

# TOWARD A RECOMMENDATION SYSTEM FOR IMAGE SIMILARITY METRICS

Peter Bajcsy, Joe Chalfoun and Mary Brady  
National Institute of Standards and Technology  
100 Bureau Drive, Gaithersburg, MD 20899, USA  
{peter.bajcsy, joe.chalfoun, mary.brady}@nist.gov

## ABSTRACT

This paper addresses the problem of mapping application specific requirements on image similarity metrics to the plethora of existing image similarity computations. The work is motivated by the fact that there is no method for choosing a similarity metric that is suitable for a given application. We approached the problem by designing a theoretical and experimental framework for creating sensitivity signatures of similarity metrics. In this paper, we outline the classifications of image similarity metrics found in the literature, the space of application parameters and requirements, derivations of similarity dependencies on application parameters, and experimentally obtained sensitivity signatures of similarity metrics using image simulations. These sensitivity signatures provide a way for users to query a reference database of sensitivity signatures and retrieve a recommendation for an image similarity metric. We illustrate the use of the prototype recommendation system by considering spectral calibration and spatial registration application requirements.

## KEY WORDS

SYMPOSIUM 1: IMAGING AND SIGNAL  
PROCESSING: MEDICAL IMAGE PROCESSING

## 1 Introduction

Many conclusions in the biomedical field are based on comparing measured and reference observations. One of the fundamental components of these comparisons is a measure of similarity. Visual inspection and the human perception of similarity play a prominent role in biomedical research. However, advances in imaging and the corresponding growth of image data lead to an increasingly high demand for the automation of visual comparisons and for a transition from expert-applied image similarity to computer-applied image similarity. The overwhelming data volumes represent our major motivation for the automation of visual comparisons and for building reference implementations of measures of image similarity.

The goal of this work is to address the mapping between application specific requirements on image similarity metrics and the plethora of existing image similarity computations. Our objective is to provide support for finding similarity metrics that closely match application requirements on image comparisons. In order to achieve the goal, several challenges need to be

addressed: (1) organize image similarity metrics according to various application criteria, (2) model the functional dependencies of image similarity measures in a large space of application parameters and requirements, (3) develop signatures of similarity metrics that can be matched against user's inputs, and (4) prototype a recommendation framework for similarity metric selection.

To address such a wide spectrum of challenges, we leverage several existing surveys of image similarities [1–9]. The surveys allow us to introduce a tree-based taxonomy of similarity measures. Furthermore, in order to manage the large number of existing image similarity metrics, we narrow down our metric implementations and experimental studies on labeled images and single band grayscale images. These particular sub-categories of images are aligned with our primary interest in using similarity metrics for evaluating reference-based image quality, accuracy, and consistency of microscopy image segmentation results, and accuracy and robustness of tracking and registration computations in cell biology.

Our main approach toward a recommendation system for choosing image similarity metrics consists of (1) consolidating classifications of image similarity metrics found in the literature, (2) identifying and simulating the space of application parameters and requirements, (3) deriving similarity dependencies on application parameters from synthetic images to create sensitivity signatures of similarity metrics, and (4) prototyping a recommendation framework for similarity metric selection.

Our unique contribution lies in a theoretical and experimental data driven characterization of similarity metrics that maps to application specific requirements on image similarity metrics. This initial work provides an opportunity for building a fully developed reference database of sensitivity signatures of similarity metrics, for providing selection recommendations of similarity metrics driven by application specific requirements, and for eventually delivering web-based image comparison services.

## 2 Previous Work on Image Similarity

Past work on image similarity metrics [1–9] focused primarily on similarity computations that are based on (1) binary image overlap, (2) image object geometry, (3) color vectors, (4) histograms of intensity, or (5) texture. The overlap and object geometry based

similarity computations are frequently researched in segmentation papers [10–15]. For example, Zhang et al. [12] focused on unsupervised segmentation methods that depend on the choice of similarity metrics. The similarity metrics are classified according to segmentation use as intra-region, inter-region and composite pixel comparisons, and include considerations of texture, entropy and various color descriptors. Cha [16] surveyed extensively histogram-based similarity metrics. Others, e.g. Sampat et al. [7], surveyed similarity metrics based on multiple image characteristics including intensity, overlap of binary images, and image object geometry. Given the difficulty of defining texture, studies of similarity metrics involving texture descriptors have always been of interest to segmentation and content-based retrieval applications [9], [17]. Many 2D image similarity metrics are also frequently used in studying 3D images (see Benhabiles [18]). Some papers not only state mathematical definitions of image similarities but also attempt to relate mathematical definitions of similarity metrics with various visual perception criteria [1], [7], [8], [10]. These studies are driven by the need to automate visual inspections and hence replicate human perception of image similarity.

The literature survey indicates a lack of a comprehensive classification of similarity computations. Furthermore, there is a lack of understanding on how to map application requirements to specific computations of similarity metrics. While there are many papers proposing new similarity metrics [7], [8], [13], the lack of a recommendation system for choosing a similarity metric according to application specific requirements motivates our work.

### 3 Approach

To address the aforementioned problems, we first classified similarity metrics in Table 1 based on the main image similarity applications and their requirements. The requirements are organized in terms of desirable invariance and sensitivity properties to image acquisition variables and any changes in an imaged field of view. A user specifies the sensitivity of a subset of high level variables in Table 1 as inputs to search for a similarity computation matching the application requirements.

#### 3.1 A Modeling Framework

In order to match user’s inputs, we created a modeling framework consisting of three sample spaces as shown in Figure 1. The original image content space  $\{I_{i,j}^{Ref}\}$  consists of any reference images that can be obtained by simulations or by controlled experiments. The reference images are represented by a variable number of bits per pixel, number of pixels, and number of image bands/channels. The index  $i$  refers to an image content generating function (e.g., random number generators,

rule-based generators, or experimental settings), and  $j$  is the parameter of the generating function (e.g., probability distribution parameters, camera aperture values). In general, experiments are characterized by varying illumination, imaging parameters or object properties, while simulations are based on developing deterministic and statistical generative models.

**Table 1: Sensitivity and invariance of similarity metrics to image acquisition variables and content changes**

Application vs. Requirements	Sensitivity	Invariance
<b>Image Spectral Calibration</b>	Intensity changes	Translation, rotation, scale, skew, symmetry
<b>Image Spatial Registration</b>	Translation, rotation, scale, skew, symmetry	Intensity and shape changes
<b>Image Segmentation</b>	Contour changes, pixel level region overlaps	Intensity changes
<b>Content-Based Image Retrieval</b>	Coarse intensity and shape changes	Translation, rotation, scale, fine intensity and shape changes
<b>Image Compression</b>	Coarse perceptual changes	Fine perceptual changes
<b>Object Recognition (e.g., navigation, security )</b>	Salient image object characteristics	Translation, rotation, scale, skew

Next, the modified image content space  $\{I_{i,j,m,n}^{Mod}\}$  represents all possible operations and their parameters applied to a reference image. In the above notation,  $i$  and  $j$  refer to a reference image, and  $m$  and  $n$  refer to the selection of a modifying operation and its parameters applied to the reference image. The entries in Table 1 correspond to operations modifying image content, and can be mapped to classes of image acquisition variables and to physical changes in an imaged field of view. Both are modeled by the operations applied to a reference image. We consider operations that lead to image content changes in terms of position, intensity, shape, and texture of an object of interest.

The comparison space of image contents  $\{c_k(I_{i,j}^{Ref}, I_{i,j,m,n}^{Mod})\}$  is the space of all similarity computations indexed by  $k$  operating on the reference and modified images. It consists of 2D image adapters/loaders, extractors of image descriptors and similarity measures applied to image descriptors. The similarity function  $c_k: R^2 \times R^2 \rightarrow R$  yields a pair-wise similarity/proximity value that is typically from one of

the three sub-intervals of  $\mathbb{R}$ :  $[0,1]$ ,  $[-1,-1]$ , or  $[0,\infty)$ . The space of software adapters, extractors and measure computations has its associated set of data containers for loaded data, image descriptors and similarity/proximity measure values, and is very large. This similarity computation-based organization allows us to include any existing similarity metric described in the previous surveys into a software framework and classify it based on its triplet (image loader/adapter, image content descriptor, similarity measure).

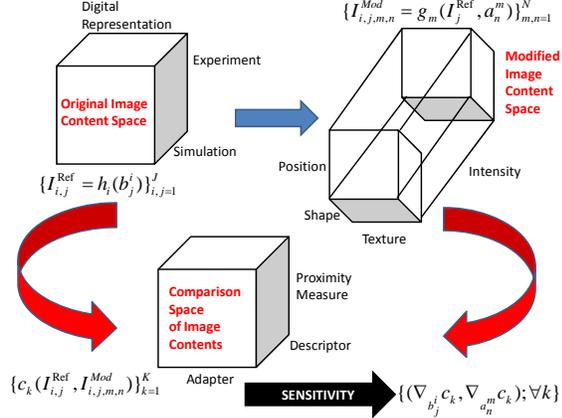


Figure 1: Overview of the modeling framework.

Finally, the sensitivity  $\vec{s}_k$  of a pair-wise image similarity is defined as the rate of change over a set of reference and modified images:  $\vec{s}_k = \left( \frac{\delta c_k}{\delta I_{i,j}^{Ref}}, \frac{\delta c_k}{\delta I_{i,j,m,n}^{Mod}} \right)$ , where the symbol  $\delta$  refers to a functional derivative. In order to compute the sensitivity  $\vec{s}_k$  analytically, one has to establish functional dependencies between the reference image generating variables and the reference image, as well as between the variables causing modifications to a reference image and the modified image, and then solve the functional derivatives. In practice, these functional dependencies cannot become users' application specific inputs since they are unknown. This problem is also computationally intractable because of the size of each considered space. The original image content space contains the number of 2D reference image instances:  $J = 2^{\frac{\#bits}{pixel} \times \#pixels}$ , where the number of bits represents all bits assigned to one or many image bands. The modified image content space has  $M \times N$  number of operations ( $M$ ) and parameters ( $N$ ), and the comparison space of image contents is composed of  $K \leq K1 \times K2 \times K3$  triples of adapters ( $K1$ ), image descriptors ( $K2$ ), and similarity measures ( $K3$ ).

We approach the problem by approximating the two differentials with two matrices shown in Equations (1) and (2). The equations assume that there exist classes of image content generative functions  $h$  and of image content modifying functions  $g$  that are characterized by operators denoted with indices  $i$  and  $m$ ,

and their respective parameters labeled as  $b_j^i$  and  $a_n^m$ . Both equations are approximations because the classes of functions and their parameters are represented only by their samples with a size equal to the number of rows and columns of each matrix. The symbol  $\delta\phi[c_k[(\cdot); \cdot]]$  represents the Frechet derivative of  $c_k(\cdot)$  with respect to  $b_j^i$  or  $a_n^m$  parameters [19]. Thus, the two equations (1) and (2) yield two matrices of sizes  $\max(i) \times \max(j)$  and  $\max(m) \times \max(n)$ , where each matrix entry is a function representing the derivative of a similarity metric functional dependency on (a) the parameter  $b$  of the image generating function  $h$  or (b) the parameter  $a$  of the image modifying function  $g$ . These two matrices represent a sensitivity signature of each similarity metric that can be matched against the users' sensitivity requirements.

$$\frac{\vec{\alpha}_k}{\vec{\alpha}_{i,j}^{Ref}} \sim S_{k1} = \left( \begin{array}{c} \delta\phi c_k(I_{i,l}^{Ref} = h_i(b_j^i), I_{i,l,m,n}^{Mod} = g_m(I_{i,l}^{Ref}, a_n^m)) b_j^i \\ \vdots \\ \delta\phi c_k(I_{i,j}^{Ref} = h_i(b_j^i), I_{i,j,m,n}^{Mod} = g_m(I_{i,j}^{Ref}, a_n^m)) b_j^i \end{array} \right)_{\max(i) \times \max(j)} \quad (1)$$

$$\frac{\vec{\alpha}_k}{\vec{\alpha}_{i,j,mn}^{Mod}} \sim S_{k2} = \left( \begin{array}{c} \delta\phi c_k(I_{i,j}^{Ref} = h_i(b_j^i), I_{i,j,l,n}^{Mod} = g_l(I_{i,j}^{Ref}, a_n^l)) a_n^l \\ \vdots \\ \delta\phi c_k(I_{i,j}^{Ref} = h_i(b_j^i), I_{i,j,mn}^{Mod} = g_m(I_{i,j}^{Ref}, a_n^m)) a_n^m \end{array} \right)_{\max(m) \times \max(n)} \quad (2)$$

From a practical perspective, end users may find specifying functional dependencies too complex. In order to simplify the format of user's inputs (values instead of curves), we bin the range of each parameter into three sub-ranges: Low, Medium and High values of parameters  $a$  or  $b$ . Furthermore, we compute the average value of each matrix entry over a sub-range of parameters and bin the average values into  $B$  categories corresponding to a range of low to high sensitivity. Finally, in order to make the user interface for entering similarity metric requirements easier visually, we rearrange the two matrices into three columns corresponding to the Low, Medium and High values of parameter ranges, and into  $\max(i) \times \max(j)$  or  $\max(m) \times \max(n)$  rows following the zig zag scan of the original matrices, as used in image processing. The presentation format of the sensitivity signature of similarity metrics is illustrated in Eq. (3) for the second matrix and denoted as  $S_{k2}^{BIN}$ . These simplifications make entering categories rather than curves more user-friendly for end users. They also make visual comparison of the signatures straight forward.

$$S_{k_2}^{BIN} = \begin{pmatrix} \bar{s}_{k_2}^L(1,1) & \bar{s}_{k_2}^M(1,1) & \bar{s}_{k_2}^H(1,1) \\ \vdots & \vdots & \vdots \\ \bar{s}_{k_2}^L(1,n) & \bar{s}_{k_2}^M(1,n) & \bar{s}_{k_2}^H(1,n) \\ \bar{s}_{k_2}^L(2,1) & \bar{s}_{k_2}^M(2,1) & \bar{s}_{k_2}^H(2,1) \\ \vdots & \vdots & \vdots \\ \bar{s}_{k_2}^L(2,n) & \bar{s}_{k_2}^M(2,n) & \bar{s}_{k_2}^H(2,n) \\ \vdots & \vdots & \vdots \\ \bar{s}_{k_2}^L(m,n) & \bar{s}_{k_2}^M(m,n) & \bar{s}_{k_2}^H(m,n) \end{pmatrix}_{\max(m) \times \max(n) \times 3} \quad (3)$$

## 4 Experimental Results

### 4.1 Creating Sensitivity Signatures of Similarity Metrics

We illustrate the process of creating sensitivity signatures of similarity metrics by focusing on the matrix  $S_{k_2}$  in Eq. (2). First, the process consists of selecting  $m=4$  operators from the modified image content space (see Figure 1) and  $n=2$  parameters per operator. The choice of operators and parameters is summarized in Table 2. Each parameter was sampled over its specific range of values, and the number of samples per parameter varied between 11 and 45. This selection defines the size of the matrix  $S_{k_2}$  to be  $4 \times 2$ . Following Eq. (3), we also created signatures with binned parameter ranges (Low, Medium, High) that equally divide the min and max value interval of a parameter. Finally, we chose to classify the averages of the 1<sup>st</sup> order derivatives over each sub-range into 9 bins which led to a matrix  $S_{k_2}^{BIN}$  per similarity metric of size  $8 \times 3$  (4 operators  $\times$  2 parameters  $\times$  3 bins – see Table 4).

Next, we sampled the comparison space of image contents by choosing  $k = 9$  similarity metrics. These similarity metrics belong to a class of (a) overlap-based metrics (Dice and Jaccard [7]), (b) clustering-based metrics (Rand Index and Adjusted Rand Index [8], [20]), and (c) histogram-based metrics (Euclidean, City Block, Chebyshev, Intersection and Divergence [16]).

### 4.2 Computing Sensitivity Signatures of Image Similarity Metrics

Table 3 shows an example of the sensitivity analyses of the histogram-based Euclidean metric for eight combinations of operators and their parameters. The columns refer to metric dependency  $c_k$  and its 1<sup>st</sup> difference in the discrete case. The difference is computed by taking the absolute value of the similarity value for the two identical images without transformation subtracted from the similarity value for the two images, one without and one with the transformation applied to it. The graphs in the last column of Table 3 are binned along the  $x$ - and  $y$ -axes,

and then presented as a matrix (a sensitivity signature of a similarity metric) following Eq. (3).

**Table 2: Samples of operators and parameters from the modified image content space. The number in parenthesis refers to the number of image samples used for each parameter**

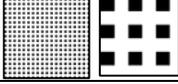
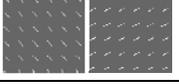
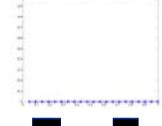
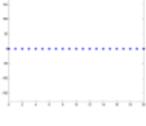
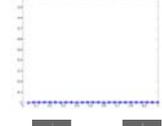
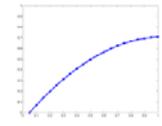
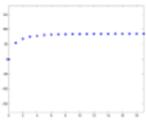
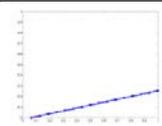
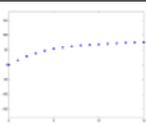
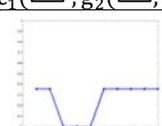
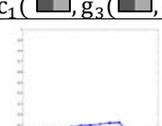
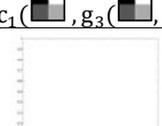
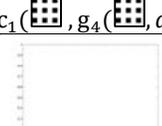
Operator	Parameter n=1	Parameter n=2
Position $m=1$	Translation Change: $a_1^1$ (21) 	Rotation Change: $a_2^1$ (45) 
Shape $m=2$	Size/scale Change: $a_1^2$ (18) 	Ellipticity Change: $a_2^2$ (16) 
Intensity $m=3$	Gamma Correction: $a_1^3$ (11) 	Blur Level Change: $a_2^3$ (15) 
Texture $m=4$	Granularity Change: $a_1^4$ (26) 	Orientation Change: $a_2^4$ (19) 

Table 4 consists of color-coded matrices described by Eq. (3) for each metric listed in the previous section. The color code legend is presented in Table 4, and the white color was introduced to refer to a non-valid computation with respect to the operator  $g_m$ .

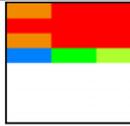
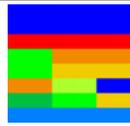
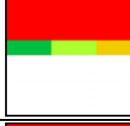
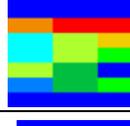
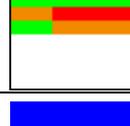
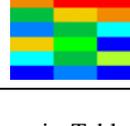
### 4.3 Examples of Application-Driven Similarity Metric Recommendation

Table 4 shows that similarity metrics have various overall sensitivities to image content changes (i.e., the distribution of blue to red colors in each signature matrix). We illustrate how the application specific requirements in Table 1 are converted into a query presented to the recommendation system and how the recommendation of a similarity metric is computed. Table 5 and Table 6 present simple color-encoded queries of the application specific requirements for image spectral calibration and image spatial registration applications (see Table 1).

**Table 3: Sensitivity analysis of the Euclidean similarity metric [16]**

Operator $\mathcal{G}_m$ : Parameter $a_n^m$	Similarity metric dependency $c_k$	1st difference of similarity metric dependency
<b>Position:</b> Translation of circle in x direction	 $c_1(g_1(a_1^1))$	
<b>Position:</b> Rotation angle of lines	 $c_1(g_1(a_1^1))$	
<b>Shape:</b> Size/scale represented by circle radius	 $c_1(g_2(a_1^2))$	
<b>Shape:</b> Ellipticity represented by flatness of ellipsoid	 $c_1(g_2(a_2^2))$	
<b>Intensity:</b> Gamma correction	 $c_1(g_3(a_1^3))$	
<b>Intensity:</b> Blur level represented by kernel size	 $c_1(g_3(a_2^3))$	
<b>Texture:</b> Granularity represented by checker size	 $c_1(g_4(a_1^4))$	
<b>Texture:</b> Orientation angle of line segments	 $c_1(g_4(a_2^4))$	

**Table 4: Sensitivity signatures of nine similarity metrics with the legend of the color codes following the blue to red schema (low to high sensitivity). White refers to the cases where similarity metrics could not be applied**

Metric	Sensitivity signature	Metric	Sensitivity signature
Dice		City_block	
Jaccard		Chebyshev	
Adjusted Rand Index		Intersection	
Rand Index		Divergence	
Euclidean		Color Legend	

The query in Table 5 describes a desirable metric that would be highly sensitive to intensity changes but not sensitive to position, shape and texture changes. Table 6 contains a query to find a metric that is highly sensitive to translation, rotation and scale but not sensitive to intensity, texture, and shape ellipticity changes.

The color-coded query is a  $8 \times 3$  matrix that can be populated by a user. The matrix entries can be specified as angles in degrees of the first difference (i.e., the slope of a line defined by the similarity value for the two identical images without transformation and the similarity value for the two images, one without and one with the transformation applied to it). The matrix entries in degrees are binned into nine uniformly distributed bins of widths 10 degrees. The binning procedure for the color-coded query is illustrated in Eq. (4).

In order to find the best match to the query, we compute first the Euclidean distance between each row of the binned input query matrix and each row of all binned sensitivity matrices in our database of sensitivity signatures. Next, we compute the average Euclidean distance over all rows with a valid entry (i.e., color-coded values different from white). Eq. (5) describes the distance computation. Finally, we recommend the similarity metric in the database with the smallest average distance to the user specified query, i.e.,

$\min_k \{d_k\}$ . The results can be visually verified in Table 5 and Table 6 (right columns).

$$S = \begin{pmatrix} \bar{s}(1,1) & \bar{s}(1,2) & \bar{s}(1,3) \\ \bar{s}(2,1) & \bar{s}(2,2) & \bar{s}(2,3) \\ \bar{s}(3,1) & \bar{s}(3,2) & \bar{s}(3,3) \\ \bar{s}(4,1) & \bar{s}(4,2) & \bar{s}(4,3) \\ \bar{s}(5,1) & \bar{s}(5,2) & \bar{s}(5,3) \\ \bar{s}(6,1) & \bar{s}(6,2) & \bar{s}(6,3) \\ \bar{s}(7,1) & \bar{s}(7,2) & \bar{s}(7,3) \\ \bar{s}(8,1) & \bar{s}(8,2) & \bar{s}(8,3) \end{pmatrix} \quad \begin{matrix} \bar{s}(p,q) \in \{1, 2, \dots, 9\} \sim \\ \sim \{[0,10], [10,20], \dots, [80-90]\} \end{matrix}$$

$$S^{USER} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 9 & 9 & 9 \\ 9 & 9 & 9 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (4)$$

$$d_k(S^{USER}, S_k^{DB-SIGNATURE}) = \frac{1}{8} \sum_{p=1}^8 \sqrt{\sum_{q=1}^3 (S^{USER}(p,q) - \bar{s}_k^{DB-SIGNATURE}(p,q))^2} \quad (5)$$

**Table 5: An example of a user defined query for the spectral calibration application requirements and its recommended similarity metric. The colors follow the color coding shown in Table 4 (red is high sensitivity [80-90]; blue refers to low sensitivity [0-10])**

Application	Change	Operation	User Query	Recommended Similarity Metric: Divergence
Spectral Calibration	Position	Translation	[High Sensitivity]	[High Sensitivity]
		Rotation		[High Sensitivity]
	Shape	Size		[Low Sensitivity]
		Ellipticity		[High Sensitivity]
	Intensity	Gamma		[High Sensitivity]
		Blur		[High Sensitivity]
	Texture	Granularity		[Low Sensitivity]
		Orientation		[High Sensitivity]

## 5 Conclusion

We described the methodology for building signatures of image similarity metrics that can aid end users in choosing similarity metrics according to their application specific requirements. The contribution of this work is in designing these sensitivity signatures. We are developing a large reference database of sensitivity signatures of similarity metrics, and plan to conduct a large scale experiment to validate any recommended similarity metric.

**Table 6: An example of a user defined query for the spatial registration application requirements and its recommended similarity metric. The colors follow the color coding shown in Table 4 (red is high sensitivity [80-90]; blue refers to low sensitivity [0-10])**

Application	Change	Operation	User Query	Recommended Similarity Metric: Divergence
Spatial Registration	Position	Translation	[High Sensitivity]	[High Sensitivity]
		Rotation		[High Sensitivity]
	Shape	Size		[Low Sensitivity]
		Ellipticity		[Low Sensitivity]
	Intensity	Gamma		[Low Sensitivity]
		Blur		[Low Sensitivity]
	Texture	Granularity		[Low Sensitivity]
		Orientation		[Low Sensitivity]

## Disclaimer

No approval or endorsement of any commercial product by NIST is intended or implied. Certain commercial software, products, and systems are identified in this report to facilitate better understanding. Such identification does not imply recommendations or endorsement by NIST nor does it imply that the software and products identified are necessarily the best available for the purpose.

## References

- [1] R. Russell and P. Sinha, "Perceptually-based comparison of image similarity metrics," *Tech Report: AI Memo 2001-014, MIT Lab, Boston, MA*, pp. 1-13, 2001.
- [2] S.-H. Cha, "Comprehensive survey on distance/similarity measures between probability density functions," *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 1, no. 4, pp. 300-308, 2007.
- [3] M. Hatzigiorgaki and A. Skodras, "Compressed domain image retrieval: a comparative study of similarity metrics," *Proceedings of SPIE.*, vol. 5150, pp. 439-449, 2003.
- [4] D. R. Martin, "An Empirical Approach to Grouping and Segmentation," *Tech Report: UCB/CSD-3-1268, Berkeley, CA*, pp. 1-150, 2003.
- [5] T. Neumuth, F. Loebe, and P. Jannin, "Similarity metrics for surgical process models," *Artificial intelligence in medicine*, vol. 54, no. 1, pp. 15-27, Jan. 2012.
- [6] M. S. Prieto and A. R. Allen, "A similarity metric for edge images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1265-1273, 2003.

- [7] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey, "Complex wavelet structural similarity: a new image similarity index," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2385-401, Nov. 2009.
- [8] R. Unnikrishnan and M. Hebert, "Measures of Similarity," *Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05)*, vol. 1, pp. 394-394, Jan. 2005.
- [9] J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Perceptual similarity metrics for retrieval of natural textures," *IEEE International Workshop on Multimedia Signal Processing*, pp. 1-5, Oct. 2009.
- [10] A. Popovic, M. Fuente, M. Engelhardt, and K. Radermacher, "Statistical validation metric for accuracy assessment in medical image segmentation," *International Journal of Computer Assisted Radiology and Surgery*, vol. 2, no. 3-4, pp. 169-181, Jul. 2007.
- [11] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Tech. Report: UCB/CSD-01-1133, Berkeley, CA*, 2001, pp. 416-423.
- [12] H. Zhang, J. Fritts, and S. Goldman, "Image segmentation evaluation: a survey of unsupervised methods," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 260-280, May 2008.
- [13] S. Takemoto and H. Yokota, "Algorithm selection based on a region similarity metric for intracellular image segmentation," in *Image Segmentation*, vol. 22, P.-G. Ho, Ed. InTech, 2011, pp. 419-435.
- [14] M. V. Droogenbroeck, "Design of statistical measures for the assessment of image segmentation schemes," *Computer Analysis of Images and Patterns, Lecture Notes in Computer Science*, vol. 3691, pp. 280-287, 2005.
- [15] M. Polak, H. Zhang, and M. Pi, "An evaluation metric for image segmentation of multiple objects," *Image and Vision Computing*, vol. 27, no. 8, pp. 1223-1227, Jul. 2009.
- [16] S.-H. Cha, "Taxonomy of nominal type distogram distance measures," in *American Conference on Applied Mathematics (MATH '08)*, 2008, no. 2, pp. 325-330.
- [17] J. Zujovic, T. N. Pappas, and D. L. Neuhoff, "Structural similarity metrics for texture analysis and retrieval," in *16th IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 2225-2228.
- [18] H. Benhabiles, J.-P. Vandeborre, G. Lavoué, and M. Daoudi, "A comparative study of existing metrics for 3D-mesh segmentation evaluation," *The Visual Computer*, vol. 26, no. 12, pp. 1451-1466, Apr. 2010.
- [19] A. Frigyik, S. Srivastava, and M. R. Gupta, "An introduction to functional derivatives," *Tech Report: UWEETR-2008-0001*, no. 206, pp. 1-8, 2008.
- [20] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "Toward objective evaluation of image segmentation algorithms," *IEEE transactions on pattern analysis and machine intelligence*, vol. 29, no. 6, pp. 929-44, Jun. 2007.