

Baum-Matching in einer Anwendung auf Gefäßsysteme im menschlichen Körper

(Andre Bläul)



Gliederung

1. Motivation
2. Der Algorithmus
 - 2.1. Schritt 1
 - 2.2. Schritt 2
 - 2.3. Die Kostenfunktionen
3. Evaluation

1. Motivation

- möchten wir zwei zu verschiedenen Zeiten aufgenommene CT-Scans von Organen automatisch vergleichen, so stehen wir vor einem Problem
- so ist die Leber ein sehr verformbares Organ
- gleiche Regionen können sich an verschiedenen Stellen befinden
- das Gefäßsystem, wie zum Beispiel das Arteriensystem, ist ein guter Indikator um bestimmte Stellen des Organs wiederzufinden
- diese haben eine sehr einfache Struktur:
 - es handelt sich im Prinzip um einen Baum
 - es gibt einen ausgezeichneten Knoten, die in das Organ eintretende Arterie, der als Wurzel fungieren darf
- im Prinzip würde es sich um Baumisomorphie handeln
- **Aber:**
- bei den CT-Scans können Knoten *verloren* gehen
- somit sind die beiden Bäume i.a. nicht isomorph
- an dieser Stelle setzt der Graphmatchingalgorithmus an

Definitionen:

- Für einen Knoten e ist $\text{out}(e)$ die Menge aller von e ausgehenden Kanten.
- $DV_L(e)$ soll den von e aufgespannten Unterbaum bis zur Tiefe L darstellen.
- H_e und H_v sind konkrete Matchings zwischen Kanten beziehungsweise Knoten.
- $H_e(e_1, e_2)$ und $H_v(v_1, v_2)$ soll heißen, dass e_1 und e_2 respektive v_1 und v_2 in dem Matching aufeinander abgebildet werden.

2. Der Algorithmus

- ausgehend von beiden Gefäßsystemen werden zwei Bäume erzeugt
- dabei soll der Eingang in das Gefäßsystem die Wurzel des Baumes sein
- wir starten bei den Wurzeln und matchen nacheinander Verzweigungen und Pfade
- ein Matching ist hier eine injektive, partielle Funktion zwischen den Knoten der beiden Bäume
- wenn im Verlauf des Algorithmus zu viele Matchings erzeugt wurden, werden die teuersten bezüglich der Kostenfunktion verworfen
- der Matchingalgorithmus ist fertig, wenn es keine Knoten mehr zum Matchen gibt

2.1 Schritt I

- Ausgangspunkt: zwei Verzweigungen v_1 und v_2 sind gegeneinander gematcht
- jetzt werden die ausgehenden Kanten von v_1 mit den ausgehenden Kanten von v_2 gematcht
- dafür werden alle möglichen Matchings zwischen $out(v_1)$ und $out(v_2)$ und deren Teilmengen überprüft
- insgesamt gibt es $\sum_{k=0}^{\min(|out_{v_1}|, |out_{v_2}|)} k! * \binom{|out_{v_1}|}{k} * \binom{|out_{v_2}|}{k}$ solche Matchings
- es ist zu beachten, dass einige ausgehende Kanten nicht gematcht werden

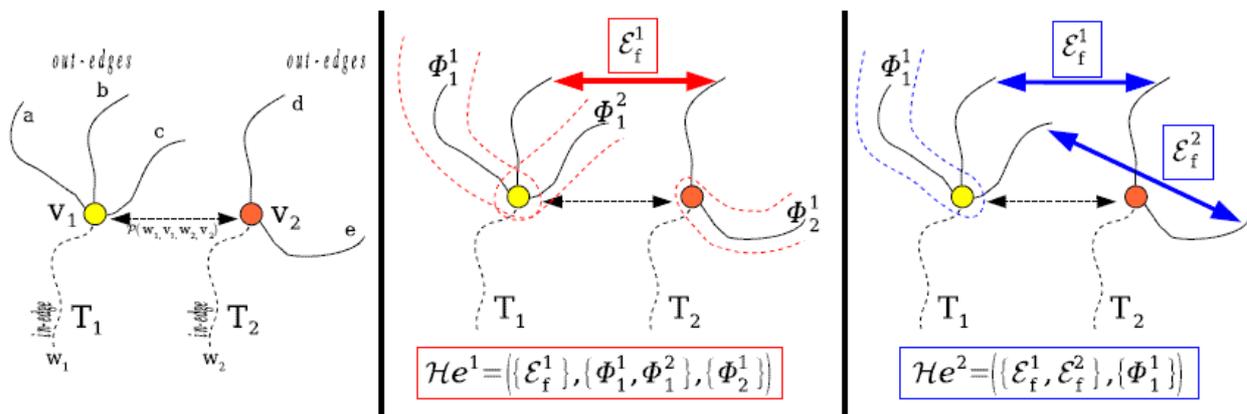


Abbildung 1: Zwei mögliche Matchings für ausgehende Kanten

2.2 Schritt II

- ausgehend von den gerade gematchten Kanten werden jetzt die zu diesen Kanten gehörenden Verzweigungen miteinander gematcht
- dazu betrachtet man den Unterbaum $DV_L(e_1)$ und $DV_L(e_2)$ und bildet je ein Matching für jede Verzweigung aus $DV_L(e_1)$ und $DV_L(e_2)$
- für den Fall, dass jede Verzweigung zwei Söhne hat, werden $(2^{L+1}-1)^2$ Matchings erzeugt

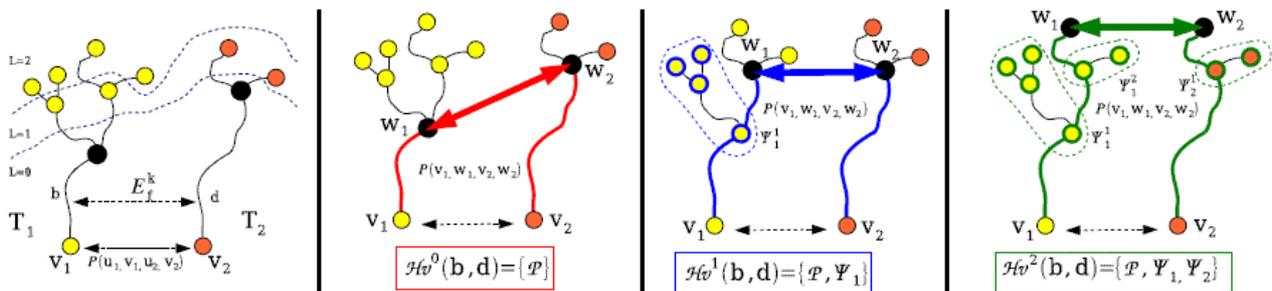


Abbildung 2: Drei mögliche Matchings für Pfade

2.3. Die Kostenfunktion

- die Kostenfunktion hat drei Bestandteile

1. Kosten für ausgehende Kanten:

für zwei miteinander gematchte Kanten e_1 und e_2 soll gelten

$$oemc(e_1, e_2) = \int_0^{l_{min}} \|e_1(t) - e_2(t) - e_1(0) + e_2(0)\|^2 dt + \gamma \int_0^{l_{min}} \left\| 1 - \frac{r_1(t)}{r_2(t)} \right\|^2 dt$$

dabei ist $e_i(t)$ die parametrische Raumkurve und $r_i(t)$ der Radius des zu e_i gehörenden Gefäßes

2. Kosten für zwei gematchte Pfade:

$$pmc(p_1, p_2) = \left(1 + \alpha \frac{l_{max}}{l_{min}} + \beta \frac{l_1 + l_2}{2} \right) \times \left(\gamma \int_0^1 \left\| 1 - \frac{r_1(t \times l_1)}{r_2(t \times l_2)} \right\|^2 dt + \int_0^1 \|p_1(t \times l_1) - p_2(t \times l_2) - p_1(0) + p_2(0)\|^2 dt \right)$$

wobei l_i die Länge und r_i der Durchmesser der zu p_i gehörenden Gefäße darstellen

3. Kosten für ungematchte Unterbäume:

diese Kosten werden folgendermaßen gebildet

$$nmtc(v) = \left(1 + \delta \frac{|T(v)|}{|T|} \right) \times pmc(e, g(e)) + \sum_{k=1}^{|\text{sons}(v)|} nmtc(w_k)$$

mit $g(e)=e$, jedoch ist $r(t)$ der minimale Radius des Gefäßes

- aus diesen drei Kostenfunktionen setzen wir je eine Kostenfunktion für die beteiligten Schritte zusammen
- um die Kosten für die Matchings nach Schritt I zu berechnen benutzen wir:

$$\text{cost}(He(v_1, v_2)) = \sum_{i=1}^{N_1} \text{oemc}(e_1^i, e_2^i) + \sum_{i=1}^{N_2} \text{nmtc}(v_1^i) + \sum_{i=1}^{N_3} \text{nmtc}(v_2^i)$$

- für die Kosten nach Schritt II benutzen wir:

$$\text{cost}(Hv(e_1, e_2)) = \text{pmc}((v_1, w_1), (v_2, w_2)) + \sum_{i=1}^{N_1} \text{nmtc}(v_1^i) + \sum_{i=1}^{N_2} \text{nmtc}(v_2^i)$$

3. Evaluation

- der Algorithmus arbeitet ausreichend schnell:
ca. 4 Minuten um 380 Knoten auf einen 1 Ghz PC zu matchen
- dafür wurde das Gefäßsystem der Leber des Visible Human Project aufgezeichnet und für Matchings verformt bzw. zufällig Knoten entfernt

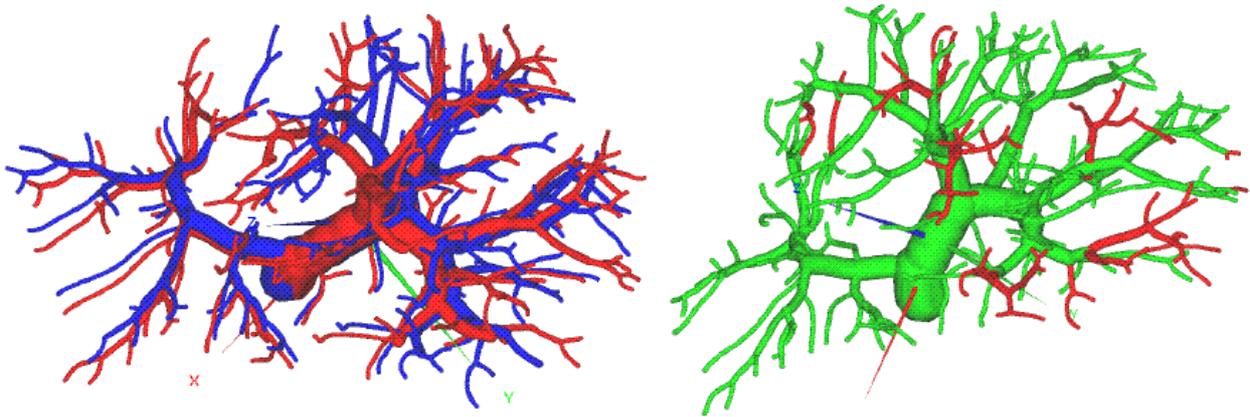


Abbildung 3: Ergebnisse aus Versuchen

- es werden ungefähr 90% der gemeinsamen Knoten miteinander gematcht

4. Quellen

- Arnaud Charnoz / Vincent Agnus / Gregoire Malandain / Luc Soler / Mohamed Tajine: Tree matching applied to vascular system. LNCS 3434 (2005), S. 183-192.
- <ftp://ftp-sop.inria.fr/epidaure/Publications/Malandain/charnoz-gbr-2005.pdf> [05.12.06]