Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

IP Management

TECHNOLOGY AND KNOWLEDGE TRANSFER

Applying graph minors and tangles to image and cluster analysis

Profs. R. Diestel and G. Whittle

AREA

- Data analysis

PROJECT KEY WORDS

- Image recognition; compression
- Cluster analysis; data mining
- Data quality assessment

DEVELOPMENT STATUS

- Mathematical theory shown to apply in principle
- Preliminary algorithmic analysis

PATENT PROCEDURE STATUS

- PCT Patent application filed

POTENTIAL FOR COOPERATION

- Licensing
- Patent right transfer

EXPOSITORY PAPER

- Tangles and the Mona Lisa, available at: http://arxiv.org/pdf/1603.06652

Ref: UHH135
July 2017

Innovation

Tangles are a novel mathematical concept designed to identify highly connected regions in a graph. Diestel and Whittle, two of the world's leading graph theorists, have generalized this idea to serve to identify highly coherent clusters in data sets. This brings the mathematical theory of graph minors, one of the deepest developments in discrete mathematics of the last 25 years, to bear on some of today's most topical computational problems: cluster analysis and image recognition. Tangles are a novel, precise, mathematical way to deal with inherently *imprecise* data sets.

Customer benefit

This innovation is now at its first, theoretical, stage. The input of first-rate mathematics is substantial, the idea to use graph minor theory in this way is entirely new.

Developing this theoretical break-through into concrete algorithms remains a non-trivial task that will require the efforts of some highly skilled computer scientists. If success-ful, the benefits in time are likely to be enormous – which makes this an attractive undertaking not only commercially, but also intellectually. We therefore expect the CS community to take the idea up in the coming years; some leading researchers have already expressed interest and enthusiasm.

Whoever owns the patent at the time this comes to fruition will benefit.

Potential applications

Image recognition and compression.

Cluster analysis. Data mining.

Data quality assessment.

Technical Description

Tangles represent a shift of paradigm in the identification of highly connected regions in a graph: rather than attempting to describe what exactly these regions consist of, in terms of vertices and edges as usual, tangles identify them by merely orienting all the low-order separations of the graph consistently: \`towards' the intended region.

All concrete highly-connected sub-structures induce such orientations. But while such concrete structures may vary, or may not be known in detail, the orientations of separations they induce are all that matters for their mathematical treatment: the two most fundamental theorems in the area, the tangle-tree theorem (which says that the highly connected sub-structures can all be separated from each other in a nested, tree-like way) and the tangle duality theorem (which says that every structure not con-taining a highly connected region has a tree structure witnessing this) can be proved at this very abstract level of separation systems. This has been carried out over the last few years by Professor Diestel and his group.

The level of abstraction is such that tangles can be defined on any data set with a meaningful notion of cutting it in two – such as lines cutting through an image. The tangles precisely identify clusters, or regions, even when these are fuzzy.

The tangle-tree theorem organizes the data sets's clusters into a structure from which its essence can be reconstructed.

The duality theorem assesses the maximum cluster coherence that the data supports.

TUTECH
TUTECH INNOVATION GMBH
IP Management
Dr. Francisco Blanco
Harburger Schlossstr. 6-12
21079 Hamburg - Germany
+49 40 76628 6587
+49 40 76629 6549
blanco@tutech.de
www.tutech.de

Gefördert durch:
Bundesministerium für Wirtschaft und Energie
aufgrund eines Beschlusses des Deutschen Bundestages
WIPANO