Quantification of the efficiency of segmentation methods on medical images by means of non-euclidean distances

J Pastore1,2, E Moler1 and V Ballarin1
1 Measurement and Signal Processing Laboratory, School of Engineering, UNMdP, J.B.Justo 4302, Mar del Plata, B7608FQD, ARGENTINA.
2 Comisión Nacional de Investigaciones Científicas y Técnicas CONICET, Rivadavia 1917, Buenos Aires, C1033AAJ, ARGENTINA.
E-mail: jpastore@fi.mdp.edu.ar

Abstract. To quantify the efficiency of a segmentation method, it is necessary to do some validation experiments, consisting generally in comparing the result obtained against the expected result. The most direct method for validation is the comparison of a simple visual inspection between the automatic segmentation and a segmentation obtained manually by a specialist, but this method does not guarantee robustness. This work presents a new similarity parameter between a segmented object and a control object, that combines a measurement of spatial similarity through the Hausdorff metrics and the difference in the contour areas based on the symmetric difference between sets.

1. Introduction
When analyzing medical images, it is very frequent to use segmentation algorithms to extract the contour of different organ or tissue images, in order to use the information provided to perform a clinical diagnostic. The objective assessment of these segmentation algorithms is a fundamental step to determine data validity and its potential clinical application. Nevertheless, few researchers devoted to medical image segmentation have embarked on the evaluation of the algorithms they have suggested.

To assess contour detection, many researchers have employed parameters deriving from the contours themselves, either as perimeters or areas [1] [2]. Fewer authors have used metrics based on distances between contours [3]. DeGraaf et al. utilized a metrical system based on the number of operations carried out to obtain the resulting segmentation [4].

When dealing with area segmentation, one of the most common ways of expressing classification accuracy is by introducing classification error as an error matrix or a confusion matrix [5]. Confusion matrices compare the relationship between reference data and the results arising from the classification.

This paper proposes a new way of measuring classification accuracy by extending Hausdorff metrics in terms of neighborhoods, and employing the concept of symmetric difference between sets.

Hausdorff distance defines a quantity that can be used as a shape comparison factor. Hausdorff distance measures the degree of similarity between two sets of points placed in a fixed position relative to each other. Besides, the symmetric difference between the sets [6] counts the number of elements in which the two sets differ.

A new parameter $\mu_a$ is presented, which combines a spatial similarity measure through Hausdorff metrics and the difference between the areas of the compared contours employing symmetric
difference between sets. This new parameter integrates a contour accuracy measure with an accuracy measure of the areas delimited by said contours.

Regarding the reference segmentation, it is defined by a manual segmentation process supervised by a field specialist. The result of this manual application process yields high degrees of variability when the image is segmented by several expert observers. This problem can be partially solved by averaging the contour sets traced by the experts, resulting in a reference contour or pattern.

2. Materials and Methods

The accuracy of every segmentation technique is determined by the implementation of a systematic process. This process aims at solving the problems arising from quantifying the difference between the final shapes obtained and the reference shape or contour. The final shape obtained describes the object to be represented or extracted employing each segmentation method.

Choosing a comparison parameter is the first step when defining a shape quantifier or comparative estimator. Hausdorff distance constitutes a parameter applicable as a shape comparison factor.

Figure 1 shows an image in which the reference contour and the segmented image contour overlap. This allows a visually assess the degree of similarity or difference between them.

2.1. Hausdorff Distance

A distance, or metric, defined in a non-empty X set, is a function \( d: X \times X \rightarrow \mathbb{R}_{\geq 0} \), where for every \( x, y, z \in X \) the following is verified:

\[
\begin{align*}
\text{a)} & \quad d(x, y) \geq 0 \\
\text{b)} & \quad d(x, y) = 0 \quad \text{si} \quad x = y \\
\text{c)} & \quad d(x, y) = d(y, x) \\
\text{d)} & \quad d(x, y) \leq d(x, z) + d(z, y) 
\end{align*}
\]

A metric space is a \((X,d)\) pair, where \( X \) is a set and \( d \) is a metrics for \( X \).

The distance from a point \( x \in X \) to a set \( B \subset X \) is defined by \( d(a,B) = \inf_{b \in B} d(a,b) \). Given two sets \( A,B \subset X \), the Hausdorff distance is defined by [7]:

\[
H(A,B) = \sup_{a \in A} h(A,B) = \sup_{a \in A} \inf_{b \in B} d(a,b)
\]

where \( h(A,B) \) and \( h(B,A) \) represent the direct distance between \( A \) and \( B \):

\[
h(A,B) = \sup_{a \in A} \inf_{b \in B} d(a,b)
\]

for any distance \( d \) defined in \( X \) (for instance the Euclidean norm or \( L_2 \)).

Function \( h(A,B) \) identifies the point \( a \in A \) furthest away from any point in \( B \). Intuitively, if \( h(A,B) = \lambda \), each point of \( A \) must be at a distance to \( B \) lesser or equal than \( \lambda \), and, in addition, at least a point in \( A \) must be exactly at distance \( \lambda \) of the set \( B \).
The Hausdorff distance measures the degree of similarity between two sets of points in a fixed position by measuring the distance between the point in $A$ furthest away from any point in $B$ and vice-versa. Thus, the notion of similarity given such distance is that every point in $A$ is near to any point in $B$ and vice versa. This metric is based on the whole sets to compute the distance.

2.2. Hausdorff Distance in a discrete plan

The contour of an object (in continuous space) can be characterized as a function in time, regarding it as a pair of signals $x(t)$ and $y(t)$. Considering a figure in a complex plane, where the real part is represented by $x(t)$ signal and the imaginary one by $y(t)$ signal, the contour can be traced by a univariate function in time. Therefore, the following function can be defined for any contour:

$$u(t) = x(t) + jy(t)$$

(4)

If $t$ is discrete then $t = t_0 + i\gamma$ and we can write $u_i = u(t) = u(t_0 + i\gamma)$, where $t = t_0, \ldots, t_0 + i\gamma, \ldots, t_0 + n\gamma$.

Given that the digital image domain is $\mathbb{Z}^2$, the contour is represented by a set $A = \{a_0, \ldots, a_{N-1}\}$ where each point $a_i \in A$ has discrete coordinates $(x_i, y_i)$.

Given two finite sets of points $A = \{a_0, \ldots, a_n\}$ and $B = \{b_0, \ldots, b_k\}$ (representing two contours) according to the definition of the Hausdorff distance in equation (2) and (3), the following equation is obtained:

$$H(A, B) = \max\left(h(A, B), h(B, A)\right) = \max\left(\max_{a \in A} \min_{b \in B} \|a - b\|, \max_{b \in B} \min_{a \in A} \|a - b\|\right)$$

(5)

if $d_p(x) = \min_{b \in B} \|x - b\|$ and $d_q(x) = \min_{a \in A} \|a - x\|$ are defined, then:

$$H(A, B) = \max_{a \in A} \max_{b \in B} d_p(a), d_q(b)$$

(6)

i.e., $H(A, B)$ can be derived calculating $d_p(a)$ and $d_q(b)$ for every point $a \in A$ and point $b \in B$, respectively. The graph of $d_p(x)$, $G(d_p(x)) = \left\{(x, d_p(x)) \mid x \in \mathbb{Z}^2\right\}$ is a surface called Voronoi surface of $B$. Such surface yields, for each point $x$, the distance from $x$ to the nearest point in $B$. The shapes obtained will be similar, provided the $H(A, B)$ distance is small.
2.3. Definition equivalent to Hausdorff distance in terms of neighborhoods

Along the lines of the Jordan curve theorem [8] which states that every simple closed curve divides the discrete plane \( \mathbb{Z}^2 \) into two connected regions, it can be ascertained that a simple closed curve constitutes the boundary of a connected region.

Before defining Hausdorff metrics in terms of neighborhoods, the following results [8] should be introduced:
- Every closed ball constitutes a connected set.
- If \( X \) and \( B \) are connected sets, the set \( X \oplus B \) is also connected, where:
  \[ X \oplus B = \{ y \in \mathbb{Z}^2 : B \cap X \neq \emptyset \} \]
  being \( B = \{ b + y, b \in B \} \) and “+” representing the vector sum.

Let’s consider the space \( \mathcal{S} \) where each point of \( \mathcal{S} \) is a non-empty compact set of \( \mathbb{Z}^2 \). If \( K_1 \) and \( K_2 \) denote two non-empty compact sets of \( \mathbb{Z}^2 \) (or, equivalently, two points of \( \mathcal{S} \)), and \( B(\varepsilon) \) the \( \varepsilon \)-radius closed ball, then:

\[
H_{\oplus}(K_1, K_2) = \inf \{ \varepsilon : K_1 \subset K_2 \oplus B(\varepsilon), K_2 \subset K_1 \oplus B(\varepsilon) \}
\]

defines a metric in \( \mathcal{S} \) known as discrete Hausdorff metric. Geometrically, \( H_{\oplus} \) is the radius of the minor closed ball \( B \), so that \( K_1 \) is contained in the dilated set \( K_2 \oplus B(\varepsilon) \) and \( K_2 \) in \( K_1 \oplus B(\varepsilon) \).

Remarks:
- Distance is remarkably affected by the isolated points.
- If \( K_1 \) and \( K_2 \) are not compact sets, it is possible that \( H_{\oplus} = 0 \) and \( K_1 \neq K_2 \). Consider \( K_1 \) (closed) and \( K_2 = K_1 \) (interior of \( K_1 \)).
- In the particular case in which \( K_1 \) and \( K_2 \) are reduced to two points, the Hausdorff distance \( H_{\oplus}(K_1, K_2) \) coincides with the Euclidean distance.

The first remark is particularly relevant when applying the Hausdorff metrics between sets in the discrete plane and in general. If we consider the particular case in which \( K_2 = K_1 \cup \{ x \} \), where \( x \in \mathbb{Z}^2 \) is a point distant from \( K_1 \), the distance \( H_{\oplus}(K_1, K_2) \) is equal to \( d(x, K_1) \), which prevents any possibility of robustness to noise, a situation that often occurs in image analysis, in the comparison of sets of points.

We propose to combine the Hausdorff metrics with the metrics defined from the symmetrical difference to make the proposed parameter robust relative the noise.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{Distance function graphs}
\end{figure}
2.4. Symmetric Difference

Let $A$ and $B$ be two sets; the symmetric difference $\Delta$ between $A$ and $B$ is defined as:

$$A \Delta B = (A \setminus B) \cup (B \setminus A)$$ (8)

Let $\xi$ be the set made up of the finite subsets of a set $X$; the function $d_\Delta$ in $\xi \times \xi$ is defined by:

$$d_\Delta(A, B) = \left| \left( A \setminus B \right) \cup \left( B \setminus A \right) \right| = |A \Delta B|$$ (9)

Function $d_\Delta$ quantifies the difference between two contours on the basis of the area associated with them. The reason to apply it is to quantify the difference between the areas not contained in both sets. In general, this metrics is neither bounded nor independent from the size of $A$ and $B$. In fact, two sets with large cardinals differing in 3 elements, or two sets with small cardinals also differing in 3 elements yield the same $d_\Delta$ value. The function $d_\Delta$ measures the relative distance to the empty set.

If $d_\Delta$ is normalized by:

$$d_\Delta(A, B) = \begin{cases} |A \Delta B| & \text{if } A \cup B \neq \emptyset \\ |A \cup B| & \text{otherwise} \end{cases}$$ (10)

then $\left( \xi, d_\Delta \right)$ is a metric space and $d_\Delta$ is bounded by $[0, 1]$. 

---

**Figure 3.** Example of the Hausdorff distance
2.5. Linear combination of these two measures

The linear combination of these two quantities results in a new one called \( \mu_\alpha \) which measures spatial similarity and the difference between the areas of the compared contours, i.e.:

\[
\mu_\alpha = \alpha H_{\oplus} + (1 - \alpha) \varepsilon_\Lambda 
\]

(11)

where:

- \( H_{\oplus}(A, B) = \frac{H_A(A, B)}{k} \), being \( k = \max\{diameter(A), diameter(B)\} \);
- \( \varepsilon_\Lambda = 1 - d_\Lambda \) the area intersection error.

The weight factor \( \alpha \) determines the behavior of the parameter \( \mu_\alpha \). In spite of the fact that this paper employed a weight factor \( \alpha = 0.5 \) so as to consider both parameters equally, a value \( \alpha > 0.5 \), would quantify more spatial similarity, whereas \( \alpha < 0.5 \), would quantify the difference between the areas not contained in both contours. It is important to point out that the quantity \( \mu_\alpha \) must be smaller than a threshold \( T \) for the segmentation method to be deemed acceptable.

3. Results

In a first stage, the parameter was tested in synthetic curves. In contrast, in a second stage we show examples of applications of the parameter \( \mu_\alpha \) on real images. Two sets of medical images were selected; in the first part, TAC images of mediastinum were obtained by a CT scan Philips Secura. In the last several years, the technological advances have allowed the detection of smaller incidental renal and hepatic tumors. On the other hand, the advancement of intrusive techniques, like criosurgery and radiofrequency ablation, prevents in some cases the need for mayor surgery. However, both criosurgery and radiofrequency ablation, through excessively high or low temperatures, kill both the tumorlike and the healthy cells. Hence, the importance of detecting tumors with extraordinary spatial precision.

In the second part, we worked with images of bone marrow biopsies where the digital images were obtained by an optical Medicux – 12 microscope with a CCD Hitachi KP-C550 camera, through a PC with a size of 640 x 480 pixels. The anatomical-pathological reports of histological cuts give percentage results of medular celularity, indicating the presence of trabeculas, fat cells and hematopeyicas cells. Those percentages allow us to evaluate the presence and/or degree of a metabolic disorder, establishing comparisons between the normal values and the pathological ones. Generally, these measurements are done by a simple visual inspection.

Both the TAC images and the ones from the biopsies are extremely complex for segmentation because they are images with a high level of texture and high variability of characteristics between patients.
Figure 5 shows results of one of the segmentations performed on TAC images of mediastinum where we observe a liver tumor. First, an automatic segmentation done by an algorithm is shown. Then in image b) we see the reference contour and the segmented image contour overlapping, which allows us to establish graphically the degree of similarity or difference between them. Image c) is simply an amplification of the previous one to better appreciate the differences between both curves, the reference one and the one segmented automatically. In image d) both curves were isolated from the background as a previous step to calculate the accuracy of segmentation. For this segmentation, the similarity parameter proposed $\mu_{a=0.5}$ gave a value of 0.0329.

Figure 6 shows the results of two segmentations performed on the microscope images of bone marrow biopsies. First, an automatic segmentation done by an algorithm is shown in images a). Then in images c) we see the reference contour and the segmented image contour overlapping. In images d) both curves were isolated from the background as a previous step to calculate the accuracy of segmentation. For these segmentations the similarity parameter proposed $\mu_{a=0.5}$ gave a value of 0.109 and 0.0293, respectively. As we can see from the resulting values for these two last images, when the area does not vary significantly, small changes in the contours are not oversized in the parameter $\mu_a$. We also observe a decrease on the influence of the isolated points.

4. Conclusion
This work presents a new similarity quantity, which combines a measurement of the spatial similarity through the Hausdorff metrics and the difference of the areas of the compared contours using symmetric difference between sets.

The reference contour is defined by a process of manual segmentation directed by an specialist in the corresponding area. The variability of this process is avoided by averaging the set of contours established by the experts, resulting in a reference contour or pattern.

Figure 5. CAT image of a mediastinum. a) automatic segmentation, b) overlapping of both contours, i.e., reference and automatic segmentation contours, c) zoom of previous image; and d) both contours isolated from the background.
Figure 6. a) Images of bone marrow biopsies, b) Automatic segmentation, c) overlapping of both contours, reference and automatic segmentation, d) both contours isolated from the background.
The objective evaluation of the segmentation of medical images is a fundamental step to establish the validity of data and their possible clinical application. By integrating in this new parameter a measurement of accuracy of contours with a measurement of accuracy in the areas surrounded by those contours, we attain a substantial improvement in those contours with a projection due to the use of the estimator of area difference.

Acknowledgements
The authors would like to thank Dr. Carlos Capiel (Jr.) from the Instituto Radiológico of the city of Mar del Plata (Radiology Institute) for supplying the CAT images used in this work.

5. Bibliography
[1] Geiser E, Conetta D, Limacher M, Stockton V, Olivier L and Jones B 1990 A second-generation computer-based algorithm for short-axis two-dimensional echocardiographic images. Accuracy and improvement in interobserver variability. Jour. AMER. Soc. Echocardiol. 3 79-90.
[2] Vandenberg B, Ruth L, Stuholmuller P, Melton H and Skorton D 1992 Estimation of left ventricular cavity area with an on-line, semi-automated echocardiographic edge detection system Cir. 86 159-166.
[3] Friedland N and Adam D 1989 Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing IEEE Trans. Med. Imag. 10 187-199.
[4] DeGraaf C, Koster A, Vincken K and Viergerver M 1992 A methodology for validation of image segmentation algorithms Proc. IEEE Symp. Computer-Based Medical Systems 17-24.
[5] Provost F and Kohavi R 1998 On applied research in machine learning. Machine Learning 30 127-132.
[6] Kolmogorov A and Fomin S 1975 Elementos de la Teoría de Funciones y del Análisis Funcional (Moscow: MIR).
[7] Deniau L and Blanc-Talon J 1996 A new nonlinear operator for multiresolution signal and image analysis Proc. of Signal Processing and its Applications 1 447-450.
[8] Serra J 1982 Image Analysis and Mathematical Morphology vol. I (London: Academic Press)