Efficient Automatic 3D-Reconstruction of Branching Neurons from EM Data
Supplemental Material

Jan Funke[1] Bjoern Andres[2] Fred A. Hamprecht[2] Albert Cardona[1] Matthew Cook[1]

[1]Institute of Neuroinformatics
University Zurich/ETH Zurich
http://www.ini.uzh.ch

[2]HCI/IWR
University of Heidelberg
http://hci.iwr.uni-heidelberg.de

1. Assignment Features

A complete list of all features and symbolic names is given in Table 1. The features can be divided into two categories: geometry features are extracted based on the shape and location of the involved segmentation hypotheses, whereas texture features pay regard to the content and context of the segmentation hypotheses.

1.1. Geometry Features

Continuation For each pair of segmentation hypotheses that is involved in a continuation assignment, the distance between the centroids projected to the x-y-plane is taken (CD). To account for the similarity in size and shape, we also measure the symmetric set difference (SE) of the pixels after moving the centroids of the segmentation hypotheses to overlap. Here, the z-coordinate of the pixels is ignored. To not punish large segmentation hypotheses more than small ones, we normalize the set difference by the sum of the pixel sizes of the two segmentation hypotheses and take this value as an additional feature (SR).

Branching The geometry features for the branching assignments are extracted in a very similar way to the continuation cases: the two segmentation hypotheses that are in one slice are treated as one segmentation hypotheses to extract the aforementioned features (CD,SE,SR).

End The only geometry feature for an end assignment is the size of the respective segmentation hypotheses.

1.2. Texture Features

Continuation For each possible assignment between two segmentation hypotheses, we carry out two normalized cross-correlations with different pre-scaling of the slices. A coarse cross-correlation is performed after Gaussian smoothing with $\sigma = 10$ pixels in a window of three times the size of the involved segmentation hypotheses. This correlation is meant to account for the context of the segmentation hypotheses. A finer cross-correlation in a window of 1.5 times the size of the involved segmentation hypotheses measures the similarity of the interior of the segmentation hypotheses. In either cases, we perform a cross-correlation in both directions (local window from one segmentation hypotheses against the neighborhood of the other one, and vice versa). The two maxima positions relative to the respective centroid are added and constitute the offset feature (CX, CY and FX, FY). By adding the relative maxima positions, we deliberately ignore any displacement between the segmentation hypotheses, since this is taken care of already with the center distance feature (CD). Consequently, two segmentation hypotheses that agree about their displacement to each other will have an offset of zero. The maximum of both normalized cross-correlation values is taken as another feature (CV and FV).

To further evaluate the interior similarity of segmentation hypotheses we also compute histograms with ten bins of the pixel intensities for each segmentation hypotheses. The bin-wise differences between two segmentation hypotheses in a continuation assignment are added to the feature vector (H0 to H9). The same is done for normalized histograms that sum up to one (N0 to N9).

Branching For branching assignments we also compute correlation and histogram features as for the continuation assignments. For the correlation features, we regard the branching as two continuations with the same source. The correlation features CX, CY, CV and FX, FY, FV are now taken as the average values of the respective continuation features. For the histogram features we simply add the histograms of the two segmentation hypotheses that are in one slice and proceed as described for the continuation features.

End The only texture features for end assignments are the plain histogram values of the involved segmentation hypothesis.

1.3. Variable Importance

The importance of the features according to the variable importance measure of the Random Forest classifier [1] can be found in Fig. 1.

Category		Feature	Description	Assignment Case
Geometry		CD	center distance	C,B
		SE	set difference	C,B
		SR	set difference ratio	C,B
		SI	size	E
Texture	coarse correlation	CX	offset x	C,B
		CY	offset y	C,B
		CV	max value	C,B
	fine correlation	FX	offset x	C,B
		FY	offset y	C,B
		FV	max value	C,B
	histogram	H0	bin 0	C,B,E
		H1	bin 1	C,B,E
		H2	bin 2	C,B,E
		H3	bin 3	C,B,E
		H4	bin 4	C,B,E
		H5	bin 5	C,B,E
		H6	bin 6	C,B,E
		H7	bin 7	C,B,E
		H8	bin 8	C,B,E
		H9	bin 9	C,B,E
	histogram normalized	N0	bin 0	C,B,E
		N1	bin 1	C,B,E
		N2	bin 2	C,B,E
		N3	bin 3	C,B,E
		N4	bin 4	C,B,E
		N5	bin 5	C,B,E
		N6	bin 6	C,B,E
		N7	bin 7	C,B,E
		N8	bin 8	C,B,E
		N9	bin 9	C,B,E

Table 1. List of all features and assignment cases they are used in: C: continuation, B: branching, E: end.

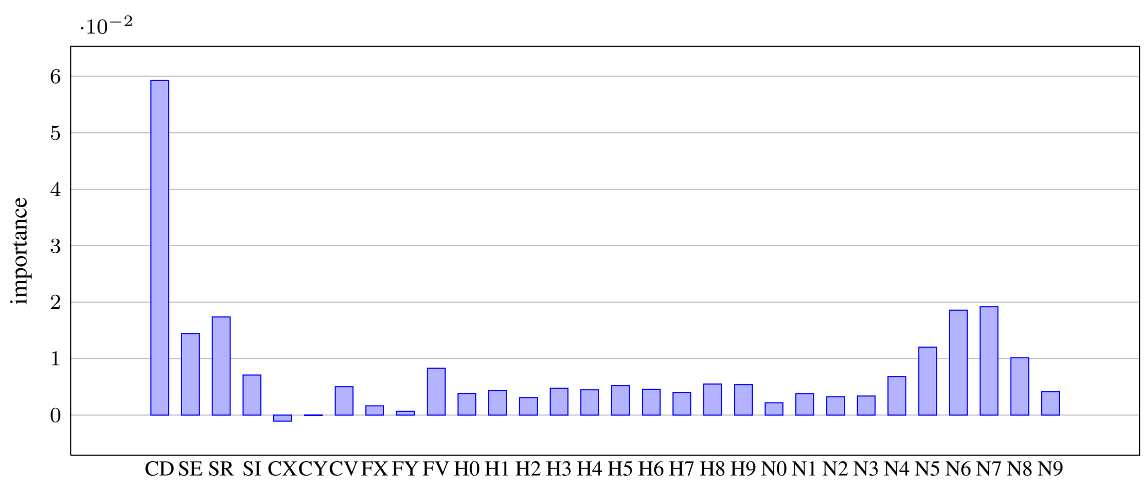

Figure 1. Variable importance of the used features as reported by the Random Forest.

References

[1] Leo Breiman, *Random Forests*, Machine Learning **45** (2001), 5–32. 10.1023/A:1010933404324. ↑1