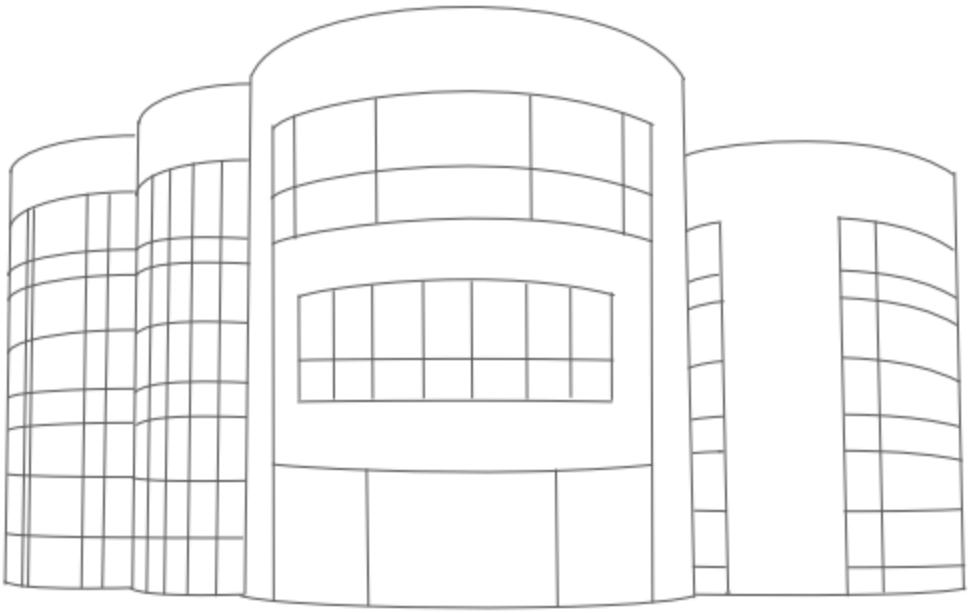
School of Computing

PhD Symposium 2015



http://phdsymp15.computing.dundee.ac.uk/

25th May 2015

Chairs				***Title***
	09:10	Registration		
	09:20	Introduction		
Enrico Pellegrini	09:25 - 10:55	Shen	Haocheng	Multimodal Multiclass Brain Tumour Image Segmentation Using Texture Features and Conditional Random Field
		Zhang	Jian	Automatic Transformation of Raw Clinical Data into Clean Data Using Decision Tree Learning Combining with String Similarity Algorithm
		Hillman	Chris	Towards real-time pre-processing of Mass Spectrometer data using a scalable cluster
	10:55 - 11:15	Coffee Break		
David Gibson	11:15 - 12:15	Sterlini	Phillipa	Guest Talk: Research Data Management and Open Access Publishing
		McCaffery	John	
	12:15 - 13:15	Lunch		
Garreth Tigwell	13:15 - 14:45	Gibson	David	SpaceWire-D Networks
		Huang	Tianjun	Reasoning about Complex Human-Object Interactions in Daily Living Activities
		Pellegrini	Enrico	Binary classification of blood vessels in arteries and veins in ultra-wide field scanning laser ophthalmoscopy
	14:45 - 15:05	Coffee Break		
Meenakshi Kesavan	15:05 - 16:05	Warburton	Chris	Machine Learning for Theory Exploration
		Farka	Frantisek	The Brave New World of Haskell Type Classes
		Vermeulen	Andre	Parallel Patterns using Heterogeneous Computing
		Janier	Mathilde	Modeling argumentative activity in mediation with Inference Anchoring Theory

26th May 2015

Chairs				***Title***
	9:20	Registration		
Frantisek Farka	09:30 - 11:00	Annunziata	Roberto	Towards Robust and Efficient Automated Curvilinear Structure Detection in Medical Images
		McNeil	Andrew	Computer-Assisted Analysis Of Arterial Narrowing In Whole-Body MRA
		Gorman	Ben	Reducing Viseme Confusion In Speech Reading
	11:00 - 11:20	Coffee Break		
Chris Warburton	11:20 - 12:05	Lawrence	John	Argument Mining
		Tigwell	Garreth	Improving Intercultural Human-to-Human Interaction through Non-verbal Visualisation
		Kesavan	Meenakshi	Discovering New Ontologies and Lemmas through Concept Blending and Reasoning
	12:05 - 12:25	Coffee Break		
	12:25 - End	Closing & Prizes		

Multimodal Multiclass Brain Tumour Image Segmentation Using Texture Features and Conditional Random Field

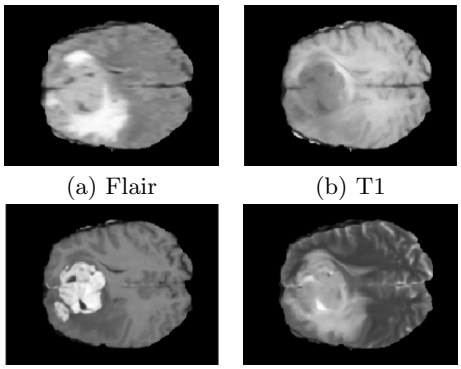
Haocheng Shen and Jianguo Zhang
Computer Vision and Image Processing Group
School of Computing, University of Dundee
Dundee, UK
hyshen@dundee.ac.uk, j.n.zhang@dundee.ac.uk

ABSTRACT

Brain tumour segmentation is an important and challenge task in monitoring the growth or shrinkage of the tumour in patient during therapy. We propose an automatic segmentation framework which applies Random Forest to classify texture features extracted from multimodal magnetic resonance (MR) images to different abnormal categories. Conditional Random Field (CRF) are then used to incorporate spatial constraints to achieve final results. The proposed method is evaluated on BRATS 2013 dataset.

1. INTRODUCTION

Brain tumour segmentation is crucial for monitoring the growth or shrinkage of the tumour in patient during therapy. However, it is a challenge problem due to the considerable intra-class variations in visual appearance of tumours from patient to patient: the structures are non-gird and they are varying in size and location. Multimodal magnetic resonance imaging (MRI) images are extensively used in brain disease diagnosis and radiotherapy due to their ability to provide complementary information for the diagnosis. The normally used modalities including T_1 -weighted MRI (T_1), T_1 -weighted MRI with contrast enhancement (T_{1c}), T_2 -weighted MRI (T_2) and T_2 -weighted MRI with fluid-attenuated inversion recovery (T_{FLAIR}) are used to enhancing different compartment of the tumour. An example of one slice in different modalities is shown in Figure 1. In current clinical practice, the delineation of tumour boundaries is still perform manually, which is time-consuming and tedious for radiologists and is also of limited use for an objective quantitative analysis. Therefore, we propose an automatic brain tumour segmentation framework using texture features from multimodal MRI and evaluate on BRATS 2013 dataset showing promising results.



(a) Flair (b) T1 (c) T1c (d) T2

Figure 1: Example of MRI multi-modal images in BRATS dataset: (a) Flair; (b) T1; (c) T1c; (d) T2.

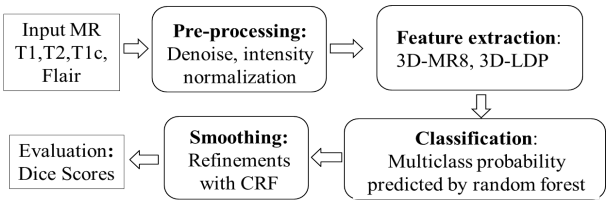


Figure 2: the proposed brain tumour segmentation framework

2. METHODS

We consider the brain tumour segmentation as a pixel-level classification problem. The framework is shown in Figure 2.

2.1 Feature Extraction

Feature extraction of voxels in the image dominates the performance of classification and segmentation. For multi-modality images, we extract feature of each voxel in each modality individually and then concatenate them together to form the final feature representation. Two types of local features are used to describe each voxel:

- **3D Maximum Response Filter (3D-MR8)**: The MR8 filter bank was proposed by [7] and has been proven robust in texture classification. It consists of 38 filters (a Gaussian and a Laplacian of Gaussian, an edge filter and a bar filter at 3 scales and 6 orientations). Measuring only the maximum response across orientations reduces the number of responses from 38 to 8 resulting in rotation invariant filters. We extend it into 3D version by calculating the original MR8 filter responses in saggital, axial and coronal views of the 3D volume separately, and then concatenate them together to form the feature representation of each voxel.

- **3D Local Difference Pattern (3D-LDP)**: LDP [6] avoid coarse quantizations of Local Binary Pattern (LBP) [5, 4] by keeping the original difference between the centre voxel and its neighbour in the patch rather than quantizing them into binary codes. We extend LDP into 3D version similarly to [8] and multi-resolution LDP patterns with radius from 1 to 3 based

on three orthogonal planes of a extracted 3D patch are calculated and then concatenated as a feature representation.

Each type of feature is encoded by Bag-of-words (BoW) and then concatenated to form the final BoW representation of the central voxel.

2.2 Classification

We choose Random Forest (RF) for classification of different abnormal tissue types in the brain. The RF is an ensemble learning algorithm that generates many weak classifiers and aggregates their results in order to make decision. Compared with SVM, RF is inherently designed for multiclass classification problem and is able to output posterior probability directly for post-processing. We rely on threshold-based entropy information gain as the split function of each node and the number of trees is set to 50.

2.3 Spatial Regularization

Classifier-based segmentation methods assume that the individual feature is independent and identically distributed (i.i.d.). Approaches based on random fields relax the i.i.d. assumption by incorporating spatial constraints and enforcing adjacent pixels belonging to the same class. Thus, we apply CRF framework similar to [2] on the probabilistic outputs of RF described above. Let $G(S, E)$ be the adjacency graph of voxels, with each voxel corresponding to a node $s \in S$, and every edge $(s_i, s_j) \in E$ indicating the neighbourhood relationship between two voxels s_i and s_j. Then CRF minimizes an energy of the form [2]:

$$-log(P(\mathbf{c}|G;w)) = \sum_{s_i \in S} \psi(c_i|s_i) + w \sum_{(s_i,s_j)\in E} \Phi(c_i, c_j|s_i, s_j) \tag{1}$$

We directly use the probability outputs $P(c_i = c|s_i)$ from RF to define the node (ψ) and edge potentials (Φ):

$$\psi(c_i|s_i) = -log(P(c_i|s_i)) \tag{2}$$

$$\Phi(c_i, c_j|s_i, s_j) = \frac{|c_i - c_j|}{1+|P(c_i|s_i) - P(c_j|s_j)|} \tag{3}$$

Where $\mathbf{s}_i$ is the feature representation of pixel i and the weight w in Eq. 1 represents the trade-off between the spatial regularization (edge-potential) and the confidence in the classification (node-potential) which is learned based on the cross-validation in the training set. We use the public library for graph-optimization [1] for the label inference.

3. RESULTS

The dataset we used is provide by MICCAI 2013 challenges on Multimodal Brain Tumour Segmentation [3]. It consists 20 high-grade glioma subjects in four modalities(T_1, T_{1c}, T_2, T_{FLAIR}) which were annotated by 3∼4 radiologists. The complete tumour region is subdivided into 4 structures: necrosis, enhancing core, non-enhancing core and edema. All 3D volumes are linearly co-registered to the T_{1c} image, skull stripped and interploated to 1mm isotropic resolution.

We used 5-fold cross validation on the dataset and the segmentation results were uploaded to the online evaluation platform and were evaluated automatically using the online evaluation tool[1]. And the evaluation was performed for three different tumour subregions: complete tumour(including all four tumour structures), tumour core(including all tumour structures except edema) and enhancing core(enhancing core only). The following Dice score is applied to measure the overlap between the predicted segmentation results(P) and the ground truth(T):

$$Dice(P, T) = \frac{|P_1 \wedge T_1|}{(|P_1| + |T_1|)/2} \tag{4}$$

The Dice score of our proposed methods for complete, core and enhancing are 0.78, 0.66, 0.61 respectively, which ranks top 10 among over 30 participated groups.

4. FUTURE WORK

In order to build automatic brain tumour segmentation system, complete pixel-level ground truth provided by radiologists are desirable. However, it is expensive and time-consuming to obtain. Our future work will focus on developing learning-based segmentation methods involving partial annotations (e.g., bounding box, clicks, curves) while obtaining competitive segmentation results compared to the ones using complete annotations.

5. REFERENCES

[1] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(11):1222–1239, 2001.

[2] B. Fulkerson, A. Vedaldi, and S. Soatto. Class segmentation and object localization with superpixel neighborhoods. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 670–677. IEEE, 2009.

[3] B. Menze, M. Reyes, and K. Van Leemput. The multimodal brain tumorimage segmentation benchmark (brats). 2014.

[4] L. Nanni, A. Lumini, and S. Brahnam. Local binary patterns variants as texture descriptors for medical image analysis. *Artificial intelligence in medicine*, 49(2):117–125, 2010.

[5] T. Ojala, M. Pietikäinen, and T. Mäenpää. Gray scale and rotation invariant texture classification with local binary patterns. In *Computer Vision-ECCV 2000*, pages 404–420. Springer, 2000.

[6] G. Sharma, S. ul Hussain, and F. Jurie. Local higher-order statistics (LHS) for texture categorization and facial analysis. In *Computer Vision–ECCV 2012*, pages 1–12. Springer, 2012.

[7] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62(1-2):61–81, 2005.

[8] G. Zhao and M. Pietikainen. Dynamic texture recognition using local binary patterns with an application to facial expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(6):915–928, 2007.

[1] http://virtualskeleton.ch/

Automatic Transformation of Raw Clinical Data into Clean Data Using Decision Tree Learning Combining with String Similarity Algorithm

Jian Zhang
j.s.zhang@dundee.ac.uk

Abstract

It is challenging to conduct statistical analyses of complex scientific datasets. Whether a scientist or a statistician, find- ing the relationships within data is a time-consuming pro- cess. The process involves pre-processing the raw data, the selection of appropriate statistics, performing analysis and providing correct interpretations, among which, the data pre-processing is tedious and a particular time drain. In plenty of data provided for analysis, there is not a stan- dard for recording the information, and some errors either of spelling, typing or transmission. Thus, there will be many expressions for the same meaning in the data, but it will be impossible for analysis system to automatically deal with these inaccuracies. What is needed is an automatic method for transforming the clinical raw data into data which is possible to process. This paper will attempt to use Decision tree learning combining with the string similarity algorithm to achieve this data transformation.

Categories and Subject Descriptors

H.3.3 Information Search and Retrieval, I.2.6Learning

Keyword

Clinic Raw data, Decision Tree Learning, String Similarity Algorithm

1. Introduction

Cancer research has become a greatly data rich environment. And several data analysis package can be used for analyzing the data like SPSS (Statistical Product and Service Solu- tions)[1], Minitab. However, before analyzing those data, it needs to be pre-processed - 'cleaned' and classified - to translate the data from 'dirty' to 'clean', making it fit-for- purpose. This process is both time-consuming and tedious. For example, a specific binary variable may require only the entries 'Yes' or 'No' in a large data column. However, 'Yes' may have been entered onto a spreadsheet as Yes, yes, Y, y, yES, 1, or been misspelled as Yed, yef, y es, y e s (note the inappropriate use of spacing), etc.. Clearly there are an infinite number of possibilities of entering this 3-letter word incorrectly, and each of these entries is treated as separate entities by a computer program.For expample,

- yes : yes, Yed, yef, Y...
- no : No, N, not...
- null : don't know, waiting for lab

There is a common set of data. Obviously, we can see that there is only three possible answers in these data like the left part of colon of the above data - Yes, No or no an- swer.However, the errors will be existence including spelling, typing or transmission. Besides, there is no an standard for scientists to collect those data and record them.Finally, the data probably will be existence like the right part of colon of the above data when the statistians get the data for analysis. It is difficult for system to directly deal with these inaccuracies for statistical analysis.

Currently statisticians will manually change these entries into the uniform string or the uniform number, which al- low the system to do further analysis, so it will save lots of time for statisticians who are short of the research time if an automated method can be used instead of manual opera- tion [2]. According to the feature of many datasets and the necessity of data classification, Decision Tree Learning has been chosen to attempt automated the processing of the raw data firstly.

2.Decision tree learning

In general,training, validation and comparision itself with training data set will be the process of making a prediction in machine learning. Decision tree learning as one of the machine learning algorithm is a method for approximating discrete-valued functions [3], which is one of the most popular inductive algorithms. Decision trees are a method to get classify different values under different attributes. It can set the rules based on the past data to classify the new raw data and then get the value, which is the same at past when the manual data pre-processing was undertaken.

ID3 was proposed by Quinlan in 1986 [4]. It is a representa- tion of an decision tree algorithm and most of the decision tree algorithms are achieved based on improvements of it.

It adopts the divide-and-conquer strategy and uses infor- mation gain as the standard to choose the attribute at the different levels of the decision tree in order that it can collect the most information of the categories about the test records in the process of testing each non-leaf node.

And its demerits are that it can only process the discrete attribute, and sometimes it is not the best to process the attribute with lots of values.

As a result of some problems of ID3 algorithms in the practical application, Quinlan proposed the C4.5 algorithm [5], strictly speaking, which is just an improved algorithm of ID3. C4.5 algorithm inherits the advantages of the ID3 algorithm, and improving some aspects of ID3 algorithm. It use information gain ration to choose attribute so that it overcomes the weakness that the system tend to choose the attribute with more values when using information gain [6], and it adds prune in the process of constructing the tree and can process the incomplete data.

3.String Similarity Algorithm

3.1 String Similarity Algorithms

In this research, comparing the string similarity of the two strings is necessary when the entry is transformed from raw data to processable data by the system. String-based similarity has a long

history. Levenshtein(Lev) proposed the edit distance, which is widely used for string similarity through calculating the minimum number of insertions, substitutions and deletions between two strings [7]. Levenshtein distance will calculate once when add, delete or substitute have been done once during transformation. It provides numeric ap- proach for transformation.

Needleman and Wunsch(N-W) [8] extended the model to al- low contiguous sequences of mismatched characters, or gaps, in the alignment of two strings, and described a general dy- namic programming method for computing edit distance.

Jaro-winkler distance(J-W) find the approximating string matching by means of calculating the number of matching characters and the number of transpositions. And a prefix scale is added in this method as well [9].

3.2 Experiments and Results

As mentioned in the introduction section, the two-category data will be undertatken in this experiment. As a result of the feature of process of transformation, the data will be divided into exist data and unknown data to analysis.The raw clinic data is provided by statistians who worked in the cancer research. The testing data have 26% different data rather than the training dataset.

Table 1: Probability of transforming the test data to clean data using different string similarity algorithms

Decision Tree	Existing Entry (%)	Unknown Entry (%)	Total (%)
C4.5	100	27	81
Lev	100	58	89
N-W	97	73	89
J-W	100	73	91

Overall, the string similarity have a higher performance for the unkown data transformation. However, it is found that the time for runing the string similarity algorithm is 7-8 times than the algorithm C4.5 during the transformation.

4.Decision Tree Learning Combing With String Similarity Algorithm

According to the two previous experiments, the algorithms C4.5 have a high performance for existing data transformation and fast whilist the string similarity algorithm have a higher performance for unkown data but is a little slower. Thus, the combination of the two algorithms is worth ex- ploring based on their features and performance.

Because the string similarity algorithm have high performance for unknown entry, the large amount of correct transformation from the unknown data can be learned by the decision tree learning to build a new and powerful decision tree for further data transformation.

The string similarity algorithm Levenshtein distance, Needleman-Wunsch, Jaro-winkler distance and a improved string sim- ilarity algorithm based on Needleman-Wunsch (NW-Len) will be undertaken to process the incorrect prediction in the following experiment combined with the algorithm C4.5. The Needleman-Wunsch algorithm is to calculate longest common substring (LCS). We can consider the effect of the length of the string in the N-W algorithm to improve it.

This experiment will undertake the same testing data set as previous experiment.

Table 2: Probability of correct prediction for differ- ent string similarity algorithms combined with deci- sion tree learning

Decision Tree	Existing Entry(%)	Unknown Entry(%)	Total (%)
C4.5	100	27	81
C4.5 + Lev	100	58	89
C4.5 + N-W	100	65	91
C4.5 + J-W	100	73	93
C4.5 + N-W + Len	100	73	93

To sum up, the results for transforming the raw clinic data to processable data have improved a lot after combination the decision tree learning algorithm with string similarity al- gorithms, especially with jaro-winkler distance and modified Needleman-Wunsch algorithm.

5.Conclusion and Future Work

Through the experiment, the decision tree combining with the string similarity algorithm can automatically transform- ing the clinic raw data into the clean data, to some ex- tent, which reduce the manual operation partially. In the next step, the correct prediction for unknown entries will be learned by the system and updated the training dataset to enhance the cleaning ability.

6.Reference

[1] A.Acock, (2005). "SAS, Stata, SPSS: A Comparison". Journal of Marriage and Family [0022-2445] Acock yr:2005 vol:67 iss:4 pg:1093 -1095.

[2] M.Dalal, and M.Zaveri,(2011). "Automatic Text Classification: A Technical Review". International Journal of Computer Applications 28(2):37-40. Published by Foundation of Computer Science, New York, USA.

[3] T. Michell, (2005). Machine Learning. Page 52.

[4] J.R.Quinlan, (1986). Induction of Decision Trees. Machine Learning. 1(1): 81-106.

[5] J.R. Quinlan(1993). C4.5 Programs for machine learning.San Mateo, CA: Morgan Kauffman.

[6] J.R. Quinlan (1996).Improved Use of Continuous attributes inC4.5. In Journal of Artificial Intelligence Research.18(4):77-90, .

[7] L.I. Levenshtein, (1996). "Binary codes capable of correcting deletions, insertions and reversals". Soviet Physics Doklady, 10, 707-710.

[8] S.B.Needleman, , and C.D.Wunsch, (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of Molecular Biology, 48, 443-453.

[9] M. A.Jaro (1989). "Advances in record linkage methodology as applied to the 1985 census of Tampa Florida". Journal of the American Statistical Association 84 (406), pp.414–20.

[10] J. Zhang, K. Petrie and T. Yu (2013). "Automatic Transformation of Raw Clinical Data into Clean Data Using Decision Tree learning". ARW 2013.

Towards real-time pre-processing of Mass Spectrometer data using a scalable cluster

Chris Hillman
School of Computing
University of Dundee

chillman@dundee.ac.uk

ABSTRACT

The human genome project was one of the largest and most well known scientific endeavors of recent times. This project characterised the entire set of genes found in human DNA. Following on from genomics, proteomics is the study of proteins, which are the products of genes found in cells. In a typical proteomics experiment, an instrument called a mass spectrometer is used to identify proteins and measure their quantities. This produces large data files that currently take many hours to process. This research is concerned with reducing the time needed to process data files to near real-time by designing a novel algorithm that can run in a parallel fashion and evaluating its performance characteristics across various cluster computing platforms that have not previously been used for this purpose.

1. INTRODUCTION

Proteomics can be defined as the large-scale study of protein properties, such as expression levels, modifications and interactions with other proteins. By studying proteins and their properties, it is possible to gain a deeper understanding of how proteins should function in healthy cells compared with diseased cells. Full proteins are in most cases too large to be measured intact by mass spectrometers, therefore a process of fragmentation breaks the proteins into constituent parts called peptides. It is these peptides that are analysed and identified by software in a data processing step after the mass spectrometer has processed a cell sample. This identification and quantification allows Life Scientists to research how different environments and compounds affect the protein expression in cells.

MapReduce is a framework used to distribute processing across clusters of computers that was first proposed by engineers working for Google [1]. It is designed to allow programming code to run on clusters of commodity hardware and abstracts the complexity of allocating, monitoring and running many parallel tasks away from the programmer. Although the basic methods are simple to understand and implement, a process will need to be re-engineered to fit into the strict system of Map and Reduce tasks.

To answer the question "Can pre-processing of Mass Spectrometer data approach real-time using a MapReduce process on a horizontally scalable cluster", parallel processing is investigated on several platforms offering different types of storage or computation frameworks.

2. RELATED WORK

A review of the current literature reveals that there are currently very few references to the use of MapReduce style parallelization for proteomics data processing. Where it is referenced it is in relation to searches for matching peak-lists with databases to identify proteins [5][6] and not the algorithms required for parallel processing or the suitability of the various types of cluster available.

3. GOALS

In this research work we investigate methods of reducing the elapsed time to process the output from Mass Spectrometers created in the course of proteomics experiments. Several cluster-based technologies that allow parallel computation are to be included.

- Hadoop – Distributed file system
- Cassandra – NoSQL database
- Teradata Aster – MPP SQL database
- Spark – In-memory distributed processing
- Storm – distributed stream processing

Benefits and drawbacks of each system are to be evaluated from a number of different points of view including ease of use and maintenance, ability to integrate with other systems into a proteomics processing pipeline, disk space requirements above and beyond that of the original data file and overall time to process the data.

4. ALGORITHM DESIGN

Much has been published regarding the process of detecting peptides in the spectral data output from mass spectrometers. The algorithm implemented by the software "MaxQuant" [2] has been chosen as the base from which to develop a novel algorithm that can be executed in parallel on a shared-nothing cluster. MaxQuant is used in many Life Sciences laboratories and supports the output from the Thermo-Fischer mass spectrometers that are used at the University of Dundee.

5. CURRENT WORK

The standard data format for the output from proteomics experiments is called mzML which is an XML file format [3] consisting of a header section containing information about the environment in which the experiment was performed, plus a section containing information about each scan performed. The mzML files are large often in excess of 5Gb and as with all XML files are difficult to distribute across a cluster and process in parallel. In order to process these files efficiently a new file format (.scmi) has been designed that contains only the information relevant to peptide identification. The parser that performs this conversion also copies the data from the source file system to the target cluster. The remainder of the information in the mzML file is still important to the complete understanding of the outcome of an experiment. A hybrid load architecture was discussed in previous work by Hillman [4] with metadata stored in

a relational database and the scan information in a distributed file system.

The data processing is broken down into several steps as described in detail below. The first step is to identify peaks in the spectral signal within individual scans, these peaks occurring along the mass dimension and are called 2D peaks. The second step is to link the 2D peaks across the scans in the time dimension, these are know as 3D peaks.

A Map task is used to process the 2D peaks and this is where the scmi file format greatly simplifies processing over the original XML, as each record in the file up to a carriage return represents a single scan from the mass spectrometer. Each scan can be processed completely independently of the others, and therefore the 2D peak-picking process can be made to operate in an embarrassingly parallel fashion. The Map task decodes the Base64 binary arrays storing the mass-to-charge and intensity data and loads them into in-memory arrays of java objects. Each peak is detected by using a slope detection algorithm with overlapping peaks introducing some degree of complexity here. In addition, noise in the signal and the way the instrument measures the peptides mean that the peaks can be shifted slightly; however, it is possible to compensate for this by calculating a theoretical peak by fitting an ideal Gaussian curve to the data. Because of the presence of carbon isotopes, it is necessary to identify peaks within an isotopic envelope that represent the same peptide. This can be done in the same Map task as the 2D peak picking. As peaks are matched within an isotopic window the charge can be calculated which allows us to deduce the final mass of the peptide.

The 2D peaks identified so far are indicators of the presence of a peptide. The mass spectrometer carries out multiple scans over time and any one peptide will take several seconds to pass through the machine. To match 2D peaks across time, the data will need to be re-distributed around the cluster and written out to disk. As the peaks detected are clustered into compact groups across the mass range, a custom partitioner function is required to ensure that correct data distribution takes place. Efficient distribution of processing and avoidance of data skew is key to a performant parallel process. Once this has been done a Reduce task is used to build the 3D peaks across time. An algorithm has been developed for this taking into account certain biological rules such as peaks occurring within a mass window (chosen to be 7 ppm in this work) and the complexity of matching peaks which have missing segments and noise in the signal.

The 3D peaks also require a de-isotoping step, for this process the techniques for 2D isotopic peak detection can be reused with a modification to ensure peaks with similar profiles are matched. At this point we have calculated the mass and intensity of molecules and can either output the results or move on to further processing such as detection of stable isotope labeling by amino acids in cell culture (SILAC) and/or database search.

In this way, the 2D and 3D peak-picking process fits well into the MapReduce programming framework, and where data needs to be redistributed, the dataset has been greatly reduced by the initial peak picking in the Map Task.

6. EVALUATION

For development and testing purposes, a virtual cluster has been constructed on a custom built server using an 8-core AMD processor with 32Gb Ram and 5 individual SSD Drives (SSD Drives were found to produce far more consistent run times than spinning disks)

Currently Virtual Clusters have been created to run Hadoop 1.3, Teradata Aster 6.1, Cassandra 2.07, Hadoop 2.4 and Spark 1.2

The output from the parallel algorithm coded in the mapreduce framework using Java has been checked against 3 current methods of processing this data.

Proteowizard, a common command line tool that is freely available and capable of peak picking in the mz domain (2D picking).

MaxQuant a GUI tool that can produce output peaks in the mz (2D) and also time (3D) domains as well as many other features not discussed here.

The Spectracus system developed in the Lamond Lab at the University of Dundee[7].

In each case the results from the mapreduce code accurately reflect that produced from the other software.

7. FUTURE WORK

Following on from the validation of the output from the parallel algorithm, research into the most efficient platform to perform the processing can begin. Note that this is an evaluation of a parallel algorithm and the MapReduce framework; hardware acceleration technologies such as GPU or FPGA are therefore not included. The platforms to be evaluated are mentioned above and include batch as well as stream processing architectures. As each claim to be linearly scalable, once experimentation into the elapsed time for the processing on different platforms is complete it will be possible to calculate the system requirements to complete the steps in a given time frame that fits within the definition of real-time in this context.

8. REFERENCES

[1] Dean, S. Jeffrey;Ghemawat 2004. MapReduce: Simplified Data Processing on Large Clusters. *OSDI '04: 6th Symposium on Operating Systems Design and Implementation*. (2004).

[2] Jurgen Cox and Matthias Mann. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantifica- tion. *Nature Biotechnology*, 26(12):1367–1372, November 2008.

[3] Deutsch, F. E.W.; Martens L; Binz P.;Kessner D.;Chambers M; Sturm M;Levander 2009. mzML: Mass Spectrometry Markup Language. (2009).

[4] Hillman, C. 2012. *Investigation of the Extraction, Transformation and Loading of Mass Spectrometer RAW Files*. University of Dundee, January 2012.

[5] Lewis, J. Steven;Csordas Attila;Killcoyne Sarah;Hermjakob Henning;Hoopmann Michael R;Moritz Robert L;Deutsch Eric W;Boyle 2012. Hydra: a scalable proteomic search engine which utilizes the Hadoop distributed computing framework. *Bioinformatics*. 13, (2012).

[6] Mohammed, M. Yassene;Mostovenko Ekaterina;Henneman Alex A.;Marissen Rob J.;Deelder Andre M.;Palmblad 2012. Cloud Parallel Processing of Tandem Mass Spectrometry Based Proteomics Data. *Journal of Proteome Research*. 11, (2012), 5101 ? 5108.

[7] Ahmad, Y; Rasheed 2014 Protein Finger Printing in Teradata. Implementation Documentation, University of Dundee

SpaceWire-D Networks

David Gibson
Space Technology Centre
University of Dundee
Dundee, DD1 4HN
dzgibson@dundee.ac.uk

Steve Parkes
Space Technology Centre
University of Dundee
Dundee, DD1 4HN
smparkes@dundee.ac.uk

ABSTRACT

SpaceWire-D is a deterministic extension to the SpaceWire standard designed to allow a single SpaceWire network to be used for command and control applications and payload data handling simultaneously. It does this by using time-division multiplexing to slice network time into a series of slots in which data-transfers are allocated allowing real-time constraints to be satisfied.

This paper provides an overview of SpaceWire-D and then describes the prototype system designed to demonstrate, test and verify the standard.

1. INTRODUCTION

SpaceWire is an onboard data-handling network used on spacecraft to connect instruments, mass-memory, onboard computers, telemetry and other subsystems [1]. SpaceWire enabled devices are connected by full-duplex data-links, providing bi-directional data-flow at transmission rates of between 2Mbit/s and 200Mbit/s.

The routing mechanism used by SpaceWire routers, wormhole routing, introduces variance in packet transmission time because of a problem called blocking which is when a packet is delayed due to a required SpaceWire link being in use by another packet. If network traffic is uncontrolled, this can mean that important traffic such as command and control packets, which have hard real-time constraints, miss their deadlines.

The following sections briefly describe the SpaceWire-D standard and how it overcomes the problems of non-determinism in regular SpaceWire networks. Following that, the prototype system used to demonstrate, test and verify the standard is described [2] [3].

2. SPACEWIRE-D

SpaceWire-D is a deterministic extension to the SpaceWire standard being designed by the University of Dundee for the European Space Agency [4]. To provide a deterministic capability, network time is divided into a series of time-slots in which RMAP [5] transactions are executed with the constraint that any transactions must finish execution before the next time-slot begins.

To synchronise time across the network, SpaceWire time-codes are broadcast by a time-code master at a rate of up to 1000Hz, giving a minimum time-slot duration of 1ms. The current time-slot is identified by a 6-bit time-value within each time-code, allowing for a schedule with 64 time-slots.

Within each time-slot, data transfers are encapsulated within a group of one or more RMAP transactions. Each transaction is an interaction between an initiator and a target, in which the initiator reads or writes to memory in the target. The initiator sends a command, the target authorizes or rejects the command and then the target sends a reply to the initiator with the status of the command and any data if relevant.

RMAP transactions are grouped into a virtual bus system, where a virtual bus consists of an initiator, one or more targets and the links in the paths between the initiator and targets. There may be more than one initiator in a SpaceWire-D network, so the schedules for each initiator have to be designed so that no virtual buses execute in the same slot if there is a chance that any of their transactions can collide and cause blocking.

The SpaceWire-D standard provides services to open, load, execute and close four types of virtual bus.

2.1 Static Bus

The static bus is allocated a single time-slot and takes a group of transactions which is executed in the same slot in every iteration of the schedule. This is the most deterministic of the virtual buses as it executes in a predictable manner.

For periodic, repeating transaction groups, this is a suitable bus.

2.2 Dynamic Bus

The dynamic bus takes a group of transactions and executes it in one of multiple time-slots. This is a less deterministic bus than the static bus as the transaction group may execute in many time-slots.

For aperiodic transaction groups with deadlines, this is a suitable bus as the time-slots can be placed at suitable intervals allowing for a deadline to be met no matter when the transaction group is loaded.

2.3 Asynchronous Bus

This asynchronous bus takes single transactions and inserts them into a prioritised queue. In the time-slot preceding one allocated to the bus, a subset of transactions is pulled from the head of the queue for execution in the next time-slot. Network information is stored in each initiator so it can accurately calculate the execution time of transactions so that the subset of transactions does not exceed the duration of a time-slot and interfere with the following time-slot.

For data transfers with no real-time constraints such as payload data, this is a suitable bus.

2.4 Packet Bus

The packet bus allows a channel to be created between an initiator and a target through which prioritised packets are transferred. The transfer of a packet takes place in three stages:

1. An initiator reads the status of a target's packet channel to see if it has a packet ready to read or a buffer ready to receive a packet.
2. One or more segments of the packet are transferred between the initiator and target via read or write RMAP transactions.
3. The initiator writes an end-of-packet (EOP) transaction to tell the target that the operation is complete.

In the time-slot preceding one allocated to the bus, a transaction group is prepared consisting of as many channel status reads, packet segment transfers and EOP transactions as will fit in a time-slot for execution in the next time-slot.

3. DEMONSTRATION SYSTEM

The system used to demonstrate, test and verify the SpaceWire-D standard consists of initiators, targets, routers, and analysis devices as shown in Figure 1.

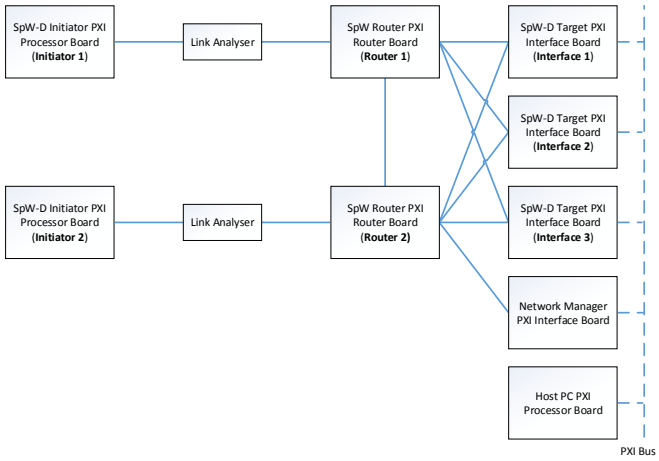


Figure 1: Demonstration System Topology

3.1 Initiators

The initiators are LEON2-FT processor boards with extensive SpaceWire support including an 8-port router and hardware RMAP engines. Each initiator runs RTEMS based embedded software [6] containing the SpaceWire-D protocol layer and a scripting engine to execute test procedures.

3.2 Targets

There are three interface boards each containing four targets and a demultiplexer to filter traffic to a target identified by a logical address.

3.3 Network Manager

The network manager writes configuration information to the initiators, targets and routers and is used to store status information and errors from the initiators.

3.4 Host PC

The host PC is used to create schedules, create virtual buses, configure devices and create scripted test procedures.

3.5 Scripting

A simple scripting language has been designed to allow the creation of time-triggered test procedures. The host PC software includes a parser for the language that compiles the high level commands into low level instructions and operands which are then written to memory in the initiator.

The scripting language includes commands to create transactions, transaction groups and packet bus operations and load virtual buses at specific points in time synchronised with the arrival of time-codes. Figure 2 shows an example of a simple script to create three transactions, a transaction group and load a static bus.

```
transaction(0, 40, read, 32, 0x8101F000, 32, 0x8101D000)
transaction(1, 40, write, 32, 0x8101F020, 64, 0x8101D020)
transaction(2, 40, read, 32, 0x8101F006, 128, 0x8101D060)
transaction_group(0, 0, 1, 2)
slot(63):
    load_sbus(1, 0, 0, 0, 0, repeat)
```

Figure 2: Example Script

4. REFERENCES

[1] European Cooperation for Space Standardisation. ECSS-E-ST-50-12C: SpaceWire, links, nodes, routers and networks. Available from: http://www.ecss.nl/, 2008.

[2] D. Gibson, S. Parkes, C. McClements, S. Mills, D. Paterson, "SpaceWire-D on the Castor Spaceflight Processor", International SpaceWire Conference, Athens, Greece, September 2014.

[3] S. Parkes, D. Gibson, A. Ferrer, "Experimental Results for SpaceWire-D", Data Systems in Aerospace Conference, Barcelona, Spain, May 2015.

[4] S. Parkes, A. Ferrer, A. Gonzalez, C. McClements, SpaceWire-D. Available from: http://spacewire.esa.int/WG/SpaceWire/, February 2014, Draft D.

[5] European Cooperation for Space Standardisation. ECSS-E-ST-50-52C: SpaceWire – Remote memory access protocol. Available from: http://www.ecss.nl/, February 2010.

[6] D. Paterson, D. Gibson, S. Parkes, "An RTEMS Port for the AT6981 SpaceWire-Enabled Processor: Features and Performance", International SpaceWire Conference, Athens, Greece, September 2014 ".

Reasoning about Complex Human-Object Interactions in Daily Living Activities

Tianjun Huang and Stephen McKenna
CVIP, School of Computing
t.huang@dundee.ac.uk

ABSTRACT

Human activity recognition is one of the important research areas in computer vision. In recent years, understanding fine-grained activities (especially activities which contain human-object interactions) in daily living has gained increasing attention due to its wide applications. However, the transformations of objects (i.e., the consequences of actions), which are strong connections between human motion and objects, have not been systematically studied in fine-grained activity recognition. In this project, we would like to explore whether it is possible to build a robust system for understanding fine-grained activities with the use of location, transformation and category information of objects and information about the motion of the person. Thus, we would not only infer what the person is doing, but also infer the current states of objects. Specifically, we mainly focus on the analysis of cooking activities in this project.

1. INTRODUCTION

As a promising and basic direction in computer vision, human activity recognition has been extensively researched in recent years. It has wide applications such as human-computer interaction, cognitive situational support, surveillance, motion capture and animation, unconstrained video search and so forth. Even though lots of attention has been paid to human activity recognition, it is still a challenging problem: heavy work for manual collection of training samples, large variations in appearance, and the definitions of different action vocabularies. Fine-grained activity recognition in daily living (e.g., cooking activities) is a challenging area in human activity recognition and is attracting increasing attention recently because of its potential applications in assisted living and child/elderly care.

Fine-grained activities in daily living usually contain a large number of interactions between human and objects. Recent works on fine-grained activity recognition are trying to extract local features for the human and associated objects to represent and classify the activity [7, 3]. However, all these methods ignored an important cue in the activities: the consequences of actions, i.e., the transformations of objects. Actions are goal oriented. The transformations which happen to the objects indicate the information of action goals and directly connect with their corresponding action classes. Motivated by the work of Yang et al. [6], we are going to utilize the information about how objects being manipulated are transforming to help solve the fine-grained human activity recognition problem. Moreover, the transformation information also can be used to infer the categories of objects: e.g., ingredients can be divided while tools cannot. This characteristic is useful in the object analysis.

2. HUMAN MOTION ANALYSIS

We separate this project into three steps: human motion analysis, non-human object analysis, and activity recognition. The transformations of objects are included in the object analysis part. In this section, we introduce our current results for the first step: human motion analysis. The 50 Salads dataset from Stein and McKenna [4] is used as the experiment dataset.

There are two phases for the human motion analysis in this project. Phase one is to estimate the foreground objects above the work surface. The second phase is built upon the result from phase one. Superpixel segmentation, an optical flow method and the location of the edge of the work surface are used to sample the possible regions of hands and arms. Features extracted from the possible regions are encoded to build a model for hands and arms. This final model was applied on the foreground regions to detect whether they are hands or arms.

2.1 Foreground-Background Estimation

From the top-down view in the kitchen worktop, the work surface usually occupies most of the camera view. Since the depth images are provided in the dataset, a 3D space can be created by setting the depth as the z coordinate. Then, we adopted the Random Sample Consensus (RANSAC) algorithm [2] to find the plane which represents the work surface in the 3D space. After the work surface plane is found in the frame, the probability of the point to be on the foreground F (i.e., $P(F|Depth)$) can be directly measured based on the signed distance between this point and the work surface plane. The logistic sigmoid function was used to map the distances to probabilities.

In order to leverage colour information to improve foreground-background estimation, we randomly sampled data points which have high $P(F|Depth)$ as the foreground training set, and data points which have low $P(F|Depth)$ as the background training set. Two Gaussian mixture models are built on the foreground training set and background training set respectively by using the Expectation Maximisation (EM) algorithm. Then, by using the GMM equation, the $P(Colour|B)$ and $P(Colour|F)$ are easily computed for each point. Assuming the prior probabilities of the point to be foreground or background are the same: $P(F) = P(B) = 0.5$, the probability of the point to be the

foreground given its colour ($P(F|Colour)$) can be computed by Bayes' theorem. The $P(F|Depth)$ and $P(F|Colour)$ are linearly combined by assigning different weightings for them to generate the final foreground probability map, as shown in Equation (1). The weighting parameter w is set to 0.7 since the depth-based probability map seems more robust than colour-based probability map.

$$P(F|Depth, Colour) = w \times P(F|Depth) + (1-w) \times P(F|Colour) \quad (1)$$

2.2 Hands and Arms

In order to capture the human motion information for the activity recognition, the movements of hands and arms above the work surface have to be detected. Firstly, we applied the Superpixels Extracted via Energy-driven Sampling (SEEDS) [5] method to over-segment the video frames into superpixels. The superpixels whose $P(F|Depth, Colour)$ are over 0.5 are labelled as foreground superpixels. Foreground superpixels are grouped into the same region if they connect with each other. A state-of-the-art optical flow algorithm [1] is used to compute the velocities for superpixels. In the foreground regions which cross the edge of the work surface, we extracted the superpixels whose velocity magnitudes exceed a threshold as the positive training set to create a colour model for hands and arms. The superpixels whose foreground probabilities are over 0.5 and not in the positive training set are used as negative training set.

The entire training set is extracted from the first 10 frames which have the positive superpixels mentioned above (experiments show that 10 frames are enough to build the model). All the colour descriptors of pixels are quantized to 128 clusters by using k-means method. Based on the 128 clusters, we built two histograms for the positive training set and negative training set respectively and call them human model (representing hands and arms) and non-human model (representing other objects). For each new superpixel, we first compute the colour descriptors of its pixels, and then quantize them to the same 128 clusters to build a histogram representation. This histogram is intersected with the human model and non-human model, and the superpixel will be labelled to the class which has a higher intersection score. The human and non-human models are updated every 10 frames by adding the histograms of these new superpixels to their corresponding classes. We applied the morphological opening operation on the final result to smooth the segmentation regions.

3. RESULTS

We randomly selected three subject videos from the 50 Salads dataset and annotated the hand and arm regions above the work surface for 20 frames in each video as ground truth. The 20 frames were selected from the start frame of the first activity in each video with an interval of five frames. We calculated the precision and recall as the measurements for each subject video. Table 1 shows the results.

The segmentation precision is relatively low compared to the recall. One of the reasons is that objects whose colours are similar to the hands and arms were also segmented since currently the human model is only based on colour information. The other problem is that objects held in the hands

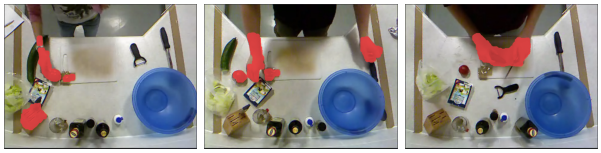


Figure 1: From left to right are the examples of hands and arms segmentation results for the first, second and third subject videos respectively.

are easily segmented with hands and arms. This issue will be solved by combining with the object analysis. Figure 1 shows examples of the segmentations in the three videos.

	First subject	Second subject	Third subject	Mean value
Precision	0.39	0.67	0.79	0.62
Recall	0.85	0.81	0.80	0.82

Table 1: Precision and recall for each subject and their mean values.

4. FUTURE PLANS

Currently, the human motion analysis result still needs to be improved. I am now trying to use graph cuts method to generate multiple hypotheses for the regions of hands and arms and other foreground objects. This will also benefit the object analysis in the future. A Bayesian network can be used as the model to represent the activities and solve the final activity recognition problem in the future.

5. REFERENCES

[1] L. Bao, Q. Yang, and H. Jin. Fast edge-preserving patchmatch for large displacement optical flow. In *Computer Vision and Pattern Recognition*, 2014.

[2] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[3] B. Ni, V. R. Paramathayalan, and P. Moulin. Multiple granularity analysis for fine-grained action detection. In *Computer Vision and Pattern Recognition*, 2014.

[4] S. Stein and S. J. McKenna. Combining embedded accelerometers with computer vision for recognizing food preparation activities. In *ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2013.

[5] M. Van den Bergh, X. Boix, G. Roig, B. de Capitani, and L. Van Gool. SEEDS: Superpixels extracted via energy-driven sampling. In *European Conference on Computer Vision*, 2012.

[6] Y. Yang, C. Fermuller, and Y. Aloimonos. Detection of manipulation action consequences (MAC). In *Computer Vision and Pattern Recognition*, 2013.

[7] Y. Zhou, B. Ni, S. Yan, P. Moulin, and Q. Tian. Pipelining localized semantic features for fine-grained action recognition. In *European Conference on Computer Vision*, 2014.

Binary classification of blood vessels in arteries and veins in ultra-wide field scanning laser ophthalmoscopy

Enrico Pellegrini
University of Dundee, Dundee, UK
e.z.pellegrini@dundee.ac.uk

ABSTRACT

Features of the retinal vasculature, such as the Artery Vein Ratio, the vessel tortuosity and vasculature fractal dimension are candidate biomarkers for systemic diseases like hypertension, arteriosclerosis and cardiovascular disease [1]. In population studies, it is important to automate the process of measuring these retinal biomarkers in order to ensure an objective quantification of their value and to avoid the time cost of having a trained observer manually annotating large numbers of images to produce such measures. Reliable retinal biomarker measures are based on an accurate segmentation of the blood vessels, a correct estimation of the blood vessel widths and a correct classification of vessels into arterioles and venules. In this paper we focus on the latter task in the particular case of ultra-wide field of view (UWFoV) scanning laser ophthalmoscope (SLO) retinal images [2]. A local classification of centreline pixels sampled along vessel segment paths is followed by a step of majority voting on each segment. The results are further improved by modelling this binary classification problem as a graph cut problem. The technique has been tested on three different dataset and preliminary results are shown.

Keywords

Retinal image, ultra-wide field of view, scanning laser ophthalmoscope, artery/vein classification, graph cut, edge classification, biomarker.

1. INTRODUCTION

Changes in morphological features associated with retinal blood vessels such as widths, tortuosity and branching angles are considered early biomarkers of systemic diseases such as stroke, hypertension, arteriosclerosis, myocardial infarction and cardiovascular disease. Devices such as the scanning laser ophthalmoscope (SLO), with an Ultra-Wide Field of View (UWFoV) of approximately 200° (compared to 30-60° with a fundus camera), can capture in a single image a larger part of the retina, allowing more extensive analysis of the associated vasculature.

The automatic quantification of morphometric vascular features is crucial, especially in population studies, where the manual annotation of large number of images would be an extremely time-consuming process. Prior to this quantification, the first step is an accurate segmentation of blood vessels. In our previous work [3], we developed a supervised segmentation technique based on a neural network classifier that achieved an accuracy value equal to 0.965 (Accuracy value of 2nd human observer = 0.970). The algorithm also proved to be the most reliable in terms of estimation of vessel widths among the state-of-the-art automated techniques used for comparison. The following step is the binary classification of blood vessels into arterioles and venules. In this paper we present a framework that combines a preliminary local classification at pixel level, a majority voting step at vessel-segment level and a separation of artery and vein trees at a global level, solved as a graph cut problem. Preliminary results are evaluated on three different datasets (*Dundee*, *WIDE* and *Tascforce*).

2. METHODOLOGY

2.1 Pixel-level classification

A set of 76 features was extracted from a set of equally spaced pixels along the centreline of the refined [4] vessels, according to the guidelines in [5], taking into consideration the substantial differences of SLO images with respect to conventional fundus camera images: the blue channel, hence the saturation, of SLO images is entirely set to 0 and the scale of vessel widths is considerably smaller due to the different image resolution. Twenty additional features representing the vessel cross-sectional profile at the pixel in question in the red and green channels of the RGB image were also added to form the final 96-feature-long feature vector. The results of a leave-one-image-out cross-validation on the three datasets using a Linear Bayes Classifier are reported in Table 1. Values of accuracy of the same analysis performed using linear and non-linear SVM's, K-Nearest Neighbour and Parzen Classifier were lower than those shown in Table 1, therefore omitted.

2.2 Vessel segment-level classification

The vessel labels were computed by thresholding the average of the artery/vein likelihood, results of the previous classification step, of all the pixels belonging to the vessel segment. Results reported in Table 1.

2.3 Global-level classification

A graph was created to represent the entire vasculature. In the graph, a node represented a vessel segment and an edge between two nodes represented a possible link between vessel segments connected to the same bifurcation or crossing point. We defined as a true link, a link between two nodes that belonged to the same artery/vein class, and as a false link, a link between a node representing an artery segment and a node representing a vein segment. Two additional nodes, both fully connected to the original nodes, were added to the graph. A representation of part of the retinal vasculature modelled as a graph can be seen in Figure 2. The problem of artery and vein trees separation was then modelled as a graph cut problem [6], creating two additional fictional nodes as source and sink.

The weight of the edges among the original nodes was defined as the likelihood of the link between the two vessel segments of being a true link. This likelihood was the output of the binary classification of links based on a set of features depending on the morphology of the vasculature at the bifurcation or crossing point and derived mainly by the number of segments attached to the point, their orientations and their widths (details omitted). The weight of the edges between the original nodes and the source node was defined as the likelihood of each vessel segment of being an artery as obtained at point 2.2. Finally, the weight of the edges between the original nodes and the sink node was defined as the likelihood of each vessel segment of being a vein.

Results of the artery vein classification after this final step are reported in Table 1.

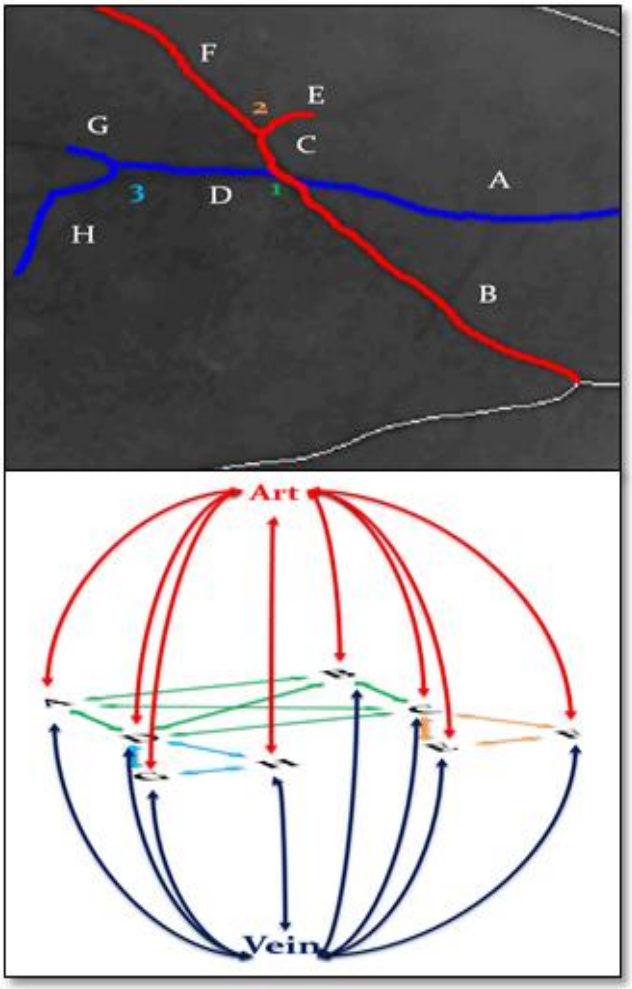


Figure 1. Representation of a small portion of the retinal vasculature (top) as a graph (botton). "Art" and "Vein" are the two additional nodes that represent source and sink of our graph cut problem.

3. EVALUATION

3.1 Materials

Three sets of images were used to test the results of this study. The first set, *Dundee*, comprising 40 UWFoV images that underwent a stereographic projection [7]. Six arterioles and six venules in zone B [8] were manually labelled by two trained observers achieving perfect agreement. . Fifteen vessels from this set were not properly segmented, then discarded from further analysis. The lack of a comprehensive ground truth for this set made it unsuitable for the last stage of our framework.

The second set, *WIDE* [9], comprising 15 UWFoW, non-reprojected, scaled down by a factor of two, images. In this case, the entire retinal vasculature was annotated by two trained observers. The images were rescaled to their original size before further processing. The third set, *Tascforce*, comprising 473 reprojected, full-resolution, UWFoW images were the entire retinal vasculature was annotated by only one observer.

3.2 Results

In Table 1, the classification accuracy of single pixels after every step of processing is shown.

Table 1. A/V classification accuracy after every step of the framework.

Dataset	Dundee	WIDE	Tascforce
Number of pixels	9490	19472	738973
Pixel-level class Acc	0.804	0.722	0.769
Segment-level class Acc	0.889	0.774	0.862
Global-level class Acc	n.a.	0.792	0.888

4. DISCUSSION

In this paper we presented a supervised technique for the classification of blood vessels in arterioles and venules in UWFoV retinal fundus images consisting of three stages: a local binary classification at pixel-level, a majority voting at vessel segment-level and a graph cut separation of vessel trees at global-level. The results are promising but the features used for the first stage of classification seems to be heavily influenced by the resolution and the quality of the images (see *WIDE* dataset results in Table 1). Additional improvement of the accuracy could be achieved by constraining the graph cut algorithm to take into consideration only the clinically plausible configurations of junctions and crossing points in the retinal vasculature.

5. REFERENCES

[1] J. J. Wang, G. Liew et al., "Retinal vascular calibre and the risk of coronary heart disease-related death," Heart 92, 1583–1587 (2006).

[2] Jones, W. and Karamchandani, G. Panoramic Ophthalmoscopy, Optomap Images and Interpretation, SLACK Incorporated (2007).

[3] Pellegrini, E. Robertson, G. et al., 'Blood vessel segmentation and width estimation in ultra-wide field scanning laser ophthalmoscopy' Biomedical Optics Express, vol 5, no. 12, pp. 4329-4337, (2014).

[4] Dashtbozorg, B. Mendonça, A. M. Et al. 'An automatic graph-based approach for artery/vein classification in retinal images', IEEE Trans Image Process. 2014 Mar;23(3):1073-83.

[5] Zamperini, A. Giachetti, A. et al., 'Effective features for artery-vein classification in digital fundus images', Proceedings of CBMS (2012).

[6] Boykov, Y. Veksler, O., et al., 'Fast Approximate Energy Minimization via Graph Cuts.', IEEE PAMI, vol. 23, no. 11, pp. 1222-1239, 2001.

[7] Croft, D. E. van Hemert, J. et al., "Precise montaging and metric quantification of retinal surface area from ultra-widefield fundus photography and fluorescein angiography," Ophthalmic Surg Lasers Imaging Retina 45, 312–317, (2014).

[8] M. M. Fraz, P. Remagnino et al., "Blood vessel segmentation methodologies in retinal images--a survey," Comput Methods Programs Biomed 108, 407–433 (2012).

[9] Estrada, R. Tomasi, C. et al., "Tree Topology Estimation", (IN PRESS), IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), (2015).

Machine Learning for Theory Exploration

Chris Warburton
University of Dundee
cmwarburton@dundee.ac.uk

Ekaterina Komendantskaya
University of Dundee
katya@dundee.ac.uk

ABSTRACT

Theory Exploration (TE) aims to automate the tedious yet necessary task of verifying Mathematicians' and programmers' work, but the blind search used by existing approaches limits them to small examples. Meanwhile, huge repositories of formal knowledge are being routinely data-mined for structure and correlation. We provide a method for guiding TE using these abundant statistics, and assess whether this hybrid approach is feasible for tackling problems of a realistic size.

1. INTRODUCTION

Small mistakes made in Mathematics and programming can cause large problems in the real world. Computer verification can reduce this risk by checking our reasoning step-by-step; but traditional software's inability to follow "obvious" arguments without explicit guidance often makes it impractical. The recent *theory exploration* approach[1] tackles this by generating a database of facts about a user-provided "theory" (eg. a software library). This database can either serve as "background knowledge" for a traditional verification tool, to help it follow more coarse-grained proof steps; or it can be queried directly to discover interesting facts about the theory.

Existing TE systems, such as QuickSpec[3] and HipSpec[2], rely on undirected, brute-force search to generate their databases. Although complete, these algorithms' exponential complexity limits their scalability to small theories with only a handful of definitions. To be practical, the technique must be able to work with theories spanning thousands of definitions, without relying on expert users to cherry-pick a sub-set.

Machine Learning (ML) offers many scalable techniques for studying large theories, such as *premise selection*[5]: choosing lemmas which are most likely to help us prove a conjecture. This is similar to our cherry-picking problem, but TE is concerned with a theory's structure rather than individual conjectures.

Theory structure has been studied by ML4PG[4], using *clustering* to find statistical similarities and hierarchies among definitions. We investigate whether such clustering information can help TE systems navigate large theories more effectively.

2. METHODOLOGY

Following the approach of QuickSpec, our "theories" are software libraries written in Haskell; a popular programming language where correctness is concerned. We divide these definitions into small clusters, using a similar technique to ML4PG. We then run QuickSpec on each cluster individually, to produce "facts" in the form of equalities relating the definitions. The facts for each cluster are then combined.

We compare the total running time and the resulting database against running QuickSpec on the whole, unclustered library. We observe the tradeoff between speed and completeness, for various library sizes.

3. RESULTS

We have developed an ML approach for analysing Haskell code, including a bespoke feature extraction method which, to our knowledge, is the first in the literature. The success of our clustering tool shows the suitability of Haskell for the same statistical approaches used to study other languages.

4. FUTURE WORK

Our tool can be improved with a more systematic consideration of the design decisions and heuristics used, but the preprocessor approach is inherently limited. A more thorough integration of our methodology into state-of-the-art TE systems would provide more opportunities for exploiting statistical information, and also to "close the loop" by using the TE database to inform the statistical algorithms.

References

[1]Buchberger, B. 2000. Theory exploration with theorema. *Analele Universitatii Din Timisoara, ser. Matematica-Informatica.* 38, 2 (2000), 9–32.

[2]Claessen, K. et al. HipSpec: Automating inductive proofs of program properties. *In workshop on automated theory eXploration: ATX 2012.*

[3]Claessen, K. et al. 2010. QuickSpec: Guessing formal specifications using testing. *Tests and proofs.* G. Fraser and A. Gargantini, eds. Springer Berlin Heidelberg. 6–21.

[4]Heras, J. and Komendantskaya, E. 2013. Statistical proof pattern recognition: Automated or interactive? *CoRR.* abs/1303.1419, (2013).

[5]Kühlwein, D. et al. 2012. Overview and evaluation of premise selection techniques for large theory mathematics. *Automated reasoning.* Springer. 378–392.

The Brave New World of Haskell Type Classes

František Farka
School of Computing, University of Dundee, and
School of Computer Science, University of St Andrews
ffarka@dundee.ac.uk

ABSTRACT

Type classes were designed as a mean of an ad–hoc polymorphism. However, elaborate hierarchies of subclasses and superclasses can also be viewed as control structures of computations — e. g. Applicative, and Monad classes represent computation with fixed and context–dependent structure respectively. These hierarchies are bound by expected behavior of instances through the typeclass laws. We present a way how to ensure some of these law automatically and also how to reduce an amount of a boilerplate code a in form of redundant instances.

1. INTRODUCTION

Haskell type classes were introduced into the Haskell[2] as a mean of an ad–hoc polymorphism, e. g., conversion to a string in form of the *Show* class. But type classes evolved into more elaborate hierarchies of control structures. We use a model hierarchy of the *Monad* class and its superclasses *Applicative*[3] and *Functor*. Type classes are subject to the expected behavior. This behavior is expressed in a form of type class laws.

2. SUPERCLASS INSTANCES

Assume the above example of Functor, Applicative, and Monad class hierarchy. McBride and Paterson [3] described the laws that relate methods of these classes:

$$\text{fmap } f\ px = \text{pure } f <\!\!*\!\!> px \tag{1}$$

$$\text{pure } x = \text{return } x \tag{2}$$

$$\text{fmap } f\ px = px \gg\!= \text{return} \circ f \tag{3}$$

When we compose laws (1) and (3) and use equational reasoning in order to modify the equation (4) such that the application of the `pure` method is replaced by the value we obtain the equation (5), which shows that behavior of any instance of *Applicative* is solely determine by the behavior of corresponding *Monad* whenever the `Monad` instance exists.

$$\text{pure } f <\!\!*\!\!> px = x \gg\!= \text{return} \circ f \tag{4}$$

$$pf <\!\!*\!\!> px = px \gg\!= \lambda x \rightarrow pf \gg\!= \lambda f \rightarrow \text{return } (f\ x) \tag{5}$$

The equation (5) is a polymorphic specification that solely defines a behavior of the `<*>` method. Based on this observation we introduce a new language feature. We modify a type class definition such that it may contain a *Default Superclass Instance*[1]. The respectful instance for equation (5) is:

```
class Applicative m ⇒ Monad m where
    ...
    default instance Applicative m where
        pure x = return x
        pf (<*>) px = px >>= λ x → pf >>= λ f →
                      return (f x)
```

The *Functor* instance can be described in the same way using the equation (3). This language construct reduces the number of instances necessary for a code using *Monad* – and any of its superclasses – only to one — the *Monad* instance itself. The other two instances are derived. Such default instance can also be provided while introducing new superclass constraint in a backward compatible manner. This makes refactoring of type class hierarchies possible.

3. RELATED WORK

There are also other approaches how to introduce a new superclass constrains. As far as we now there is no summary on these approaches besides our own previous work [1]. The obvious question is how to resolve conflicts of ordinary instances and default instances or multiple instances in case of non–linear hierarchies. However, the answer to the question is beyond the scope of this abstract and we provide the answer also in our previous work[1].

4. CONCLUSIONS

We have introduced a new Haskell language construct that reduces amount of instances that need to be provided manually. The polymorphic instance in class definition provides stronger guarantees on instance consistency as the polymorphic definition is directly devised from the type class laws. Our mechanism also allows simpler refactoring of class hierarchies as the superclass constraint can be introduced directly while maintaining backward compatibility.

5. REFERENCES

[1] F. Farka. Maintainable type classes for haskell. *Submitted to ACM SIGPLAN Haskell Symposium*, 2015.

[2] C. V. Hall, K. Hammond, S. L. P. Jones, and P. Wadler. Type classes in haskell. In D. Sannella, editor, *Programming Languages and Systems - ESOP'94, 5th European Symposium on Programming, Edinburgh, U.K., April 11-13, 1994, Proceedings*, volume 788 of *Lecture Notes in Computer Science*, pages 241–256. Springer, 1994.

[3] C. Mcbride and R. Paterson. Applicative programming with effects. *J. Funct. Program.*, 18(1):1–13, Jan. 2008.

Parallel Patterns using Heterogeneous Computing

Mr Andreas Vermeulen
University of St Andrews
Saint Andrews, Fife KY16 9AJ
University of Dundee
Nethergate,Dundee DD1 4HN
a.f.vermeulen@dundee.ac.uk

Dr Vladimir Janjic
University of St Andrews
Saint Andrews, Fife KY16 9AJ
vj32@st-andrews.ac.uk

Mr Andy Cobley
University of Dundee
Nethergate, Dundee DD1 4HN
acobley@computing.dundee.ac.uk

ABSTRACT
An enhancement of a Research Information Factory using heterogeneous computing and parallel knowledge-extraction patterns.

Categories and Subject Descriptors
H.4 [**Information Systems Applications**]: Miscellaneous

General Terms
Theory, Framework, Application, Research, Hardware

Keywords
knowledge-extraction, patterns, information factory, RIF, RIFF, RIFC, heterogeneous computing, parallel patterns, cassandra, spark, opencl, fastflow, cuda, 3D torus network

1. INTRODUCTION
The increasing demand for data into knowledge conversion requires more volume, variety, velocity and veracity [8] in processing solutions with energy and natural resources requirements that are undesirable.

Research goal is to develop effective processing patterns with less overall energy cost.

2. BACKGROUND
Heterogeneous computing systems [8] uses central processing units (*CPU*), graphical processing units (*GPU*) and field programmable gate arrays (*FPGA*) to enable low energy processing.

Parallel Patterns are common libraries like (CUDA [11], OpenCL [9, 5, 13, 14], FastFlow [2] and ZeroMQ [6]) to guide the processing.

Efficiency and Energy-awareness is controlling efficiency of processing [15] in Floating-point Operations Per Second per Watt (*FLOP/S/W*) to achive energy requirements adviced by new euopean energy laws.

3. PROPOSED SOLUTION
The research uses parallel patterns for knowledge extraction, mechanisms for storing and extracting data while using minimum amounts of energy. It covers three basic stages:

3.1 Heterogeneous systems.
The research will study the fundamental behavior of heterogeneous computing components using a nVidia Jetson TK1 development kit. [12] and Tilera TILE-Mx100 processor [10].

3.2 Research Information Factory Framework
The framework (*RIFF*) [1] uses a parallel processing pattern via **R**etrieve-**A**ssess-**P**rocess-**T**ransform-**O**rganise-**R**eport rules.

3.3 Research Information Factory Cluster
The cluster (*RIFC*) is a 3D torus appliance [1] using Cassandra database [3, 4] and Spark Engine [4, 7] for data processing.

4. CONCLUSION
Heterogeneous systems with parallel patterns is the optimum option to achieve the goal. The Research Information Framework is a new set of guidelines to achieve the research goal is to process with less energy.

5. REFERENCES
[1] Yuichiro Ajima, Shinji Sumimoto, and Toshiyuki Shimizu. Tofu: A 6d mesh/torus interconnect for exascale computers. *Computer*, (11):36–40, 2009.
[2] Marco Aldinucci, Marco Danelutto, Peter Kilpatrick, and Massimo Torquati. Fastflow: high-level and efficient streaming on multi-core.(a fastflow short tutorial). *Programming multi-core and many-core computing systems, parallel and distributed computing*, 2011.
[3] Cassandra. Apache cassandra.
[4] Datastax. Getting started with apache spark and cassandra.
[5] Khronos OpenCL Working Group et al. The opencl specification, version 1.2, 15 november 2011. *Cited on pages*, 18(7):30.
[6] Pieter Hintjens. Ømq-the guide. *Online: http://zguide. zeromq. org/page: all, Accessed on*, 23, 2011.
[7] Apache Incubator. Spark: Lightning-fast cluster computing, 2013.
[8] Ashfaq A Khokhar, Viktor K Prasanna, Muhammad E Shaaban, and Cho-Li Wang. Heterogeneous computing: Challenges and opportunities. *Computer*, 26(6):18–27, 1993.
[9] Khronos. Opencl.
[10] Timothy G Mattson, Rob Van der Wijngaart, and Michael Frumkin. Programming the intel 80-core network-on-a-chip terascale processor. In *Proceedings of the 2008 ACM/IEEE conference on Supercomputing*, page 38. IEEE Press, 2008.
[11] nVidia. Cuda toolkit.
[12] nVidia. The world's first embedded supercomputer.
[13] Kavya Subraya Shagrithaya. Enabling development of opencl applications on fpga platforms. 2012.
[14] John E Stone, David Gohara, and Guochun Shi. Opencl: A parallel programming standard for heterogeneous computing systems. *Computing in science & engineering*, 12(1-3):66–73, 2010.
[15] Qiang Wu, Yajun Ha, Akash Kumar, Shaobo Luo, Ang Li, and Shihab Mohamed. A heterogeneous platform with gpu and fpga for power efficient high performance computing. In *Integrated Circuits (ISIC), 2014 14th International Symposium on*, pages 220–223. IEEE, 2014.

Modeling argumentative activity in mediation with Inference Anchoring Theory

Mathilde Janier
m.janier@dundee.ac.uk

1. INTRODUCTION

Our research focuses on studying the argumentative process i.e. how the arguments exchanged between conflicting parties form a reasonable discussion in order to lay the foundations for a computational tool which could be used in dispute mediation, either to make sessions more effective or to help students during their training. A close analysis of mediation discourse is therefore needed: for this we turn to the Inference Anchoring Theory (IAT) approach [1]. IAT describes dialogue dynamics in a precise way, and elicit how sequences of utterances work together to form arguments in a dialogical context. IAT has already proven useful to study discourse in mediation [3].

2. ANALYZING WITH IAT

We focus here on one type of source of impasse, the trickiest moments of a mediation session [4, 2]. In the following dialogue, Eric is a party and Mildred is the mediator. Here, Eric's claims challenge his opponent's competence.

Eric: *I'm just a bit reluctant to hand over to Viv at this early stage, because of the complexity and if you make a mistake, you waste such a lot of time. But I don't know whether Viv thinks that she's up to it or whether you think you could handle that project.*

Mildred: *What about if we perhaps separate it, had a bit of time and we spoke with each of you to look at the finance project and just see our different expectations and what you would see dealing with that project and then perhaps when we had a picture from both of you, if both of you came back to discuss your different pictures. Do you think that would work?*

Due to space limitations, the IAT analysis of this dialogue, realized on OVA+ is not presented here. OVA+ is a web-based interface for the graphical representation of arguments. The argument map is however available online[1].

The IAT analysis shows that Eric casts doubts about Viv's capacities and provides arguments for this. However, it does not make the discussion move forward since the other party does not respond. The mediator detects this source of impasse and uses an assertive question to change the topic of the discussion.

3. MODELING MEDIATION DISCOURSE

To model the mediation discourse we take into account the sequences of moves that represent a particular argumentative feature of the dialogues. Here is the resulting table for the impasse presented in Section 2. The first column of the tables represent the locutions. The transitions between locutions appear in the third column. The illocutionary forces corresponding to the locutions and the ones anchored in the transitions appear in the fourth column. Finally, the letters in the third column symbolize the propositional contents of each locution (a different letter for each different propositional content).

LOC.	PART.	TRANS.	IF	PROP CONTENT
Loc_1	$party_1$		A	a
		Loc_1;Loc_2	arguing	default inference([b,c];a)
Loc_2	$party_1$		A	b
		Loc_2;Loc_3	ø	ø
Loc_3	$party_1$		A	c
Loc_4	mediator		AQ	d

Table 1: Negative collateral implications and redirecting

The Table shows the source of impasse: only one of the two parties is talking (see the two Assertion in the fourth column) and arguing. The Table also shows the mediator's tactic: asking an assertive question (AQ) that allows her to shift the topic (no transition between her locution and the previous locutions).

4. CONCLUSION

This type of analysis can be extended to most of the mediation discourse in order to come up with a clear image of the argumentative process. IAT analyses have only been carried on one transcript of mediation, that is why the next step of the research is to verify if e.g all the analyses of this type of source of impasse in other mediation session present the same particularity, namely that the mediator poses an assertive question when there is only one party arguing.

5. REFERENCES

[1] K. Budzynska and C. Reed. Whence inference. Technical report, University of Dundee, 2011.

[2] S. Jacobs and M. Aakhus. What mediators do with words: Implementing three models of rational discussion in dispute mediation. *Conflict resolution quarterly*, 20(2):177–203, 2002.

[3] M. Janier, M. Aakhus, K. Budzynska, and C. Reed. Games mediators play: Empirical methods for deriving dialogue structure. In *MET-ARG workshop*, December 2014.

[4] S. G. Morasso. *Argumentation in dispute mediation*. John Benjamins Publishing Company, 2011.

[1] `http://ova.arg-tech.org/analyse.php?url=local&plus=true&corpusid=110&aifdb=4596`

Towards Robust and Efficient Automated Curvilinear Structure Detection in Medical Images

R. Annunziata, E. Trucco
School of Computing,
University of Dundee,
Dundee, UK.

P. Hamrah, A. Kheirkhah
Department of Ophthalmology,
Harvard Medical School,
Boston, USA.

ABSTRACT

Segmenting dendritic trees and corneal nerve fibres is challenging due to their uneven and irregular shape. Our first contribution is a novel ridge detector, SCIRD, which is rotation, scale and curvature invariant [2]. Then, we present a novel feature boosting method combining hand-crafted appearance features with learned context filters. Experimental results on 3 challenging and diverse datasets show that our methods outperform state-of-the-art approaches.

INTRODUCTION

Automated systems analysing curvilinear structures in the medical domain would allow cost-effective large screening programs aimed at early diagnosis. To measure morphometric properties such as tortuosity [1, 6], these structures have to be segmented accurately. Many solutions have been proposed to cope with problems such as low signal to noise ratio at small scales, confounding non-target structures and non-uniform illumination [3, 4, 7, 9]. Most approaches are based on a local *tubularity* measure estimated via hand-crafted features (HCFs) [3, 4, 9], or learned from training data [8]. While HCF methods are fast, they are based on assumptions that are violated by highly fragmented and tortuous structures, i.e. continuous and locally straight tubular shapes. While discontinuity can be addressed by elongated kernels (e.g., Gabor [9]), no *hand-crafted* model for non-straight tubular shapes has been proposed so far. So, our first contribution is a novel HCF ridge detector, SCIRD (Scale and Curvature Invariant Ridge Detector) which is rotation, scale and curvature invariant. Recently, combining HCFs with learned filters has proven successful as it exploits the efficiency of a fast HCF approach to reduce the amount of learned filters [7]. Although this method outperforms state-of-the-art approaches such as [3, 4], our second contribution is motivated by its modest success in segmenting fragmented structures such as corneal nerve fibres and neurites. Integrating *context* information with *appearance* features has been recently found to reduce these shortcomings, by learning multiple discriminative models sequentially [10]. In this paper, we propose a novel approach for combining HCFs (i.e. appearance) with learned context filters which improves results in [7] at the *same* computation cost (i.e. learning a single discriminative model).

METHODS

Curved-Support Gaussian Models

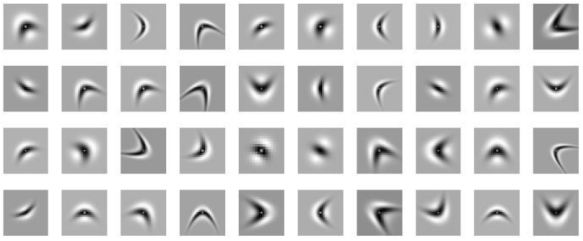


Figure 1: A subset of SCIRD filters.

We model a 2-D curvilinear structure using curved-Gaussian models [5]:

$$\Gamma(x_1, x_2; \boldsymbol{\sigma}, k) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{x_1^2}{2\sigma_1^2}} \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{\left(x_2 + k x_1^2\right)^2}{2\sigma_2^2}}. \quad (1)$$

where (x_1, x_2) is a point in the principal component coordinate system of the target structure, (σ_1, σ_2) control the elongation of the shape ("memory") and its width, respectively; in fact, the first term of $\Gamma(x_1, x_2; \boldsymbol{\sigma}, k)$ controls the longitudinal Gaussian profile of the model, while the second controls the cross-sectional Gaussian profile. Importantly, we add a new parameter, k, to control the curvature of the Gaussian support.

Scale and Curvature Invariant Ridge Detector

Various detectors of tubular image structures compute the contrast between the regions inside and outside the ridge (e.g., [3, 9]). We extend this idea to curved-support Gaussian models by computing the second-order directional derivative in the gradient direction at each pixel. To improve efficiency, we formulate the problem of tubularity estimation (I_{ww}) as a filtering operation:

$$I_{ww}(x, y; \boldsymbol{\sigma}, k) = I(x, y) * K_{ww}(x, y; \boldsymbol{\sigma}, k), \quad (2)$$

where $I(x, y)$ represents the grey-level of a monochrome image at the location (x, y) and

$$K_{ww} = (\tilde{\Gamma}_x \Gamma_{xx} + \tilde{\Gamma}_y \Gamma_{yx})\tilde{\Gamma}_x + (\tilde{\Gamma}_y \Gamma_{xy} + \tilde{\Gamma}_y \Gamma_{yy})\tilde{\Gamma}_y. \quad (3)$$

Here, $\tilde{\Gamma}$ is a curved-support Gaussian model with a constant longitudinal profile. Single subscript indicates first derivative (e.g., $\tilde{\Gamma}_x$), double subscripts is used for second derivatives (e.g., $\tilde{\Gamma}_{xx}$). To achieve scale and curvature invariance, we create a filter bank generated by making σ_2 and k span scale and curvature range for the specific appli-

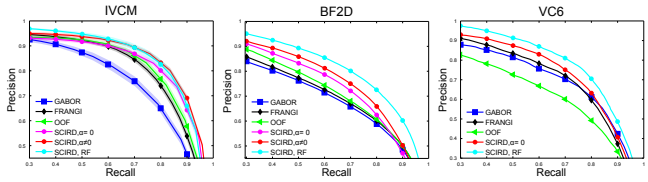


Figure 2: P-R curves for SCIRD and baselines.

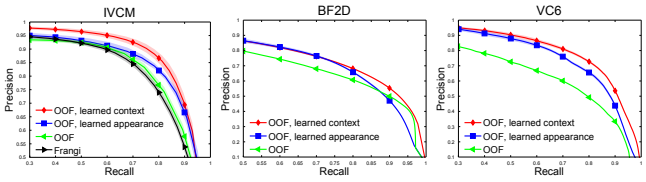


Figure 3: P-R curves for our feature boosting method and baselines.

cation at hand. Rotation invariance is obtained adding rotated filters replicas. Fragmented structures are dealt with by tuning the "memory" parameter σ_1. Figure 1 shows some of the kernels used in our experiments. Finally, SCIRD selects the maximum response over all kernels. Notice, SCIRD can be employed for both segmentation (after thresholding) *and* centreline detection (after non-maxima suppression).

Learning Context Filters

Auto-context [10] is a well-known method to take context information into account. However, it is based on learning multiple discriminative models in a sequential pipeline, making this approach impractical for high resolution images or large datasets. Our goal is to exploit context information without increasing computational cost with respect to the solution in [7]. This is achieved by learning a *single* discriminative model which takes as input both *appearance* (i.e. likelihood computed on the original image [10]) and context information (i.e. relations between objects [10]). We use the Optimally Oriented Flux (OOF) feature [4] as HCF, shown to outperform other HCFs on the datasets we use for our tests in [7]. Then, we learn *context* filters applying the K-means algorithm to the patches of OOF feature maps. Intuitively, we learn a set of probing filters capturing key configurations actually present in the OOF filtered maps. Notice that our context feature learning includes HCF into the filter learning process, while appearance feature learning as proposed in [7] does not. So, while the latter could potentially learn redundant filters, i.e. filters reconstructed by a combination (linear or non-linear) of the HCF already modelling the appearance, our context filters learning method is complementary.

RESULTS AND DISCUSSION

We evaluate our method and baselines performance on 3 very diverse datasets, showing corneal nerve fibres and neurites (IVCM, BF2D and VC6). First, we compare SCIRD against 3 HCF state-of-the-art ridge detectors: Frangi [3], Gabor [9] and Optimally Oriented Flux (OOF) [4]. The results in Figure 2 show that SCIRD outperforms the other methods on all datasets suggesting that our filters behave better than others at low resolution and low SNR when dealing with tortuous and fragmented structures. The low number of false positives when false negatives are low, implies that SCIRD selects target structures with higher confidence.

Second, we compare our method of feature boosting (i.e. context learning) against the method proposed in [7] (i.e. appearance learning). As Figure 3 shows, our method outperforms the baselines on all datasets. In particular, boosting OOF with learned context features outperforms the boosting technique proposed in [7] at the *same* computational cost (i.e. learning a single discriminative model). In terms of efficiency, learning a filter bank of 100 filters using K-means algorithm is significantly faster than sparse coding. In fact, less than 30 seconds are typically required to learn our filter banks which compares favourably with several days reportedly needed to learn 121 filters using sparse coding [7].

CONCLUSION

We addressed the problem of curvilinear structure segmentation by proposing a novel HCF (SCIRD) and a new method combining HCFs with learned context filters. Experimental results show that our methods outperform state-of-the-art approaches. Our future work will investigate the combination of SCIRD with learned context filters. Part of this work (i.e., SCIRD) will be presented at MICCAI 2015 [2].

REFERENCES

[1] R. Annunziata, A. Kheirkhah, S. Aggarwal, B. M. Cavalcanti, P. Hamrah, and E. Trucco. Tortuosity classification of corneal nerves images using a multiple-scale-multiple-window approach. In *OMIA Workshop , MICCAI*, 2014.

[2] R. Annunziata, A. Kheirkhah, P. Hamrah, and E. Trucco. Scale and curvature invariant ridge detector for tortuous and fragmented structures. In *MICCAI*, 2015, (In press).

[3] A. Frangi, W. Niessen, K. Vincken, and M. Viergever. Multiscale vessel enhancement filtering. In *MICCAI*. 1998.

[4] M. W. Law and A. C. Chung. Three dimensional curvilinear structure detection using oof. In *ECCV*. 2008.

[5] J. K. Lin and P. Dayan. Curved gaussian models with application to the modeling of foreign exchange rates. In *Computational Finance*. MIT Press, 1999.

[6] A. Lisowska, R. Annunziata, and E. Trucco. An experimental assessment of five indices of retinal vessel tortuosity with the ret-tort public dataset. In *IEEE EMBS*, 2014.

[7] R. Rigamonti and V. Lepetit. Accurate and efficient linear structure segmentation by leveraging ad hoc features with learned filters. In *MICCAI*. 2012.

[8] A. Sironi, V. Lepetit, and P. Fua. Multiscale centerline detection by learning a scale-space distance transform. In *CVPR*, 2014.

[9] J. V. Soares, J. J. Leandro, R. M. Cesar, H. F. Jelinek, and M. J. Cree. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE TMI*, 2006.

[10] Z. Tu and X. Bai. Auto-context and its application to high-level vision tasks and 3d brain image segmentation. *IEEE TPAMI*, 2010.

Computer-Assisted Analysis Of Arterial Narrowing In Whole-Body MRA

Andrew McNeil
1.07, School of Computing
University of Dundee
a.y.mcneil@dundee.ac.uk

Emanuele Trucco
School of Computing
University of Dundee

Graeme Houston
Clinical Radiology
Ninewells Hospital

ABSTRACT

Contrast-enhanced Whole-Body Magnetic Resonance Angiography (WBMRA) is performed by injecting a contrast agent into the vein and acquiring images using an MRI scanner as the agent passes through the arteries of interest. This technique generates high contrast in the channel within the artery where the blood is flowing (the lumen), thereby providing a non-invasive, comprehensive imaging method for assessing cardiovascular disease (CVD) throughout the entire body. Analysing the resulting large datasets, however, is very labour-intensive and thus there is a great need for robust, automated, quantitative analysis tools to help stage the disease from WBMRA examinations.

The main aim of this project is to created an automated software tool which can detect and grade stenoses from WBMRA datasets, highlighting problematic areas for clinicians.

It has been found, through a close examination of a lumen calibre measurement system called "GroBa", that in most cases simple measures of vessel calibre are not stable enough, or contain enough contextual information, to provide accurate stenosis detection or severity gradings, particularly over the large range of vessel sizes found in WBMRA. Current work is therefore focussed on developing a machine-learning-based approach, utilising ground truth data which is being collected at Ninewells Hospital.

1. INTRODUCTION

Each year CVD causes over 4 million deaths in Europe; 47% of all deaths [1]. Due to the large burden CVD places on both public health and the economy (costing the EU an estimated €196 billion a year due to health care costs and productivity losses [1]), there is therefore much interest in the early staging of CVD (identifying its severity and distribution, similar to what is done for cancer) to improve patient outcomes.

This project is a collaboration between the CVIP group at Dundee University, the Cardiovascular and Diabetes Imaging Research Unit at Ninewells Hospital and Toshiba Medical Visualization Systems in Edinburgh. The key aim is to develop an automated software solution for accurately segmenting the main arterial tree and detecting and grading stenoses (narrowing of the lumen) in WBMRA examinations, the results of which should be presented in a way which integrates well with current diagnostic approaches, thereby aiding clinicians in their diagnoses.

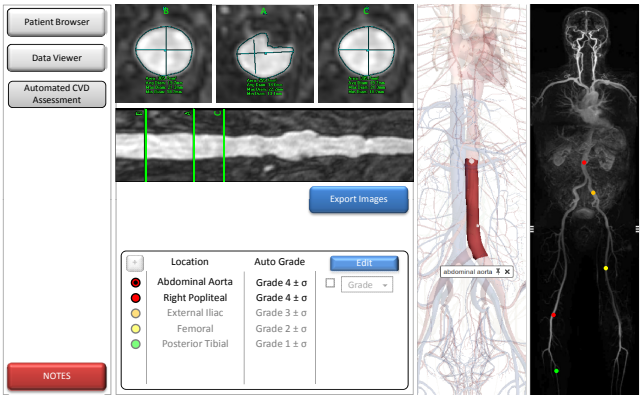


Figure 1: Example screen from the interface prototype of a proposed WBMRA vascular analysis program, showing the automated CVD assessment display.

First we will describe the outcomes of an interface prototyping exercise which helped define the clinical problem we are targeting, and the technologies which must be developed in order to meet that need. We will then briefly describe the typical approach for measuring lumen diameters and detecting stenoses, using the outcomes of an evaluation of the "GroBa" system to provide the motivation for the current machine-learning-based approach currently being pursued.

2. PROTOTYPE INTERFACE

In close collaboration with clinicians from Ninewells Hospital, and members from Toshiba, an interface "paper prototype" was constructed, aiming to bring together all the key functionalities required for the diagnostic task, allowing us to visualise the final clinical application. This not only served as motivation for the direction and emphasis of our future research, but is also an important step towards transferring our research out of the lab and into general use, laying the groundwork for future software development through our partnership with Toshiba.

Figure 1 shows an example screen, demonstrating the patient the automated CVD assessment window. Based on our discussions with the clinicians, a number of desirable features and interactive functionalities were identified, such as the ability to visualise and process MRA datasets alongside other relevant information (blood tests, kidney time-activity

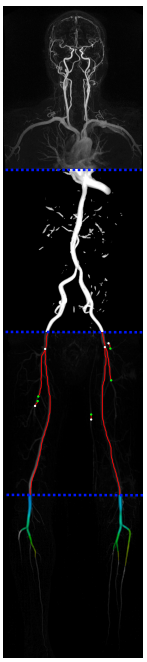


Figure 2: Example whole body MRA dataset, showing the the enhancement, centreline, and measurement stages of the GroBa system.

curves, etc.), and to highlight problematic areas visually on the scans themselves.

In order to meet these requirements, a number of key technologies were identified. The main focus of this project was therefore determined to be developing the automated stenosis detection algorithm itself.

3. TYPICAL LUMEN MEASUREMENT & STENOSIS DETECTION

"GroBa" is a lumen calibre measurement technique based on growing "balloons" inside a segmented arterial lumen [2]. The balloon is initialised as a single voxel on the centreline, which is iteratively grown until the length is twice the calculated width. The vessel calibre is then estimated from the equation of a cylinder, using the calculated balloon volume and it's half-diagonal.

A thorough evaluation of the MATLAB implementation of the GroBa system was carried out, testing it against both synthetic lumina, and real WBMRA datasets provided by Ninewells Hospital. The fully automatic system begins by first applying a "vesselness" filter, enhancing vessel-like structures by analysing the eigenvalues of the Hessian matrix. The resulting "vesselness map" is thresholded to produce a binary vessel map. The centrelines of each vessel is then calculated via a skeletonization method based on fast-marching techniques [3]. Finally the lumen calibre throughout each segmented vessel is measured using the "balloon" method described above, with the calibre information presented as an HSV colour-space overlay on the maximum intensity projection (MIP) (see Figure 2).

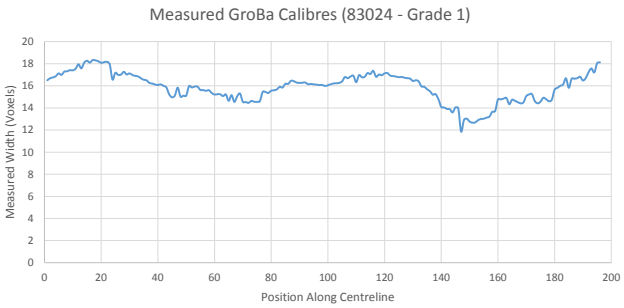


Figure 3: Graph showing the estimated lumen calibres in the aorta of a diseased patient, highlighting the difficulties of detecting stenoses from a single measure (the true 30% stenosis is around position 150)

Figure 3 shows a typical plot of lumen calibre, where stenoses are detected as a local max-min-max in the measurements. However when comparing with ground-truth annotations, we find that in most cases the measures themselves are not stable enough, or contain enough contextual information, to provide accurate stenosis detections or severity gradings, particularly over the large range of vessel sizes found in WBMRA.

4. CONCLUSIONS & FUTURE WORK

The main aim of this project is to develop an automated system for detecting and grading stenoses from WBMRA, a key indicator of cardiovascular disease severity. A prototyping exercise was undertaken with clinicians at Ninewells Hospital, which identified the visualisation and analysis technologies which are required in order to meet these demands.

Through an examination of the GroBa system for lumen calibre measurement, it was found that using a single measure such as this gives poor results when compared to ground-truth stenosis measurements carried out by a trained clinician. Our current work is therefore focussed on developing a machine-learning approach, leveraging appropriate machine learning techniques (such as decision trees or Gaussian processes) with a number of local features (calibre, area, shape, curvature, etc.) to train a stenosis detection and grading system using manual ground truth annotations currently being collected at Ninewells Hospital.

5. REFERENCES

[1] M. Nichols, N. Townsend, R. Luengo-Fernandez, J. Leal, A. Gray, P. Scarborough, and M. Rayner. *European Cardiovascular Disease Statistics, 2012 Edition*. European Heart Network & European Society of Cardiology, 2012.

[2] A. Perez-Rovira, E. Trucco, J. Weir-McCall, and G. Houston. Groba: Growing balloons for calibre measurement on stenotic lumens. In *Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on*, pages 1–6, 2012.

[3] R. Van Uitert and I. Bitter. Subvoxel precise skeletons of volumetric data based on fast marching methods. *Medical Physics*, 34(2):627–638, 2007.

Reducing Viseme Confusion In Speech Reading

Benjamin M. Gorman
University of Dundee
Dundee, Scotland
b.gorman@dundee.ac.uk

ABSTRACT

Speech reading is an invaluable technique for people with hearing impairments or those listening in adverse conditions. Speech reading is limited as several speech sounds are produced from identical mouth shapes. This decreases comprehension and causes confusion during conversation. This research introduces a visualisation technique named PhonemeViz, which aims to allow the user to determine which phoneme is hidden behind a particular viseme. PhonemeViz will be evaluated alongside six other techniques from related work to determine if it improves speech reading accuracy.

1. INTRODUCTION

Speech reading (often called lip reading) refers to using information from a speakers lips, facial expressions and gestures to understand conversation. Speech reading is commonly used by those who are deaf and those with impaired hearing. However, it has also been shown that those with typical hearing subconsciously make use of visual cues [1]. When speech reading, each speech sound (phoneme) corresponds to a facial and mouth position (viseme). For example, acoustically speaking in English, /l/ and /r/ can be quite similar (especially in clusters, such as 'grass' vs. 'glass'), yet visual information can show a clear contrast. A phoneme is a unit of a language's phonology, which is combined with other phonemes to form words. For example in English, the difference between the words mat and bat occur due to the exchange of the phoneme /m/ for the phoneme /b/. A viseme is a facial and mouth position which several speech sounds may correspond to. Therefore, speech reading is limited in that many phonemes share the same viseme and thus are impossible to distinguish from visual information alone.

2. RELATED WORK

Speech readers born or who have grown up deaf may have never heard spoken language and are unlikely to be fluent users of it, which makes speech reading more difficult. There are a number of techniques which have been used in order to overcome the challenges presented by speech reading.

Cued Speech is a phonemic-based system which makes traditionally spoken languages accessible by using a small number of hand shapes, known as cues, (representing consonants) in different locations near the mouth (representing vowels), as a supplement to speech reading. However, like sign language in order for it to be effective it relies on both parties of a conversation to be fluent, and therefore is not a complete solution to the limitations of speech reading.

Lip Assistant [8] is a system which uses a video synthesizer to generate video-realistic mouth animations. These mouth animations are superimposed on the original video as an assistant for hearing impaired viewers to make a better understanding on the audio-visual contents. Experimental results show that lip assistant can significantly improve speech intelligibility.

Closed captioning displays the audio portion of a television programme as text on the TV screen, providing access to news, entertainment and information for individuals who are deaf or hard-of-hearing. Whilst live captioning is becoming more prominent, it is still prone to errors and it can never be real-time as the system needs to wait until the entire word has been spoken until it can determine which word it is.

A spectrogram is a visual representation of the spectrum of frequencies in a sound wave over time. Spectrograms are used to identify words phonetically. However, becoming competent can take considerable training [2]. Watanabe et al [7] introduce a visualisation that creates readable patterns by integrating different speech features into a single picture, with the final pattern resembling a spectrogram which has been coded by patterns, colours and labels. However, this system was only evaluated by three participants with extensive experience of reading spectrograms.

VocSyl [3] is a software package that provides real-time visual feedback in response to vocal pitch, loudness, duration, and syllables. The authors believe that a reinterpretation of voice will allow for both a new understanding of one's vocalisation by audio/visual feedback. Pietrowicz and Karahalios [6], build upon this work by introducing visualisations which show phonetic detail. The colour mappings in the visualisation represent phonological detail with a distinct colour representing high closed vowels, mid vowels, low open vowels, diphthongs, liquids, nasals/glides and fricatives/affricates/stops.

iBaldi [5], is an iOS application which transforms speech into visual cues to supplement speechreading. The cues are three coloured discs; nasality (red), friction (white), and voicing (blue), which appear when a phoneme from a corresponding group is presented. The cues are located near a computer generated face's mouth. However like reading a spectrogram, the technique requires a large amount of training in order to become effective.

3. PHONEMEVIZ

After reviewing the literature, three key features of an effective visualisation to support speechreading were outlined. A spectrogram demonstrates that persistence is important, as the ability to "catch up" on what has been said provides a better chance of achieving comprehension. This can be demonstrated further by the shortcomings of iBaldi, as the coloured discs appear too fast for the user to notice. Many of the previously discussed techniques also require a large amount of training in order to be effective. Speechreading is already a high work load task and therefore it was clear that an effective visualisation should have a low work load and a low amount of training in order for a user to become competent quickly. Many of the visualisations discussed also rely on colour information. However, for the visualisation to be usable by as many users as possible colours are not the best approach, as users with colour vision deficiency would not be able to discriminate the subtle differences [1].

PhonemeViz focuses on reducing viseme confusion which occurs at the start of words. PhonemeViz places consonant phonemes in a semi circular arrangement, an arrow beginning from the centre of this semi-circle points at the last heard consonant phoneme. PhonemeViz is positioned at the side of a users face, beginning at the forehead and ending at the chin. This allows the user to follow the speaker's eyes and lip movements whilst watching for changes in the arrow's position with their peripheral vision.

4. EVALUATION

PhonemeViz will be evaluated against six visualisation techniques: Captions, Lip-Magnification, Spectrogram, iBaldi and VocSyl and no visualisation which will act as the control condition. Captions will be presented as white text on a black bar at the bottom of the screen. Lip-Magnification will be a modified version of the technique introduced in [8], instead of animating a set of computer generated lips, a magnification of the speakers lips will be presented at the right hand side of the speakers face. The spectrogram will draw from right to left of the screen and adopt a rainbow colour palette. An implementation of iBaldi will follow its description in [5]. The implementation of VocSyl will follow the description outlined in [3], and will adopt the colour palette introduced in [6]. Each technique will be implemented as an openFrameworks application and made use of an open source add-on called ofxTimeline (`github.com/YCAMInterlab/ofxTimeline`).

For each technique participants will view a sequence of videos (with muted audio) in which a speaker is saying a group of selected words that can cause confusion due to visemes. The participant will be told to press the spacebar when they think that the speaker is saying a particular word.

4.1 Procedure

This study uses a repeated measures design and participants will attend two study sessions. There will be two days between each session. Techniques and word groups are to be counterbalanced across participants. Participants will need to be over 18 and have good or corrected vision. Participants who answer 'yes' to any of the questions on the demographics form that relate to frequent use of speech reading will be eligible to take part in our post-evaluation analysis.

To account for variance in speech reading proficiency, we will record participants' ability to speech read. In total 40 words were chosen from the British National Corpus, and were selected from each quarter of the list ordered by frequency in daily speech. Each word was recorded with a female speaker facing the camera from the shoulders up. The participant is in control of when the test proceeds and has two viewings of each word. The participant will be given a randomised list of 10 words at a time has to match the words with the number in which they think are presented.

The words for the evaluation were chosen by looking at the phoneme to viseme table in [4]. Three words were chosen for each viseme group that were similar apart from the initial consonant phoneme at the beginning. In total the evaluation uses five groups /p/,/t/,/k/,/ch/, /k/ with two (/p/ and /t/) being repeated but with different words.

4.2 Results

Each participant will complete 36 trials for each of the 7 techniques. For each trial we record the accuracy of their response; True Positive, False Positive, True Negative or False Negative. Precision will be calculated by TP / (TP + FP). Recall will be calculated by TP / (TP + FN). An F1 Score will be calculated for each participant with each technique where F1 = (2 * (Precision * Recall)) / (Precision + Recall). F1 Score will be one of our dependent variables. For each technique participants will also complete a NASA Task Load Index (NASA-TLX). This is a subjective, multidimensional assessment tool that rates perceived workload in order to assess a task, system, or team's effectiveness or other aspects of performance. Participants will also rank each technique in order of preference giving reasons for their ordering.

5. REFERENCES

[1] E. Goldstein. *Sensation and perception*. Cengage Learning, 2013.

[2] B. G. Greene, D. B. Pisoni, and T. D. Carrell. Recognition of speech spectrograms. *JASA*, 76(1):32–43, 1984.

[3] J. Hailpern, K. Karahalios, L. DeThorne, and J. Halle. Vocsyl: Visualizing syllable production for children with asd and speech delays. In *Proc. ASSETS '10*, pages 297–298. ACM.

[4] P. Lucey, T. Martin, and S. Sridharan. Confusability of phonemes grouped according to their viseme classes in noisy environments. In *Proc. of Australian Int. Conf. on Speech Science & Tech*, pages 265–270, 2004.

[5] D. W. Massaro, M. M. Cohen, W. Schwartz, S. Vanderhyden, and H. Meyer. Facilitating speech understanding for hearing-challenged perceivers in face-to-face conversation and spoken presentations.

[6] M. Pietrowicz and K. Karahalios. Sonic shapes: Visualizing vocal expression. In *ICAD 2013*.

[7] A. Watanabe, S. Tomishige, and M. Nakatake. Speech visualization by integrating features for the hearing impaired. *IEEE Trans. Speech Audio Process.*, 8(4):454–466, 2000.

[8] L. Xie, Y. Wang, and Z.-Q. Liu. Lip assistant: Visualize speech for hearing impaired people in multimedia services. In *Proc. SMC'06*, volume 5, pages 4331–4336. IEEE.

Argument Mining

John Lawrence
School of Computing
University of Dundee
Dundee
j.lawrence@dundee.ac.uk

Chris Reed
School of Computing
University of Dundee
Dundee
c.a.reed@dundee.ac.uk

1. INTRODUCTION

Argument Mining is the automatic identification of the argumentative structure contained within a piece of natural language text. By automatically identifying this structure and its associated premises and conclusions, we are able to tell not just *what* views are being expressed, but also *why* those particular views are held. One of the first approaches to argument mining, is the work carried out by Moens et al. beginning with [3], where text is first split into individual sentences and then a machine learning technique used to classify these sentences as being part of an argument or not before being additionally classified as premise or conclusion. This is built upon in [1] which looks at classifying the argument scheme of an argument that has already undergone successful extraction of premises and conclusions.

2. ARGUMENT MINING TECHNIQUES

In order to improve on these existing techniques, we have developed a range of methods for determining argument structure based around the manual analysis process. Firstly, we have looked at using the presence of discourse indicators, linguistic expressions of the relationship between statements, to determine relationships between the propositions. We have also developed a topic based approach [2], investigating how changes in the topic being discussed relate to the argumentative structure being expressed. Finally, we have implemented a machine learning approach based on argumentation schemes [5], enabling us to not only identify premises and conclusions, but to determine how exactly these argument components are working together.

Based on the results from the individual implementations, we have combined these approaches, taking into account the strengths and weaknesses of each to improve the accuracy of the resulting argument structure.

3. CURRENT RESULTS

The results for discourse indicators show that when they are present in the text, they give a strong indication of the connection between propositions (precision of 0.89); however, the low frequency with which they can be found means that they fail to help identify the vast majority of connections (recall of 0.04).

For the topic based approach, again, precision (0.82) is higher than recall (0.56) suggesting that although this method may fail to find all connections, those that it does find can generally be viewed as highly likely.

Finally, by considering the features of the individual types of premise and conclusion that comprise an argumentation scheme, we achieved similar performance (F-scores between 0.75 and 0.93) to that of [1], where the occurrence of a particular scheme was identified with accuracies of between 62.9% and 90.8%. However, these previous results only considered spans of text that were already known to contain a scheme of some type and required a prior understanding of the argumentative structure contained within the text.

Although there are strengths and weaknesses to each of these approaches, by using them in combination we can achieve results that are remarkably close to a manual analysis of the same text. The accuracy we have achieved for determining connections between propositions (f-score of 0.83) compares favourably with other results from the argument mining field. For example, in [4] sentences were classified as either premise (F-score, 0.68) or conclusion (F-score, 0.74). In the case of our combined results, not only are we able to determine the premises and conclusion of an argument, but its schematic structure and the precise roles that each of the premises play in supporting the conclusion.

4. REFERENCES

[1] V. W. Feng and G. Hirst. Classifying arguments by scheme. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 987–996. Association for Computational Linguistics (ACL), 2011.

[2] J. Lawrence, C. Reed, C. Allen, S. McAlister, and A. Ravenscroft. Mining arguments from 19th century philosophical texts using topic based modelling. In *Proceedings of the First Workshop on Argumentation Mining*, pages 79–87, Baltimore, Maryland, June 2014. Association for Computational Linguistics (ACL).

[3] M.-F. Moens, E. Boiy, R. M. Palau, and C. Reed. Automatic detection of arguments in legal texts. In *Proceedings of the 11th international conference on Artificial intelligence and law*, pages 225–230. ACM, 2007.

[4] R. M. Palau and M.-F. Moens. Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law*, pages 98–107. ACM, 2009.

[5] D. Walton, C. Reed, and F. Macagno. *Argumentation Schemes*. Cambridge University Press, 2008.

Improving Intercultural Human-to-Human Interaction through Non-verbal Visualisation

Garreth Tigwell
University of Dundee
Dundee, Scotland
gwtigwell@dundee.ac.uk

1. ABSTRACT

The intended message behind non-verbal behaviour can differ between cultures. This leads to communication breakdown and misunderstandings between people. To overcome this, I will develop wearable technology to augment the user's perception of culturally different non-verbal cues. The wearable technology will be developed with a focus on user-centred design and user experience. Evaluations from in-lab and real-world user studies will iteratively improve the technology. By allowing people to understand the non-verbals of others, our technology has the potential to reduce the difficulties of intercultural communication worldwide, improving business, travel, and deepening our interpersonal experience.

2. INTRODUCTION

Non-verbal behaviour is a component of human communication. Non-verbal cues have been shown to help when judging rapport between speakers [2], they assist in communicating ideas and managing interactions [4], as well as improving the efficiency of communication by disambiguating any intended meaning (e.g., sarcasm) [5]. Whilst the importance of non-verbal cues during conversation between people from the same culture often go unnoticed, the different meanings of those behaviours could have a substantial negative influence on the understanding between people from different cultures. For example, nodding your head in the UK can indicate agreement or understanding, however, this meaning is not universal and in other cultures it indicates the person hears you talking or signals "no" [9]. To address the challenges associated with intercultural human-to-human interaction, I will develop non-verbal visualisations that will provide people (e.g., a group of students and international students) with a tool to help them identify the meanings behind various non-verbal cues when talking to somebody from a different culture. This will reduce the chances of communication breakdown, as well as misunderstanding, hostility, and social isolation or exclusion that could result.

3. RELATED WORK

It has been demonstrated that there is a need to consider the effect a person's cultural background can have. Kayan, Fussell and Setlock [3] suggest features offered by Instant Messaging clients should be selectable by users to highlight their preferred method of communication for improved intercultural communication. Similarly, Reinecke and Bernstein [8] explain that by changing the interface according to the cultural background of the user, there is an overall better user experience reported.

4. OBJECTIVES

There will be three stages during the course of this research to develop a wearable technology solution, which will augment the user's perception of non-verbal cues, thus allowing the user to successfully identify and interpret culturally different non-verbal cues. The first stage requires development of a visual feedback system which will inform users of the intended meaning of the non-verbal components of conversation and social interaction that differ between cultures. These abstract visualisations can be evaluated qualitatively and quantitatively with participants. The second stage will involve producing a method for detecting non-verbal cues that are culturally distinct by adapting current image processing techniques (e.g., for identifying gestures), as well as speech processing techniques (e.g., to analyse vocal tones). Various factors can be explored such as looking for immediate cues (e.g., head-shaking or pointing), as well as determining how feasible the detection of those would be. Additional sensors (e.g., physiological sensors) could be used if the system's automated detection is inadequate. The third and final stage will combine the visualisation and detection components from the previous two stages to create a working prototype that will be evaluated with user groups consisting of people from different cultural backgrounds. The solution can be refined by feedback from the user groups, thereby improving it even further.

5. GENERALISATIONS

Cultural differences for the meaning of non-verbal cues is only one example of a reason for augmenting a user's perception of non-verbal cues. There is an opportunity to explore how this proposed feedback system assists those with Autism Spectrum Disorder (ASD). ASD typically involves problems with nonverbal social interaction and communication [7]. Furthermore, I can investigate the effectiveness of using the system as a tool for self-monitoring during situations where the other person is sensitive to nonverbal behaviour [1], but is also unable to wear or use the tool themselves. An example of this is a carer for a person with dementia who could use a reflective form of the tool to provide feedback to help improve communication [6]. Alternatively, the feedback mechanism can be changed (e.g., providing vibration or tactile feedback in place of visualisations). This would benefit people with a visual impairment who cannot see subtle body language or facial expressions.

6. REFERENCES

[1] I. Damian, C. S. S. Tan, T. Baur, J. Schöning, K. Luyten, and E. André. Augmenting social interactions: Realtime behavioural feedback using social signal processing techniques. In *Proc. CHI'15*, pages 565–574.

[2] J. E. Grahe and F. J. Bernieri. The importance of nonverbal cues in judging rapport. *J. Nonverbal behav.*, 23(4):253–269, 1999.

[3] S. Kayan, S. R. Fussell, and L. D. Setlock. Cultural differences in the use of instant messaging in asia and north america. In *Proc. CSCW'06*, pages 525–528.

[4] M. Knapp and J. A. Hall. *Nonverbal communication in human interaction*. Cengage Learning, 7th edition, 2010.

[5] J. Kruger, N. Epley, J. Parker, and Z.-W. Ng. Egocentrism over e-mail: Can we communicate as well as we think? *J. Pers. Soc. Psychol.*, 89(6):925, 2005.

[6] P. McCallion, R. W. Toseland, D. Lacey, and S. Banks. Educating nursing assistants to communicate more effectively with nursing home residents with dementia. *The Gerontologist*, 39(5):546–558, 1999.

[7] P. Mundy, M. Sigman, J. Ungerer, and T. Sherman. Defining the social deficits of autism: The contribution of non-verbal communication measures. *J. Child Psychol. Psychiatry*, 27(5):657–669, 1986.

[8] K. Reinecke and A. Bernstein. Knowing what a user likes: A design science approach to interfaces that automatically adapt to culture. *MIS Q.*, 37(2):427–454, June 2013.

[9] L. A. Samovar, R. E. Porter, E. R. McDaniel, and C. S. Roy. *Communication Between Cultures*. Cengage Learning, 2012.

Discovering New Ontologies and Lemmas through Concept Blending and Reasoning

Meenakshi Kesavan
School of Computing
University of Dundee
Dundee DD1 4HN
m.kesavan@dundee.ac.uk

ABSTRACT
The process of examination and application of ontologies is currently being tried in multiple fields. Previously, the research interests were mainly focussed on the development of ontologies. But now, the subject is shifting towards the usage of these ontologies to improve machine intelligence.

1. AIM
My research aim is to manipulate ontologies to discover new concepts, relationships and rules by utilizing a concept known as "Concept Blending"

The research can proceed in two different ways:
- To build ontologies by abstracting the concepts from a particular domain of interest and reasoning them
- To discover new relationships using Conceptual Blending

2. DOMAIN CONSIDERED
Hindustani Classical Music

3. MOTIVATION
Knowledge representation in the area of Indian Music, particularly Hindustani music is a burgeoning area. Although some work has been done in Carnatic music, Hindustani music remains unexplored. Such an exclusive knowledge base for the same, does not exist at the moment.

4. INDIAN MUSIC ONTOLOGY
It provides a formal framework for dealing with music-related information using Semantic web,
- To design and analyse domain knowledge
- To enable knowledge sharing and
- To enable reuse of the domain knowledge

Music mainly relies upon Raga (the basis of Indian Melody) and Tala (Indian Rhythmic pattern). In order to classify a composition, the features of Raga plays a major role.

5. CONCEPT BLENDING
"Producing novel ideas through unfamiliar combinations of familiar ideas". Using conceptual blending, unexplored musicological relations between different combination of ragas can be discovered. The taxonomic classification/ ontology helps to identify related ragas and the relationships between them. This can be further extended to discover a new raga by unleashing the potential of Concept Blending.

6. KEY CHALLENGES
- The structure of a raga is highly complex and requires domain expertise for understanding.
- The musical notations are difficult to understand and analyse
- Classifying and exploring non-linear ragas requires expertise in music domain.

7. FUTURE WORK
- To try this method using statistical machine learning for lemma discovery in theorems [2],
- To explore connections with analogy work in theorem proving [2],
- To explore swara-pattern recognition and emotion recognition of ragas.
- To integrate audio information to aid in the classification of the raga

8. REFERENCES
1. Kutz Oliver, Neuhaus Fabian, Mossakowski Till, and Codescu Mihai (2014), "Blending in the Hub - Towards a computational concept invention platform."

2. Jónathan Heras, Ekaterina Komendantskaya, Moa Johansson, Ewen Maclean (2013), "Proof-Pattern Recognition and Lemma Discovery in ACL2."

3. Marco Schorlemmer; Alan Smaill; Kai- Uwe Kuhnberger; Oliver Kutz; Simon Colton; Emilios Cambouropoulos; Alison Pease (2014) , "COINVENT: Towards a Computational Concept Invention Theory."

4. Natalya F. Noy and Deborah L. McGuinness, Ontology Development 101: A Guide to Creating Your First Ontology

Differentially Private Bayesian Programming

Gian Pietro Farina
g.p.farina@dundee.ac.uk

Marco Gaboardi
m.gaboardi@dundee.ac.uk

ABSTRACT

There has always been friction, due to privacy issues, between who devises and uses bayesian learning algorithms and who provides sensitive data from where to learn. With the establishment of differential privacy as the standard definition of privacy, former are now asked to prove to the latter that their algorithms meet this definition. Technical proofs are, most of the times, obscure to data owners. Moreover, the proofs are complicated and very error prone. A functional probabilistic programming language whose type system captures bayesian differentially private computations is proposed to ease this burden.

Keywords

Differential Privacy, Type Theory, Automatic Verification, Bayesian Inference

1. INTRODUCTION

By probabilistic programming language we mean one that a) treats random variables as first class types b) allows realizations of random variables and c) allows conditioning of random variables. Differential Privacy [3], on the other hand, gives strong guaranteed bounds on the increase of the harm to an agent when his data is part of the input w.r.t when it is not. Abstracting the property of a particular value forming part of the input to a general relation of neighboring on the input space, we get the following definition. Let $\epsilon \in \mathcal{R}^+$, $M : A \to B$ be an always terminating randomized program and $\phi \subseteq A \times A$ be a neighboring relation: we say that M is ϵ-differentially private w.r.t ϕ iff $\forall S \subseteq B. \forall a, a' \in A. a \phi a' \implies \Pr[M(a) \in S] \leq e^{\epsilon} \Pr[M(a') \in S]$. Intuitively what this definition means for an algorithm M is that: whatever breach in the privacy of an agent is considered, the probability of that breach happening increases only of a small factor when the agent's data is part of the input w.r.t when it's not; indeed ϵ should be thought as a small constant. Once a neighboring relation on the input space and a distance on the output space of the program M are defined, differential privacy can be seen as a relational property [2] of M. This allows automatic verification theory to help in proving differential privacy of M.

2. THE LANGUAGE

The programming language proposed is based on $Hoare^2$ [1]. $Hoare^2$ is built on approximate relational refinement types which are also adopted in this work. With this sort of types it is possible to reason about relational properties of programs, and hence capture differentially private computations. In this work, $Hoare^2$ is extended with the primitive **observe** in order to cope with conditioning of random variables to handle Bayesian inference. The informal operational semantics of **observe** is the following: given a prior probability distribution and an observation, **observe** should provide an updated posterior probability distribution. The language deals with both discrete and continuous probability distribution.

3. EXPERIMENTAL RESULTS

Two classical tasks in Bayesian learning are analyzed in the differential privacy framework and encoded in the language. The first one is the process of inferring in a differential private way the bias of a bernoulli random variable given a set of realizations of that random variable, the second one is differential private bayesian linear regression given a set of points in the plane. *Well typedness* of the programs guarantees the privacy of the single agent's input data. A standard way to make an algorithm differentially private is to add statistical noise to the output. Adding noise to the input is also an option. Both ways are analyzed and typechecked for the two tasks.

4. METHODS

Noise is added proportionally to the *sensitivity* of the process. The sensitivity of a function f is an upper bound on how much f can magnify the distance over neighbor inputs. The *accuracy* of the process is another important parameter: how far apart from the real answer is the result of a differentially private algorithm? Upper bounds on accuracy and sensitivity are provided for both the inference algorithms. For instance, it is proved that in the first tasks the sensitivity of the process cannot be made arbitrarily small as the size of the input grows. Similar upper bounds are proved for the bayesian linear regression process.

5. FUTURE WORK

A proof of soundness w.r.t language semantics for the typing rule of the **observe** is to be provided. More examples coming from machine learning literature should be analyzed, made differentially private and encoded in the language. An important issue which is planned as future work is the verification of the accuracy of a differentially private algorithm and in general of other properties such as robustness.

6. REFERENCES

[1] G. Barthe, M. Gaboardi, E. J. G. Arias, J. Hsu, A. Roth, and P. Strub. Higher-order approximate relational refinement types for mechanism design and differential privacy. In *Proceedings of the 42nd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2015*, pages 55–68.

[2] G. Barthe, B. Köpf, F. Olmedo, and S. Zanella Béguelin. Probabilistic relational reasoning for differential privacy. In *ACM SIGPLAN Notices*, volume 47, pages 97–110. ACM, 2012.

[3] C. Dwork. Differential privacy. In *Encyclopedia of Cryptography and Security*, pages 338–340. Springer, 2011.