Cartography and Geographic Information Science
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcag20

River network completion without height samples using geometry-based induced terrain

Tsz-Yam Lau [a] & W. Randolph Franklin [b]
[a] Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA
[b] Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, 12180, USA
Published online: 29 Apr 2013.

To cite this article: Tsz-Yam Lau & W. Randolph Franklin (2013): River network completion without height samples using geometry-based induced terrain, Cartography and Geographic Information Science, DOI:10.1080/15230406.2013.780785

To link to this article: http://dx.doi.org/10.1080/15230406.2013.780785

Cartography and Geographic Information Science, 2013
http://dx.doi.org/10.1080/15230406.2013.780785

River network completion without height samples using geometry-based induced terrain

Tsz-Yam Lau[a]* and W. Randolph Franklin[b]

[a]*Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY 12180, USA;* [b]*Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA*

(*Received 9 September 2012; accepted 6 February 2013*)

We present an improved method for connecting broken river segments by incorporating segment geometry into the induced terrain approach. The reconnection problem is important in aerial photography. Canopies and clouds cover parts of rivers, and a complete river network is necessary for transportation, land-use planning, and floodplain control. Our previously presented induced terrain approach guarantees that the global network topology is hydrologically consistent. However, it assumes many reliable height samples, which may be unavailable because of extensive view obstacles or gentle surfaces, to reconstruct the induced terrain surface. In this paper, we exploit river segment geometries to generate the induced terrain. We assign lowest heights to the known river locations, and higher and higher heights for locations further and further away. Since areas radiating from segment tips tend to be where a river is broken, we assign relatively small heights to these areas, so as to favor water flow toward them in the subsequent river derivation. On average, such a tip-biased model improves reconnection accuracy by five percentage points (which is 40% of what we would have been corrected if we had had rich height samples covering 10% of the terrain cells) over the intuitive baseline model in which the height of a location increases with its distance from the respective nearest river location, without considering the potential effect of segment tips as the tip-biased approach.

Keywords: river segments; river network completion; height awareness; segment geometry; hydrology

1. Introduction

This paper presents how river segment geometry can be combined with the induced terrain approach (Lau and Franklin 2011) to improve river segment reconnection when rich terrain height samples are unavailable. The river network completion problem is important since many transportation issues, such as ship routing, pollutant monitoring, and floodplain control, rely on full network information.

Due to their capability to cover a wide area within a short time, aerial surveys are now the main source of river network data. The extraction of river networks involves a classification of terrain features according to the multi-spectral data collected from the spatial domain. In a few locations, we fail to detect any signal due to view obstacles such as tree canopies and clouds. At a few other locations, we may have difficulty determining whether the spectrum represents a water body or some other terrain feature (Zhang 2000). The classification result thus consists of a disconnected river network containing gaps so as to avoid false positives. This is not a problem if complementary ground surveys are performed, or ample time is provided to complete the reconnections manually. Otherwise, say if the river network is on an extra-terrestrial planet, or we are monitoring a large quantity of such multispectral images for potentially changing hydrography in real time (for example, those in a tropical rain forest), automatic bridging of those broken segments is needed.

Figure 1, left visualizes the river segment reconnection problem on an $m \times m$ *river location grid* $\mathbf{R}$. In each of these m^2 cells, $R_{i,j}$ is supposed to have a binary value, indicating if it is on a river or not (1 for a river location and 0 otherwise), but we miss the values of some of these cells. Our task is to recover those missing river locations so as to reproduce the original complete river network.

Figure 1, right, shows the workflow of the induced terrain approach. Realizing that height plays an important role in determining water flow physically, we first reconstruct a terrain surface consistent with the given height samples and other available height-affecting terrain attributes such as known river locations (Lau and Franklin 2011), water flow directions (Lau, Li, and Franklin 2011), and known non-river locations (Lau and Franklin 2012). Then based on that induced terrain, we derive a river network that is consistent with our knowledge of the global network topology (Lau and Franklin 2011), also known as hydrological consistency, to reconnect the segments. Using additional terrain knowledge to eliminate invalid reconnections, this approach features higher accuracy than conventional morphological approaches which treat the issue as a typical line-joining problem. We will detail this approach in Section 2.

*Corresponding author. Email: rpi.laut@gmail.com

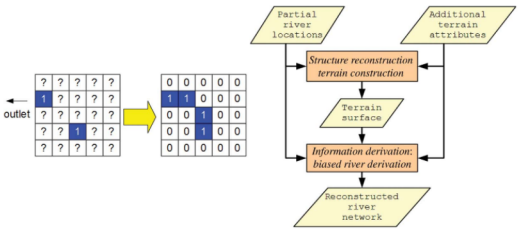

Figure 1. The river segment reconnection problem on a 5 × 5 river location grid **R** (left). The induced terrain approach workflow (right).

Previous discussions of the induced terrain approach assume rich height samples (density around 10%, which means 10% of the terrain grid cells have height samples) that are typically available. This paper, in contrast, works on situations with insufficient reliable height samples. Similar to river network identification, there will be no such problem if ample time is allowed for data collection. Otherwise, say if we work on extra-terrestrial or real-time height data, the height samples collected through LiDAR aerial survey could be too sparse if the terrain is under deep fog or heavy tree canopies at the time of data collection. The situation can also happen with a relatively flat terrain. For example, in Amazonia the slope may be just 2.7 cm/km (Irion and Kalliola 2010). In both cases, little noise in a few measured height values could affect the topological relationship between neighboring heights in the induced terrain, which could eventually change the topology of the reconstructed river network drastically. Nonetheless, we would like to preserve the use of the induced terrain approach because it is the only known scheme to embrace other relevant local and global terrain knowledge in the reconnection process. To enable the approach to still work in this situation, we propose inducing the terrain from the river segment geometry instead. Section 3 will summarize a few heuristics that are often used by humans when reconnecting broken segments, look into a previous morphological attempt called *axis-oriented linking* which incorporates those heuristics and finally discuss how those heuristics could be integrated into the induced terrain approach with a *baseline terrain model* variant called the *tip-biased terrain model*.

2. Induced terrain approach

2.1. Structure reconstruction

The induced terrain approach starts with reconstructing a full height grid **Z** according to the given height samples and other additional terrain attributes. We emphasize the use of heights because the topological relationship among the neighboring locations determines how water is routed: water at a location flows to its lower neighbors. A hill sitting between two river segments could act as an obstacle to block water of one segment to flow to the other segment.

With evenly distributed height samples, we may first reconstruct a preliminary terrain from the given partial heights using a general surface reconstruction scheme such as *natural neighbor interpolation* (NN) (Sibson 1981). We then lower the heights of given river locations by a *trench amount*, a technique known as *stream burning* (SB) (Hutchinson 1989), to help those river locations trap water there. We name this combination *natural neighbor-stream burning* (NN–SB). If we have additional information, we may further refine the terrain. For example, known stream directions may be used to correct heights that are not decreasing along the given river segments (Lau, Li, and Franklin 2011). We may also raise the heights of those locations known to have no river flow to a level unreachable by any water. This practice prevents them from being river locations, thus increasing the opportunity of obtaining correct reconnections (Lau and Franklin 2012).

2.2. Information derivation

To reconnect the segments, we compute a river network from the above reconstructed terrain using a river derivation algorithm biased toward the given river locations. Such a river derivation algorithm first computes the water flow directions of all the locations based on their heights and then finds the amount of water passing through each cell based on those directions. After that, each given river location is offered an initial water amount that is exactly the critical amount (accumulation cutoff). This practice makes sure that the location has sufficient water to be identified as a river location even if it receives no inflow from other locations due to the unavoidable

imperfect induced terrain generation. Meanwhile, all other locations are allocated zero initial water. They have to receive water from a given river location for river flow. Next, we compute the total amount of water passing through each location. If, at the end, we need to trim the river network to one-cell wide, we protect the given river locations from being removed. This practice ensures that the resulting river network passes through, and hence reconnects, the entire given river locations.

The use of the above river derivation algorithm helps enforce hydrological consistency in the reconnected river network. Hydrological consistency refers to the set of global constraints that governs the topology of the complete river network. Similar to the completion of road networks, enforcing these global restrictions helps eliminate invalid reconnections and thus give way to the correct ones (Steger, Mayer, and Radig 1997). One common constraint governs how water is routed away from local sinks. One class of river derivation algorithms, such as `r. watershed` (Ehlschlaeger 2008) available in a geographic information systems (GIS) software called Geographic Resource Analysis Support System (GRASS), assigns flow directions such that water flows "uphill" and escapes the sink by following a *least-cost path*, where the costs are the elevations of the locations (Morris and Heerdegen 1988). Some others, such as *Terraflow* (Arge, Mitasova, and Toma 2002), choose to fill all sinks by *flooding* so that all cells can then be assigned flow directions that never go uphill (Jenson and Domingue 1988). Another common constraint is whether the network should consist of a number of tree branching structures. For example, the single-flow direction (SFD) version of `r. watershed` guarantees that every location is assigned a single direction to route its water to a terrain edge or a sink in a loop-free manner. To allow braids or loops in the river network, the multiple-flow direction (MFD) version of `r. watershed` allows distribution of water from the river location to two or more neighbors and hence river branching. Indeed, we are still seeing the development of faster river derivation schemes that satisfy different sets of hydrological consistency constraints (Metz, Mitasova, and Harmon 2011; Magalhaes et al. 2012). One should pick the river derivation algorithm that matches expectations about the global network topology.

3. Segment geometry

3.1. Baseline model

As described in Section 1, there are river reconnection scenarios without sufficient reliable height samples. Knowing that the given river locations are where water is trapped and that the water at other locations is expected to move to the respective nearest river locations, we may model each known river location as a local minima to guide water flowing as such. Without any information on the heights of the river locations, we arbitrarily set all their heights to a value z_{river}. All the other locations are higher than z_{river}. The further away such a location is from the respective nearest river location, the higher it is. This leads to the baseline model illustrated in Figure 2. Mathematically, define $\sigma > 0$ to be the *height growth rate* indicating the per-pixel height increase on moving away from the nearest river location in the known river location grid **R**. The model evaluates the height of a location (k, l), $z_{k,l}$, with the following formula.

$$z_{k,l} = \min_{(i,j):R_{i,j}=1} \left(z_{i,j} + \sigma\delta(p_{i,j}, p_{k,l})\right) \tag{1}$$

where $z_{i,j} = z_{\text{river}}$ and $\delta(p_{i,j}, p_{k,l})$ is the pixel distance between (k, l) and the respective nearest given river location (i, j).

$$\delta(p_{i,j}, p_{k,l}) = \sqrt{(i-k)^2 + (j-l)^2} \tag{2}$$

Under this terrain model, pair of river locations which are further apart are separated by a higher crest higher than the pair which is closer. This implies a higher cost (for the least-cost path approach of river derivation) or more

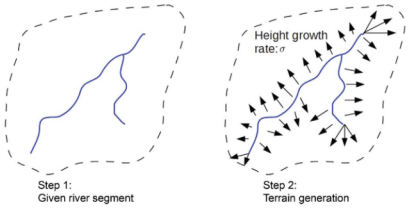

Figure 2. Baseline terrain model.

flooding (for the flooding approach of river derivation). As a result, the model favors *shortest straight-line reconnections* between segments, a major heuristic that people often use when joining segments.

3.2. *Power of the baseline model*

To evaluate the performance of the baseline model, we perform a test with the 12, 400 × 400 digital elevation models (DEMs) shown in Figure 3. These DEMs are extracted from two 30-m cells and two 90-m cells, as described in Table 1. They are all mountainous and hilly terrains whose missing heights are hard to reconstruct in contrast to those in gentle surfaces. We derive the respective full river networks with the complete elevation data since we need the accurate ground-truth river networks for comparison with the reconnection results. Real partial observations rely on humans to complete the missing parts, which may mean errors. We run `r. watershed` with an accumulation cutoff threshold = 200 and initial water amount at each location = 1 over these 12 DEMs. We pick `r. watershed` due to its readiness to accept different initial water amounts at different cells, which is necessary for biased river derivation. We obtain the eight-connected river networks.

We sample for observed river locations as follows: first we divide the whole grid into 20 × 20 subgrids. In each subgrid, we randomly pick a point and mask an area of 12 × 12 around it. This practice mimics the occlusions in real aerial photos. Figure 4 shows the resulting full and partial river locations of the *mtn*3 data set.

To emulate height samples of various densities, we hide different percentages of heights in the ground-truth elevation grid. The trench amount used in NN–SB induced terrain generation is arbitrarily set to 30 because the precise value above some floor value is unimportant (Callow, Niel, and Boggs 2007). For the subsequent biased river derivation, we use the same accumulation threshold as when we derive the theoretical river networks.

Figure 5, left, shows the *mtn*3 terrain reconstructed with the baseline terrain model. We compute the *correct adjacent segment reconnection rate* which is the

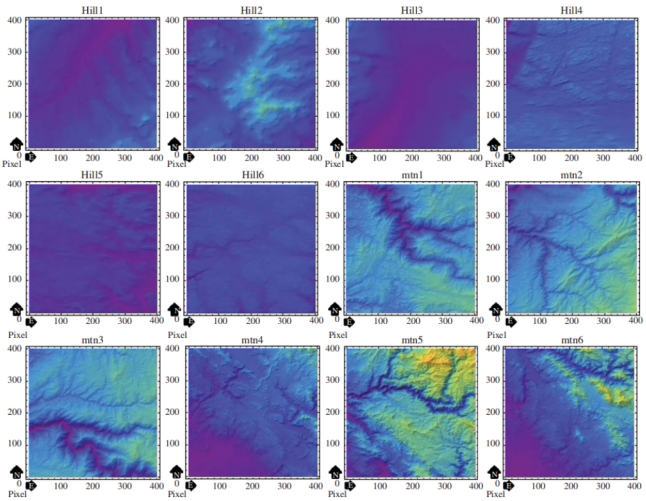

Figure 3. Test DEMs.

Table 1. Test DEM information.

DEM	Type	Name	Range (pixel unit)	Elevation in meters		
				Mean	Std. Dev.	Range
Hill1	SRTM1	W111N31	401:800,1:400	1251	79	1105:1610
Hill2	SRTM1	W111N31	401:800,401:800	1548	134	1198:1943
Hill3	SRTM1	W111N31	1:400,1:400	1309	59	1199:1699
Hill4	SRTM3	W60N52	401:800,401:800	441	62	140:583
Hill5	SRTM3	W60N52	401:800,401:800	486	38	368:598
Hill6	SRTM3	W60N52	801:1200,801:1200	447	32	229:536
*mtn*1	SRTM1	W121N38	1201:1600,1201:1600	712	146	219:1040
*mtn*2	SRTM1	W121N38	2801:3200,801:1200	847	152	330:1283
*mtn*3	SRTM1	W121N38	3201:3600,401:800	723	161	233:1021
*mtn*4	SRTM3	W121N37	1:400,401:800	294	143	45:883
*mtn*5	SRTM3	W121N37	1:400,801:1200	857	281	239:1767
*mtn*6	SRTM3	W121N37	401:800,801:1200	415	234	72:1297

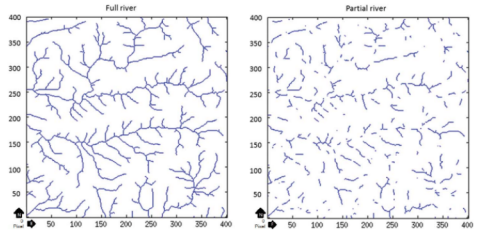

Figure 4. *mtn*3 full river network (left) and partial river network (right).

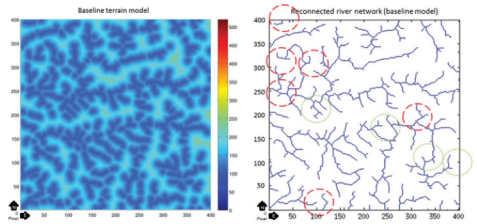

Figure 5. Reconstructed *mtn*3 terrain with baseline terrain model (left) and the respective reconnected river network (right). Some wrong reconnections are circled.

proportion of river segments which are connected back to their respective adjacent downstream segment.

According to Table 2, when the height density is 10%, the mean correct adjacent segment reconnection using the height-dependent NN–SB scheme reaches 84%. In contrast, the results are just 72% when we use the height-independent baseline model instead. There is a 12% margin from NN–SB and the difference is statistically

Table 2. Correct adjacent segment reconnection rates with diffrerent terrain reconstruction schemes discussed in this paper: baseline, tip-biased, and NN–SB.

			NN–SB with height sample density				
DEM	Baseline (%)	Tip-biased (%)	10% (%)	5% (%)	1% (%)	0.5% (%)	0.1% (%)
Hill1	65.07	70.33	78.47	74.64	66.99	61.72	47.85
Hill2	76.62	81.39	84.42	83.98	72.29	62.34	48.48
Hill3	56.25	63.24	60.29	58.46	58.09	52.94	55.88
Hill4	67.52	73.36	76.74	72.63	55.84	48.18	37.59
Hill5	72.87	78.54	86.23	80.16	63.16	53.44	40.49
Hill6	75.00	81.25	85.55	81.25	63.28	53.13	48.05
*mtn*1	77.92	77.08	92.08	85.42	74.17	61.25	47.50
*mtn*2	75.22	82.30	95.13	88.05	72.12	63.72	51.33
*mtn*3	72.32	78.57	90.18	85.27	72.32	60.71	49.55
*mtn*4	70.41	76.19	80.27	74.83	60.2	54.42	40.48
*mtn*5	72.97	79.15	93.82	86.87	69.11	53.67	43.63
*mtn*6	76.71	79.92	92.37	84.74	69.88	59.04	48.19
Mean	71.57	76.78	84.62	79.69	66.45	57.05	46.59

significant. When the height density drops, the margin diminishes. The NN–SB results are still superior when the sample density is 5%. However, when the height sample density drops to 1%, the baseline model performs better than the NN–SB counterpart.

3.3. *Shortcomings of the baseline model*

While the above baseline model matches the shortest straight-line reconnection heuristic, we do observe a few reconnections that are odd and wrong indeed, such as those circled in Figure 5 (right). They look strange since they violate the following heuristics.

- *Extend from tips.* Humans tend to extend segments from the tips rather than the other parts of the segment, because they expect that a river is more likely to have been blocked from view starting there. For example, in Figure 6 (left) we would rather take a long reconnection emerging from the segment tips rather than a short link that emerges from the middle of the segments. For real examples,

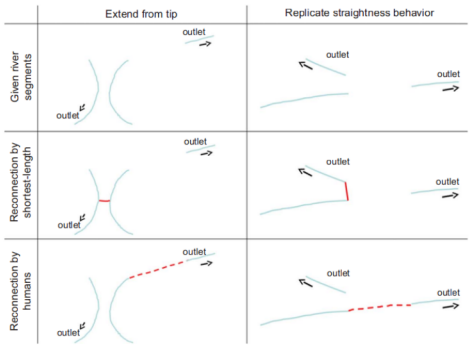

Figure 6. Odd reconnections using the shortest straight-line reconnection principle.

see those indicated with green dotted circles in Figure 5, right.

- *Replicate straightness behavior.* Humans tend to expect a straight segment to have straight extrapolation, and vice versa. For example, in Figure 6 (right), the shortest-length reconnection does not look usual as it inserts a great bending between the two straight reconnected segments. It appears much more natural for the left segment to go straight and connect to the straight segment on the right even though it is longer. For real examples, see those indicated with red dashed circles in Figure 5 (right). In fact, the shape of a river tends to be locally similar due to the closeness in water speed, soil erodibility, vegetation, etc.

A few morphological methods pay attention to these reconnection heuristics. A notable example is the *axis-oriented linking* as shown in Figure 7 (Zhang 2000).

The first heuristic is realized in how the algorithm grows a segment: a variable operation window is moved across all pixels of the river location grid **R** repeatedly. A segment (axis) is extended n pixels at a time (step length) when its tip is swept (Figure 7, left).

The second heuristic is reflected in how the algorithm determines the extension direction: when the center of the operation window falls on a starting pixel of a segment P_1, the algorithm first finds the nearest starting pixel of another segment P_3 from within a search window of the segment's forward direction. If such a P_3 is found, the algorithm extends the current segment toward the nearest segment to P_4 such that P_4 lies next to P_1 while it is closest to P_3 (Figure 7, middle). Otherwise, it grows the segment according to its exhibited straightness behavior. It searches for the pixel P_2 which lies on the same segment as P_1 and is most distant from P_1 in the window. The slope between P_1 and P_2 is computed and the segment is extended according to that slope (Figure 7, right).

3.4. Tip-biased model

To utilize the above two additional heuristics in the induced terrain approach, we adopt the *tip-biased model* shown in Figure 8. This new model features two parameters: θ which tells the extent of bending we accept for privileged reconnections at the segment tips, and σ' which suggests the extent that we favor river growing at the segment tips spanned by θ over the other parts of the segments. The process first estimates the *forward direction* of each segment tip with the vector that goes from the other end of the segment or the nearest branching point to the tip location. Then we construct a binary matrix **I** which indicates the *privileged regions* that are radiating from the forward direction of the segment tip by an angular range $\pm\theta/2$. They are so called because we are applying a height growth rate σ' smaller than σ in these regions.

$$z_{k,l} = \begin{cases} \min_{(i,j):R_{i,j}=1} \left(z_{i,j} + \sigma\delta(p_{i,j}, p_{k,l})\right) \text{ if } I_{k,l} = 1 \\ \min_{(i,j):R_{i,j}=1} \left(z_{i,j} + \sigma'\delta(p_{i,j}, p_{k,l})\right) \text{ otherwise.} \end{cases} \quad (3)$$

The *mtn*3 terrain generated with such a tip-biased model is shown in Figure 5 (bottom right). As you can see, the heights of these privileged areas are smaller than those outside these regions in general. As a result, for locations which are of the same distance away from the nearest river locations, the ones inside the privileged regions are of lower elevation than the counterparts outside. The former is thus more likely to trap water, become river locations and hence parts of the reconnections.

3.5. Parameter settings

θ tells the extent of bending we accept for privileged connections. A smaller θ tends to make the process selective, favoring connections according to segment straightness. This is similar to the step in axis-oriented linking in which the current tip is asked to look for another segment to connect in a forward search space only.

σ' determines to what extent we favor river growing in the privileged regions over other parts of the segment. A small σ' relative to σ tends to favor extensions at the segment tips because for some two locations which are of the same distance from the respective nearest river locations, the one inside the privileged region has a lower height than the counterpart outside. In other words, we offer locations inside the privileged region with additional distance allowance. If accompanied with a small θ, we further favor extensions according to segment straightness and hence reconnections of facing segments even if they are far apart. This is similar to the step length parameter n in axis-oriented linking. The larger is the value of n, the longer a segment is grown in each step.

The choices of θ and σ' depend clearly on one's prior knowledge on the sinuosity of the river network. However, there are cases in which we have no such prior knowledge. Therefore, it is interesting to see what works in an average case.

In an attempt to find such a θ, we first fix σ' to 0.50σ and set θ to four different levels: $\pi/4$, $\pi/2$, $3\pi/4$, and π. Table 3, left displays the respective segment reconnection results. Results with $3\pi/4$ and π are statistically significantly less accurate than those with $\pi/4$ and $\pi/2$. This means reconnections are expected to be within $\pm\pi/4$ of the forward direction.

Then we fix θ to $\pi/2$ and set σ' to three different levels: 0.25, 0.50, and 0.75 of σ. Setting σ' to 0.50σ gives statistically significantly better results than the other settings, as shown in Table 3, right. It appears that such a moderate setting compensates for the effect of segment geometry over distance from nearest river locations well.

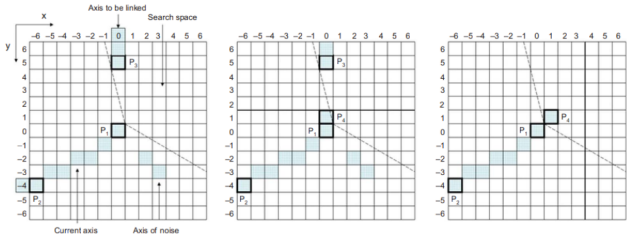

Figure 7. Axis-oriented linking. Operation window and the search space (left). A starting pixel of another segment is found in the search space (middle). No starting pixel of another segment is found in the search space (right) (Zhang 2000).

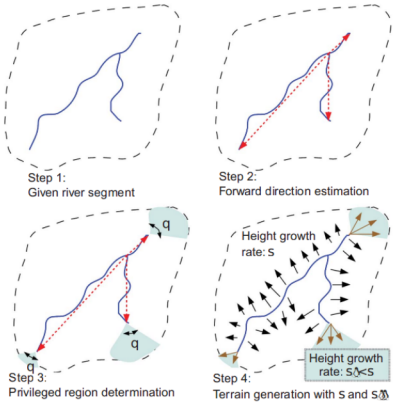

Figure 8. Tip-biased terrain model.

3.6. *Improvement over baseline model*

We compare the results computed by the tip-biased terrain model using the best parameter setting ($\sigma' = 0.50\sigma$, $\theta = \pi/2$) with those by the baseline terrain model and NN–SB. As shown in Table 2, NN–SB unsurprisingly offers the highest correct reconstruction rate in almost all the DEMs under test. When height data are not available, our tip-biased terrain model performs 5.20% better (on average) than the unbiased baseline model, which is 40% of the improvement which would have been achieved with rich height samples of density 10%. All the wrong reconnections circled in dotted lines in Figure 5 (right) are corrected. We acknowledge a few new incorrect reconnections on switching from the baseline model to the tip-biased model. However, the few reconnections that are made wrongly are generally outnumbered by the corrected

Table 3. Correct adjacent segment reconnection rates with the tip-biased terrain model of different θ and σ values. The best reconnection rates for each DEM are bolded.

	$\sigma' = 0.5\sigma$				$\theta = \pi/2$		
DEM	$\theta = \pi/4$	$\theta = \pi/2$	$\theta = 3\pi/4$	$\theta = \pi$	$\sigma' = 0.25\sigma$	$\sigma' = 0.50\sigma$	$\sigma' = 0.75\sigma$
Hill1	**70.81%**	70.33%	64.11%	64.11%	**72.72%**	70.33%	67.46%
Hill2	79.65%	**81.39%**	77.92%	76.19%	**82.68%**	81.39%	79.22%
Hill3	62.87%	**63.24%**	56.62%	56.62%	**68.01%**	63.24%	58.09%
Hill4	**75.91%**	73.36%	70.07%	67.15%	70.80%	**73.36%**	70.44%
Hill5	**81.38%**	78.54%	77.73%	74.49%	**79.76%**	78.54%	77.73%
Hill6	79.30%	**81.25%**	77.34%	73.83%	79.68%	81.25%	**81.64%**
*mtn*1	**81.25%**	77.08%	77.92%	76.67%	74.58%	77.08%	**80.00%**
*mtn*2	75.66%	**82.30%**	79.65%	76.55%	75.66%	**82.30%**	79.20%
*mtn*3	77.68%	**78.57%**	77.68%	72.77%	76.34%	**78.57%**	77.23%
*mtn*4	72.79%	**76.19%**	77.49%	72.11%	74.49%	**76.19%**	72.45%
*mtn*5	76.83%	**79.15%**	75.68%	69.11%	75.29%	**79.15%**	78.38%
*mtn*6	79.12%	**79.92%**	78.31%	77.51%	79.12%	**79.92%**	78.31%
Paired diff. (θ)	$\{\pi/8\}\,\Delta\{\pi/4\}$	$\{\pi/8\}\,\Delta\{\pi/2\}$	$\{\pi/4\}\,\Delta\{\pi/2\}$	$\{3\pi/4\}\,\Delta\{\pi/2\}$	$\{3\pi/4\}\,\Delta\{\pi/4\}$	$\{3\pi/4\}\,\Delta\{\pi/2\}$	
Average	Δ0.67%	4.68%	5.35%	2.54%	Δ2.82%	Δ2.14%	
95% CI min	Δ1.41%	3.94%	4.74%	2.05%	Δ3.36%	Δ2.95%	
95% CI max	0.07%	5.42%	5.96%	3.02%	Δ2.27%	Δ1.34%	
Paired diff. (σ')	$\{0.25\sigma\}$–$\{0.50\sigma\}$	$\{0.25\sigma\}$–$\{0.75\sigma\}$	$\{0.50\sigma\}\,\Delta\{0.75\sigma\}$				
Average	Δ1.02%	0.75%	1.76%				
95% CI min	Δ1.78%	Δ0.31%	1.24%				
95% CI max	Δ0.25%	1.80%	2.29%				

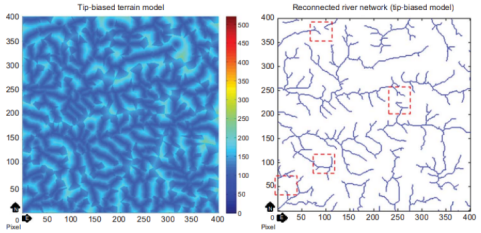

Figure 9. Reconstructed *mtn3* with tip-biased terrain model (left) and the respective reconnected river network (right). Some new wrong reconnections are enclosed with dashed rectangles.

counterparts. This supports the use of our tip-biased model over the baseline model.

4. Conclusion

We have presented a method using segment geometry that can be used instead of the unreliable or unavailable partial height data in the induced terrain approach of reconstructing hydrological consistent river networks from broken segments. This is important because the induced terrain approach embraces other available local and global terrain knowledge, such as known non-river locations, water flow directions, and global river network topology, in the reconnection process, which improves accuracy. Instead of growing heights uniformly from the segments in all directions as the intuitive baseline model does, we propose the tip-biased model in which we reduce the height growth rate of locations radiating from the segment tips.

Embracing a few additional segment reconnection heuristics, this alternative model improves the correct reconnection rate by around five percentage points on average, which is around 40% of what we would have been corrected if we have had rich height samples covering 10% of the terrain cells and used natural neighbor interpolation to generate the induced terrain. For future work, first note that in Section 3.3, we are talking about universal θ and σ^l values that fit the whole region. However, we realize that different local regions may improve best at different settings. It may be interesting to adjust the parameters at individual tips. For example, our tip-biased model assumes that missing channels tend to originate at the segment tips but not from the mid-segment area, which is highly probable (that is why humans bias reconnections there). If it is no longer the case (such as when canopies are likely to obscure small tributaries that join large river in mid-segment) in some local regions, σ^l should approach σ there. Second, the highly-twisted portions of the river segments appear likely to be a river branch, as shown by a few wrong reconnections that emerge with the use of the tip-biased model in Figure 9 (right). We may consider these locations as tips as well. Third, current privileged regions span the same angle ($\theta/2$) for both the left and right sides of the segment forward directions. In case the stream is highly meandered, we may need to span different angles for the two sides so as to favor reconnections from one side but not the other as implied by the trend of the given stream. Fourth, we do not use any height samples in the process. We are interested in developing a scheme that incorporates both segment geometry and height samples, because they are potentially complementing with each other. Last but not least, although the tip-biased terrains achieve better river segment reconnection results, they do not look natural. Future work could improve this situation.

Ultimately, we would like to port the same solution framework to complete some other 2D or 3D networks, such as fragmentary dendrite networks. Preliminary work includes studying how similar the river reconnection domain is to those problems, and figuring out the appropriate clues in those new domains.

Acknowledgements

This research was partially supported by NSF grants CMMI-0835762 and IIS-1117277. We also appreciate the comments of the reviewers on the initial draft.

References

Arge, L., H. Mitasova, and L. Toma. 2002. "r.terraflow." Accessed February 18, 2011. http://www.cs.duke.edu/geo*/terraflow/r.terraflow.html.

Callow, J. N., K. P. V. Niel, and G. S. Boggs. 2007. "How does Modifying a DEM to Reflect Known Hydrology Affect Subsequent Terrain Analysis?" *Journal of Hydrology* 332: 30–39.

Ehlschlaeger, C. 2008. "GRASS GIS: r.watershed." Accessed June 23, 2010. http://grass.itc.it/gdp/html_grass63/r.watershed.html. Last Visit 6/23/2010.

Hutchinson, M. F. 1989. "A New Procedure for Gridding Elevation and Stream Line Data with Automatic Removal of Spurious Pits." *Journal of Hydrology* 106: 211–232.

Irion, G., and R. Kalliola. 2010. "Long-Term Landscape Development Processes in Amazonia." In *Amazonia, Landscape and Species Evolution: A Look into the Past*, edited by C. Hoorn, and F. Wesselingh, 185–200. 1st ed. Oxford: Wiley-Blackwell.

Jenson, S. K., and J. O. Domingue. 1988. "Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis." *Photogrammetric Engineering and Remote Sensing* 54: 11.

Lau, T.-Y., and W. R. Franklin. 2011. "Completing Fragmentary River Networks via Induced Terrain." *Cartography and Geographic Information Science* 38 (2): 162–174.

Lau, T.-Y., and W. R. Franklin. 2012. "Better Completion of Fragmentary River Networks with the Induced Terrain Approach by Using Known Non-River Locations." In *15th International Symposium on Spatial Data Handling*, Bonn, August 22–24.

Lau, T.-Y., Y. Li, and W. R. Franklin. 2011. "Joining Fragmentary River Segments with Elevations and Water Flow Directions Using Induced Terrain." In *21st Annual Fall Workshop on Computational Geometry* (Extended Abstract), City College of New York, NY, November 4–5.

Magalhaes, S. V. G., M. V. A. Andrade, W. Randolph Franklin, and G. C. Pena. 2012. "A New Method for Computing the Drainage Network Based on Raising the Level of an Ocean Surrounding the Terrain." In *Bridging the Geographic Information Sciences*, Lecture Notes in Geoinformation and Cartography, edited by J. Gensel, D. Josselin, and D. Vandenbroucke, 391–407. Berlin: Springer.

Metz, M., H. Mitasova, and R. S. Harmon. 2011. "Efficient Extraction of Drainage Networks from Massive, Radar-Based Elevation Models with Least Cost Path Search." *Hydrology and Earth System Sciences* 15 (2): 667–678.

Morris, D., and R. Heerdegen. 1988. "Automatically Derived Catchment Boundary and Channel Networks and their Hydrological Applications." *Geomorphology* 1: 131–141.

Sibson, R. 1981. "A Brief Description of Natural Neighbor Interpolation." In *Interpreting Multivariate Data*, edited by V. Barnett, 21–36, Chapter 2. New York: John Wiley & Sons.

Steger, C., H. Mayer, and B. Radig. 1997. "The Role of Grouping for Road Extraction." In *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, edited by A. Gruen, E. P. Baltsavias, and O. Henricsson, 245–256. Basel: Birkh Auser Verlag.

Zhang, Y. 2000. "A Method for Continuous Extraction of Multispectrally Classified Urban Rivers." *Photo grammetric Engineering and Remote Sensing* 66 (8): 991–1000.