise border localization based on a variable support method (Tombari et al., 2007) for retrieving depth discontinuities. However, for images with a large range of disparities, their accuracy decreases. LCVB-DEM (Martins et al., 2015) introduces a trained binocular neuronal population to learn how to decode disparities. With the help of monocular cells, they encode both line and edge information which are critical for persevering discontinuities. However, due to the fact that mechanism of cell responses is complex, their accuracy is far from what is desirable.

Algorithm (Nasrabadi and Choo, 1992) based on the variance equations for each direction is less robust with limited application. To enhance the robustness, methods (Lee et al., 1994) (Huang and Wang, 2000) (Achour and Mahiddine, 2002) (Laskowski et al., 2015) formulated their respective energy

---

* Corresponding author

functions which contain the constraints of similarity, smoothness and uniqueness to work well on noisy images. Although the above Hopfield neural network (HNN) based matching methods choose different constraints and obtain a desirable result for regions of interests, they neglect the matching of discontinuities. Neurons in HNN are interconnected and therefore, the change of a neuron state will affect all other input neurons. HNN converges to a stable state by updating the neurons from the activation function. This implies that it can obtain the global optimization result automatically. Hence, we use the HNN model to solve our optimization problem.

Neurons in our method are based on the disparity space rather than the pixels in the left image and the right image as the aforementioned methods. For the formation of objective function, we use discontinuity information obtained from our novel 3D space to calculate matching cost, and constraints of uniqueness and position to reduce the search for a solution. To show the performance of our algorithm, we test it on Middlebury benchmark to present our accuracy and show the improvement of state-of-the-art methods by raising their accuracy in discontinuities.

3. THE METHOD

This section describes a new objective function for stereo matching, and shows the derivation of the proposed optimization method.

3.1 Energy Function

Stereo matching is about assigning the best disparity for each pixel. Figure 1 is the disparity space. Different methods relate to various cost paths from the first row to the last row. Result should be smooth on continuities and preserve discontinuities on disparity map. The key problems are the energy function and the optimization method.
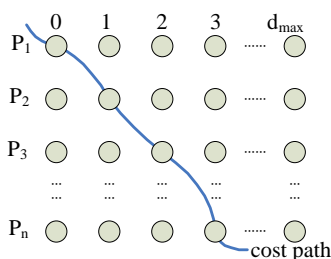

Figure 1. Disparity space. There are n pixels, namely P_1, P_2, ..., P_n, and their disparities are ranged from 0 to d_{max}. The cost path relates to the assignment of disparities for each pixel.

The radial basis function gives an expression on the relationship between pixels and its neighbour (Buhmann, 2004). Thus, we calculate the surrounding information of a pixel P by

$$S(P) = \sum e^{\left(-\frac{(I_{P'}-I_P)^2}{2\sigma^2}\right)} \quad (1)$$

In Eq.(1), I_P is the intensity value of a pixel P, and σ is the variance of neighbors of P. The neighbor pixels are selected by a window with a size of N. Figure 2(a) is the left image, and its surrounding information is shown from Figure 2(b) to 2(c) in

different N. A small N results sharp discontinuities and a large N causes smooth discontinuities.

We describe a pixel P in a 3D space XYZ as (x, y, S) as shown in Figure 3(a) rather than only intensity. S is the surrounding information of the pixel at (x, y) in the image. The discontinuities are outstandingly visible and the continuities are smooth as shown in Figure 3(b). We call this space as RBI (Radial Basis Information) space and following is our objective function for the matching.
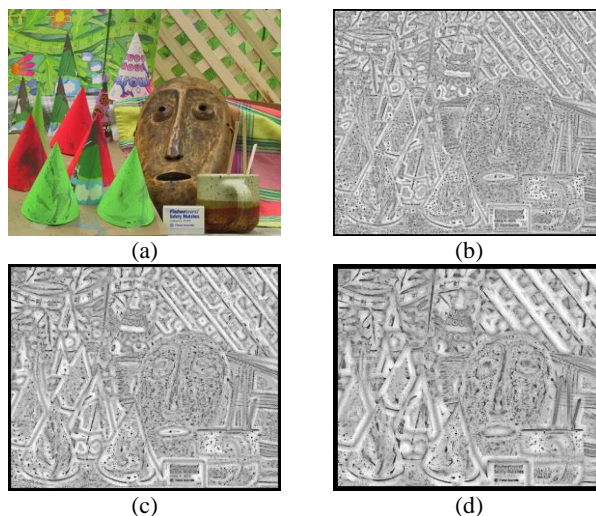

Figure 2. Surrounding information of Cones. (a) Original left image. (b) N=3. (c) N=5. (d) N=8.
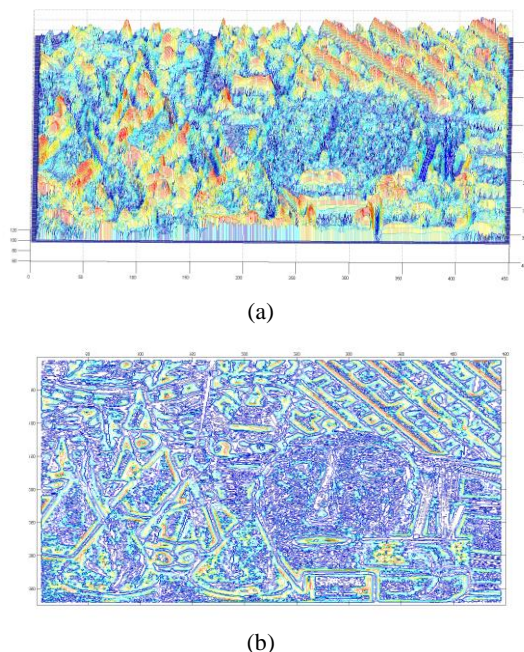

Figure 3. Display of the RBI space. (a) Cones in the 3D RBI space. (b) The contour of Cones in the RBI space.

Analog to the gradient of intensity in 2D obtained by Sobel (Farid and Simoncelli, 1997), the gradient of surrounding information G is

$$G \approx (x, y, z)J(x, y, z)^T \geq 0 \quad (2)$$

In Eq.(2), **J** is the Jacobian matrix obtained by the gradient of S in the direction of axis, namely S_x, S_y, and S_z, and defined as

$$\begin{pmatrix} S_x^{\,2} & S_x \cdot S_y & S_x \cdot S_z \\ S_x \cdot S_y & S_y^{\,2} & S_y \cdot S_z \\ S_x \cdot S_z & S_y \cdot S_z & S_z^{\,2} \end{pmatrix}$$

The eigenvectors of J show the direction of a pixel in RBI space, and eigenvalues show the scale. We use Principal Component Analysis (PCA) (Jolliffe, 2002) to calculate the principle direction of each pixel as shown in Figure 4.
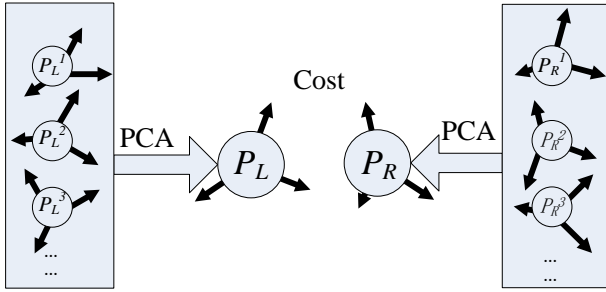

Figure 4. Calculation of the cost between left image pixel P_L and right image pixel P_R. Each pixel has different gradients in the direction of Z, and we use PCA to obtain the principal direction of the pixel.

Denote the eigenvectors and eigenvalues of P_L as $\boldsymbol{L_x}$, $\boldsymbol{L_y}$, $\boldsymbol{L_z}$ and α_L, β_L, γ_L, respectively. Similarly, $\boldsymbol{R_x}$, $\boldsymbol{R_y}$, $\boldsymbol{R_z}$ and α_R, β_R, γ_R correspond to P_R. The matching cost for P_L and P_R is

$$Cost = \left| left - right \right|^2 \tag{3}$$

where $Left = \alpha_L \cdot L_x + \beta_L \cdot L_y + \gamma_L \cdot L_z$
$$Right = \alpha_R \cdot R_x + \beta_R \cdot R_y + \gamma_R \cdot R_z$$

The matched pixels should have the similar surrounding information, which is guaranteed by Eq.(3). In addition, pixels are unlikely to be matched from different areas, such as from discontinuities to continuities. Thus, there should be a regional restriction term. We replace $Cost$ by adding the difference of surrounding RBI S^L of P_L and S^R of P_R as

$$Cost' = w_1 \cdot \sum \left(S_p^L - S_p^R \right) + w_2 \cdot Cost \tag{4}$$

These two weighting coefficients w_1 and w_2 are chosen as constants in our algorithm. The optimization is in the disparity space, and the whole cost E for matching is

$$E = \sum_{i=1}^{n} \sum_{k=0}^{d_{max}} Cost'_{ik} \, P_{ik} \tag{5}$$

In Eq.(5), $Cost'_{ik}$ is the matching cost between P_i in the left image and P_{i+k} in the right image. $Cost'_{ik}$ implies the cost of assigning k as the disparity of P_i in the disparity space. P_{ik} is 1 or 0 which indicates the disparity of P_i is k or not. The optimal solution is the assignment of disparities with the minimal E.

Suppose there are two pixels $P(x_L, y_L)$ and $P(x_L', y_L')$ in the left image, and their matching pixels are $P(x_R, y_R)$ and $P(x_R', y_R')$ in the right image, respectively. If x_L' is less than x_L, x_R' should not be much larger than x_R. We use a sigmoid function to keep the position in a limited changed after matching and obtain $Cost_{ikjl}$ as

$$Cost_{ikjl} = \left(\frac{Cost'_{ik} + Cost'_{jl}}{2} \right) \times \frac{1}{\eta} \tag{6}$$

where $\eta = \left(\dfrac{1}{1 + e^{-\mu \cdot x}} + \dfrac{1}{1 + e^{-\mu \cdot y}} \right)$
$$x = \left(x_i' - x_j' \right)\left(x_i - x_j \right)$$
$$y = \left(y_i' - y_j' \right)\left(y_i - y_j \right)$$

If x is negative, η is small, and $Cost'_{ijkl}$ goes to large. μ is used to control threshold-likeness or threshold-dislikeness of the output. Each time, we choose two pixels P_i and P_j in the left image together for the matching, and add the position constraint for the objective function as

$$E = \sum_{i=1}^{n} \sum_{k=0}^{d_{max}} \sum_{j=1}^{n} \sum_{l=0}^{d_{max}} Cost_{ikjl} \, P_{ik} P_{jl} \tag{7}$$

3.2 The optimization model

This is a NP-hard problem which may not be solved in a polynomial time. We need to reduce the search for a solution. The proposed objective function lacks the constraint of the number of matches. Each pixel is allocated with a unique disparity. Thus, our final energy E' is calculated as

$$E' = \sum_{i=1}^{n} \sum_{k=1}^{d_{max}} \sum_{j=1}^{n} \sum_{l=1}^{d_{max}} Cost_{ikjl}' \, P_{ik} P_{jl} + U \tag{8}$$

where $U = \sum_{i=1}^{n} (1 - \sum_{k=1}^{d_{max}} P_{ik})^2$

Expand U, we have

$$\begin{aligned}
U &= \sum_{i=1}^{n} \left(1 - 2 \times \sum_{k=1}^{d_{max}} P_{ik} + \left(\sum_{k=1}^{d_{max}} P_{ik} \right)^2 \right) \\
&= \sum_{i=1}^{n} 1 - 2 \times \sum_{i=1}^{n} \sum_{k=1}^{d_{max}} P_{ik} + \sum_{i=1}^{n} \left(\sum_{k=1}^{d_{max}} P_{ik} \right)^2 \\
&= n - 2 \times \sum_{i=1}^{n} \sum_{k=1}^{d_{max}} P_{ik} + \sum_{i=1}^{n} \left(\sum_{k=1}^{d_{max}} P_{ik} \sum_{l=1}^{d_{max}} P_{il} \right) \\
&= n - 2 \times \sum_{i=1}^{n} \sum_{k=1}^{d_{max}} P_{ik} + \sum_{i=1}^{n} \sum_{k=1}^{d_{max}} \sum_{j=1}^{n} \sum_{l=1}^{d_{max}} P_{ik} P_{jl} \delta_{ij}
\end{aligned} \tag{9}$$

When i is equal to j, δ_{ij} is 1, otherwise δ_{ij} is 0. Substitute Eq.(9) for U in Eq.(8) to obtain Eq.(10) where E'' is $(E'-n)/2$.

$$E'' = \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{d_{max}}\sum_{j=1}^{n}\sum_{l=1}^{d_{max}}(Cost_{ikjl} - \delta_{ij})P_{ik}P_{jl} - \sum_{i=1}^{n}\sum_{k=1}^{d_{max}}P_{ik} \quad (10)$$

Let the weight matrix **W** be

$$W_{ikjl} = 1 - \frac{2}{1+e^{\lambda(Cost_{ikjl}' - \delta_{ij} - K)}} \quad (11)$$

The graph of Eq.(11) is shown in Figure 5. K describes the cost for the exact matching which is increased with the noise in the images. If P_L is in discontinuities, we need a large λ to prevent P_L from mismatching. However, when P_L is in a large continuous area, a small λ is required to search the optimal matching. We calculate λ by Eq.(12) where S_{mean}^{L} is the mean of S^{L}.
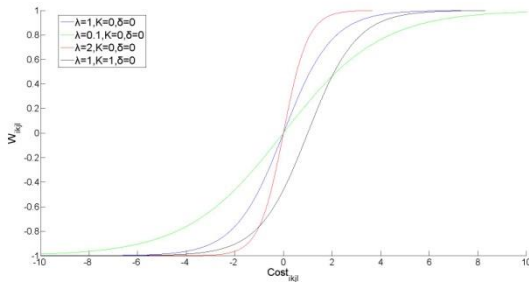
$$\lambda = \frac{S_{mean}^{L}}{S(P)} \quad (12)$$

Figure 5. Graph of **W** with different parameters.

The graph of λ in different areas of Teddy is shown in Figure 6. A large λ which relates to discontinuities is brighter than a small one.

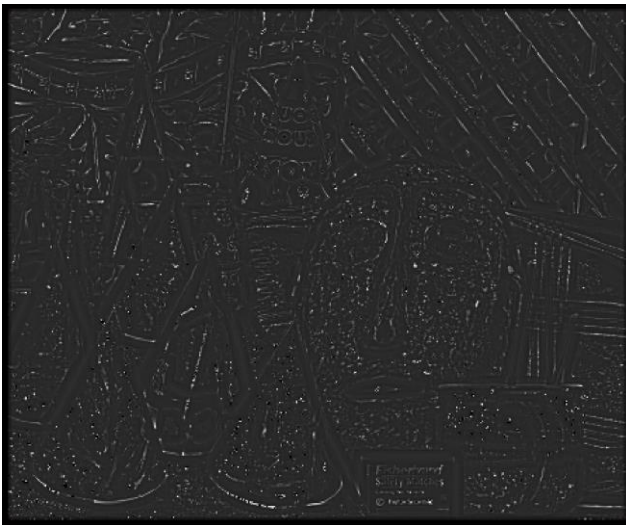Figure 6. λ in different areas of Teddy.

Now E'' is calculated by Eq.(13) which is the objective function for optimization in Hopfield neural network (Hopfield, 1982). Neurons are based on the disparity space. Different connections of neurons indicate different assignments of disparities for each pixel. Neurons can not be self-connected and hence W_{ikik} is 0. The state of a neuron P_{ik} is 0 or 1 which is calculated by the net.

P_{ik} set to 1 implies that this neuron is active and the disparity of the pixel P_i is k. When the state of a neuron changes from P_{mn} to P_{mn}', the energy E'' changes from E_{old} to E_{new} as shown in Eq.(14) and the differnce is ΔE as calculated in Eq.(15).

$$E'' = (\frac{1}{2})\sum_{i=1}^{N}\sum_{k=1}^{d_{max}}\sum_{j=1}^{N}\sum_{l=1}^{d_{max}}W_{ikjl}P_{ik}P_{jl} - \sum_{i=1}^{N}\sum_{k=1}^{d_{max}}P_{ik} \quad (13)$$

$$
\begin{aligned}
E_{new} &= \frac{1}{2}\sum_{\substack{i=1\\\&\\i\neq m}}^{N}\sum_{\substack{k=1\\\&\\k\neq n}}^{d_{max}}\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{ikjl}P_{ik}P_{jl} - \sum_{\substack{i=1\\\&\\i\neq m}}^{N}\sum_{\substack{k=1\\\&\\k\neq n}}^{d_{max}}P_{ik}\\
&+ \frac{1}{2}\sum_{\substack{k=1\\\&\\k\neq n}}^{d_{max}}\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{mkjl}P_{mk}P_{jl} + \frac{1}{2}\sum_{\substack{i=1\\\&\\i\neq m}}^{N}\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{injl}P_{in}P_{jl}\\
&+ \frac{1}{2}\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{mnjl}P_{mn}'P_{jl} + \frac{1}{2}\sum_{\substack{i=1\\\&\\i\neq m}}^{N}\sum_{\substack{k=1\\\&\\k\neq n}}^{d_{max}}W_{ikmn}P_{ik}P_{mn}'\\
&+ \frac{1}{2}W_{mnmn}P_{mn}'P_{mn}' - \sum_{k=1\&k\neq n}^{d_{max}}P_{mk} - \sum_{i=1\&i\neq m}^{N}P_{in} - P_{mn}'\\
&= \frac{1}{2}\sum_{\substack{i=1\\\&\\i\neq m}}^{N}\sum_{\substack{k=1\\\&\\k\neq n}}^{d_{max}}\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{ikjl}P_{ik}P_{jl} - \sum_{\substack{i=1\\\&\\i\neq m}}^{N}\sum_{\substack{k=1\\\&\\k\neq n}}^{d_{max}}P_{ik}\\
&+ \frac{1}{2}\sum_{\substack{k=1\\\&\\k\neq n}}^{d_{max}}\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{mkjl}P_{mk}P_{jl} + \frac{1}{2}\sum_{\substack{i=1\\\&\\i\neq m}}^{N}\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{injl}P_{in}P_{jl}\\
&+ \sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{mnjl}P_{mn}'P_{jl} - \sum_{k=1\&k\neq n}^{d_{max}}P_{mk} - \sum_{i=1\&i\neq m}^{N}P_{in} - P_{mn}'
\end{aligned} \quad (14)
$$

$$\Delta E = E_{new} - E_{old}$$

$$= \left[\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{mnjl}(P_{mn}' - P_{mn})P_{jl}\right] - \left(P_{mn}' - P_{mn}\right)$$

$$= \left[\sum_{\substack{j=1\\\&\\j\neq m}}^{N}\sum_{\substack{l=1\\\&\\l\neq n}}^{d_{max}}W_{mnjl}P_{jl} - 1\right]\left(P_{mn}' - P_{mn}\right) \quad (15)$$

$$= net_{mn}\Delta P_{mn}$$

In Eq.(15), net_{mn} is the output of HNN for the neuron P_{mn}. If ΔP_{mn} is 0, ΔE is 0 implying that E'' is unchanged; if ΔP_{mn} is a positive (implying P_{mn}' is 1) and net_{mn} is a negative, therefore ΔE is a negative; if ΔP_{mn} is a negative (implying P_{mn}' is 0) and net_{mn} is a positive, therefore ΔE is a negative. Thus, E'' is always falling during the updating of neuron states. The rule of updating is Eq.(16) (Nasrabadi and Choo, 1992). After a number of iterations, we obtain the minimum E' which relates to

the required optimal disparity map.

$$\begin{cases} \left[\sum_{j=1}^{N}\sum_{l=1}^{d_{max}}W_{ikjl}P_{jl} - 1\right] < 0, \quad P_{ik}' = 1 \\ \left[\sum_{j=1}^{N}\sum_{l=1}^{d_{max}}W_{ikjl}P_{jl} - 1\right] > 0, \quad P_{ik}' = 0 \quad (16) \\ \left[\sum_{j=1}^{N}\sum_{l=1}^{d_{max}}W_{ikjl}P_{jl} - 1\right] = 0, \quad P_{ik}' = P_{ik} \end{cases}$$

4. EXPERIMENTS AND EVALUATION

The workflow of our method is shown in Figure 7. First, we extract the surrounding information S of the input images and form our 3D RBI space to describe each pixel. Second we calculate the gradient of S and use PCA to obtain the principal directions of each pixel. Third, we calculate the matching cost, and add the constraints of the region, position and uniqueness to obtain the objective function. Finally, we solve the optimization problem by HNN. The output of HNN is the global minimum E which relates to the disparity map.
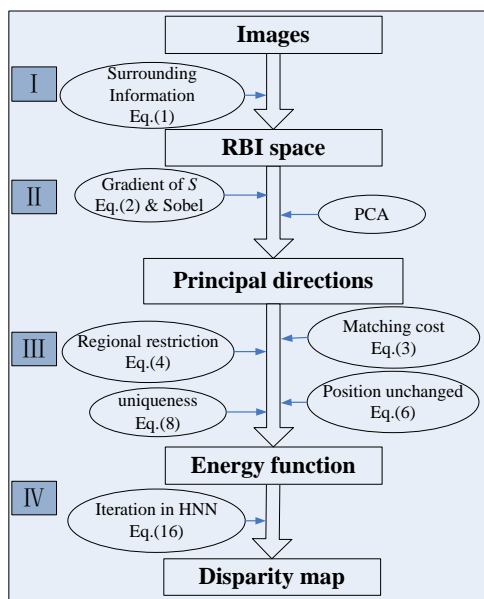

Figure 7. The workflow of our method.

Images used for experiments are Tsukuba (384×288 pixels), Venus (434×383 pixels), Cones (450×375 pixels) and Teddy (450×375 pixels) from the Middlebury stereo benchmark (Scharstein and Szeliski, 2002). Figure 8 consists of the ground truth, discontinuities, and continuities for test dataset. The size of window N for the neighbor pixels in Eq.(1) depends on the size of featureless areas. In our experiment, N is 5 which balances the accuracy and time consumed. For the gradient of S, we extend the Sobel into 3D. We choose three principal directions from PCA to calculate the Eq.(3). For Eq.(4) and Eq.(12), the S^L and S^R are the surrounding RBI of P_L's and P_R's adjacent pixels, respectively. We want to highlight the function

of discontinuities, thus w_1 is 1 while w_2 is 100. For η in Eq.(6), μ is 1. Different images need different parameters, while for aforementioned images, these parameters work effectively.

The size of window N for the neighbor pixels in Eq.(1) depends on the size of featureless areas. In our experiment, N is 5 which balances the accuracy and time consumed. For the gradient of S, we extend the Sobel into 3D. We choose three principal directions from PCA to calculate the Eq.(3). For Eq.(4) and Eq.(12), the S^L and S^R are the surrounding RBI of P_L's and P_R's adjacent pixels, respectively. We want to highlight the function of discontinuities, thus w_1 is 1 while w_2 is 100. For η in Eq.(6), μ is 1. Different images need different parameters, while for aforementioned images, these parameters work effectively.

We conducted two experiments to validate the proposed method. First is to show the accuracy of the test images and the second is to illustrate the improvement of state-of-the-art methods by our algorithm.

The accuracy of compared methods and ours are shown in Table 1. In Table 1, *con*, *all* and *dis* denote the continuities, all areas and discontinuities, respectively. Results show that our method works pretty well, especially in discontinuities.

We compare our method with four discontinuities-preserving methods, including LLR, imprNLCA, Borders, and LCVB-DEM. We employ our method to improve other methods. This is convenient with less complexity, because the input for our method is just the discontinuous areas. As shown in Figure 9, all these methods are improved by our algorithm with respect to the accuracy of *all* areas and therefore, the bottleneck problem has been addressed effectively. Figure 10 shows the disparity maps of LLR, imprNLCA, Borders, LCVB-DEM and our results. From Figure 10, our results have an unambiguous discontinuous areas and less unmatched pixels which are shown as the 'black areas' on the disparity map.

5. CONCLUSIONS

This paper proposes an algorithm to preserve discontinuities in the disparity map. We form an energy function based on our novel 3D space. The discontinuity information is obtained from radial basis information. We calculate the difference of direction vectors for the cost of matching, and use the constraints of the region, position and uniqueness to reduce the search for a solution. To obtain the minimum cost, this function is deduced as an optimization problem which can be solved by HNN. By updating the neurons from the activation function, the global solution is generated automatically which is the optimal disparity map for the input images.

To evaluate our method, we show our results on the Middlebury benchmark. Our method outperforms state-of-the-art discontinuities-preserving methods, especially in terms of discontinuities. Apart from this, it can be used to augment other state-of-the-art approaches by preserving their discontinuities to solve the bottleneck of stereo matching.
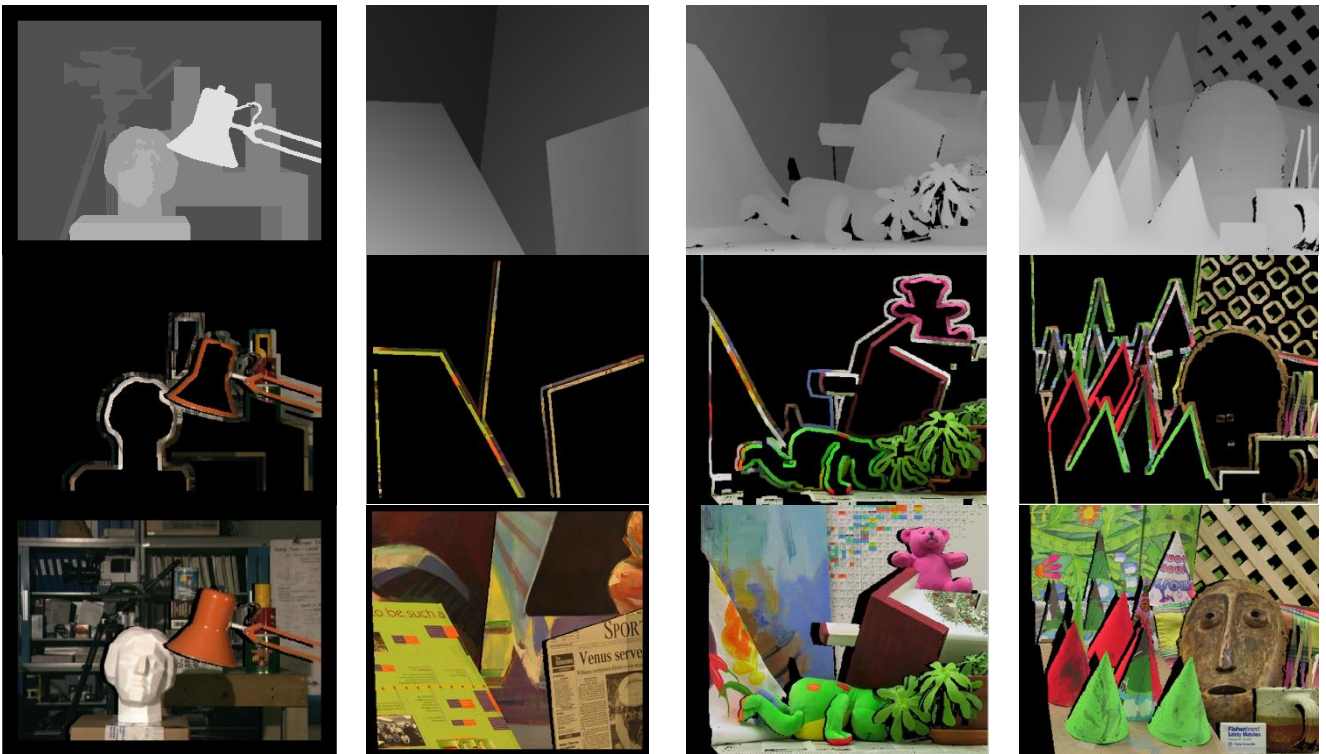
Figure 8. Ground truth, discontinuities, and continuities. Ground truth (first row), discontinuities (second row), continuities (third row).

Method	Tsukuba			Venus			Teddy			Cones		
	con	*all*	*dis*	*con*	*all*	*dis*	*con*	*all*	*dis*	*con*	*all*	*dis*
LLR	1.05	1.65	5.64	0.29	0.81	3.07	4.56	9.81	12.2	2.17	8.02	6.42
imprNLCA	1.38	1.83	7.38	0.21	0.41	2.26	5.99	11.5	14.3	2.85	6.68	7.93
Borders	1.29	1.71	6.83	0.25	0.53	2.26	7.02	12.2	16.3	3.90	9.85	10.2
LCVB-DEM	4.49	5.23	21.3	1.32	1.67	11.5	9.99	16.3	26.1	6.56	13.6	18.2
our method	1.18	1.79	4.89	0.20	0.47	1.91	4.82	9.84	11.3	6.07	9.47	6.34

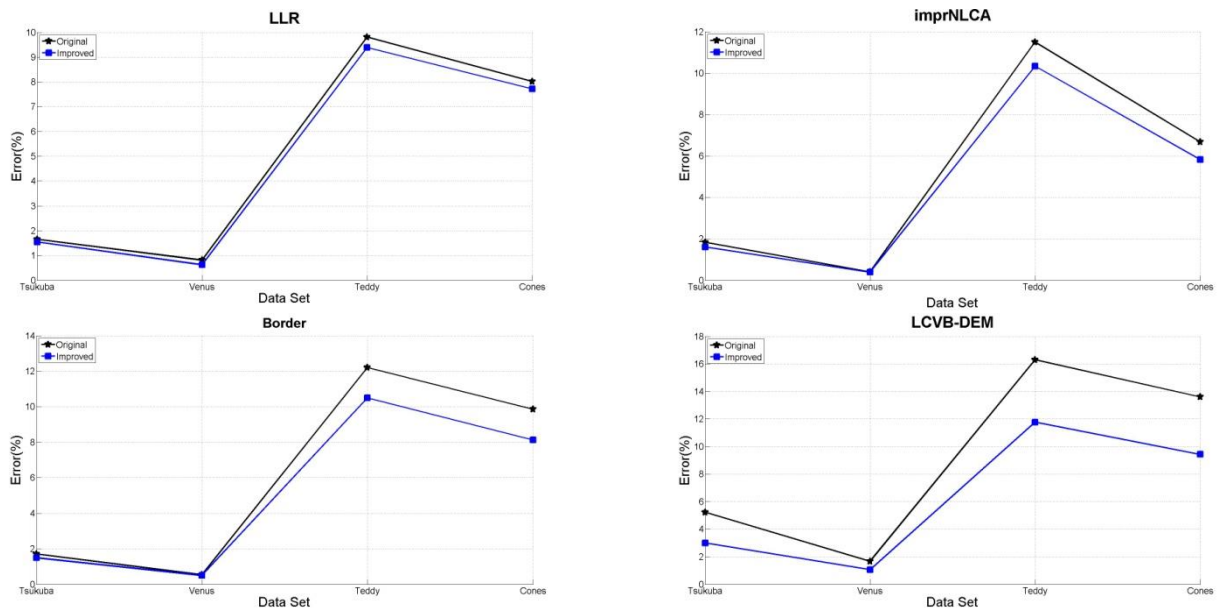Table 1. Comparison of results with error threshold 1.0

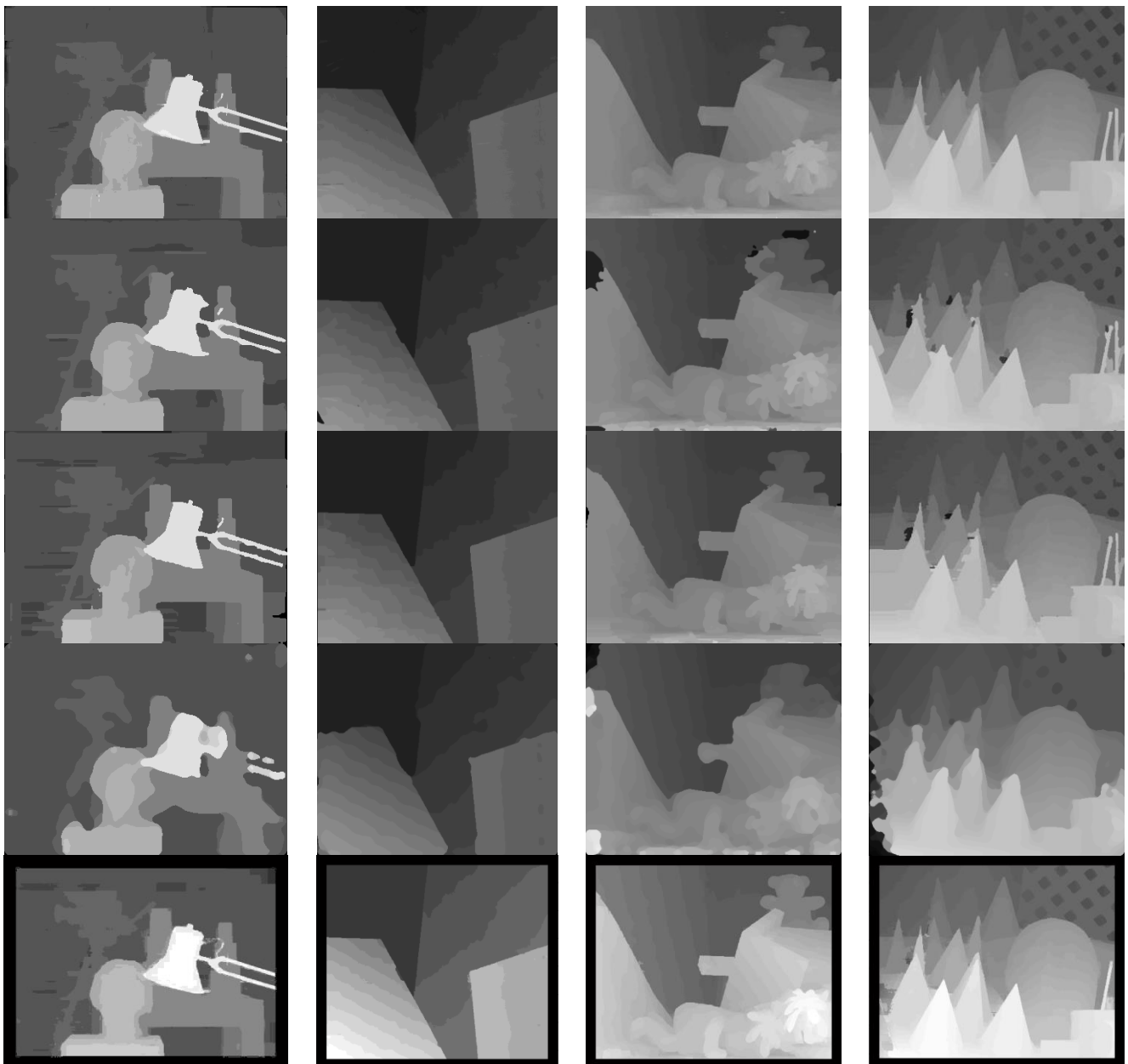Figure 9. The improvement of start-of-the-art methods on *all* areas.

Figure 10. Results of different methods. LLR (first row), imprNLCA (second row), Borders (third row), LCVB-DEM (forth row) and our method (last row).

REFERENCES

Achour, K. and Mahiddine, L., 2002. Hopfield neural network based stereo matching algorithm. *Journal of Mathematical Imaging and Vision*,16(1), pp. 17–29.

Boykov, Y., Veksler, O. and Zabih, R., 2001. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11), pp. 1222–1239.

Buhmann, M. D., 2004. Radial basis functions: theory and implementations. *Cambridge Monographs on Applied and Computational Mathematics*, 12, pp. 147–165.

Chen, D., Ardabilian, M., Wang, X. and Chen, L., 2013. Animproved non-local cost aggregation method for stereo matching based on color and boundary cue. *IEEE International Conference on Multimedia and Expo*, pp. 1–6.

Cheng, F., Zhang, H., Sun, M. and Yuan, D., 2015. Cross-trees, edge and super-pixel priors-based cost aggregation for stereo matching. *Pattern Recognition*, 48(7), pp. 2269–2278.

Farid, H. and Simoncelli, E. P., 1997. Optimally rotation equivariant directional derivative kernels. *Computer Analysis of Images and Patterns*, pp. 207–214.

Hopfield, J. J., 1982. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8), pp. 2554–2558.

Huang, C.-h. andWang, T.-h., 2000. Stereo correspondence using hopfield network with multiple constraints. *IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 2, pp. 1518–1523.

Laskowski, Ł., Jelonkiewicz, J. and Hayashi, Y., 2015. Extensions of hopfield neural networks for solving of

stereomatching problem. *Artificial Intelligence and Soft Computing*, pp. 59–71.

Lee, J. J., Shim, J. C. and Ha, Y. H., 1994. Stereo correspondence using the hopfield neural network of a new energy function. *Pattern Recognition*, 27(11), pp. 1513–1522.

Martins, J. A., Rodrigues, J. M. and du Buf, H., 2015. Luminance, colour, viewpoint and border enhanced disparity energy model. *PloS one*, 10(6), pp. 1-24.

Mattoccia, S., Tombari, F. and Di Stefano, L., 2007. Stereo vision enabling precise border localization within a scanline optimization framework. *Asian Conference on Computer Vision*, pp. 517–527.

Nasrabadi, N. M. and Choo, C. Y., 1992. Hopfield network for stereo vision correspondence. *IEEE Transactions on Neural Networks*, 3(1), pp. 5–13.

Ogale, A. S. and Aloimonos, Y., 2005. Shape and the stereo correspondence problem. *International Journal of Computer Vision*, 65(3), pp. 147–162.

Scharstein, D. and Szeliski, R., 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1-3), pp. 7–42.

Sun, J., Zheng, N.-N. and Shum, H.-Y., 2003. Stereo matching using belief propagation., *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(7), pp. 787–800.

Tombari, F., Mattoccia, S. and Di Stefano, L., 2007. Segmentation-based adaptive support for accurate stereo correspondence. *Advances in Image and Video Technology*, pp. 427–438.

Yang, Q., 2012. A non-local cost aggregation method for stereo matching. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1402–1409.

Zhu, S., Zhang, L. and Jin, H., 2012. A locally linear regression model for boundary preserving regularization in stereo matching. *European Conference on Computer Vision*, pp. 101–115.