Reconstructing Evolving Tree Structures in Time Lapse Sequences by Enforcing Time-Consistency - Appendix

In this appendix, we first formally derive the objective function we minimize in Section 4.3 of the main paper. We then provide additional results and justify the parameter choices made to obtain them.

A Deriving the Objective Function

In Section 4, we described briefly how temporal consistency could be enforced by incorporating into the objective function of Eq. 2 a penalty term involving the flow variables used to solve the QMIP problem of Eq. 3. In this Section, we justify the form of this additional term in more detail.

We first define the constraints that must be imposed on the flow variables to ensure temporal consistency and then derive the complete objective function accordingly.

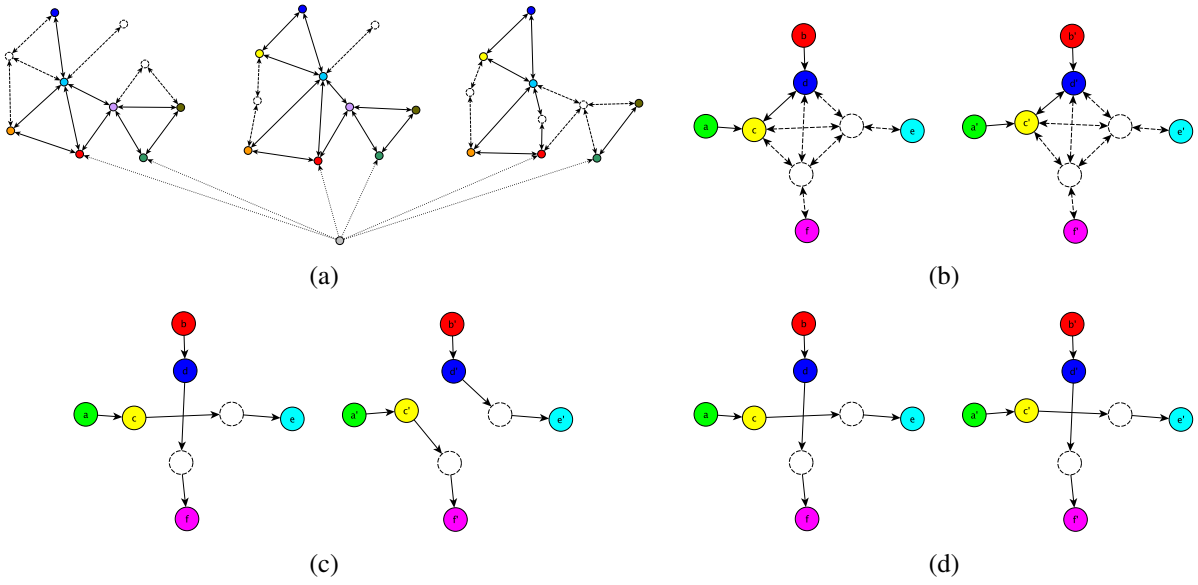
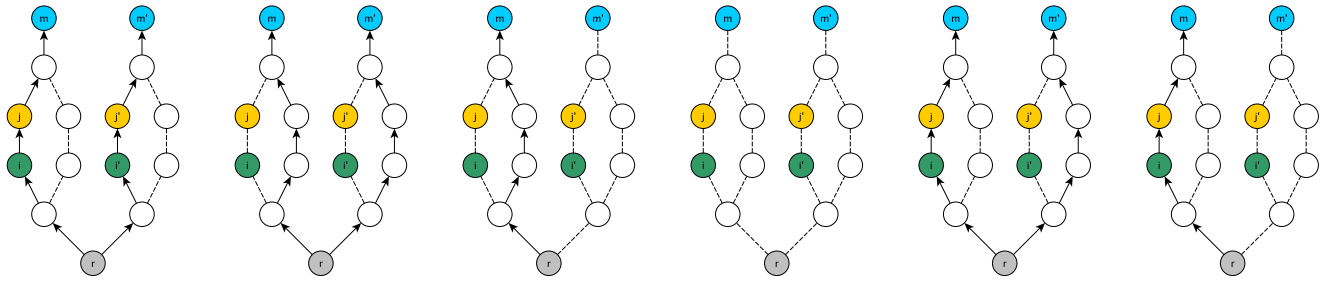


Fig. A.1. Spatio-temporal graphs and consistency constraints. (a) The imaginary root vertex $\mathbf{x}_r$ introduced in Section 4 is shown in gray at the bottom. Each of the three time points contain two manually annotated physical roots marked in red and green. The vertices for which correspondences where not found in adjacent time points are represented by white circles with dashed borders. The temporal edges $\mathcal{E}_t$ are not explicitly shown. Instead, corresponding vertices are shown with matching colors. The dotted arrows represent the imaginary edges from $\mathbf{x}_r$ to the physical roots. The double-sided arrows connecting vertices represent the two oppositely directed *spatial edges* between neighbouring vertices. *Spatial edges* that are part of the corresponding edges set $\bar{\mathcal{E}}_t$ are denoted by solid lines and the others by dashed lines. (b) A small spatio temporal graph for two trees at two different times. The root vertices are labelled a, b and a', b' respectively but the imaginary root vertex is not shown. The solid circles represent vertices for which correspondences have been successfully established and the corresponding vertices are represented with matching colors. The dashed circles represent vertices without no corresponding ones. The edges joining corresponding vertices appear as solid lines, the others as dashed ones. (c,d) Two potential trees, with only the one on the left being temporally consistent. The other one is disallowed by our temporal consistency constraint because $1 = f^e_{ac} \neq f^{e'}_{a'c'} = 0, \quad 0 = f^e_{bd} \neq f^{e'}_{b'd'} = 1, \quad 0 = f^f_{ac} \neq f^{f'}_{a'c'} = 1, \quad 1 = f^f_{bd} \neq f^{f'}_{b'd'} = 0.$



(c1) (c2) (c3) (c4) (i1) (i2)

Fig. A.2. These six example graphs illustrate all possible temporal consistency situations for a pair of corresponding vertices $(\mathbf{x}_m, \mathbf{x}_{m'})$ and a pair of corresponding edges $(\mathbf{e}_{ij}, \mathbf{e}_{i'j'})$ at two consecutive times. (c1-c4) Consistent flows. (i1-i2) Inconsistent flows.

a.1 Using Flow Variables to Model Temporal Consistency

Recall from Section 3 that we build spatio-temporal graphs such as the one depicted by Fig. A.1(a). Let us assume that for any spatially-connected pair of vertices $\mathbf{x}_i^n$, $\mathbf{x}_j^n$ in image I^n and another pair of spatially connected vertices $\mathbf{x}_{i'}^{n+1}$, $\mathbf{x}_{j'}^{n+1}$ in image I^{n+1} the spatial edge $\mathbf{e}_{ij}^n$ corresponds to the spatial edge $\mathbf{e}_{i'j'}^{n+1}$ provided that $(\mathbf{x}_i^n \leftrightarrow \mathbf{x}_{i'}^{n+1}), (\mathbf{x}_j^n \leftrightarrow \mathbf{x}_{j'}^{n+1}) \in \mathcal{E}_t$. In other words, if for two spatial edges in two consecutive images both endpoints of the one edge correspond to the endpoints of the other edge according to $\mathcal{E}_t$ then we assume that those edges correspond to each other.

Let $\bar{\mathcal{E}}_t = \{(\mathbf{e}_{ij}^n, \mathbf{e}_{i'j'}^{n+1}) | \mathbf{e}_{ij}^n, \mathbf{e}_{i'j'}^{n+1} \in \mathcal{E}_s \wedge \mathbf{e}_{ii'}^{n,n+1}, \mathbf{e}_{jj'}^{n,n+1} \in \mathcal{E}_t\}$ be the set of corresponding edges; $\mathbf{x}_m^n$ and $\mathbf{x}_{m'}^{n+1}$ two corresponding vertices in two consecutive images I^n and I^{n+1}, that is, $(\mathbf{x}_m^n \leftrightarrow \mathbf{x}_{m'}^{n+1}) \in \mathcal{E}_t$; $\mathbf{e}_{ij}^n$ and $\mathbf{e}_{i'j'}^{n+1}$ two corresponding edges in two consecutive images, that is, $(\mathbf{e}_{ij}^n, \mathbf{e}_{i'j'}^{n+1}) \in \bar{\mathcal{E}}_t$. We consider $\mathbf{x}_m^n$ and $\mathbf{x}_{m'}^{n+1}$ to be connected to the imaginary root in a temporally consistent way with respect to $\mathbf{e}_{ij}^n$ and $\mathbf{e}_{i'j'}^{n+1}$ if one of the four following conditions holds:

c1) Both $\mathbf{x}_m^n$ and $\mathbf{x}_{m'}^{n+1}$ are part of the solution and the paths connecting them to $\mathbf{x}_r$ traverse the edges $\mathbf{e}_{ij}^n$ and $\mathbf{e}_{i'j'}^{n+1}$ respectively, which implies $f_{ij}^m = f_{i'j'}^{m'} = 1$. See Fig. A.2(c1).

c2) Both $\mathbf{x}_m^n$ and $\mathbf{x}_{m'}^{n+1}$ are part of the solution and neither of the paths connecting them to $\mathbf{x}_r$ traverses the edges $\mathbf{e}_{ij}^n$ and $\mathbf{e}_{i'j'}^{n+1}$, which implies $f_{ij}^m = f_{i'j'}^{m'} = 0$. See Fig. A.2(c2).

c3) Exactly one of the vertices $\mathbf{x}_m^n$ and $\mathbf{x}_{m'}^{n+1}$ is part of the solution but the path connecting it to $\mathbf{x}_r$ does not traverse the respective one of the corresponding edges $\mathbf{e}_{ij}^n$ and $\mathbf{e}_{i'j'}^{n+1}$, which implies $f_{ij}^m = f_{i'j'}^{m'} = 0$. See Fig. A.2(c3).

c4) Neither of the vertices $\mathbf{x}_m^n$ and $\mathbf{x}_{m'}^{n+1}$ is part of the solution, which implies $f_{ij}^m = f_{i'j'}^{m'} = 0$. See Fig. A.2(c4).

By contrast, the connection is inconsistent, potentially indicating a topology change, if one of the two following situations arises:

i1) Both $\mathbf{x}_m^n$ and $\mathbf{x}_{m'}^{n+1}$ are part of the solution but only one of the paths connecting them to $\mathbf{x}_r$ traverses the respective one of the corresponding edges $\mathbf{e}_{ij}^n$ and $\mathbf{e}_{i'j'}^{n+1}$, which implies either $f_{ij}^m = 1, f_{i'j'}^{m'} = 0$ or $f_{ij}^m = 0, f_{i'j'}^{m'} = 1$. See Fig. A.2(i1).

i2) Exactly one of the vertices $\mathbf{x}_m^n$ and $\mathbf{x}_{m'}^{n+1}$ is part of the solution and the path connecting it to $\mathbf{x}_r$ traverses the respective one of the corresponding edges $\mathbf{e}_{ij}^n$ and $\mathbf{e}_{i'j'}^{n+1}$, which implies either $f_{ij}^m = 1, f_{i'j'}^{m'} = 0$ or $f_{ij}^m = 0, f_{i'j'}^{m'} = 1$. See Fig. A.2(i2).

Since changes are relativey infrequent but nevertheless possible, we introduce a temporal consistency parameters $q > 0.5$ that one of the four consistency constraints holds. Figs. A.1(b-d) depict a toy example in which these constraints favor one interpretation over another.

a.2 QMIP Formulation

In Section 3 we wrote the posterior distribution of Y given the spatial edges $\mathcal{E}_s$ and the temporal edges $\mathcal{E}_t$ as

$$P(Y = y | \mathcal{I}, \mathcal{X}, \mathcal{E}_s, \mathcal{E}_t) \propto P(\mathcal{I}, \mathcal{X}, \mathcal{E}_s | Y = y) \, P(Y = y | \mathcal{E}_t) \ . \tag{A.1}$$

Assuming the image data and the spatial edges to be conditionally independent from the temporal edges given Y yields

$$P(\mathcal{I}, \mathcal{X}, \mathcal{E}_s | Y = y) = \prod_{\mathbf{e}_{ij}^n, \mathbf{e}_{jk}^n \in \mathcal{E}_s} \left(\frac{p_{ijk}}{1 - p_{ijk}} \right)^{y_{ij} y_{jk}} \ , \tag{A.2}$$

$$P(Y = y | \mathcal{E}_t) = \prod_{(\mathbf{x}_m^n, \mathbf{x}_{m'}^{n+1}) \in \mathcal{E}_t} \prod_{(\mathbf{e}_{ij}^n, \mathbf{e}_{i'j'}^{n+1}) \in \bar{\mathcal{E}}_t} \left(\frac{q}{1 - q} \right)^{2 f_{ij}^m f_{i'j'}^{m'} - f_{ij}^m - f_{i'j'}^{m'}} . \tag{A.3}$$

We derive the image term of Eq. (A.2) as in [1]. and model the prior term as a tree structured Bayesian network that captures temporal relationships between edges in $\bar{\mathcal{E}}_t$. We can therefore write

$$
\begin{aligned}
&P\left(Y=y|\mathcal{E}_t\right)= \\
&=\prod_{(\mathbf{x}_m^n,\mathbf{x}_{m'}^{n+1})\in\mathcal{E}_t}\prod_{(\mathbf{e}_{ij}^n,\mathbf{e}_{i'j'}^{n+1})\in\bar{\mathcal{E}}_t} P\left(F_{i'j'}^{m'}=f_{i'j'}^{m'}|F_{ij}^m=f_{ij}^m\right)\prod_{\mathbf{e}_{ij}^0\in\mathcal{E}_0}P\left(F_{ij}^0=f_{ij}^m\right) && \text{(A.4)}\\
&\propto\prod_{(\mathbf{x}_m^n,\mathbf{x}_{m'}^{n+1})\in\mathcal{E}_t}\prod_{(\mathbf{e}_{ij}^n,\mathbf{e}_{i'j'}^{n+1})\in\bar{\mathcal{E}}_t}\left(\frac{P\left(F_{i'j'}^{m'}=f_{i'j'}^{m'}|F_{ij}^m=1\right)}{P\left(F_{i'j'}^{m'}=f_{i'j'}^{m'}|F_{ij}^m=0\right)}\right)^{f_{ij}^m}P\left(F_{i'j'}^{m'}=f_{i'j'}^{m'}|F_{ij}^m=0\right) && \text{(A.5)}\\
&\propto\prod_{(\mathbf{x}_m^n,\mathbf{x}_{m'}^{n+1})\in\mathcal{E}_t}\prod_{(\mathbf{e}_{ij}^n,\mathbf{e}_{i'j'}^{n+1})\in\bar{\mathcal{E}}_t}\left[\frac{P\left(F_{i'j'}^{m'}=f_{i'j'}^{m'}|F_{ij}^m=1\right)}{P\left(F_{i'j'}^{m'}=f_{i'j'}^{m'}|F_{ij}^m=0\right)}\right]^{f_{ij}^m}\left[\frac{P\left(F_{i'j'}^{m'}=1|F_{ij}^m=0\right)}{P\left(F_{i'j'}^{m'}=0|F_{ij}^m=0\right)}\right]^{f_{i'j'}^{m'}}P\left(F_{i'j'}^{m'}=0|F_{ij}^m=0\right) && \text{(A.6)}\\
&\propto\prod_{(\mathbf{x}_m^n,\mathbf{x}_{m'}^{n+1})\in\mathcal{E}_t}\prod_{(\mathbf{e}_{ij}^n,\mathbf{e}_{i'j'}^{n+1})\in\bar{\mathcal{E}}_t}\left[\frac{P\left(F_{i'j'}^{m'}=1|F_{ij}^m=1\right)}{P\left(F_{i'j'}^{m'}=1|F_{ij}^m=0\right)}\right]^{f_{ij}^m f_{i'j'}^{m'}}\left[\frac{P\left(F_{i'j'}^{m'}=0|F_{ij}^m=1\right)}{P\left(F_{i'j'}^{m'}=0|F_{ij}^m=0\right)}\right]^{f_{ij}^m(1-f_{i'j'}^{m'})}\left[\frac{P\left(F_{i'j'}^{m'}=1|F_{ij}^m=0\right)}{P\left(F_{i'j'}^{m'}=0|F_{ij}^m=0\right)}\right]^{f_{i'j'}^{m'}} && \text{(A.7)}\\
&\propto\prod_{(\mathbf{x}_m^n,\mathbf{x}_{m'}^{n+1})\in\mathcal{E}_t}\prod_{(\mathbf{e}_{ij}^n,\mathbf{e}_{i'j'}^{n+1})\in\bar{\mathcal{E}}_t}\left(\frac{q}{1-q}\right)^{f_{ij}^m f_{i'j'}^{m'}}\left(\frac{1-q}{q}\right)^{f_{ij}^m-f_{ij}^m f_{i'j'}^{m'}}\left(\frac{1-q}{q}\right)^{f_{i'j'}^{m'}} && \text{(A.8)}\\
&\propto\prod_{(\mathbf{x}_m^n,\mathbf{x}_{m'}^{n+1})\in\mathcal{E}_t}\prod_{(\mathbf{e}_{ij}^n,\mathbf{e}_{i'j'}^{n+1})\in\bar{\mathcal{E}}_t}\left(\frac{q}{1-q}\right)^{2f_{ij}^m f_{i'j'}^{m'}-f_{ij}^m-f_{i'j'}^{m'}}, && \text{(A.9)}
\end{aligned}
$$

where $\mathcal{E}_0\subset\mathcal{E}_S$ denotes the set of spatial edges associated to the first image at $n=0$. We assume uniform prior for these edges and drop the $P\left(F_{ij}^0=f_{ij}^0\right)$ terms from Eq. A.4. Eq. A.7 is obtained by simple algebraic manipulations and dropping the constant terms that do not depend on the f_{ij}^m or $f_{i'j'}^{m'}$ variables. Finally, we substitute the persistent probabilities $P\left(F_{i'j'}^{m'}=0|F_{ij}^m=0\right)$ and $P\left(F_{i'j'}^{m'}=1|F_{ij}^m=1\right)$ with q, and probabilities $P\left(F_{i'j'}^{m'}=1|F_{ij}^m=0\right)$ and $P\left(F_{i'j'}^{m'}=0|F_{ij}^m=1\right)$ with $(1-q)$ in Eq. A.8, and rearrange common terms in Eq. A.9.

The *maximum a posteriori* estimate is the value of y that that maximizes the product of the probabilities $P\left(\mathcal{I},\mathcal{X},\mathcal{E}_s|Y=y\right)$ of Eq. A.2 and $P\left(\mathcal{I},\mathcal{X},\mathcal{E}_s|Y=y\right)$ of Eq. A.3. By taking the negative log of this product, we can write that it also is

$$
y^*=\arg\min_{y\in\mathcal{Y}}\sum_{\mathbf{e}_{ij}^n,\mathbf{e}_{jk}^n\in\mathcal{E}_s}w_{ijk}y_{ij}^n y_{jk}^n-\sum_{(\mathbf{x}_m^n,\mathbf{x}_{m'}^{n+1})\in\mathcal{E}_t}\sum_{(\mathbf{e}_{ij}^n,\mathbf{e}_{i'j'}^{n+1})\in\bar{\mathcal{E}}_t}w_p\left(2f_{ij}^m f_{i'j'}^{m'}-f_{ij}^m-f_{i'j'}^{m'}\right),
$$

where $w_p=\log\frac{q}{1-q}$ and q is the probability that one the four temporal consistency conditions introduced above is satisfied.

B Additional Results

b.1 Runner Bean

We used a nine-frame time-lapse sequence of a growing runner bean depicted by Fig. A.3. This is relevant because monitoring the growth of a plant has many uses. They include testing different environmental conditions, getting to understand the influence of specific pesticides or other agricultural products, and evaluating models of plant development and growth [2]. Again, we picked six of the images for training the path classifier and the other three images constituted the testing sequence.

Our algorithm correctly reconstructs the evolving structure and the important topological changes are automatically found. More specifically, in Fig. A.3(g), one can see that there is nonlinear deformation between the structures over time. Initially the plant is partially bent and then straightens. Since the GPR allows for nonlinearity, the correct correspondences between the tree structures are nevertheless found and the tree reconstructions and registration are achieved accurately.

b.2 Road Results

Fig. A.4 depicts the ground truth delineations for the roads of Figs. 4 and 5 in the main paper. We now revisit some of the results obtained on the latter to study the influence of various parameter choices in our approach, both in terms of establishing correspondences and providing roots for our trees.

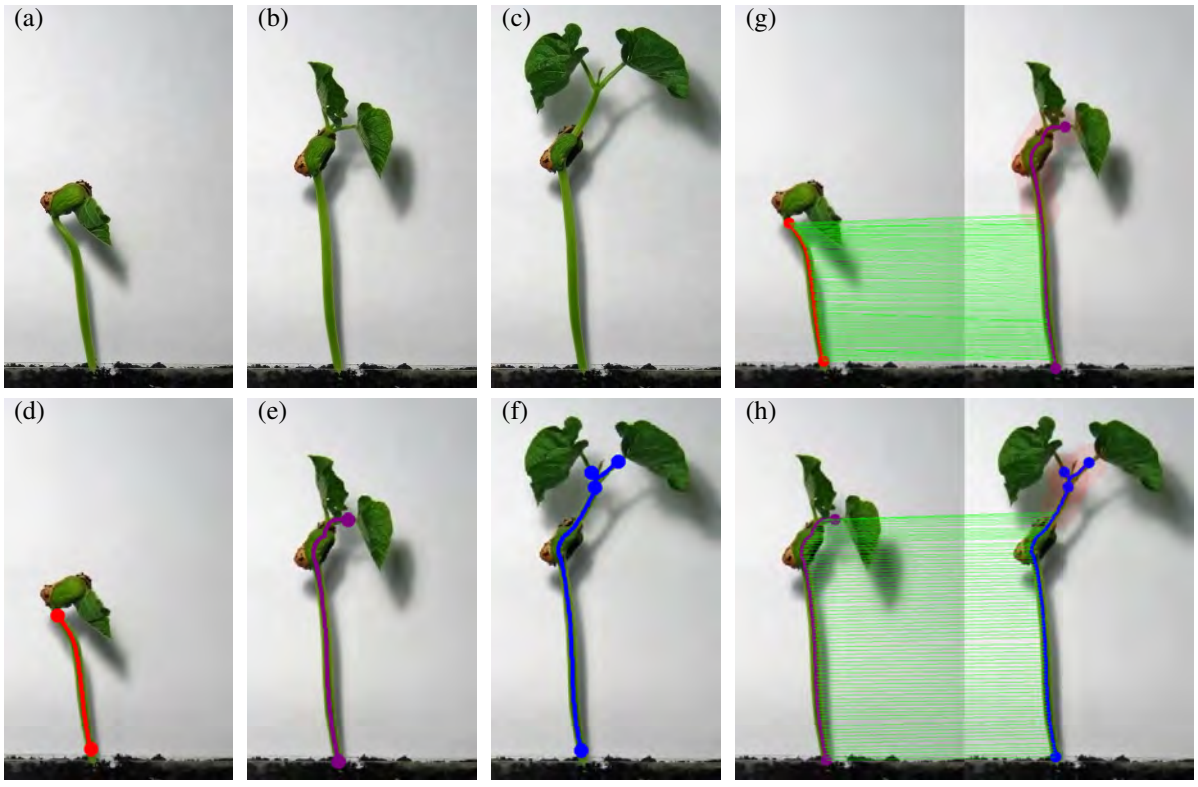


Fig. A.3. Reconstruction and automatic change detection using a time-lapse sequence for growing runner bean. (a, b, c) Original images. (d, e, f) Reconstructed trees in each one of them. (g, h) The horizontal green lines represent temporal edges between their vertices. This figure, as well as most of the subsequent ones, is best viewed in color.

Fig. A.4. Ground truth delineations for the road images of Figs. 4 and 5 in the main paper.

b.2.1 Correspondence Parameters

The temporal correspondences we establish between graph vertices found at different times are an essential building block of the spatio-temporal graphs we introduced in Section 3 of the main paper. Computing them requires setting three parameters, the size of the neighborhood used to compute the Normalized Cross Correlation (NCC), the maximum Mahalanobis distance between corresponding points, and the minimum tubularity value for a centerline point to be considered.

As shown in Table 1, using too small correlation neighborhoods degrades the results by making NCC insufficiently

Fig. A.5. Tree root selection. **Top row.** Roots selected either manually (circles) to produce the results depicted by Fig. 4 or automatically (crosses). **Bottom row.** Delineations using the automatically selected roots. They are very similar to those of Fig. 4 except for a small jog in the 2016 image in the immediate vicinity of the root.

size	Image #1	Image #2	Image #3	Average
5 x 5	0.29	0.72	0.47	0.49
15 x 15	0.54	0.72	0.71	0.66
25 x 25	0.38	0.31	0.27	0.32

TABLE 1

DIADEM score of *Roads* dataset for different sizes of neighborhood used to calculate NCC.

discriminative and allowing too many false correspondences. Similarly, using too large neighborhoods yields corresponding vertices that may be away from the actual centerline because the correlation decreases too slowly with distance to the centerline. In practice, neighborhoods of size half the vertex sampling distance, which we take be either or 20 or 30, yield the best results.

In Table 2, we report the influence of the Mahalanobis distance threshold. Choosing a very low one results in many true correspondences being rejected with an attending performance loss. When it becomes sufficiently large, enough correspondences are detected and they tend to be true positives as only points whose tubularity is larger than 0.15 times the maximum tubularity in the image are considered. In practice, we use a threshold of 4, which reduces the search area without impacting performance.

b.2.2 Root Selection

Ideally, the 2D location of the root nodes should be the same in all images and such that the linear structure is visible in all of them. However, under this constraint, their exact root placement has only a very local influence if any, on the final result.

In Fig. A.5, we show the delineations obtained by picking as the root the highest tubularity point in the first image and then transferring its location to the other ones, instead of introducing a manually selected one, as we did in Fig. 5 of the main paper. Because the corresponding node is now missing, this slightly changes the result in its neighborhood in one of the images but has otherwise almost no impact on the overall result.

Square distance	Image #1	Image #2	Image #3	Average
10	0.53	0.63	0.26	0.47
15	0.54	0.72	0.71	0.66
50	0.54	0.70	0.70	0.64
100	0.54	0.72	0.71	0.66

TABLE 2

DIADEM score of *Roads* dataset for different values of the square of the maximum allowable squared Mahalanobis distance between two correspondences.

REFERENCES

[1] E. Turetken, F. Benmansour, B. Andres, P. Glowacki, H. Pfister, and P. Fua, "Reconstructing Curvilinear Networks Using Path Classifiers and Integer Programming," *PAMI*, vol. 38, no. 12, pp. 2515–2530, 2016. 3

[2] P. Prusinkiewicz, "Modeling of Spatial Structure and Development of Plants: a Review," *Scientia Horticulturae*, vol. 74, no. 1–2, pp. 113–149, 1998. 3