GRAPH-BASED RGB-D IMAGE SEGMENTATION USING COLOR-DIRECTIONAL-REGION MERGING

Xiong Pan[1], Zejun Zhang[1,*], Yizhang Liu[1], Changcai Yang[1],
Qiufeng Chen[1], Li Cheng[2], Jiaxiang Lin[1], Riqing Chen[1]
{[1]the Digital Fujian Institute of Big Data for Agriculture and Forestry, College of Computer and Information Sciences; [2]Jinshan college}, Fujian Agriculture and Forestry University, Fuzhou, China
zjzhang_fafu@163.com

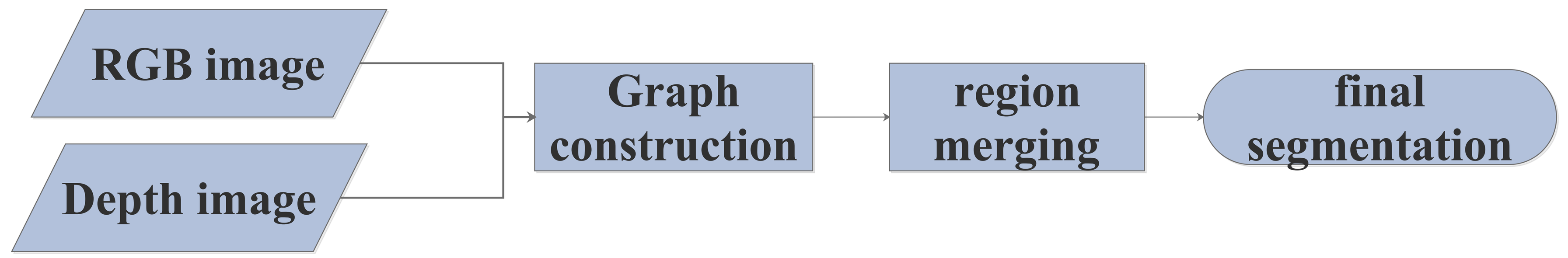
1. Motivation

1) Image segmentation is an important task in image and video processing and has many applications.

2) Based on the RGB-D image data, the quality of segmenting scenes can by improved via exploiting color and depth information simultaneously.

So, we proposed a novel graph-based RGB-D image segmentation algorithm for indoor scene.

2. Outline of the method



3. The Proposed Method

- Dissimilarity of two adjacent pixels

The dissimilarity of two adjacent pixels, p_i and p_j, consists of two terms, color- and normal-variation between these two pixels, and they are linearly combined using adaptive data-driven weight. The dissimilarity of p_i and p_j is given by

$$\kappa(p_i,p_j)=\frac{w(p_i,p_j)\cdot c(p_i,p_j)+v(p_i,p_j)\cdot f(p_i,p_j)}{w(p_i,p_j)+v(p_i,p_j)}$$

where $c(p_i, p_j)$ and $f(p_i, p_j)$ are color- and normal-variation between p_i and p_j, respectively. $w(p_i, p_j)$ and $v(p_i, p_j)$ are data-driven weight calculated through using color information of pixels around these two pixels.

The color variation $c(p_i, p_j)$ is from [1]

$$c(p_i,p_j)=\sqrt{r(p_i,p_j)^2+g(p_i,p_j)^2+b(p_i,p_j)^2}$$

and the normal variation is defined as

$$f(p_i,p_j)=1-\cos\left(\left\langle \vec{n}_{p_i},\vec{n}_{p_j}\right\rangle\right)$$

where $\vec{n}_{p_i}$ and $\vec{n}_{p_j}$ denote the normal vectors at pixel p_i and p_j, respectively.

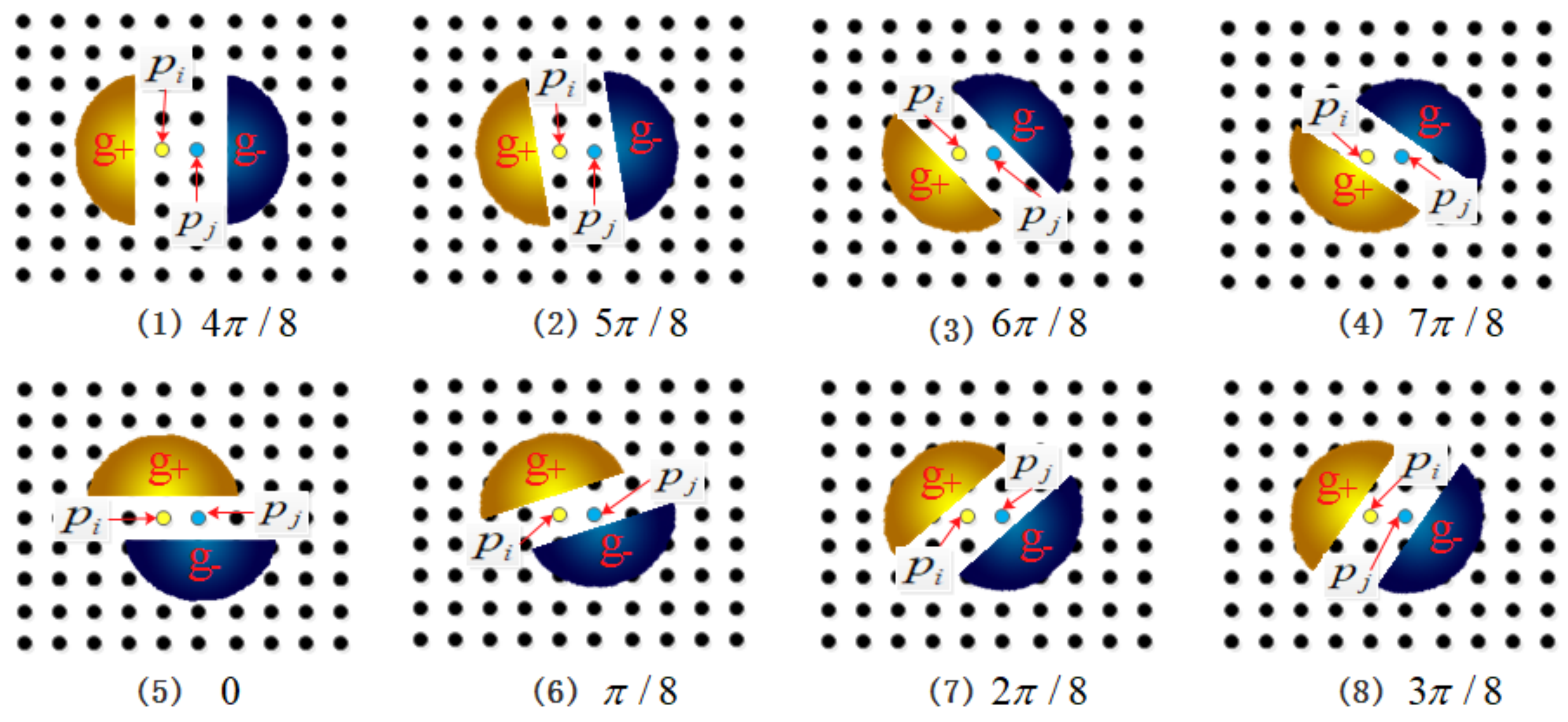


Fig. 1 Bi-semi-Gaussian functions.

The weight $w(p_i, p_j)$ and $v(p_i, p_j)$ are defined as

$$w(p_i,p_j)=\max_{\theta}\left\{\nabla(p_i,p_j;\theta)\right\}$$

$$v(p_i,p_j)=\exp\left(w(p_i,p_j)-\alpha\right)$$

and $\nabla(p_i,p_j;\theta)$ is defined as

$$\nabla(p_i,p_j;\theta)=\left|m_{g_+}(p_i,\theta)-m_{g_-}(p_j,\theta)\right|$$

α is a constant, the bi-semi-circle regions at direction θ, as shown in Fig. 1.

- Region merging

Inspired by [1], the proposed criterion consists of three aspects, as shown in Fig. 2.

The termination threshold value is defined as:

$$T(\Omega_i;\Omega_j)=\left[\varepsilon-\varphi(\Omega_i,\Omega_j)\right]/N_i$$

$$T(\Omega_j;\Omega_i)=\left[\varepsilon-\varphi(\Omega_i,\Omega_j)\right]/N_j$$

where ε is a constant; N_i and N_j are the number of pixels in the regions, respectively. $\varphi(\Omega_i,\Omega_j)$ is defined as

$$\varphi(\Omega_i,\Omega_j)=1-\cos\left\langle \vec{n}_{\Omega_i},\vec{n}_{\Omega_j}\right\rangle$$

where $\vec{n}_{\Omega_i}$ and $\vec{n}_{\Omega_j}$ are average normal in regions Ω_i and Ω_j, respectively.

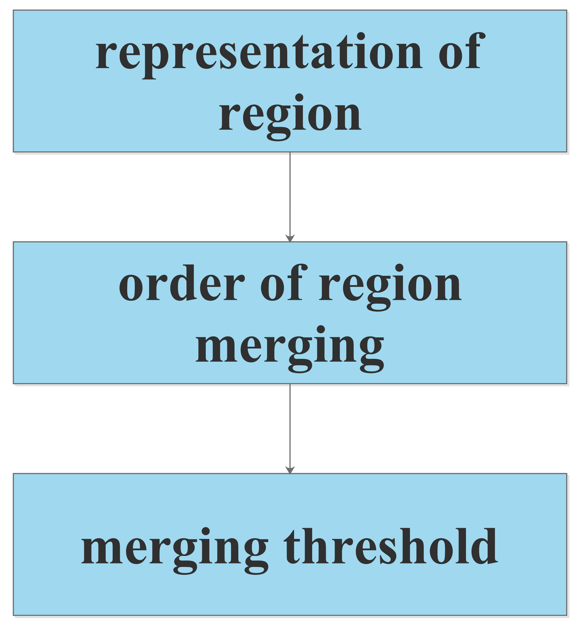


Fig. 2 Three aspects

REFERENCES

[1] P.F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation", International Journal of Computer Vision, vol. 59, no. 2, pp. 167-181, 2004.

4. Experiments

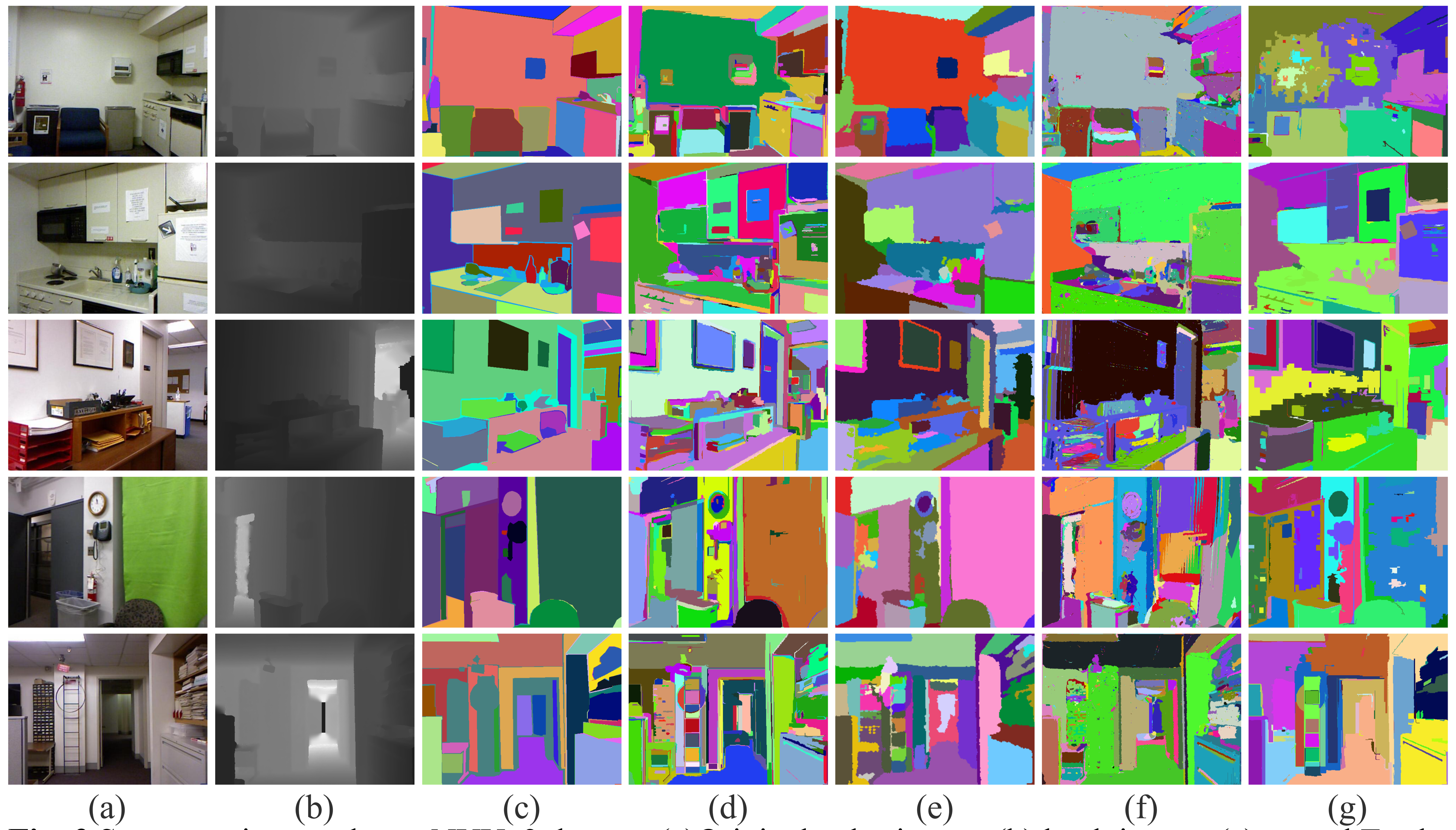


Fig. 3 Segmentation results on NYUv2 dataset. (a)Original color image, (b)depth image, (c)ground Truth, (d)our method, (e)JSCD-RM, (f)PIS, and (g)GB-RGBD.

Table 1. Comparing with the state-of-the-art methods.

	GTRC			PRI		VI		P	R	BFM
	ODS	OIS	Best	ODS	OIS	ODS	OIS			
PIS	0.43	-	-	0.88	-	3.16	-	0.37	0.69	0.48
GB-RGBD	0.44	0.49	0.56	0.87	0.89	2.46	2.34	0.40	0.60	0.48
JCSD-RM	**0.55**	-	-	**0.91**	-	**2.12**	-	**0.56**	0.43	0.49
OUR METHOD	0.50	**0.55**	**0.61**	0.90	**0.91**	2.42	**2.25**	0.38	**0.76**	**0.51**

ODS: a universal fixed scale OIS: a fixed scale per image
Best: from any level of the hierarchy or collection Ground truth: extracted from [5]

Table 2. Comparing with the proposed and JCSD method.

	GTRC	PRI	VI	BFM
JCSD	0.46	0.87	2.68	0.46
OUR METHOD	**0.50**	**0.90**	**2.42**	**0.51**