

# Image Processing Verification Tool–IPVT

Dusan Heric, Bozidar Potocnik

University of Maribor  
Faculty of Laboratory of Electrical Engineering and Computer Science  
System Software Laboratory  
Smetanova 17, 2000 Maribor, Slovenia

## ABSTRACT

Validation is an important part of image processing. Automated extraction of structures from medical images like MRI or CT is becoming a routine process for diagnosis. Therefore, the accurate region detection is necessary. The proposed Image Processing Verification Tool–IPVT is as a tool for statistical validation of the image processing results. The IPVT tool statistically evaluates the segmentation ability of the segmentation algorithm based on a comparison with references provided by the experts. This tool comprises a graphical users interface, a 2D and 3D image processing verification process, and report generation.

## 1. INTRODUCTION

Segmentation extracts the information from image. The correctness of information depends on the developer's additional work. Unfortunately, a validation of results is non-transparent, because the researchers did not use identical evaluation procedure. Some of them made assessment with one expert's annotations and did not take into account intra-expert and inter-expert variability. On the other hand, the other researchers made subjective "ad hoc" assessment without standard measures. Two types of evaluation can be considered for the assessment of segmentation quality, depending on the availability of reference segmentation (*ground truth*) [2]. When the ground truth is known, we deal with non-problematic relative evaluation. On the other hand, if we are dealing with a standalone evaluation we have difficulties to quantify ground truth. That problem is usual in medical images where the absence of exact location of the observed object is due to different noises, artefacts very common. MR-images are example of such images where optimum-quality acquisition is unfeasible [3].

Different metrics have been proposed to compare segmentation results, including volume measures, spatial overlap measures [4], and boundary measures (e.g. Mean Absolute Distance–MAD). Metrics proposed in [5] are useful if ground truth segmentation results are available, while the work [6] evaluates segmentation methods quantitatively, if ground truth is not available. A good validation of image segmentation and expert quality is indicated in [1], where the reference annotations are calculated from a group of expert segmentations.

## 2. IPVT FUNCTIONALITY

The IPVT tool to some extent follows ideas proposed in work [1]. The IPVT tool is generic tool for a statistical comparison of segmentation results with references obtained by one or several experts.

Fig. 1. depicts a scheme of the IPVT processing chain. The computer's segmentation results denoted also as the apprentice, are compared to the experts' annotations. In the IPVT tool these annotations are overlaid onto the original image for visual comparison. The experts' references introduce certain amount of inter-observers variability in the evaluation process [8]. To minimise this variability, the IPVT tool supports two approaches to calculate a "mean observer", which are Polar Histogram and Big-Small regions methods proposed in work [7]. Mean observer is treated as an expert, whose annotations are completely correct. Afterwards, a comparison between the computer segmentation results (apprentice) and this "mean observer" is executed. The result of execution is set of statistical measures, which are finally grouped in a HTML document.

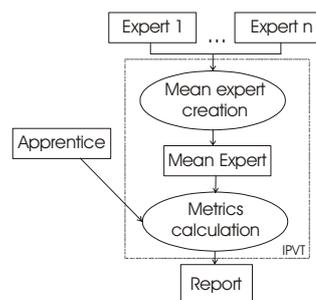


Fig. 1. IPVT verification processing chain

The IPVT tool distinguishes between two types of verification. The first is 2D verification, where the 2D image is compared with reference 2D image. The other type is 3D verification, where the volume object is assessed in entirety.

The IPVT tool supports several actions. There are four most important:

- Import/export annotations—the user imports experts' and apprentice annotations. The input data can be stored either as the Matlab matrix or as the MR document generated by the IST tool [8].

- b) Mean observer creation process calculates from experts' annotations the ground truth. The ground truth can be calculated by the Polar histogram or Big-Small regions methods [7]. The main purpose of this action is to ensure quality and reliable references.
- c) Custom scenario—the analyzer creates a custom scenario and, thus, constructs his own workflow of verification. Thus he defines the verification type, statistical measures, and how many experts will be included in the study. The IPVT tool supports *Default scenario*, which calculates all statistical measures on all images for all annotations; while a *custom scenario* is executed only on selected slices (in case of MR-image) and annotations. This flexible definition of scenarios also enables the quality assessment of manual annotations provided by experts.
- d) Report generator—executes the custom or default scenario. The output with calculated statistical measures is written in a HTML document. By using calculated statistical measures, the analyzer/user can easily and efficiently take a decision about the quality and efficiency of the segmentation algorithm.

### 3. STATISTICAL MEASURES

The IPVT tool evaluates the segmentation accuracy according to five statistical measures; namely ratios  $r_1$  and  $r_2$ , Hausdorff distance, mean absolute distance, and our own measures called spherical distance. By 2D verification the comparison is done on slice to slice basis. On the other hand, 3D verification evaluates the volume object entirely.

#### 3.1 Ratios $r_1$ and $r_2$

Two quantitative indexes, denoted by  $r_1$  and  $r_2$ , for measuring the algorithm's segmentation efficiency were introduced in [9]. Efficiency is calculated upon an intersection between the segmented region and the reference object (i.e. the expert's reading). The ratio  $r_1$  is the ratio between the area of intersection and reference object, while the ratio  $r_2$  is the ratio between the area of intersection and segmented region (1):

$$r_1 = \frac{area_{inter}}{area_{ref}} \quad \text{and} \quad r_2 = \frac{area_{inter}}{area_{region}}. \quad (1)$$

Fig. 2. depicts a scheme for calculation of ratios  $r_1$  and  $r_2$ .

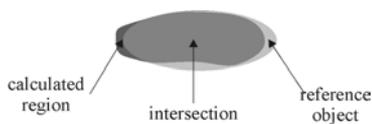


Fig. 2. Scheme for calculation of ratios  $r_1$  and  $r_2$ .

Both indexes give 1 for perfect agreement and 0 for complete disagreement between calculated and reference object.

#### 3.2 Mean absolute distance

The mean absolute distance between curves  $A$  and  $B$  is defined [10]:

$$d(A, B) = \frac{1}{2} \left( \frac{1}{n} \sum_{i=1}^n d(\mathbf{a}_i, B) + \frac{1}{m} \sum_{i=1}^m d(\mathbf{b}_i, A) \right), \quad (2)$$

where  $A = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$  and  $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m\}$ . Elements  $\mathbf{a}_i$  and  $\mathbf{b}_i$  denote the points from the curves. The distance of point  $\mathbf{a}_i$  ( $\mathbf{a}_i \in A$ ) to the closest point on the curve  $B$  is denoted by  $d(\mathbf{a}_i, B)$ , and is defined as

$$d(\mathbf{a}_i, B) = \min_j \|\mathbf{b}_j - \mathbf{a}_i\|. \quad (2)$$

#### 3.3 Hausdorff distance

The Hausdorff distance measures the maximum distance between two curves. The distance from the curve  $A$  to the nearest point on the curve  $B$  is measured and vice versa. The distance between curve  $A$  and curve  $B$  is defined as

$$h(A, B) = \min \{d(A, y); y \in B\}, \quad (3)$$

where  $d(A, y)$  is determined by (2). Finally, the Hausdorff distance  $H(A, B)$  is defined as a maximum distance between the distances  $h(A, B)$  and  $h(B, A)$ :

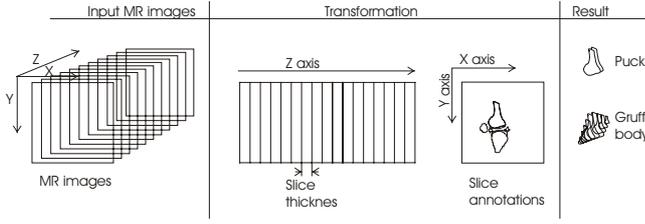
$$H(A, B) = \max \{h(A, B), h(B, A)\}. \quad (4)$$

#### 3.4 Spherical distance

In contrast to the 2D estimation, in 3D verification process the references are bodies (i.e. 3D models) and the final assessment thus covers whole observed object. The volume estimation process is similar to the well-known 2D estimation process, but an input is either 3D model or annotations for every slice in the image sequence. If input is in a form of slice annotations then a transformation process from slice annotations to 3D body is required, otherwise a spherical distance between bodies can be immediately measured.

If input into 3D verification process is in the form of 2D slice annotations, then these annotations must be first transformed into a 3D body. For each annotation just two dimensions are available, while the third—volume information is obtained from a slice thickness (Fig. 3.). The slice thickness is one of the MR scanner device parameters [7]. Afterwards, the annotations are presented into the 3D world as “pucks”. The puck is defined by annotation's boundary (XY plane) and slice thickness (Z

axis). Finally, the pucks are put together on puck by puck basis into a gruff body.

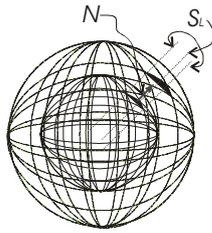


**Fig. 3.** The process chain of puck and gruff body creation.

Spherical distance is a volume measure. It is independent of form and expanse of the surface of the assessed bodies. The surface is connected with patches. Each patch is defined by at least three edge points (nodes, vertexes), where these points lie in a plane. Fig. 4 depicts two spheres—small and large. Small sphere is inside the bigger one. The coloured triangles represent two patches—the red lies on the small sphere, while the blue lies on the bigger sphere. An orange arrow is a normal on the red patch. Green straight lines through patch edge points have a same direction as the normal.

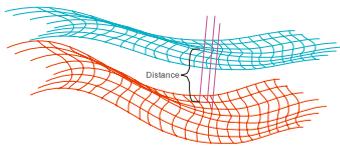
Before distance between two bodies (i.e. volumes) is calculated centres of gravity for both bodies need to be aligned (matched). After this translation, the spherical distance is calculated according to our five-step procedure:

- 1) Calculate normals—for each patch from assessed surface (body) the normal is calculated.
- 2) Create straight lines—straight lines through the patch edge points. They have same direction vector as the normal of patch (see Fig. 4.).



**Fig. 4.** Two spheres, patch, normal on the patch, and straight line

- 3) Calculate Euclidian distance—this distance is measured between the edge point and nearest point where the straight line pierce through a reference surface (see Fig. 5.).



**Fig. 5.** Euclidian distance between two surfaces.

- 4) Calculate statistical factors—for each surface the average Euclidian distance and their standard deviation is calculated
- 5) Finally, the spherical distance is defined as the minimum between the averages of Euclidian distances between all patches on both bodies.

### 3.4.1 Mathematical formalisation

The inputs into spherical distance calculation process are two surfaces  $S_1$  and  $S_2$ . The surface is a data structure, which consists of vertices and patches. The vertex is defined by three points

$$V = \{x, y, z; x, y, z \in \mathfrak{R}\}. \quad (5)$$

Patch  $P_a$  is a data structure, and is same as the plane  $P$  defined by three vertices that are non-collinear, but bounded by straight lines  $l$  between vertices:

$$\begin{aligned} P &= \{V_1, V_2, V_3; V_i \in V\} \\ P_a &= \{x; x \in P(V_1, V_2, V_3) \wedge \\ & x \leq l(V_1, V_2) \wedge x \leq l(V_2, V_3) \wedge x \leq l(V_1, V_3)\} \end{aligned} \quad (6)$$

The surface is defined as a union of three vertices and one patch:

$$S = \{\cup\{\{V_1, V_2, V_3\}, \{P_a(V_1, V_2, V_3)\}\}; V_i \in V\}. \quad (7)$$

For spherical distance calculation, first, patch normals are calculated. The normal  $N$  is calculated as

$$N = \{a \times b; a = V_1 - V_2, b = V_1 - V_3; V_i \in P_a\}. \quad (8)$$

Afterwards, straight lines through patch edges are constructed. Straight line  $S_L$  has the same direction factor as the patch normal and is defined as

$$S_L = V_1 + t * \vec{d}_N; V_1 \in P_a \wedge \vec{d}_N, \quad (9)$$

where  $\vec{d}_N$  is a direction of normal. Then a distance between surfaces  $A$  and  $B$  is calculated as follows

$$E_d = \sqrt{V_1^2 + T_2^2}; T_2 \in S_L \cap B \wedge V_1 \in A,$$

where  $T_2$  is the nearest point where the straight line  $S_L$  pierce through a reference surface  $B$ . Finally, the spherical distance  $D_s$  is calculated as a minimum between the averages of Euclidian distances  $E_d(A, B)$  and  $E_d(B, A)$ :

$$D_s = \min\left\{\frac{1}{n} \sum_{\forall A \in S_1} E_d(A, B), \frac{1}{m} \sum_{\forall B \in S_2} E_d(B, A)\right\}; \quad (10)$$

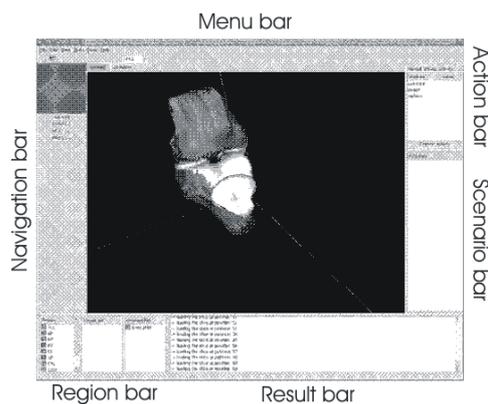
$$A \wedge B \in S$$

where  $n$  and  $m$  are number of patches on surface  $S_1$  or  $S_2$ , respectively.

Our proposed spherical distance is interpreted as the distances between assessed and expert's surface. This distance points out how much two patches with regard to their normal direction differ from the other surface and how tightly all distances are clustered around this distance.

#### 4. GRAPHICAL USER INTERFACE

The primary goal of the IPVT application is validation of image segmentation results, but it is also applicable as a training system. This application is interactive and has an advanced, efficient, and user-friendly graphical user interface (GUI). It supports the several useful actions, which are: a) simple MR image navigation, b) zoom functions, c) 3D model presentation and visualization, d) statistical image processing verification (contour based in 2D plane and 3D volume measurements), e) report generation in the form of HTML documents, f) conversion routines, g) scenario generation and h) drag and drop facility.



**Fig. 6.** The IPVT main widget.

Fig. 6. depicts the IPVT main widget, which is composed of the following major graphical components: a) *Menu bar* offers typical application commands, such as open, import or export the expert's or estimated annotations, scale up/down image, etc. b) *Navigation bar* enables navigation with the "next" and "previous" buttons, previewing the MR-slices, and the slice loading. c) *Region bar* is separated in three portions: the regions' names, and lists of experts and estimates. "Regions' names" form a list of, already known, grouped regions of annotations (e.g. bone femur). The "expert list" contains the labels of imported experts, while the "estimate list" contains the labels of estimates. d) *Action bar* is used for scenario prescription. e) *Scenario bar* is just a list of custom defined actions, which prescribe the scenario. f) *Result bar* is an output window where the results of the scenario execution are displayed. g) *Tab-image-view* provides a stack of the loaded images slice widgets. The analyzer selects an MR image by clicking on the widget tab and,

thus, visually and/or by using the integrated statistical tool calculates the differences between segmentation results and expert's annotations.

The IPVT tool usage is conformable because it is consistent with standard common applications. Therefore, it does not require a lot of time to become familiar with it. The final report is HTML based, thus it is possible to present the estimation data immediately via Internet.

The IPVT tool was the main validation tool in the scope of the SimBio project [7]. Automated detection of important structures in human knee joint in MR images were assessed. The IPVT toll proved as powerful and reliable utility tool.

#### 5. CONCLUSIONS

The IPVT application with statistical measures—ratios  $r_1$ ,  $r_2$ , mean absolute distance, Hausdorff distance and spherical distance are presented in this paper. Proposed metrics have been used to compare ground truth with the apprentice annotations. But, the final decision about segmentation algorithm quality is made by the analyzer, who takes into consideration IPVT final report, domain knowledge, and visual inspection.

#### REFERENCES

- [1] S. K. Warfield, K. H. Zou, M. R. Kaus and W. M. Wells, "Simultaneous Validation of Image Segmentation and Assessment of Expert Quality", *International Symposium on Biomedical Imaging IEEE* pp.1494-1498, 2002.
- [2] P.L.Correia and F. Pereira, "Objective Evaluation of Video Segmentation Quality," *IEEE Transaction on Image Processing* vol. 12 no. 2 pp. 850-863, 1993.
- [3] J. C. Russ, "The Image Processing Handbook", second ed., CRC Press, 1995.
- [4] A.P. Zijdenbos, B.M. Dawant, R.A. Margolin and A.C. Palmer, "Morphometric analysis of white matter lesions in MR images: method and validation", *IEEE Transactions on Medical Imaging* vol. 13 no. 4 pp. 716-724, 1994.
- [5] C.E. Erdem and B.Sankur, "Performance evaluation metrics for object-based video segmentation", *Im Proc. X European Signal Processing Conference* vol. 2 pp. 917-920, 2000.
- [6] C.E. Erdem, A.M. Tekalp and B. Sankur, "Metrics for performance evaluation of video object segmentation and tracking without ground-truth", *Image Processing Proceedings* pp. 69-72 2001.
- [7] SimBio Consortium, SimBio - A Generic Environment for Bio-numerical Simulation, <http://www.simbio.de/~simbio>
- [8] D. Heric and B. Potonik, "An Image Segmentation Tool (IST)", 4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications vol. 1 pp. 483-489 2003.
- [9] B. Potočnik and D. Zazula, "Assessing the efficiency of segmentation algorithms for ultrasound images", *Electrotechnical review* pp. 97-104 vol. 69 ed. 2-3 2001.
- [10] I. Mikic, S. Krucinski and J. D. Thomas, "Segmentation and tracking in echocardiographic sequences—active contours guided by optical flow estimates", *IEEE Trans. med. imag.* vol. 17 no. 2 pp. 274-284 1998.