# MINIMUM SPANNING DISTANCE FOR IMAGE SEGMENTATION

*Chao-Te Chou*[†] *Wei-Chih Tu*[‡] *Shao-Yi Chien*[†‡]

National Taiwan University, Taipei, Taiwan
[†]Department of Electrical Engineering [‡]Graduate Institute of Electronics Engineering

## ABSTRACT

In this paper, we review the design of Minimum Barrier Distance and propose a new path-wise distance metric called Minimum Spanning Distance (MSD). Unlike most existing distance metrics, which only define distance between two pixels on gray-scale images, the proposed distance metric conceptually estimates the color space spanned by the colors on the path of interest. Therefore, the MSD takes into consideration the three channels on color images at the same time to compute distance. Compared with other distance metrics, MSD can not only achieve the highest numerical scores but also produce visually good segmentation maps in our experiment of interactive segmentation on the Gulshan dataset.

***Index Terms***— Distance Transform, Interactive Segmentation

## 1. INTRODUCTION

Distance transform aims to find the minimum distance for each pixel to a set of target pixels (or called seeds). Based on different applications and input modality, many distance metrics have been studied over the years. In general, distance metrics can be roughly classified into two categories. One is the point-wise method, which measures distances between two pixels directly. The commonly used Euclidean distance belongs to this category. The other class is the path-wise method, which measures distance depending on the spatial path between two pixels. The geodesic distance [1] and barrier distance [2] lie in this category.

Both distance metrics have their merits in different applications. In this paper, we consider image segmentation as an application. Fig. 1 illustrates the case of interactive object segmentation. The user provides some scribbles in foreground object and background regions. We can perform object segmentation by computing distance transform to the drawn foreground pixels and background pixels respectively. We then classify all pixels as foreground or background according to their distance to the two sets of targets. To this end, the path-wise distance usually has favorable performance against point-wise distance since point-wise distance measurement usually ignores connectivity between pixels. Previously, Criminisi *et al.* [3] have exploited geodesic distance transform in a segmentation system. Another application that path-wise distance is favored against point-wise distance is salient object detection. Wei *et al.* [4] proposed a useful assumption that boundary pixels are mostly background. They perform geodesic distance transform to measure the connectivity of a pixel to its nearest boundary and use the distance as the saliency value.

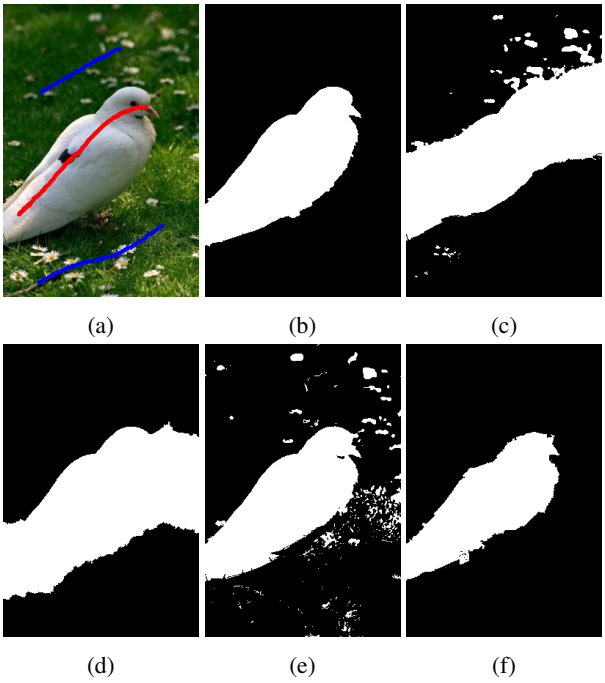
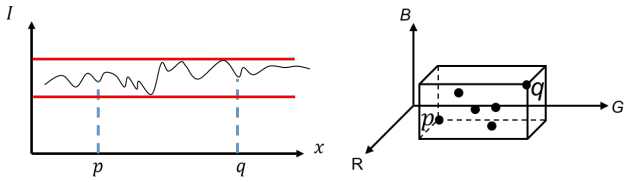


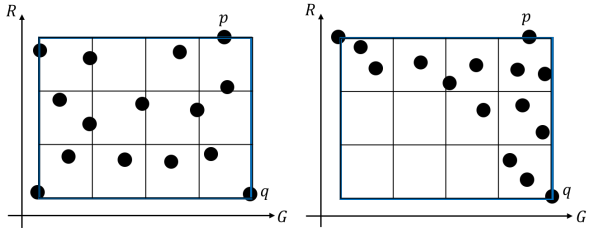

**Fig. 1**: We show the results of interactive image segmentation with different distance transforms. (a) Input image and user specified scribbles (b) Groundtruth (c) Euclidean distance (d) Geodesic distance (e) MBD (f) The proposed MSD

Recently, a new path-wise distance metric, Minimum Barrier Distance (MBD) [2, 5], was shown to have low sensitivity to noise, blur and the position of seeds. Recent works [6, 7] in salient object detection followed the boundary prior assumption [4] and used MBD as the distance metric to achieve state-of-the-art performance. They also noticed the superiority of barrier distance over geodesic distance.




(a) Geodesic and barrier distance (b) Vectorial MBD

**Fig. 2**: Illustrations of path-wise distance metrics




(a) Path 1 (b) Path 2

**Fig. 3**: We show a color space example that VMBD fail to properly measure distance. (a) and (b) illustrate two different color distributions (ignore B axis for better visualization) while they have the same VMBD value.

Despite the success of MBD in saliency detection, however, we notice the limitation of MBD as it is defined only in grayscale images. Previous works applied MBD to color images by computing MBD for each color channel separately and average them to get the final distance map. We argue that this simple integration could work in saliency detection with boundary prior since image boundary (background region) is usually simple and homogeneous, but it is not the case for segmentation. When we consider the case of scribble-based segmentation as described earlier, the color distribution of a scribble might be diverse (*i.e.* target seeds have many color appearance) and it can cause the distance transform noisy, resulting many false segmentation as shown in Fig. 1(e).

In this paper, we review the design of MBD and generalize the philosophy to color space as a new path-wise distance metric. The proposed distance metric conceptually estimates the color space spanned by the colors on the path of interest and finds the minimum among all candidate paths, so we term this metric as the Minimum Spanning Distance (MSD). We perform evaluation on the Gulshan dataset [8] and the results show that MSD achieves better performance in both cross-entropy and weighted $F_\beta$ [9] than MBD.

## 2. BACKGROUND

We briefly review the path-wise distance metrics before we introduce the proposed MSD. For simplicity, consider an image as a 4-connected planar graph. A path between two pixels of interest, $p$ and $q$, is a consecutive pairs of pixels and can be expressed as $\pi = \{\pi(0), ..., \pi(n)\}$, where $p = \pi(0)$, $q = \pi(n)$, and consecutive pixels are adjacent neighbors. Given a path-wise distance metric $f(\pi)$, the distance transform for pixel $p$ with respect to a seed set $S$ can be written as

$$D(p) = \min_{\pi \in \Pi_{p,S}} f(\pi), \tag{1}$$

where $\Pi_{p,S}$ is the set of all paths between $p$ and seeds in $S$.

The geodesic distance [1] takes the accumulated changes of all traversed pixels as distance. It can be written as

$$f_{Geo}(\pi) = \sum_{i=0}^{n-1} |I(\pi(i+1)) - I(\pi(i))|. \tag{2}$$

The barrier distance [2], on the other hand, takes the difference of maximum and minimum values (*i.e.* the intensity barrier) on the path as distance. It can be expressed as

$$f_{MBD}(\pi) = \max_{i=0}^{n} I(\pi(i)) - \min_{i=0}^{n} I(\pi(i)). \tag{3}$$

Both $f_{Geo}$ and $f_{MBD}$ were designed to measure the variation along a path. Fig. 2(a) shows a 1D illustration of these two distance metrics. The length of the curve captures the geodesic distance and the margin spanned by the two red lines is the barrier distance of pixel $p$ and $q$. This example also simulates small intensity fluctuation of weak texture or noise in images. We can also see why barrier distance is more robust here. The geodesic distance tends to accumulate all small changes along the path. The barrier distance is robust to such textural or noisy signal better as the barrier size keeps almost the same within texture or noisy regions.

MBD in Eq. 3 only defines in grayscale images. However, most of the time we deal with color images. The idea of barrier distance for higher dimensional data is not clear. Previous works [6, 7] exploit MBD by computing distance transform for each channel separately and average them as final distance. Such simple integration treats each channel as independent signal and the paths traversed in different channels may be different, which sometimes causes difficulty to separate an object from cluttered background. Previously, the Vectorial Minimum Barrier Distance (VMBD) [10] was presented to extend MBD to higher dimensional data. For color images, the VMBD computes the following

$$f_{VMBD}(\pi) = B_r(\pi)B_g(\pi)B_b(\pi), \tag{4}$$

where $B_c(\pi), c \in r, g, b$ computes the barrier distance along the $c-th$ channel. In the geometric view, VMBD computes the volume of the minimal axis-aligned bounding box enclosing all colors on the path as illustrated in Fig. 2(b). We argue that VMBD is not a good generalization for MBD in color space. Again, consider paths between pixel $p$ and $q$. Fig. 3 depicts two possible color distributions resulting in the same VMBD value. Yet, it is clear that the color distribution on

path 1 is more diverse, so we expect the measured distance on path 1 should be larger than that on path 2. However, VMBD can not discriminate the two cases properly.

## 3. MINIMUM SPANNING DISTANCE

Based on the discussion in previous section, we think MBD in color space is not as simple as finding the barrier along each axis and compute the product. The spirit of MBD as well as other path-wise distance metrics is essentially to measure the appearance variation along a path. By definition in Eq. 3, MBD is computed as difference of the maximum value and the minimum value on a path. In another view, it finds the length on the intensity axis occupied by the pixel values visited along a path. With the concept in mind, we proposed to generalize the idea of MBD to color space by using the space spanned by the colors visited along a path as a distance metric. In this way, we can successfully distinguish the difference of the two paths in Fig. 3. For a pixel of interest $p$, we compute the spanning distance for all paths connecting $p$ and seeds in $S$ and find the minimum distance among them. Therefore, we term the proposed distance as the Minimum Spanning Distance (MSD).

In practice, we compute MSD by counting the number of colors visited along a path. Human eyes can not discriminate minor difference among similar colors, which allows us to quantize the color space into discrete boxes and count the number of boxes instead of colors. We consider 8-bit dynamic range (0-255) for each color channel. In order to divide the space into boxes with identical volume, we choose the box size to be $2^k$. For example, if the box size is 16, then each axis is divided into 16 sections and there will be $16^3$ boxes in total. For convenience, we introduce $B_p$ as the box index of the color possessed by pixel $p$. We also maintain an index list $L(\pi)$ for each pixel. $L_p(\pi)$ stores the indices of visited boxes along path $\pi$ connecting $p$ to $S$. The proposed MSD can be simply written as

$$f_{MSD}(\pi) = len(L(\pi)), \tag{5}$$

where $len(.)$ returns the number of indices in a list.

Despite the quantization helps reduce the number of colors to track, computing exact MSD for all pixels is still challenging. Inspired by the fast geodesic distance transform [1] and approximate MBD [6], we also use raster scan to compute MSD. The raster scan technique traverse an image in forward and backward passes iteratively. Consider an image as a 4-connected graph, in the forward pass, we update each pixel from its upper and left neighbors. In the backward pass, we update each pixel from its lower and right neighbors.

Let $p$ be the current pixel in the updating step and $q$ be an adjacent pixel of $p$. We assume that $q$ has been visited in current iteration. We check the index list $L_q$ holding by $q$ to see if updating from $q$ will result in smaller distance for $p$. We first compute a temporary index list,

$$L_t = L_q \cup B_p. \tag{6}$$

Then, we compare $L_t$ with $L_p$ and update the distance as

$$D(p) = \min\{len(L_t), len(L_p)\}. \tag{7}$$

If $D(p)$ is updated by $L_t$, we will also update $L_p = L_t$.

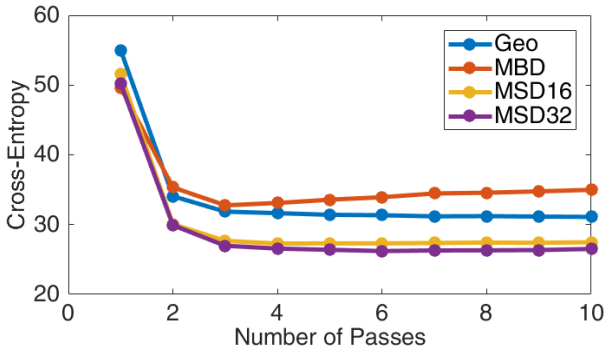



**Fig. 4**: Convergence analysis for raster-scan based distance transform.

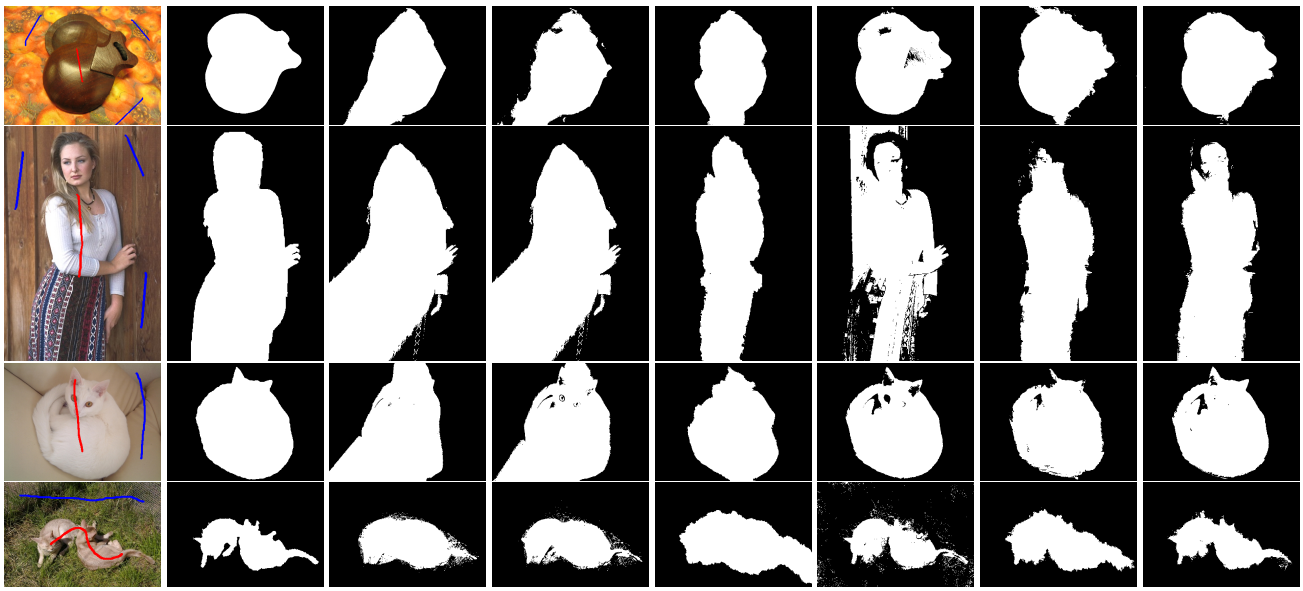
## 4. EXPERIMENTS

We perform evaluations in interactive segmentation using the Gulshan dataset [8]. In this dataset, 151 natural images are provided with groundtruth object mask along with computer-generated scribbles indicating foreground and background regions. For all distance metrics, we perform segmentation simply by running distance transform with respect to foreground or background seeds. Denote the distance maps as $D_F$ and $D_B$. We obtain the final segmentation map by

$$S_i = \begin{cases} 0, & \text{if } D_F(i) < D_B(i) \\ 1, & \text{othwise} \end{cases}$$

The target of interactive objection segmentation is to classify pixels into two classes, foreground or background, so we use cross-entropy as one of our performance metrics. Lower cross-entropy value indicates better segmentation results. Furthermore, we also perform evaluation using weighted $F_\beta$ [9]. This metric can better evaluate the segmentation map since it takes into consideration the location of errors in the predicted maps. The higher weighted $F_\beta$ means better segmentation performance.

We denote our method using box size 16, 32, 64 as MSD16, MSD32, and MSD64 respectively. Note that MSD128 is meaningless as box size 128 is too large and box size smaller than 8 will cause large memory consumption. We run above MSD variants for iterations to convergence. As we show on the right of Table. 1, MSD64 has poor performance due to over quantization. MSD16 and MSD32 are



(a) Input (b) Groundtruth (c) EUC1 (d) EUC3 (e) Geo (f) MBD (g) MSD16 (h) MSD32

**Fig. 5**: Visual comparison of our MSD with other distance metrics in image segmentation results.

**Table 1**: Numerical evaluation on Gulshan dataset.

| Distance metrics | EUC1 | EUC3 | Geo | MBD | MSD16 | MSD32 | MSD64 |
|---|---|---|---|---|---|---|---|
| Cross-Entropy | 29.27 | 28.64 | 31.34 | 33.93 | 27.31 | **26.21** | 32.81 |
| Weighted $F_\beta$ | 0.6631 | 0.6706 | 0.6469 | 0.6166 | **0.6821** | 0.6807 | 0.5812 |

considered as suitable settings. We compare with other path-wise distance metrics geodesic distance (Geo) and MBD as well as point-wise Euclidean distance. All path-wise distance metrics are implemented using 4-connected graph. For color images, Geo and MBD are performed channel by channel and then we average the maps as final distance. For Euclidean distance, we can also compute distance channel by channel and average afterwards. We denote it as EUC1. We can also compute pixel-wise distance for all channels at the same time. We denote this variant as EUC3.

We have conducted iteration analysis to determine the number of passes needed for the raster-scan based distance transform. We present the results in Fig. 4 and it suggests that 6 iterations are sufficient for all methods to converge. The experimental results based on 6 scan passes (3 forward and 3 backward) are shown in Table. 1. Our MSD16 and MSD32 consistently achieve the top two in both cross-entropy and weighted $F_\beta$. We also show some visual comparison in Fig. 5. The first row shows an image with simple foreground object and complex background region, our MSD32 method can produce the best segmentation result. The the second row shows an woman whose hair color is similar to the background wood color. It shows the weakness of both the our method and MBD. When the color of foreground object and background region are too similar, both method can not produce favorable segmentation result. However, MSD will not over extend the region of foreground object to background region. The images in the third and forth rows show that even with only one foreground scribble and one background scribble MSD can still produce good quality segmentation result. Furthermore, the forth row shows that MBD sometimes produces segmentation with some white dots noise in the background region. In conclusion, the proposed MSD has better performance both numerically and subjectively.

## 5. CONCLUSION

In this paper, we proposed the Minimum Spanning Distance (MSD), whcih estimates the color space spanned by the colors on the path of interest. By simply quantizing the color space into discrete boxes and counting the number of boxes on a path, our method takes into consideration the three channels on color images at the same time to compute distance. Therefore, it does not simplify the three channels information into one dimension as other distance metrics defined on grayscale image do. Evaluating MSD in the task of interactive segmentation, it can achieve the highest numerically score and provides better visual results.

**Acknowledgements.** This work is partially supported by Ministry of Science and Technology of Taiwan, under Grant MOST 106-2221-E-002-215-MY3.

## 6. REFERENCES


[1] Pekka J Toivanen, "New geodosic distance transforms for gray-scale images," *Pattern Recognition Letters*, vol. 17, no. 5, pp. 437–450, 1996.

[2] Robin Strand, Krzysztof Chris Ciesielski, Filip Malmberg, and Punam K Saha, "The minimum barrier distance," *Computer Vision and Image Understanding*, vol. 117, no. 4, pp. 429–437, 2013.

[3] Antonio Criminisi, Toby Sharp, and Andrew Blake, "Geos: Geodesic image segmentation," *European Conference on Computer Vision*, 2008.

[4] Yichen Wei, Fang Wen, Wangjiang Zhu, and Jian Sun, "Geodesic saliency using background priors," *European Conference on Computer Vision*, 2012.

[5] Krzysztof Chris Ciesielski, Robin Strand, Filip Malmberg, and Punam K Saha, "Efficient algorithm for finding the exact minimum barrier distance," *Computer Vision and Image Understanding*, vol. 123, pp. 53–64, 2014.

[6] Jianming Zhang, Stan Sclaroff, Zhe Lin, Xiaohui Shen, Brian Price, and Radomir Mech, "Minimum barrier salient object detection at 80 fps," in *IEEE International Conference on Computer Vision*, 2015.

[7] Wei-Chih Tu, Shengfeng He, Qingxiong Yang, and Shao-Yi Chien, "Real-time salient object detection with a minimum spanning tree," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[8] Varun Gulshan, Carsten Rother, Antonio Criminisi, Andrew Blake, and Andrew Zisserman, "Geodesic star convexity for interactive image segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2010.

[9] R. Margolin, L. Zelnik-Manor, and A. Tal, "How to evaluate foreground maps," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014.

[10] Andreas Kårsnäs, Robin Strand, and Punam K Saha, "The vectorial minimum barrier distance," in *IEEE International Conference on Pattern Recognition*, 2012.