Natarajan Meghanathan
Dhinaharan Nagamalai (Eds)

Computer Science & Information Technology

Second International Conference on Computer Science and Information
Technology (COSIT 2015) March 21~22, 2015,
Geneva, Switzerland

AIRCC

Volume Editors

Natarajan Meghanathan,
Jackson State University, USA
E-mail: nmeghanathan@jsums.edu

 Dhinaharan Nagamalai,
Wireilla Net Solutions PTY LTD,
Sydney, Australia
E-mail: dhinthia@yahoo.com

Typesetting: Camera-ready by author, data conversion by NnN Net Solutions Private Ltd., Chennai, India

Preface

Second International Conference on Computer Science and Information Technology (COSIT-2015) was held in Geneva, Switzerland, during March 21~22, 2015. International Conference on Software Engineering (SEC-2015), Second International Conference on Signal and Image Processing (SIGL - 2015), and Second International Conference on Artificial Intelligence and Applications (AIAPP-2015) were collocated with the COSIT-2015. The conferences attracted many local and international delegates, presenting a balanced mixture of intellect from the East and from the West.

The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding computer science and information technology and to establish new collaborations in these areas. Authors are invited to contribute to the conference by submitting articles that illustrate research results, projects, survey work and industrial experiences describing significant advances in all areas of computer science and information technology.

The COSIT-2015, SEC-2015, SIGL-2015, AIAPP-2015 Committees rigorously invited submissions for many months from researchers, scientists, engineers, students and practitioners related to the relevant themes and tracks of the workshop. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer review process which comprised expert reviewers. These reviewers were selected from a talented pool of Technical Committee members and external reviewers on the basis of their expertise. The papers were then reviewed based on their contributions, technical content, originality and clarity. The entire process, which includes the submission, review and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Committees led to an exciting, rich and a high quality technical conference program, which featured high-impact presentations for all attendees to enjoy, appreciate and expand their expertise in the latest developments in computer network and communications research.

In closing, COSIT-2015, SEC-2015, SIGL-2015, AIAPP-2015 brought together researchers, scientists, engineers, students and practitioners to exchange and share their experiences, new ideas and research results in all aspects of the main workshop themes and tracks, and to discuss the practical challenges encountered and the solutions adopted. The book is organized as a collection of papers from the COSIT-2015, SEC-2015, SIGL-2015, AIAPP-2015

We would like to thank the General and Program Chairs, organization staff, the members of the Technical Program Committees and external reviewers for their excellent and tireless work. We sincerely wish that all attendees benefited scientifically from the conference and wish them every success in their research. It is the humble wish of the conference organizers that the professional dialogue among the researchers, scientists, engineers, students and educators continues beyond the event and that the friendships and collaborations forged will linger and prosper for many years to come.

Natarajan Meghanathan
Dhinaharan Nagamalai

Organization

General Chair

Natarajan Meghanathan   Jackson State University, USA
Dhinaharan Nagamalai   Wireilla Net Solutions PTY LTD, Australia

Program Committee Members

Abdelhamid Abdelhadi Mansor   University of Khartoum, Sudan
Abdul Kadhir Ozcan   KTO karatay University, Turkey
Adam Przybylek   Gdansk University of Technology, Poland
Adnan Albar   King AbdulAziz University, Saudi Arabia
Afshari   Islamic Azad University, Iran
Aftab Alam   Kingdom of Saudi Arabia,Saudi Arabia
Ali Chaabani   National School of Engineering, Tunisia
Ali Hussein   Alexandria University, Egypt
Anand Bajpai   Al Sharq Studies Institute, Dubai
Andrews Samraj   Mahendra Engineering Colleges, India
Archana Kakade   G. H. Raisoni College of Engineering, India
Arti Khaparde   Maharashtra Institute of Technology,Pune
Baghdad Atmani   University of Oran 1, Algeria.
Balwant Sonkamble   Pune Institute of Computer Technology, India
Banu Gunel   ODTU Graduate School of Informatics, Turkey
Barbaros Preveze   Cankaya University,Turkey
Chih-Lin Hu   National Central University, Taiwan
Chun-Yi Tsai   National Taitung University, Taiwan
Dac-Nhuong Le   Vietnam National University, Vietnam
Dean Wu   Chung Jung Christian University, Taiwan
Deema Alathel   King Saud University,Saudi Arabia
Deepti Shrimankar   VNIT Nagpur, India
Deepti   G.H.Raisoni College of Engineering, India
Dias N. G. J   University of Kelaniya, Srilanka
Dong Huang   Chinese Academy of Sciences, China
Dorgi Netto   Universidade Federal da Paraíba, Brazil
Dossari, Khalid M   Saudi Aramco, Saudi Arabia
Eva Shalin   Pentecost University College, Ghana
Faheem Ahmed   Thompson Rivers University, Canada
Farah Jemili   University of Sousse, Tunisia
Fatih Korkmaz   Çankiri Karatekin University, Turkey
Habib Rasi   Shiraz University of Technology, Iran
Hania Aoudia   University Paris, France
Héldon José   Integrated Faculties of Patos (FIP), Brazil
Himanshu Monga   Shiva Institute of Engg & Technology, India
Isa Maleki   Islamic Azad University, Iran
Islam Atef   Alexandria University, Egypt

Israashaker Alani   GaziAntep University, Turkey
Izzat Alsmadi   Boise State University, USA
Jeremy (Zheng) Li   University of Bridgeport, USA
Kuldeep Kumar   National University of Singapore, Singapore
Laudson Souza   Integrated Faculties of Patos (FIP),Brazil
Mohamed-K HUSSEIN   Suez Canal University, Egypt
Mohammad Khubeb Siddiqui   Salman bin Abdulaziz University, Saudi Arabia
Mohammed A. AlGhamdi   Umm AlQura University, Saudi Arabia
Mohammed Erritali   Sultan Moulay Slimane University, Morocco
Mohammed   Salman Bin Abdulaziz Univ. , Saudi Arabia
Mujiono Sadikin   Universitas Mercu Buana, Indonesia
Ned   Al Hosn University, UAE
Neela Narayanan V   VIT University Chennai, India
Neeli R. Prasad   CTIF, Denmark
Nik Adilah Hanin Zahri   University of Perlis, Malaysia
Noureddine Hassini   University of Oran 1, Algeria
Omid Pavand   Queueing System and Stochastic Process, Iran
Prayla Shyry S   Sathyabama University, India
Rahil Hosseini   Azad University, Iran
Rajdeep Singh   Aerospace Research Institute, Iran
Rajput H   Indian Institute of Technology, India
Rohail Hassan   Universiti Teknologi Petronas (UTP), Malasiya
Roka imareef   Sebah University of Libya, Libya
Saad Darwish   University of Alexandria, Egypt
Seyyed Reza Khaze   Islamic Azad University, Iran
Shahrouz moaven   Sharif university of technology, Iran
Subil Mathew Abraham   IBM Global Solution Center,USA
Thillaikarasi   Annamalai University, India
Usman Ali Khan   The Islamia University of Bahawalpur ,Pakistan
Wajeb Gharibi   Jazan University, Saudi Arabia
Yahya Slimani   University of Manouba, Tunisia
Yemna Sayeb   Manouba Universit, Tunisia
Zan zika   University of Central Europe, Skalica, Slovakia

Technically Sponsored by

Networks & Communications Community (NCC)

Computer Science & Information Technology Community (CSITC)

Digital Signal & Image Processing Community (DSIPC)

Organized By

Academy & Industry Research Collaboration Center (AIRCC)

TABLE OF CONTENTS

Second International Conference on Computer Science and Information Technology (COSIT 2015)

International Conference on Software Engineering (SEC 2015)

Second International Conference on Signal and Image Processing (SIGL 2015)

Second International Conference on Artificial Intelligence and Applications (AIAPP 2015)

ROBUST REAL TIME FACE RECOGNITION AND TRACKING ON GPU USING FUSION OF RGB AND DEPTH IMAGE

Narmada Naik[1] and Dr.G.N Rathna[2]

[1]Department of Electrical Engineering,
Indian Institute of science, Bangalore, India
`nnreema22@gmail.com`
[2]Department of Electrical Engineering,
Indian Institute of science, Bangalore, India
`rathna@ee.iisc.ernet.in`

ABSTRACT

This paper presents a real-time face recognition system using kinect sensor. The algorithm is implemented on GPU using opencl and significant speed improvements are observed. We use kinect depth image to increase the robustness and reduce computational cost of conventional LBP based face recognition. The main objective of this paper was to perform robust, high speed fusion based face recognition and tracking.

The algorithm is mainly composed of three steps. First step is to detect all faces in the video using viola jones algorithm. The second step is online database generation using a tracking window on the face. A modified LBP feature vector is calculated using fusion information from depth and greyscale image on GPU. This feature vector is used to train a svm classifier. Third step involves recognition of multiple faces based on our modified feature vector.

KEYWORDS

GPU, OpenCL, Face detection, Tracking, LBP, Histogram, SVM, Depth Camera (Kinect).

1. INTRODUCTION

Face recognition is a challenging field in pattern recognition. Face recognition appears to offer more advantages than other biometrics [1-2]. Each biometric has its own merits and demerits. Face recognition can be done implicitly i.e. the face image can be captured from certain distance with less expensive equipment's, whereas iris recognition can't be done implicitly and requires expensive equipment's for recognition. Speech recognition is susceptible to noise, creating distortion on the original speech, whereas face recognition is less susceptible to noise. In the past, various methods of face recognition has been proposed and survey can be found in [3-6]. With the development of stochastic machine learning, there are lots of face recognition algorithms based on local features [7].

After being integrated with mobile services and social networking for real time application face recognition has received significant attention in recent years [8]. Despite progress in face recognition field, exact face recognition remains a challenging task in real time environment. In videos the face expression, lighting conditions and location varies from frame to frame [9-10].

The facial expression and pose are the most important aspect for video based face recognition systems. With change in expression or pose the accuracy of the recognition system decreases. In this paper we have done face recognition based on fusion of RGB and depth image. Fusion based region of interest removes undesired effects due to pose variations and illumination changes [11]. A typical video based face recognition system consist of face detection, face tracking, feature extraction, training and classification. Here, each individual need to be tracked from different view angle and need to be processed in real time. This real time face recognition system thereby process huge amount of data and requires to be fast. Face recognition algorithms are suitable for parallel processing. Therefore, in this paper the implementation was done on GPU, which is a parallel device using heterogeneous computing language OpenCL.

In this paper, we present online video based face recognition of multiple faces, using Kinect sensor. In the detection stage the face was detected on RGB camera and corresponding depth map of face comes from depth image. Rest of the paper is arranged as follows: Section 2 is overview of the algorithm, where we briefly discuss about detection of face using rgb camera, face tracking, feature extraction using rgb and depth image combined and SVM classification for multiple face recognition. In section 3 we elaborate GPU-CPU based implementation of Local Binary Pattern(LBP) feature extraction. Section 4 is Experimental results and we conclude our work in section 5.

2. OVERVIEW OF ALGORITHM

2.1 Face detection mapping of rgb image depth

Face detection is not straightforward because it has lot of variation, such as pose variation, image orientation, illumination condition and facial expression [12]. Many novel approaches have been already proposed. In this paper, face detection with RGB image was done using viola jones algorithm [13]. Then the corresponding depth image of the detected face in RGB was calculated using camera transformation and affine transformation. Camera transformation was implemented by applying one to one mapping from RGB detected face enclosed in a rectangle to its depth coordinate system by using Kinect sensor. The result obtained with this mapping gives us a rhombus enclosed detected face because of camera transformation [14]. Since we need a rectangle region of the detected face for LBP feature extraction, we performed affine transformation.

Affine transformation is defined as a function between affine spaces which keeps points, straight lines and planes [15]. It is defined as,

p'=x.u+y.v+t (1)

Where u and v are basis vectors and t is point, and u, v and t (basis and origin) forms a frame for an affine space, p' represent the affine transformation**.**

Pseudocode for RGB to Depth transformation.

```
For i< srcDepth.rows and
For j< srcDepth.cols do
depth=srcCopyD(i,j)
NuiImageGetColorPixelCoordinatesFromDepthPixelAtResolution()
if(x>0 && y>0 && x<grayImage.cols && y<grayImage.rows) then
if grayImage.(x,y)==255 then
grayImageDepth.(x,y)=255
push to a vector point
findContours();
drawContours();
fitEllipse(contour);
Rect r= boundingRect (Mat (contour));
```
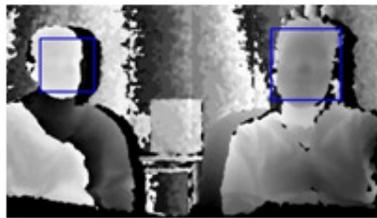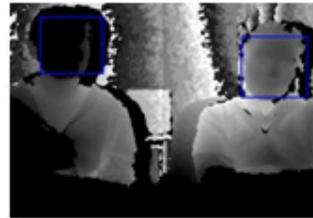

Fig.1(a)   Fig.1(b)

Fig.1(a) is depth image of corresponding RGB image and its face detected region, when object was far away from camera and Fig.1(b) when object was near to camera.

2.2 Face Labelling

Face labelling gives the proper identity of each individual in recognition. In this paper, each detected face is given a label. So in recognition part the face along with its label get displayed, which will help to distinguish between faces. Since, this paper presents multiple face recognition. Here, each individual face is given a particular label as shown in the Fig.2(a) and Fig.2(b).

Fig.2(a) Label for personA.   Fig.2(b) Label for personB.
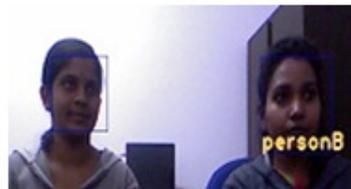
2.3 Face Tracking

The detected face is tracked in order to get multiple data set for single face which increases the accuracy of online face recognition. Person specific face tracking is a challenging task due to changes in environment, such as illumination, and also due to occlusions. Person with specific label or identity is to be tracked for a certain time period [9]. In this paper, tracking algorithm was based on nearest centre of the rectangle which encloses the face in the next frame. If c1 is the centre of rectangle as shown in Fig.4 where the detected face is enclosed, and c2 is the rectangle with detected face in next frame of same personA, whereas c3 and c4 are rectangles of the detected faces of person and personC, then the face of personA is tracked by calculating the nearest center among c2, c3 and c4 of latest frame to c1 in the previous frame. Thereby, all the detected faces are tracked. The data set at each frame gets updated during tracking.
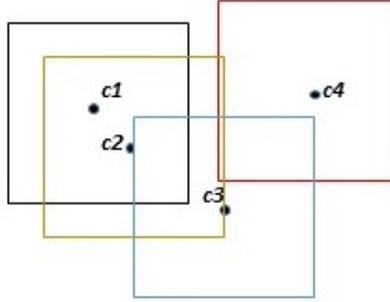
Fig.3 Rectangles obtained during face detection

Fig.3(a)Face tracking for personA

Fig.3(b)Face tracking for personB.

2.4 Feature extraction

Feature extraction is a form of reducing the dimension. Many holistic face descriptor are present such as, PCA(Principal Component Analysis), LDA(Linear Discrimant Analysis)[15], etc. In this paper LBP operator is used for feature extraction. The LBP operator was originally proposed by ojala[16]. LBP operator is one of the best texture descriptors. Its key advantage is, its invariant to monotonic grey level changes. The LBP operator initially threshold the 3x3 neighbourhood pixel based on the centre pixel, and get the decimal value as shown in **Fig.4**. The LBP operator is

defined as : $LBP_{P,R}(X_C, Y_C) = \sum_{p=o}^{P-1} S(g_p - g_c)2^p$ (2)

Is the intensity of the image at the Pth sample point where P is the total number of the sample point at a radius of R denoted by(P,R)[17]. The P spaced sampling points of the window are used to calculate the difference between centre gth and its surrounding pixel. Histograms reduces the feature vector length but still retaining 90% of data[17]. Histogram calculation method is given by:

Histogram calculation method is given by: $H_s(p,k) = \sum_{i=1}^{I}\sum_{j=1}^{J} f(LBP(P,R),P)$ (3)

where k is an integer to represent the sub-histograms i.e. obtained from each image k=1,2...K. K is the total no of histograms ,and $f(x,y) = \{ {}_{0,otherwise}^{1,x=y}$ where f(x, y) is the LBP calculated value at pixel(x, y).

Fig.4 LBP Calculation

Output of Viola Jones algorithm gives a rectangular window enclosing the face in image. This window has the face as well as some part of background. We use the depth image for this window to segment out face from background. At boundary between face and background depth changes abruptly. Thus using the depth image we can segment face from background. We assume that face is nearest object to the camera.

2.5 Training and classification

With the increase of face recognition application in real world, the accuracy of recognition and training plays a crucial role. Facial pose causes much difficulty in recognition. Therefore, in this paper we considered the fusion of RGB and depth image, where undesired effects of the faces image are removed. Here, each detected face after the feature has been extracted was trained using support vector machine [18]. Support vector machine is a classifier defined by a hyperplane. SVM is basically finding a hyperplane that gives the largest minimum distance to the training dataset.

In this paper while training personA was considered as class faces and personB face was considered as class nonfaces. Similarly, while training for personB it was considered as class faces and personA as class nonfaces. Classification result is shown in **Fig.5**.

Fig.5 Multiple recognized faces.

3. IMPLEMENTATION

The main goal of this paper was online multiple face recognition using both color and depth image information from Kinect sensor (Kinect camera). The implementation methodology outlines key activities are: Face detection, Labelling and Tracking, RGB to depth transformation, Training, Classification. The complete system of face recognition is shown in **Fig.6** Here, implementation of the algorithm is done using both OpenCL and OpenCV. LBP feature extraction is computationally expensive, OpenCL is used to accelerate the speed of the feature extraction for multiple data sets.

Here, the face recognition system uses both color and depth image information for detection, tracking, feature extraction. The detected face in color space is resized to 200x200 pixels and is mapped to its depth coordinate system. The detected face is labelled and tracked. Similarly, labelling and tracking for other faces are also done and simultaneously the database gets updated every time a new face is labelled.

After tracking and updating the database, the LBP feature extraction of the detected face and histogram was calculated using OpenCL in GPU. After LBP feature extraction histogram is calculated ranging from 0 to 255, for each face in GPU. Now, the histogram were used to for SVM training. Then it was tested, both training and testing was done on CPU
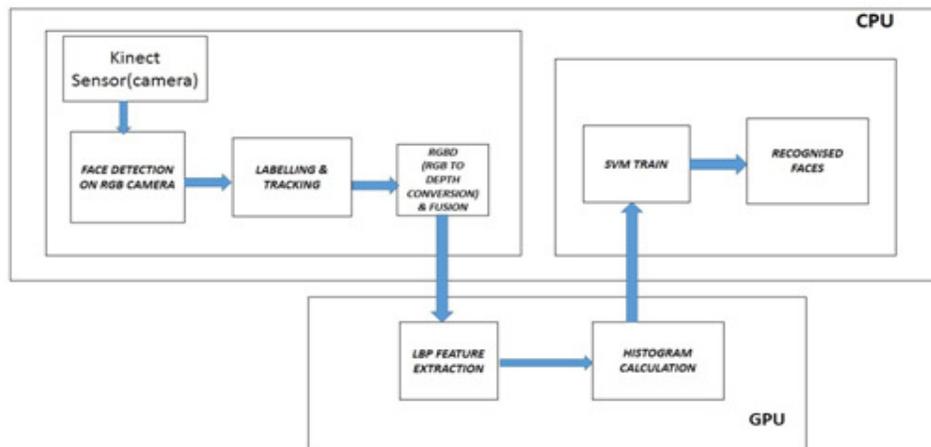
Fig.6 Block diagram of the Implementation of algorithm discussed.

4. EXPERIMENTAL RESULTS

Results of the experiments with tracking and without tracking using RGB only, and using RGB and depth for real time multiple face recognition is shown in **Table.1**. Based on the experimental results fusion based real time multiple face recognition on GPU is a robust face recognition system.

Table.1 Comparison Table

INPUT FOR FEATURE EXTRACTION	PERCENTAGE OF ACCURACY(WITH TRACKING)	PERCENTAGE OF ACCURACY(WITHOUT TRACKING)
RGB Image	70-75%	40-45%
FUSION Image	94-98%	50-60%
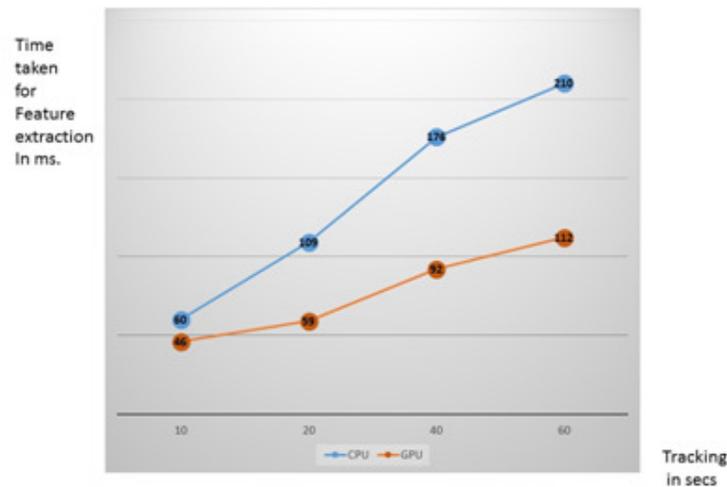
Fig.7 Performance chart for GPU and CPU for feature extraction, with tracking in secs.

In this paper, LBP feature extraction was implemented both on CPU and GPU to compare the performance. Experiment Results are shown in Fig.7.

5. CONCLUSIONS

This paper discuss implementation of fusion based robust real time multiple face recognition tracking and LBP feature extraction on GPU. All of the modules are the integral and important part in this algorithm for multi face recognition, because accurate face classification depends on this. The modules shown in Fig.6 was implemented on AMD 7670M GPU, i5 3rd generation CPU. Experiments have been performed and results indicate that our system is efficient and accurate for real time multiple face recognition.

REFERENCES

[1]   Bakshi, Urvashi, Rohit Singhal, and Monika Malhotra. "Biometric Technology: A Look and Survey at FaceRecogntion."

[2]   Bakshi, Urvashi, and Rohit Singhal. "A SURVEY ON FACE DETECTION METHODS AND FEATURE EXTRACTION TECHNIQUES OF FACE RECOGNITION."

[3]   J. Yang, D. Zhang, A.F. Frangi, J.-Y. Yang, Two-dimensional PCA: a new approach to appearance-based face representation and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 26 (1) (2004) 131–137

[4]   M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neurosci. 3 (1) (1991) 71–86.

[5]   Y. Xu, D. Zhang, J. Yang, J.-Y. Yang, An approach for directly extracting features from matrix data and its application in face recognition, Neurocomputing 71 (2008) 1857–1865.

[6]   J. Yang, D. Zhang, Y. Xu, J.-Y. Yang, Two-dimensional discriminant transform for face recognition, Pattern Recognition 38 (7) (2005) 1125–1129.

[7]   Vijayakumari, V. "Face Recognition Techniques: A Survey." World Journal of Computer Application and Technology 1.2 (2013): 41-50.

[8]   Dave, Guillaume, Xing Chao, and Kishore Sriadibhatla. "Face Recognition in Mobile Phones."

[9]   Kim, Minyoung, et al. "Face tracking and recognition with visual constraints in real-world videos." Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008.

[10] Shan, Caifeng, Shaogang Gong, and Peter W. McOwan. "Facial expression recognition based on local binary patterns: A comprehensive study." Image and Vision Computing 27.6 (2009): 803-816.

[11] Anapu, Srinivas Kishan, and Srinivasa Rao Peri. "Fusion of RGB and Depth Images for Robust Face Recognition using Close-Range 3D Camera."

[12] Wagner, Andrew, et al. "Toward a practical face recognition system: Robust alignment and illumination by sparse representation." Pattern Analysis and Machine Intelligence, IEEE Transactions on 34.2 (2012): 372 -386.

[13] Viola, Paul, and Michael J. Jones. "Robust real-time face detection." International journal of computer vision 57.2 (2004): 137-154.

[14] Henry, Peter, et al. "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments." The International Journal of Robotics Research 31.5 (2012): 647-663.

[15] Kannappan, Karthikeyan, and Charles C. Stearns. "Method for 2-D affine transformation of images." U.S. Patent No. 5,475,803. 12 Dec. 1995.

[16] Bhatt, Bhumika G., and Zankhana H. Shah. "Face feature extraction techniques: a survey." National conference on recent trends in engineering & technology. 2011.

[17] T. Ojala, M. Pietika¨inen, and D. Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions,"Pattern Recognition,vol. 29, no. 1, pp. 51-59, 1996.

[18] Ahonen, Timo, Abdenour Hadid, and Matti Pietikäinen. "Face recognition with local binary patterns." Computer vision-eccv 2004. Springer Berlin Heidelberg, 2004. 469-481.

[19] Naik, Narmada, and G. N. Rathna. "Real Time Face Detection on GPU Using OpenCL." Computer Science (2014).

Supervised Classification of Satellite Images To Analyze Multi-Temporal Land Use and Coverage: A Case Study For The Town of Marabá, State of Pará, Brazil

Priscila Siqueira Aranha[1], Flavia Pessoa Monteiro[1], Paulo André Ignácio Pontes[5], Jorge Antonio Moraes de Souza[2], Nandamudi Lankalapalli Vijaykumar[4], Maurílio de Abreu Monteiro[3] and Carlos Renato Lisboa Francês[1]

[1]Federal University of Pará (UFPA), Pará, Brazil
{priscilasa, flaviamonteiro, rfrances}@ufpa.br
[2]Federal Rural University of Amazonia (UFRA), Pará, Brazil
jorge.souza@ufra.edu.br
[3]Federal University of South and Southeast of Pará (UNIFESSPA), Pará, Brazil
maurilio.monteiro@unifesspa.edu.br
[4]National Institute for Space Research (INPE), São Paulo, Brazil
vijay.nl@inpe.br
[5]Federal Institute of Education, Science and Technology of Pará (IFPA),
Pará, Brazil
paulo.pontes@ifpa.edu.br

ABSTRACT

Amazon has one of the most diversified biome of the planet. Its environmental preservation has an impact in the global scenario. However, besides the environmental features, the complexity of the region involves other different aspects such as social, economic and cultural. In fact, these aspects are intrinsically interrelated, for example, cultural aspects may affect land use/land cover characteristics.

This paper proposes an innovative methodology to investigate changes of critical factors in the environment, based on a case study in the 26 de Março Settlement, in the city of Marabá, in the Brazilian Amazon. The proposed methodology demonstrated, from the obtained results, an improvement of the efficiency of the classification technique to determine different thematic classes as well as a substantial enhancement in the precision of classified images. Another important aspect is the automation in the process with a minimum computational effort in performing the required analyses

KEYWORDS

Image Classification, LULC, LULCC, LANDSAT 5 satellite, 26th March Settlement, TerraClass, Decision Tree, CART, MPL.

1. INTRODUCTION

In Amazon, actions of political nature have major social, economical and ecological significance in the rural settlement. Families that have been settled tell stories of success and failure[1]. Nowadays, in Legal Amazon there are 3554 settlement projects of the Brazilian Land Reform where 752,000 families live in 767 million hectares. 392 of these settlements have damaged forest coverage reaching around 50% to 75% of the lots, thus, affecting permanent preservation areas (APP) as well as areas of legal reservation (ARL). This has a tremendous impact. The Brazilian Government estimates that 75% of emitted CO_2 is due to change in land use of and forests. Fearnside points out that a part of the Amazon forest turning into pasture and agriculture is responsible, in the 1990s, for an annual emission of 1.20 Pg CO_2 [1].

Deforestation rates in Amazon are quite high and it is controversial to determine those that are responsible for such actions. According to IMAZON (Institute of Citizen and Environment of Amazon), 37% of the settlement areas have been deforested [2]. INCRA (Institute for Settlements and Land Reform) admits the problem and the situation is much worse in the State of Para.

Most of the research work shows that practices and processes in land use in Amazon, with emphasis on forest degradation, may affect regional and global climate and ecosystems [3].
So, the study of the dynamics of the land use and coverage in the settlements of the Amazon region turns out to be an important tool to determine and understand the reasons that lead to environmental degradation of such areas.

The work described in this paper proposes a methodology, using supervised classification, to determine land use and coverage (LULC) and changes in land use and coverage (LULCC) that took place during a defined period. The work uses multi-temporal analysis of LANDSAT 5 satellite images (TM Sensor). In order to demonstrate the efficiency of the approach, a case study used the area of 26th March Settlement of Marabá city, in the State of Pará (Brazilian Amazon).

2. RELATED WORK

Due to continental dimensions of the Amazon, remote sensing is indispensible to monitor the environment, detect fire and deforestation actions. In this context, methods and tools to classify satellite images have been in the spotlight to study the land use and coverage as well as changes occurred along these years, so that some public policies can be put into place to efficiently deter deforestation by aiding settled families to work on this issue but without losing their sustenance. Classification techniques can be supervised and non-supervised. In supervised classification, one has to know a priori a domain specialist information, while in the second, this is not required and may be obtained from existing maps or field work [4].

A methodology that takes advantage of supervised and non-supervised classifications is proposed in [5]. On the other hand, [6] and [7] compared Support Vector Machine-SVM, Artificial Neural Networks-ANN and Maximum Likelihood Classification-MLC to infer which could be the best technique to classify images. All the three showed to be equivalent with a 94% of precision to classify land coverage in coastal areas [6]. [7] showed that ANN had a better performance than SVM and MLC but requires a deeper understanding to clearly define in which situations ANN is in fact better as the authors expected that SVM would perform better. In [8], Classification and Regression Trees–CART, ANN and SVM were compared and showed that SVM could be more generalized than others when using small training samples. Precision obtained was between 70% and 80% for samples between 20 and 800 pixels. ANN produced around 67% to 76% of precision while CART 62% to 73%. MLC was applied to study land use and coverage of Yamuna River

bed in Delhi [9] and obtained a precision of around 77.1% and 88.6% with Kappa coefficient of 0.71 to 0.86.

MLC and Decision Tree were employed [10] to classify an area in Victoria, Australia and results showed a precision of 74% to 78% with Kappa coefficient of 0.70 and 0.75 respectively. Decision Tree algorithm had a better performance. [11] analyzed territorial dynamics of land coverage in the State of Rondônia, Brazil, also recognized as deforested area. The authors employed supervised classification based on Bhattacharya algorithm associated to image segmentation by Spring software. The study mapped seven different classes of land use and coverage besides identifying changes occurred in the area.

A study on Tourist State Park in Alto Ribeira, São Paulo State (also in Brazil) investigated the performance of different classification methods to map land coverage [12]. Two methods – hybrid classification per pixel (ERDAS 9.1 software) and classification based on objects (eCognition version 5 software) - were employed. Object-based classification performed better with a Kappa coefficient of 0.8687 (hybrid classification obtained a Kappa coefficient of 0.2224). The expected quality was obtained due to domain specialist knowledge during the classification process.

ANN was evaluated by [13] to classify images to analyze land use and coverage. It showed to be efficient, especially, when that input data were not normally distributed. For complex mapping applications, the authors recommend to employ supervised MLP (Multi Layer Perceptron) networks that could classify images with better precision. They also recommend non-supervised SOM (Self-Organizing Mapping) to analyze spectral features between and within the classes.

3. METHODOLOGY

A methodology has been defined to qualify the deforestation and then classify the satellite images by distributing each pixel in thematic classes. Maps from TerraClass project from 2008 to 2015 [14] were used.

The main objective of TerrClass [14] is to qualify deforestation in Legal Amazon based on deforestation mapped and published by PRODES project [INPE, 2008] and by satellite images. This is done by mapping land use and coverage and also by evaluation this dynamics. It consists of four main stages, executed manually and automatically. The first refers to data pre-processing based on acquiring satellite images. Then, GIS (Geographic Information System) is used to compose bands and geometric correction. Lastly, region of interest is extracted. They are manually performed in sequence for all the obtained images. Once the first phase is over, images are stored in a repository so that they go through classification as shown in Figure 1.

The next phases are automatically executed by an algorithm written in Python. It follows the steps from Figure 1 starting from executing the transformation phase to data normalization as well as definition of input variables to be used during the classification. Results from the classification go through pre-processing and presented in tabular and graphic forms with the percentage values of each area based on thematic class. Percentage figures of each area organized its thematic class are also shown along with the classified image. Details of the phases described in the Figure follow.
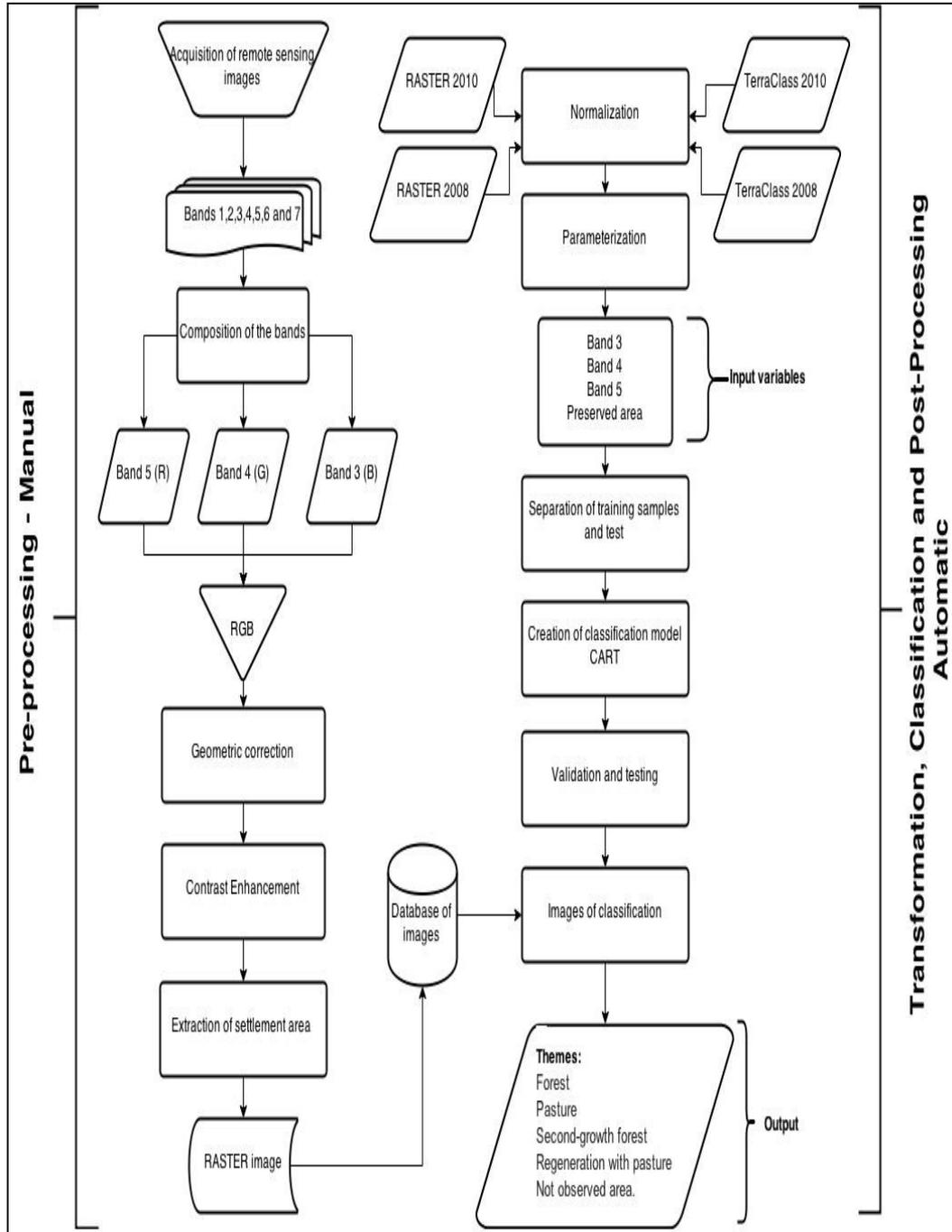
Figure 1. Model of the methodology

3.1. Study Area and Data

The study area corresponds to PA 26th March, created on 19/December/2008. It used to belong to Cabeceiras farm located in Marabá, meso-region of the Southeast of Pará (Figure 2). Its area is 11,919.36 ha with 6 clusters of houses and two headquarters besides 207 family lots [15].
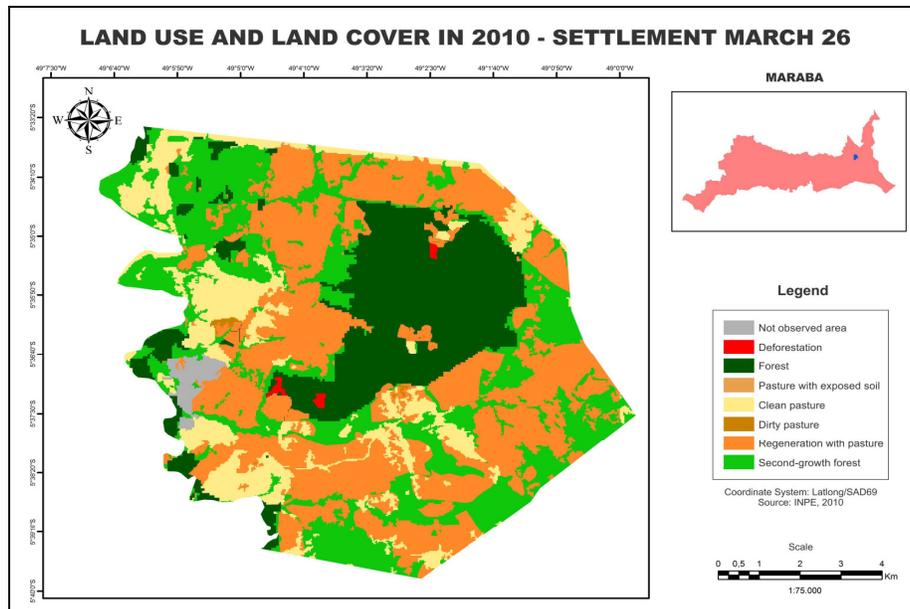
Figure 2. Study Area

In order to perform this study, LANDSAT 5 satellite images, TM sensor, with a geometric resolution of 30 meters for the bands 1, 2, 3, 4, 5 e 7 e 120 meters have been used. After the geo-processing, a database was created adopting a 1:25.000 scale with geographic coordinates SIRGAS 2000. Images were obtained from Brazilian Institute for Space Research (INPE). The collected images were selected according to the percentage of cloud coverage not going beyond 5% threshold. With this criterion, 12 images between 1986 and 2010 were selected also following an interval of 2 to 3 years. In order to minimize seasonal effects of the changes on the land use and coverage, the selected images are from the months between July and September.

3.2. Pre-processing

The original images obtained haven't been pre-processed and present noise due to atmospheric influence and geometric distortions. So, they have to be adjusted in order to avoid problems with quality and validity of extracted information. Once such inconsistencies are sorted out, multi-temporal datasets obtained can now be used in the normalization phase with much more reliability. So, the pre-processing is critical and influences on the quality/validity of the extracted information.

During the pre-processing, ArcGIS, version 10.0 was used. Geometric and radiometric corrections were applied to deal with the positioning and intensity of the pixel values that usually are frequent due to variation of altitude, velocity and atmospheric conditions in the satellite sensors [16]. With an objective of improving the quality of the image, contrast highlighting was applied to distinguish the elements of the image. Empirically, Minimum-Maximum was opted to manipulate the histograms. The values of 45 and 205 were considered as minimum and maximum respectively.

Spatial resolution had to be uniform in 30 m. So, resampling based on closest neighbor was used fixing the error to 0.5 pixel that might generate discontinuity on the resampled image, but without changing the radiometric values of the original image. To minimize the effects due to the combined use of different bands, the same number was used for all the images. In order to compose the bands, 5, 4 and 3 were selected, as they were more representative to distinguish

vegetation coverage and to map land use. ArcGIS software was employed to collect training samples and testing samples for the classification phase. TerraClass 2010 was the basis to select the samples. The idea was to qualify deforestation, mapping the land use and coverage in the Legal Amazon from LANDSAT 5 TM satellite images divided into their respective orbit-points.

TerraClass maps all the Legal Amazon and it is expected that the classification may present some inconsistencies when the region of interest is very small as it is the case for Settlement 26th March. Figure 2 shows an example of such an inconsistency. Highlighted area of image A refers to class Non Observed Area but it was classified as Regeneration with Pasture. This could have happened due to high intensity of the spectral return of this class. So, to minimize such distortions, samples were extracted in polygon shapes of those regions whose correspondence between the original and classified images was true.
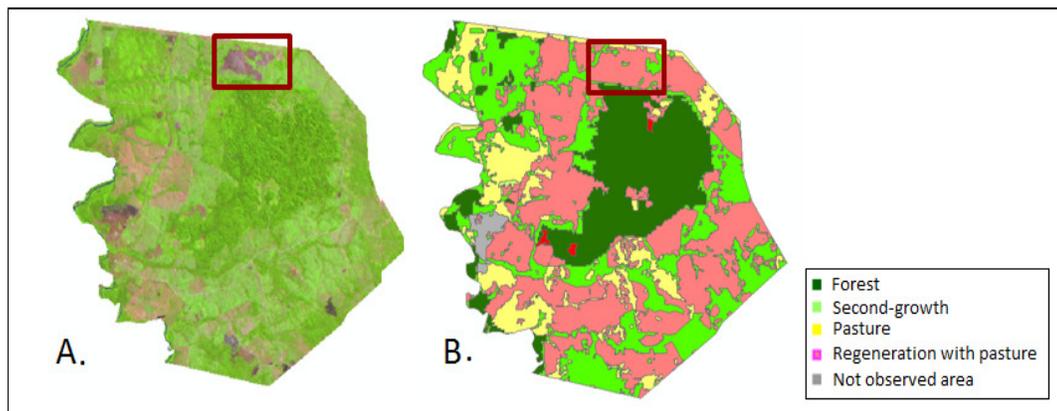
Figura 2. (A) Original Image   (B) Classified Image from TerraClass 2010

3.3. Data Transformation

As pointed out in the methodology, data transformation process is divided into two stages: normalization and parameterization. After the pre-processing phase, it is necessary to normalize so that information in the images may be understood during the classification.

For the normalization process, CSV formatted files were created containing information of the positioning of the pixel and gray levels of 3, 4 and 5 bands for those images considered for training and tests. For the parameterization, input values and output thematic classes were defined. Output classes were forest, secondary vegetation, pasture, regeneration with pasture and non-observed area (clouds and forest regional fires) due to their significance for the region of interest. Input values were the gray levels of bands 3, 4 and 5, and the variable "preserved area", of binary type with 0 or 1. To classify an image of the year x, for example, all the pixels of classified image for the year $x-1$ are verified. If the pixel corresponds to forest, 1 is assigned to the variable, otherwise 0. Its main objective is not to allow that deforested areas are classified as forest.

3.4. Supervised Classification

CART (Classification and Regression Trees) algorithm-based Decision Tree (DT) technique was employed to generate supervised classification model. This is because it associates three important characteristics: speed, precision and simplicity. Besides, it is easy to understand and basically depends on rules based on *if-then*.

Decision Trees represent knowledge and they are an efficient way to build classifiers to predict classes based on attribute values of a set of data. In a Decision Tree, knowledge is tested and represented in each node that conducts a search to its offspring, going down right from the root towards the leaves [17]. CART, proposed in [19] searches for relationships among data besides generating simple decision tree that is readable. To validate the results from employing DT, MLC (Maximum-Likelihood Classification) was compared as it is widely known and applied in several remote sensing problems [18]. Besides, most of the Geo-Processing tools include this algorithm.

3.5. Precision Evaluation

It is important to evaluate the quality of the classification results as it is concerned with extracting intended patterns from the input data. Here, confusion matrix and Kappa coefficient were used to express the precision.. Confusion matrix enables understanding the behavior of the classifier, reacting to effects of wrong prediction. Rows represent expected values and columns real ones. As per Kappa coefficient, it evaluates how good is the model. It is a discrete multi-variate technique that determines the level of the precision varying between 0 and 1. Closer to 1, better is the classification result.

Agreement Level	Values
Poor	0.20 <
Weak	0.21 – 0.40
Moderate	0.41 – 0.60
Good	0.61 – 0.80
Excellent	0.81 - 1

Table 1. Kapp Coefficient Values [19]

The test samples obtained the following results: general precision of the classification technique of DT was 97.7%, with Kappa coefficient 0.97 and the time to classify was 11 seconds. MLC technique yielded 82.75% with Kappa coefficient of 0.78 and within 4 minutes.

Class (%)	Forest		Secondary Vegetation		Pasture		Regeneration with Pasture		Non Observed Area	
	AD	MLC	AD	MLC	AD	MLC	AD	MLC	AD	MLC
Forest	98.75	83.53	0	0	0.76	28.57	0	0	0	0
Secondary Vegetation	0	1.18	97.81	76.67	0	1.50	0	4.35	3.97	7.94
Pasture	1.10	15.29	0	1.67	97.79	66.92	0.84	4.35	0.37	0
Regeneration with Pasture	0.15	0	0	15	1.39	3.01	98.38	91.30	1.08	0
Non Observed Area	0	0	2.19	6.67	0.06	0	0.78	0	94.58	92.06

Table 2. Confusion Matrix from Classification Techniques

Table 2 shows that MLC was inferior to DT with an emphasis on "Pasture" class that was wrongly classified with only 69.2% of correct classification. The difference is of 30.87% when compared to DT. For the "Non-Preserved Area", MLC was successful in 92.06% of the cases while DT 94.58%, smallest rate of success.

4. RESULTS

For this study, five classes were defined: forest, pasture, regeneration with pasture, secondary vegetation and non-observed area. After employing DT and MLC, it is clear that there is fragility in classifying small areas from TerraClass. It is important to point out that TerraClass is an excellent to tool to deal with areas of continental dimensions, such as Legal Amazon. DT showed to be much more efficient with respect to MLC obtaining a general precision of 97.7% and Kappa coefficient of 0.97. MLC obtained 82.75% and Kappa coefficient of 0.78. Figure 3 shows the results of the classification for the years 2008 and 2010 based on TerraClass, DT and MLC.
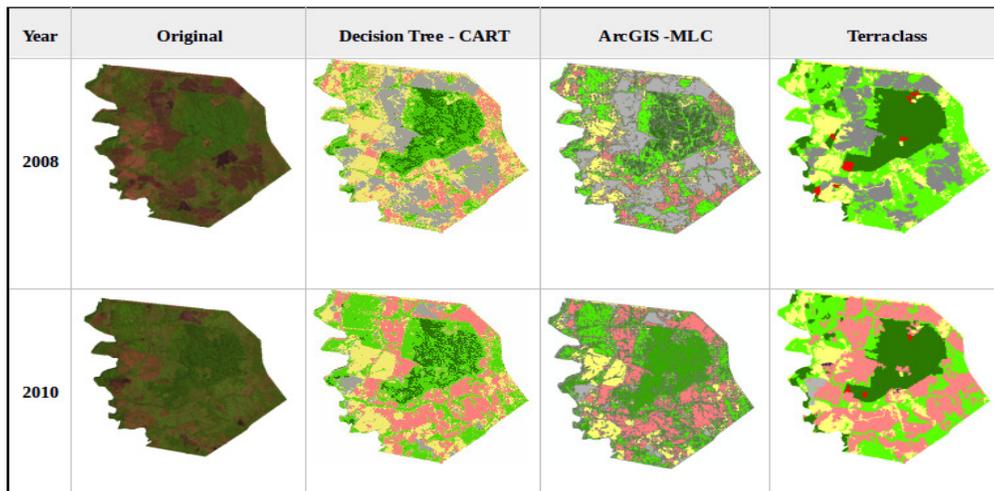
Figure 3. Original Test Image resulted from DT and TerraClass Classification from 2008 and 2009

Based on the classification from DT, one can realize that vegetation coverage of 26th March Settlement suffered a significant degradation, giving place to large areas of pasture. Such dynamics may be explained from the struggle for land ownership between the MST-Movement of Workers without Land and owners of Cabeceiras farm that lead to cutting down a great part of the native forest thus trying to show that the land was productive to avoid expropriation by government in the land reform program. In 2008, when the settlement was created, those that were settled received areas that were highly anthropic. But, deforestation remains, at a slower pace, of course. This occurs due to lack of resources of the settled families that are unable to maintain the pastures clean, leading to, time to time, deforesting new areas. The abandoned pastures give place to regeneration with pasture and at a later stage to secondary vegetation as can be observed in Figure 4.
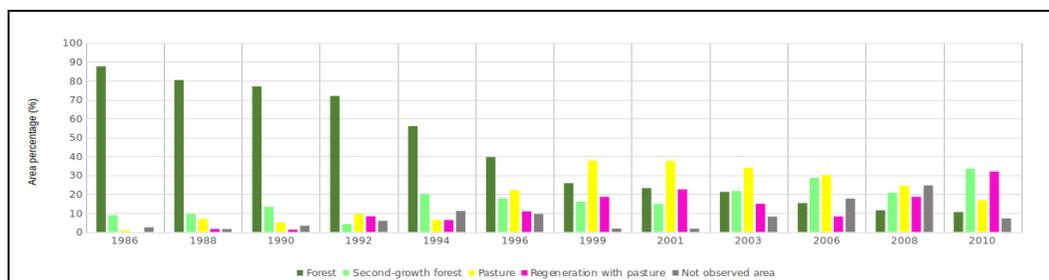
Figure 4. Dynamics of Use and Occupation by Settlement 26th March.

5. CONCLUSION

The study showed that Decision Tree classification technique is more precise than Maximum Likelihood Classification used by ArcGIS 10 tool. The entire process after pre-processing is automatic without the need to train the model every time an image goes through a classification. This approach optimizes time and computational effort. By employing TerraClass, a domain specialist was not required by using thematic classes (resulted from TerraClass) as outputs to construct supervised classification model. Based on the obtained results from DT, it was possible to observe that the landscape of 26th March Settlement is dominated by cultivated pastures and degradation of forest areas increased and directly related to this type of practice. The work also demonstrated with respect to [10] that got a precision of 78% with DT. The work is expected to continue by including more information, such as, other sensors to measure land quality, use of UAV (Unmanned Aerial Vehicle) and socio-economic information of the local population.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Fearside, P. M, (2000)"Greenhouse gas emissions from land use change in Brazil's Amazon region", In: Lal, R.; Kimble, J.M.; Stewart, B.A. Ed. Global climate change and tropical ecosystems. Boca Raton: CRC Press, 2000. p. 231-249. Advances in Soil Science.

[2]   Amintas Brandão Jr., Carlos Souza Jr.,Andréia Pinto & Paulo Amaral, (2013) "Situação dos Desmatamentos nos Assentamentos de Reforma Agrária no Estado do Pará.".

[3]   Gabriel de Oliveira e Elisabete Caria Moraes, (2003) "Classificação Digital do Uso e Cobertura da Terra na Amazônia com Base em Dados Orbitais ASTER.".

[4]   Sílvia Maria Dias Pedro Rebouças, (2011) "Metodologias de Classificação Supervisionada para Análise de Dados de Microarrays." Doutoramento em Estatística e Investigação Operacional (Especialidade de Probabilidades e Estatística)

[5]   Hossein Saadat, Jan Adamowski, Robert Bonnell, Forood Sharifi, Mohammad Namdar, Sasan Ale-Ebrahim, (2011) "Land use and land cover classification over a large area in Iran based on single date analyses of satellite imagery." ISPRS Journal of Photogrammetry and Remote Sensing.

[6]   Brian W. Szuster, Qi Chen, Michael Borger, (2011) "A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones." Applied Geography.

[7]   Prashant K. Srivastava, Dawei Han, Miguel A. Rico-Ramirez, Michaela Bray, Tanvir Islam, (2012) "Selection of classification techniques for land use/land cover change investigation." Advances in Space Research (a COSPAR publication).

[8]   Yang Shao, Ross S. Lunetta, (2012) "Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points." ISPRS Journal of Photogrammetry and Remote Sensing.

[9]   Rucha R., Mangesh Warthe, Sharad Dwivedi, Ritesh Vijay and Tapan Chakrabarti, (2011) "Monitoring changes in lad use land cover of Yamuna riverbed in Delhi: a multi-temporal analysis" International Journal of Remote Sensing.

[10] Liang ZHAI, Jinping SUN, Huiyong SANG, Gang YANG, Yi JIA, (2012) "Large area Land Cover Classification with LANDSAT ETM+ Images based on Decision Tree"

[11] Gabriel de Oliveira e Elisabete Caria Moraes, (2013) "Classificação Digital do Uso e Cobertura da Terra na Amazônia com Base em Dados Orbitais ASTER".

[12] T. Rittl, M. Cooper, R. J. Heck and M. V. R. Ballester, (2013) "Object-Based Method Outperforms Per-Pixel Method for Land Cover Classification in a Protected Area of the Brazilian Atlantic Rainforest Region."

[13] Hui Yuan, Cynthia F. Van Der Wiele, Siamak Khorram, (2009) "An Automated Artificial Neural Network System for Land Use/Land Cover Classification from Landsat TM Imagery." Remote Sens., 1, 243-265; DOI:10.3390/RS1030243.

[14] Lillesand, T.M. and Kiefer, R.W. (2000) "Remote Sensing and Digital Image Interpretation", Wiley, New York, 724 p.

[15] Allison Reynaldo da Costa Castro* Orlando dos Santos Watrin. (2013) "Análise espacial de áreas com restrição legal de uso do solo em projeto de assentamento no sudeste paraense." Geoinformação e Sensoriamento Remoto em Geografia. Geografia Ensino & Pesquisa, vol. 17, n. 2, maio./ago. DOI: 10.5902/22364994/10779

[16] Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). "Classification and Regression Trees. Wadsworth".

[17] Garcia, S.C, (2000)"O uso de árvores de decisão na descoberta de conhecimento na área da saúde." In: SEMANA ACADÊMICA, 2000. Rio Grande do Sul: Universidade Federal do Rio Grande do Sul.

[18] Landis, J. R. and Koch, G. G. (1977) "The measurement of ob- server agreement for categorical data." Biometrics. 33: 159– 174.

AUTHORS

Studying currently Masters sandwich at the National Institute for Space Research - INPE in Sao Jose Dos Campos.Master's Degree in Electrical Engineering-Applied Computing at the Federal University of Para - UFPA(2013).Has experience in wireless sensor networks and embedded systems.

Master students in Electrical Engineering in the field of Applied Computing at Federal University of Para.Graduated in Computer Engineering from the Community College of the Amazon(IESAM).Has experience in application development for Digital TV, Web systems development using Java and PHP and the study of network simulations using the Network Simulator 3

Graduated in History (1991),with a Masters in Development Planning(1996),both from the Federal University of Para(UFPA).She holds a doctorate in Sustainable Development in the Humid Tropics(2001),also by UFPA, institution where he is a professor since 2002.He has conducted research in which addresses economic, social and spatial implications of the conversion of natural resources into commodities,with special attention to the Amazon Brazil's east

Graduated in Data Processing Technology in the University of the Amazon(1996) and master's degree in Computational Mathematics from the University of Sao Paulo(2000).He is currently a doctoral student at the Federal University of Para and assistant professor in the Amazon Federal rural University.He has experience in computer science,with emphasis on data communication technologies,mainly in the following themes:performance evaluation and wireless sensor networks

PhD in Electronic Engineering(area:Computer) from the Technological Institute of Aeronautics(1999).Participated in the Post - doctoral program at University College Cork(UCC) in Coastal and Marine Resources Centre(CMRC),Cork, Ireland in the area of computer modeling coastal environment. He has experience in computer Science, with emphasis on performance Evaluation, Software Testing with Formal Specification, Time Series Analysis and Computational Modelling of coastal environment

Effective Teacher of the Federal Institute of Education, Science and Technological Para - IFPA. Have Graduation in Technology in Data Processing by the Para State University Center - CESUPA (2003). Specialist in Application Development for internet by the Federal University of Para- UFPA (2004). Master in Electrical Engineering in computing applied by UFPA (2013). PhD in Electrical Engineering in Computing applied by UFPA (2014)

Graduated in Bachelor of Computer Science from the Federal University of paru(1995), master`s degree in Computer Science and Computational Mathematics from the University of Sao Paulo (1998) and a PhD in Computer Science and Computational Mathematics from the University of Sao paulo (2001). Has experience in computer science, with emphasis on performance evaluation, acting on the following topics: Digital TV, Access Technologies, performance Markov models and simulation, correlation techniques (Bayesian networks)and optimization.

INTENTIONAL BLANK

USING SPECTRAL RADIUS RATIO FOR NODE DEGREE TO ANALYZE THE EVOLUTION OF SCALE-FREE NETWORKS AND SMALL-WORLD NETWORKS

Natarajan Meghanathan

Jackson State University, 1400 Lynch St, Jackson, MS, USA
natarajan.meghanathan@jsums.edu

ABSTRACT

In this paper, we show the evaluation of the spectral radius for node degree as the basis to analyze the variation in the node degrees during the evolution of scale-free networks and small-world networks. Spectral radius is the principal eigenvalue of the adjacency matrix of a network graph and spectral radius ratio for node degree is the ratio of the spectral radius and the average node degree. We observe a very high positive correlation between the spectral radius ratio for node degree and the coefficient of variation of node degree (ratio of the standard deviation of node degree and average node degree). We show how the spectral radius ratio for node degree can be used as the basis to tune the operating parameters of the evolution models for scale-free networks and small-world networks as well as evaluate the impact of the number of links added per node introduced during the evolution of a scale-free network and evaluate the impact of the probability of rewiring during the evolution of a small-world network from a regular network.

KEYWORDS

Eigenvalue, Spectral radius, Scale-free network, Small-world network, Node degree

1. INTRODUCTION

Network analysis and visualization of large complex real-world networks, ranging anywhere from social networks [1][2], co-authorship networks [3], Internet [4], World wide web [4] to biological networks [5] and etc is an actively researched area in recent years. The strength of network analysis is to abstract the complex relationships between the members of the system in the form of a graph with nodes (comprising of the constituent members) and edges (weighted or unit-weight as well as directed or undirected, depending on the nature of the interactions) and study the characteristics of the graph with respect to one or more metrics (like node degree, diameter, clustering coefficient and etc). The adjacency matrix A(G) of the network graph G essentially captures the presence of edges between any two vertices. For any two vertices i and j in graph G, the entry in the ith row and j^{th} column of A(G) is 1 if there is an edge from vertex i to vertex j and 0 otherwise. This paper focuses on undirected graphs (edge i-j exists in both the directions: i-->j as well as from j-->i) and node degree as the metric under study.

Spectral decomposition is a method of projecting the characteristics of a network graph in n-dimensions or directions (that are mutually perpendicular) where n is the number of vertices in

the graph. The projection in each direction is represented in the form of a scalar value (called the eigenvalue) and its corresponding vector with entries for each vertex (called the eigenvector). Though the number of dimensions in the spectrum is the number of vertices in the graph, most of the variations could be captured in the first few dimensions of the coordinate system represented by the eigenvalues and the eigenvectors. The largest eigenvalue of the projection is called the principal eigenvalue and the corresponding eigenvector is called the principal eigenvector (could be determined by executing the power iteration algorithm [6] on the adjacency matrix of a network graph). The principal eigenvalue and its corresponding eigenvector capture maximum amount of variability in the data (in the case of a network graph, the data are the edges connecting the vertices). In this paper, we make use of the principal eigenvalue (also called the spectral radius) of the adjacency matrix of complex network graphs to analyze the variations in node degree and correlate with the coefficient of variation of node degree. We specifically use the spectral radius ratio for node degree as the basis to evaluate the evolution of two commonly studied complex network models, such as the scale-free networks and small-world networks. To the best of our knowledge, we could not come across any related work that uses spectral radius as the basis to analyze the evolution of scale-free networks and small-world networks.

The rest of the paper is organized as follows: In Section 2, we introduce the power iteration method that was used in this research to calculate the spectral radius of a network graph. Sections 3 and 4 validate our hypothesis on scale-free networks and small-world networks respectively. Section 7 concludes the paper.

2. POWER ITERATION METHOD TO CALCULATE SPECTRAL RADIUS

The power iteration method can be used to calculate the principal eigenvalue (i.e., spectral radius) and the corresponding principal eigenvector of a graph based on its adjacency matrix. The eigenvector X_{i+1} of a network graph at the end of the $(i+1)^{\text{th}}$ iteration is given by: $X_{i+1} = AX_i / \|AX_i\|$, where $\|AX_i\|$ is the normalized value of the product of the adjacency matrix A of a given graph and the tentative eigenvector X_i at the end of iteration i. The initial value of X_i is [1, 1, ..., 1], a column vector of all 1s, where the number of 1s correspond to the number of vertices in the graph. We continue the iterations until the normalized value $\|AX_{i+1}\|$ converges to that of the normalized value $\|AX_i\|$. The value of the column vector X_i at this juncture is declared the principal eigenvector of the network graph and the normalized value to which the iterations converge is the principal eigenvalue (i.e., the spectral radius). Figure 1 illustrates an example for computation of the spectral radius on a network graph. The converged normalized value of 2.21 is the spectral radius of the graph (denoted $\lambda_{\text{sp}}(G)$). If k_{min}, k_{avg} and k_{max} are the minimum, average and maximum node degrees, then, $k_{min} \leq k_{avg} \leq \lambda_{\text{sp}}(G) \leq k_{max}$ [7].

Iteration 1

$$
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \\ 2 \end{bmatrix} \equiv \begin{bmatrix} 0.213 \\ 0.426 \\ 0.426 \\ 0.639 \\ 0.426 \end{bmatrix}
$$

Normalized Value = 4.69

Iteration 2

$$
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.213 \\ 0.426 \\ 0.426 \\ 0.639 \\ 0.426 \end{bmatrix} = \begin{bmatrix} 0.426 \\ 0.852 \\ 1.065 \\ 1.278 \\ 1.065 \end{bmatrix} \equiv \begin{bmatrix} 0.195 \\ 0.389 \\ 0.486 \\ 0.584 \\ 0.486 \end{bmatrix}
$$

Normalized Value = 2.19

Iteration 3

$$
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.195 \\ 0.389 \\ 0.486 \\ 0.584 \\ 0.486 \end{bmatrix} = \begin{bmatrix} 0.389 \\ 0.779 \\ 1.07 \\ 1.361 \\ 1.07 \end{bmatrix} \equiv \begin{bmatrix} 0.176 \\ 0.352 \\ 0.484 \\ 0.616 \\ 0.484 \end{bmatrix}
$$

Normalized Value = 2.21

Iteration 4

$$
\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0.176 \\ 0.352 \\ 0.484 \\ 0.616 \\ 0.484 \end{bmatrix} = \begin{bmatrix} 0.352 \\ 0.792 \\ 1.100 \\ 1.320 \\ 1.100 \end{bmatrix}
$$

Normalized Value = 2.21 converges

Figure 1: Example to Illustrate the Execution of the Power Iteration Method to Compute the Spectral Radius of a Network Graph

3. SIMULATION AND ANALYSIS OF SCALE-FREE NETWORKS

A scale-free network is a network wherein a majority of the nodes have a very small degree and very few nodes (but appreciable number of nodes) have a very high degree [4]. An airline network is typically a scale-free network with certain airports being used as high-degree hubs with connections to other hubs as well as connections to nearby low-degree nodes/airports. The simulations for scale-free networks were conducted as follows: We use the Barabasi-Albert (BA) model [8] for generating scale-free networks. According to this model, to start with, the network is assigned an initial set of nodes (3, in our simulations): the links between which are chosen arbitrarily, as long as there is one link per node (accordingly, we consider each node for link inclusion and connect the node to a randomly chosen node, other than itself, from the initial set of nodes). We then introduce nodes, one at a time. At each time step, a new node is introduced to the network with m number of new links (varied from 2, 3, 4, 5, 10, 20, 30, 40, 50 and 90) connecting to the nodes that are already exist in the network (at most one new link per node). If the number of new links that could be added to the network is less than the number of nodes in the network, then the number of new links is set to the number of nodes in the network (in such cases, the newly introduced is basically connected to every node that already exist in the network). Once the number of new links added per node becomes less than the number of nodes that already exist in the network, then the node to pair with is chosen according to the probability formula that denotes the probability at which an already existing node i with $\prod(k_i^t) = \dfrac{k_i}{\sum\limits_{j=1}^{t-1} k_j}$ degree k_i at time instant t (corresponds to the introduction of node with ID t) acquires a link. We compute such probabilities for the nodes that were introduced prior to time instant t (i.e., node IDs 1,..., t-1). Note that the denominator in the probability formula sums only the degrees of the nodes that already exist in the network before the introduction of the new node and its links (i.e., the new node and its impact on the node degrees is not considered in the denominator portion of the formula). The values reported in Figures 2-5 and Tables 1-2 for the Spectral Radius Ratio for Node Degree (ratio of the spectral radius of the adjacency matrix and the average node degree) and the coefficient of variation of node degree (ratio of the standard deviation to the average node degree) are average values obtained from 100 runs of the simulation code for a particular condition (initial number nodes, total number of nodes and number of new links added per node introduction).

Table 1: Spectral radius ratio for node degree vs. coefficient of variation of node degree after the introduction of the last node in the network (Scale-free model for network evolution - initial # nodes: 3, total # nodes: 1000)

	Number of new links added per node at the time of node introduction									
	2	3	4	5	10	20	30	40	50	90
$\lambda_{sp}(G)/k_{avg}$	2.63	2.35	2.09	2.09	1.83	1.64	1.55	1.48	1.43	1.31
k_{SD}/k_{avg}	1.31	1.19	1.07	1.08	0.92	0.79	0.73	0.69	0.64	0.54

Table 2: Spectral radius ratio for node degree vs. coefficient of variation of node degree after the introduction of the last node in the network (Scale-free model for network evolution - initial # nodes: 3, total # nodes: 100)

	Number of new links added per node at the time of node introduction									
	2	3	4	5	10	20	30	40	50	90
$\lambda_{sp}(G)/k_{avg}$	1.75	1.53	1.43	1.45	1.28	1.15	1.09	1.05	1.03	1.00
k_{SD}/k_{avg}	0.97	0.80	0.69	0.69	0.55	0.39	0.31	0.23	0.17	0.02

The spectral radius ratio for node degree could be used as a measure to determine the appropriate number of new links per node (that could be added to the network upon introduction of a node into the network) to achieve a desired threshold level of variation in node degree. As we divide the spectral radius by the average node degree, the spectral radius ratio for node degree can be used as a measure to judge the variation of node degree across networks of any size. Tables 1 and 2 display the spectral radius ratio for node degree and coefficient of variation of node degree after the introduction of the last node in the network. As we can see, for a fixed value of the number of new links added per node and the initial number of nodes, the smaller the value of the total number of nodes to be introduced to the network, the lower the variation in node degree. As we increase the number of new links added per node introduction, the distribution of links is relatively more even across all the nodes, especially when the total number of nodes in the network is lower.

Table 3: Correlation coefficient between spectral radius ratio for node degree vs. coefficient of variation of node degree as a function of time - with the introduction of nodes in the network (Scale-free model for network evolution - initial # nodes: 3)

Total # Nodes	Number of new links added per node at the time of node introduction									
	2	3	4	5	10	20	30	40	50	90
1000	0.974	0.987	0.986	0.969	0.978	0.976	0.974	0.972	0.969	0.967
100	0.981	0.989	0.987	0.973	0.958	0.942	0.866	0.796	0.734	0.934
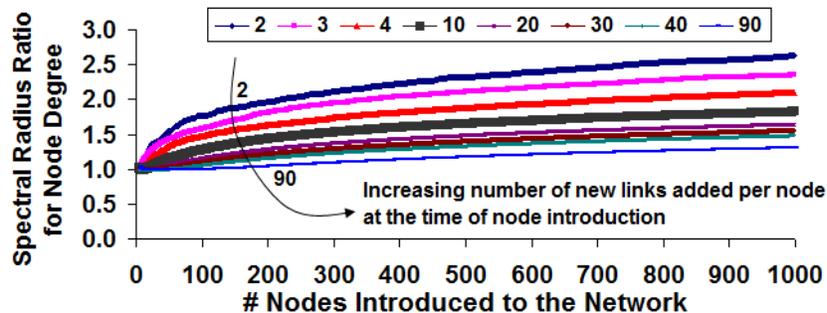
Figure 2: Distribution of the spectral radius ratio for node degree under the scale-free model for network evolution (initial # nodes: 3; total # nodes: 1000)
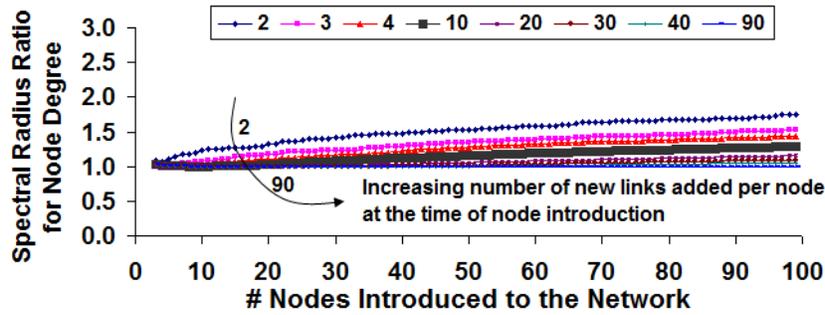
Figure 3: Distribution of the spectral radius ratio for node degree under the scale-free model for network evolution (initial # nodes: 3; total # nodes: 100)

Figures 2-3 and 4-5 respectively capture the distribution of the spectral radius ratio for node degree and the coefficient of variation of node degree as a function of time (upon the introduction of each new node into the network). After the initial set of nodes are assigned at least one link (per node), we add new nodes to the network, one at a time, as per the probability formula for an existing node to get a new link as a function of its degree. With passage of time, nodes that have been around in the network for a longer time incur a larger degree compared to nodes that were recently added; such a variation in node degree increases with time. Thus, we could see both the spectral radius ratio for node degree and the coefficient of variation of node degree to exhibit a very high correlation and similar pattern of distribution with time.

For a fixed value of the number of new links added per node and the initial number of nodes, the larger the value for the total number of nodes to be introduced to the network, the larger the space of distribution (range of values) for the two measures as a function of time as well as larger the correlation coefficient between the two measures. In other words, larger the scale-free nature of the network, the larger the correlation coefficient between the spectral radius ratio for node degree and the coefficient of variation of node degree. Table 3 displays the values for the correlation coefficient computed using the distribution of the spectral radius ratio for node degree and the coefficient of variation [9] of node degree as a function of time, with the introduction of new nodes into the network (for fixed values of the number of new links added per node introduction and the initial number of nodes). Apparently, when then the spectral radius ratio for node degree approaches close to 1 (scenario of total of 100 nodes with 90 new links per node introduction), the correlation coefficient between the two measures (that was henceforth decreasing with increase in the number of new links added per node) bounces back and approaches 1.
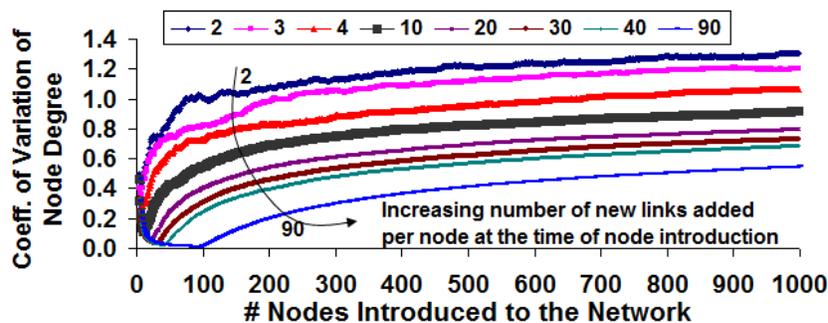
Figure 4: Distribution of the coefficient of variation of node degree under the scale-free model for network evolution (initial # nodes: 3; total # nodes: 1000)
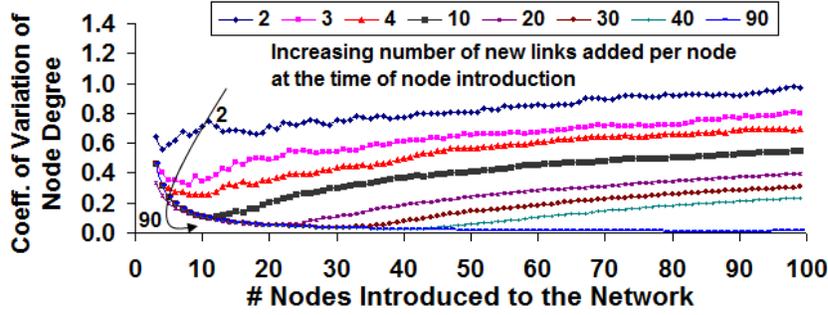
Figure 5: Distribution of the coefficient of variation of node degree under the scale-free model for network evolution (initial # nodes: 3; total # nodes: 100)

4. SIMULATION AND ANALYSIS OF SMALL-WORLD NETWORKS

A small-world network is a type of graph in which most nodes are not neighbors of one another, but most nodes can be reached from every other by a small number of hops [4]. An example for a small-world network is the road network of a country, wherein most of the cities are not neighbors of each other, but can be reached from one another by going through a fewer number of intermediate cities. The simulations for small-world networks are conducted according to the Watts and Strogatz model [10], explained as follows.

We start with a regular 1-dimensional ring network comprising of two concentric rings, each with the same number of nodes (say, N; hence, the total number of nodes in the network is referred to as $2N$). The nodes are initially organized in a ring network as shown in the left side of Figure 6. The number of neighbors per node is 4. For a node i in the outer ring, its neighbors in the outer ring are chosen as follows: $(i + 1) \% N$ and $(i - 1 + N) \% N$. For a node i in the inner ring, its neighbors in the inner ring are chosen as follows: $N + ((i + 1) \% N)$ and $N + ((i - 1 + N) \% N)$. In the initial regular ring network, as each node has 4 neighbors, the average number of neighbors per node is 4; with a total of $2N$ nodes, the total number of links in the initial regular ring network is (average node degree * total number of nodes / 2) = (4 * 2N / 2) = 4N = 2*total number of nodes in the network.
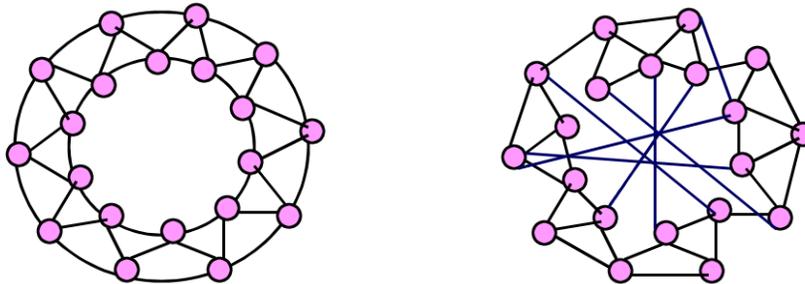
Regular Network (before rewiring)   Small-World Network (after rewiring)
Figure 6: Transformation of a regular network to small-world network

We consider each link in the initial set of links for a possible rewiring with a probability β. For each link (u-v), where $u < v$ (in the order of the numerical IDs), we generate a random number (uniform-randomly distributed from 0 to 1) and if the random number is less than or equal to β, the node with the lower ID u is paired with a randomly chosen node (other than itself, node v and

the current neighbors) and the newly created link is not considered for any rewiring. This way, the average number of neighbors per node at any time is still the initial value for the average number of neighbor per node (which is 4 in our simulations); but, variation in the distribution of the node degree starts creeping in with rewiring. The network at the right side of Figure 6 displays a sample network at the end of rewiring.
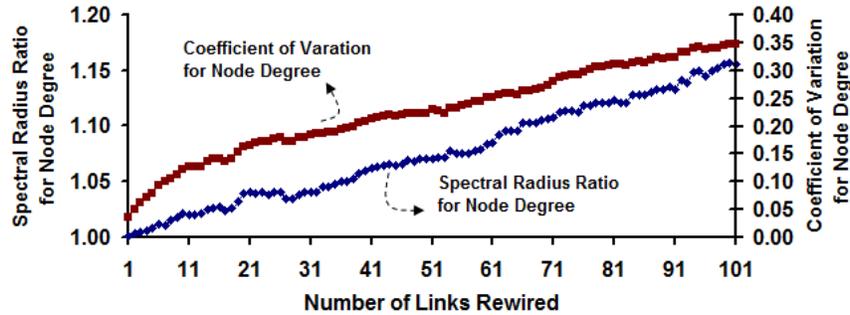
Figure 7: Distribution of variation in node degree during the process of rewiring (total # nodes: 100; rewiring probability: 0.5)
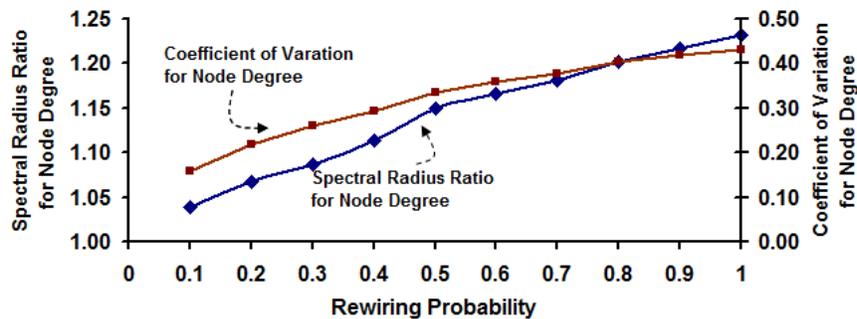
Figure 8: Distribution of variation in node degree as a function of the rewiring probability (averaged across the total # nodes from 100 to 1000)

We calculated the spectral radius of the network during the course of the rewiring of the links in the network (i.e., whenever a link is rewired). For a fixed total number of nodes and rewiring probability, as more links get rewired, the spectral radius of the network increased and the coefficient of variation in the node degree also increased (the correlation coefficient always remained above 0.95 during the process of rewiring). Figure 7 illustrates the distribution of the spectral radius for node degree and the coefficient of variation of node degree during the process of rewiring for a network with a total of 100 nodes (i.e., # nodes per ring N is 50) and a rewiring probability of 0.5 - about 50% of the total number of links (0.5 * 2 * 100 = 100) got rewired and the figure displays the increase in the two metrics with rewiring.

For a fixed probability of rewiring, the variation in the node degree is about the same when the value for the total number of nodes in the network is varied (from 100 to 1000 in the simulations). In other words, for any fixed value of the rewiring probability, the spectral radius ratio for node degree observed at the end of rewiring a network of 100 nodes is about the same as the spectral radius ratio for node degree observed at the end of rewiring a network of 1000 nodes. This is a significant observation as the total number of nodes we start with does not impact the variation in the node degree during the course of rewiring (for a fixed probability of rewiring).
For a fixed value of the total number of nodes, smaller the value for the rewiring probability, smaller the variation in the node degree, as reflected in Figure 8 (averaged for each rewiring

probability, based on the observations for the total number of nodes from 100 to 1000, as we do not see significant variation in node degree with increase in the number of nodes for a fixed probability of rewiring). We observe the spectral radius ratio for node degree to increase from 1 to 1.15 as we increase the rewiring probability from 0 to 0.5; but, as we increase the rewiring probability from 0.6 to 1 (the network being completely random), the spectral radius ratio for node degree increased only from 1.15 to 1.25.

5. CONCLUSIONS

The high-level contribution of this research is the identification of a positive correlation between the ratio of the spectral radius to the average node degree to that of the coefficient of variation of node degree for networks simulated from theoretical models (for scale-free networks and small-world networks). The positive correlation has been observed for networks of different sizes, ranging from networks of 5 nodes, to a few hundred nodes, all the way to networks of thousands of nodes. With this observation, we can now confidently say that the closer the value of the spectral radius ratio for node degree to 1.0, the smaller the variation among the degrees of the nodes in the network. The variation in the node degrees for any two networks can also be simply compared based on the spectral radius ratio for node degrees observed for the two networks, rather than requiring to compute the standard deviation of the node degrees for the two networks. As explained in Sections 3 and 4, the operating parameters of the theoretical models can be tuned to obtain a desired variation in the degree of the nodes in the network, measured in the form of the spectral radius ratio for node degree.

REFERENCES

[1]   D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, "The Bottlenose Dolphin Community of Doubtful Sound Features a Large Proportion of Long-lasting Associations," Behavioral Ecology and Sociobiology, vol. 54, pp. 396-405, 2003.
[2]   W. W. Zachary, "An Information Flow Model for Conflict and Fission in Small Groups," Journal of Anthropological Research, vol. 33, no. 4, pp. 452-473, 1977.
[3]   M. E. J. Newman, "The Structure of Scientific Collaboration Networks," Proceedings of the National Academy of Sciences of the United States of America, vol. 98, no. 2, pp. 404-409, January 2001.
[4]   M. Newman, Networks: An Introduction, 1st ed., Oxford University Press, May 2010.
[5]   M. Girvan and M. E. J. Newman, "Community Structure in Social and Biological Networks," Proceedings of the National Academy of Sciences of the United States of America, vol. 19, no. 12, pp. 7821-7826, April 2002.
[6]   G. Strang, Linear Algebra and its Applications, Cengage Learning, 4th edition, July 2005.
[7]   B. G. Horne, "Lower Bounds for the Spectral Radius of a Matrix," Linear Algebra and its Applications, vol. 263, pp. 261-273, September 1997.
[8]   A-L. Barabasi and R. Albert, "Emergence of Scaling in Random Networks," Science, vol. 286, pp. 509-512, October 1999.
[9]   J. Benesty, J. Chen, Y. Huang, I. Cohen, "Pearson Correlation Coefficient," Springer Topics in Signal Processing: Noise Reduction in Speech Processing, vol. 2, pp. 1-4, March 2009.
[10] D. J. Watts and S. H. Strogatz, "Collective Dynamics of 'Small-World' Networks," Nature, vol. 393, pp. 440-442, June 1998.

CLUSTERING ASSISTED FUNDAMENTAL MATRIX ESTIMATION

Hao Wu and Yi Wan

School of Information Science and Engineering, Lanzhou University, China
`wuhao11905@163.com, wanyijs@163.com`

ABSTRACT

In computer vision, the estimation of the fundamental matrix is a basic problem that has been extensively studied. The accuracy of the estimation imposes a significant influence on subsequent tasks such as the camera trajectory determination and 3D reconstruction. In this paper we propose a new method for fundamental matrix estimation that makes use of clustering a group of 4D vectors. The key insight is the observation that among the 4D vectors constructed from matching pairs of points obtained from the SIFT algorithm, well-defined cluster points tend to be reliable inliers suitable for fundamental matrix estimation. Based on this, we utilizes a recently proposed efficient clustering method through density peaks seeking and propose a new clustering assisted method. Experimental results show that the proposed algorithm is faster and more accurate than currently commonly used methods.

KEYWORDS

Fundamental Matrix, Clustering, Density Peaks

1. INTRODUCTION

Fundamental matrix (F-matrix) is an important tool often used in many computer vision tasks. It reflects the corresponding relationship between two pictures shot at the same scene but taken from different viewpoints. It is widely used in tasks such as camera tracking, image rectification, and 3D reconstruction (e.g., [1], [2]). The fundamental matrix constrains the coordinates between two 2D images from two different viewpoints. It contains information about the cameras' focal lengths, the positions of optical centers, and the rotation and translation between the two cameras [1].

In order to calculate the F-matrix, using the fact that it contains 7 free variables, Hartley used 7 pairs of corresponding points, calculated 3 possible F-matrices, and called it 7-Point algorithm [2]. Tsai and Huang proposed a linear algorithm [3], called the 8-Point algorithm for computing the F-matrix. It uses 8 pairs of matching points and estimates the F-matrix by solving a set of linear equations and making the F-matrix subject to a rank-2 constraint. There are also some other nonlinear methods. For example, Mohr and others assumed that the errors of different corresponding points are mutually independent and have the same standard deviation. They calculated the spatial position of the 3D points and the F-matrix by minimizing the error between the observed values and the fitted values of projected 2D coordinates from the 3D points [4].

A problem with current methods is that there are usually noise (error of 1 or 2 pixels) and point pairs that are wrongly matched. But all the algorithms mentioned above need to be calculated by

the least square method, which requires that the noise mean is 0, and that all matched data points are correct when estimating the F-matrix. Thus the above mentioned algorithms are susceptible to noise and incorrectly matched point pairs.

Some robust algorithms have been proposed to solve these problems. Zhang put forward an algorithm called M-Estimators that can distinguish inliers from outliers. This algorithm could adapt to noise but not the wrongly-matched points [5]. Rousseeuw and Leroy proposed a Least Median of Squares (LMeDS) algorithm. It randomly selects 7 pairs of corresponding points to calculate the fundamental matrix and the error function, then uses an adaptive threshold related to error function to find the outliers. The result is shown to be quite encouraging [6].

2. PROPOSED ALGORITHM

In this section we propose a new F-matrix computation method. We briefly review some preliminary topics in the first three subsections for completeness, then uses another three subsections to present the proposed algorithm.

2.1. Epipolar Geometry and Fundamental Matrix

Epipolar geometry describes the relationship between two images taken by two cameras C and C' pointing at the same scene. It is independent of the structure of the scene and is only related to the intrinsic and extrinsic parameters of the cameras.

As illustrated in Fig. 1, I and I' are the image planes of camera C and C' respectively. Their coordinate systems are set up as is commonly done. Specifically, origin point is located at the top-left of the image while the x-axis points to the right and the y-axis points to the bottom. The points $m = (x, y, 1)^T$ and $m' = (x', y', 1)^T$ are the image points in plane I and I' of the same 3D point M. Because C, M and m are on a line while C', M, m' are on another line, the six points are on a plane π, which is commonly called the epipolar plane. The image point of C on the plane I' is called the epipole e' while e is the image point of C' on the plane I. e and e' are also on the plane π. The line $l : ax + by + c = 0$ which crosses m, e and the line $l' : a'x + b'y + c' = 0$ are called the epipolar lines.

The fundamental matrix F is a 3*3 matrix with rank 2. It satisfies the following constraint [1]

$$m'^T Fm = 0, \tag{1}$$

which means the point m' lies on the epipolar line $l' = Fm$ while the point m lies on the epipolar line $l = F^T m'$ as well.
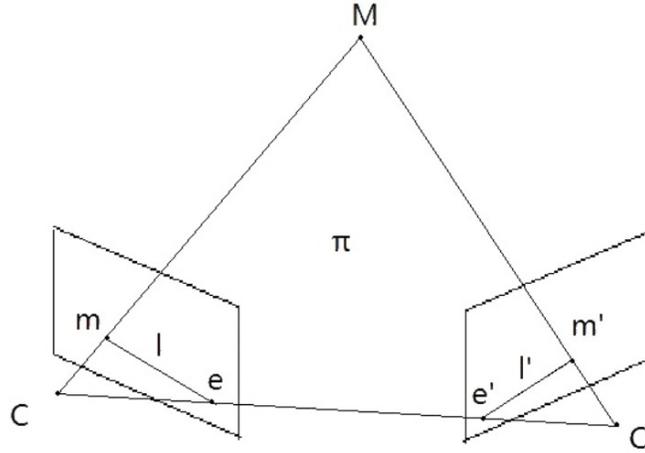
Figure 1. Illustration of the epipolar geometry.

2.2. Review of the 8-Point Algorithm

Suppose the F-matrix takes the form of

$$F = \begin{pmatrix} f_{00} & f_{01} & f_{02} \\ f_{10} & f_{11} & f_{12} \\ f_{20} & f_{21} & f_{22} \end{pmatrix}. \tag{2}$$

Also suppose that at least eight pairs of points $m_n = (x_n, y_n, 1)^T$ and $m'_n = (x'_n, y'_n, 1), n = 1, 2, ..., N$ with $N \geq 8$ are given.

Then from (1), we obtain the following set of linear equations

$$Af = (a_1 a_2 ... a_N)^T f = 0, \tag{3}$$

In which

$$a_n = (x_n x'_n, y_n x'_n, x'_n, x_n y'_n, y_n y'_n, y'_n, x_n, y_n, 1)^T, \tag{4}$$

$$f = (f_{00}, f_{01}, f_{02}, f_{10}, f_{11}, f_{12}, f_{20}, f_{21}, f_{22})^T. \tag{5}$$

Without regard to the rank 2 constraint tentatively. When the F-matrix is multiplied by a scalar, it still satisfies (1). Now the F-matrix has 8 independent variables. We can utilize the constraint [3]

$$(f_{00})^2 + (f_{01})^2 + ... + (f_{22})^2 = 1. \tag{6}$$

After solving the Eq. 2, the F-matrix still does not yield to the rank 2 constraint, so we can use the SVD decomposition [3]

$$F = UDV^T = U \begin{pmatrix} \sigma_0 & 0 & 0 \\ 0 & \sigma_1 & 0 \\ 0 & 0 & \sigma_2 \end{pmatrix} V^T, \sigma_0 .. \sigma_1 .. \sigma_2 .. 0. \tag{7}$$

Since F is of rank 2, we have $\sigma_2 = 0$. Then the final estimate F' of the 8-point algorithm is given as

$$F' = U \begin{pmatrix} \sigma_0 & 0 & 0 \\ 0 & \sigma_1 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T. \tag{8}$$

2.3. Review of RANSAC

Given any statistical model, Random Sample Consensus (RANSAC) distinguishes the outliers from a data set by using a threshold of error and then finds the best parameters which describes the data set well.

Before running RANSAC to estimate the F-matrix, we need a given th threshold (which is the distance between a point and its epipolar line) and the confidence p of the model. Then the number of iterations k and the number of inliers d must satisfy $1 - p = (1 - (d/N)^8)^k$ [9]. The process is shown in Fig. 2.
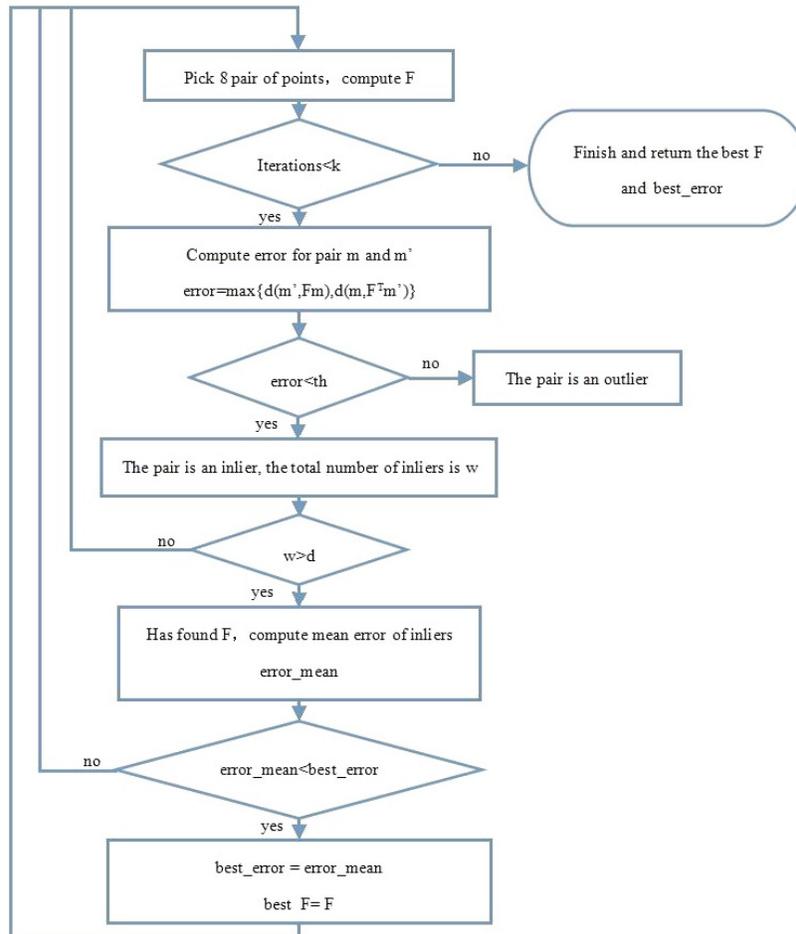
Figure 2. Flowchart of the general RANSAC algorithm

2.4. The 4D Space of Matching Points

For any pair of feature points $m = (x, y, 1)^T$ and $m' = (x', y', 1)^T$, the coordinates of the points describe the positions of the points in the image planes I and I'. The difference between the two points describes the transformation between the two viewpoints. Generally, in an image, the feature rich areas tends to have more concentrated feature points to form the matching pairs for F-matrix estimation. In order to study the distribution and transformation of the matching points, we construct a 4D space in which the vectors are of the form $q_n = (x_n, y_n, x_{n'}, y_{n'})^T$. The 4D space contains the information about both the 2D feature point distribution and the transformation between the matching points. It can help find useful characteristics of the matching points to estimate the F-matrix.

Figure 3. Feature point distribution illustration. The feature points are found by SIFT. Areas with more concentrated features points are marked out in the boxes.

Fig. 3 shows an example of the distribution of the feature points in a general natural image. After using SIFT to find the feature points in a real image pair, we get $N = 411$ pairs of matching points. We can see from the Fig. 3 that most of the feature points locate in some small areas shown in the boxes. This observation provides us with the hint that by doing a cluster analysis on this 4D vectors, we may be able to identify those feature point pairs that are more reliable than those isolated ones. In the next subsection we describe a practically efficient clustering method.

2.5. Cluster Analysis

Clustering analysis that classifies the data is a common method used to analyze multidimensional data. Data points that are dense in a multidimensional area and far from other dense area tend to belong to the same cluster. The clustering result describes the distribution of the data in multidimensional space. When clustering, we commonly treat the point that is closest to the center of an area as the center of the cluster. The center point of a cluster decides the basic property of the cluster. These ideas are recently used by Rodriguez and Laio in their paper "Clustering by fast search and find of density peaks". The key steps are described as follows.

(1) Computation of the local density ρ_i.

$$\rho_i = \sum_j \chi(d_{ij} - d_c),$$

(9)

$$\chi(x) = \begin{cases} 1, x < 0 \\ 0, x \geq 0 \end{cases}$$

(10)

where d_{ij} is the Euclidean distance of data point i and j, d_c is a parameter which makes the mean value of ρ_i equal to $2\%N$. N is the total number of data points. Larger ρ_i means the point i is more likely to be the center of a cluster.

(2) Computation of δ_i.

$$\delta_i = \max_{j:\rho_j > \rho_i} (d_{ij})$$

(11)

If ρ_i is a global max, let $\delta_i = \max_j(d_{ij})$. δ_i is the minimum distance between the points which have bigger ρ than ρ_i and the point i, a large value of δ_i means the point i is far away from other dense area.

(3) Drawing the decision figure.

Using ρ as the horizontal axis and δ as the vertical axis, a decision figure of all the data points can be drawn. Then those points with large product values of $\rho * \delta$ will be the cluster centers.

2.6. The Complete Proposed F-matrix Estimation Algorithm

Given any image pair, we can use the SIFT algorithm to first find the N pairs of corresponding feature points $m_n = (x_n, y_n, 1)^T$ and $m_n = (x_{n'}, y_{n'}, 1)^T, (n = 1, 2, ..., N)$. Then the complete F-matrix estimation algorithm can be summarized as follows:

(Step 1) Use the N pairs of points to construct the 4D vectors $q_n = (x_n, y_n, x_{n'}, y_{n'})^T$.
(Step 2) Compute of the local density ρ_i using (9) and (10),
where d_{ij} is now the Euclidean distance between vectors q_i and q_j.
(Step 3) Compute the distance δ_i using (9) and (11).
(Step 4) Obtain the decision figure and find the cluster centers, which are treated as inliers.
(Step 5) Use RANSAC on the inliers of step 4, the F-matrix of the randomly chosen points is computed by the 8-Point method.

As will be shown later, because the clustering step can generally produce more reliable inliers than the traditional methods, the final estimate of the F-matrix will be more accurate. In addition, because of the same reason, less iterations are needed, which means faster running time. We pick the point whose $\delta_i * \rho_i$ is lower than $\alpha * \delta_{max} * \rho_{max}$ as an outlier. Point i whose $\delta_i * \rho_i$ is higher than that is decided as a clustering center in [7]. It represents the basic transformation of the local area. α affects the total number of outliers which are picked out that larger α will result in more outliers. Meanwhile, α also influences the reliability of the inliers. With a large amount of points

and g vectors located among an inlier q_i, it usually indicates that g_i is quite close to the local mean value and represents the actual transformation between two images at the point i.

As is described above, feature points often lie in areas that have obvious characteristics. In a static scene, we consider that these dense points are often eliminated from false matching and are more reliable than the points that are isolated when being used in estimating the F-matrix because of the presence of noise. Besides, the coordinates of the matching points which are corresponding to the same object in 3D space and are located in the same area in the image plane, often change slowly along with the position in image plane. If there is noise whose mean value is 0, we can refer to the other points in the same area to decrease the influence of noise.

In the decision figure, a false corresponding pair q_i has big δ_i and small ρ_i in common.

The center of a cluster of q has big δ and ρ. The points in a cluster excluding its center have small δ_i but big ρ_i.

As is described in [8], inliers that all come from a small area make no sense because they can hardly represent the transformation between the image pairs in the whole range of the image. However, our optimization method easily overcomes this by choosing the inlier point with a large δ.

3. EXPERIMENTS

Ten pairs of images (5 pairs of real images and 5 pairs of synthetic images) captured from ten scenes are used. Then the F-matrix is estimated utilizing the proposed method. We compare the proposed method with the classical methods which are completed by directly calling the functions in OpenCV. We only show the results of the estimation of F-matrix of one pair of real images and one pair of synthetic images here. While the results of the proposed method in the ten pairs of images are all encouraging.

3.1. Real Images

When $th = 2.2$, experimental results of the two images captured by a camera are shown in Fig. 4 which is the decision figure in which the curve represents $\alpha = 0.011$ and the bigger circles are outliers found by the conventional RANSAC method. Besides, Fig. 5 shows the matching pairs which are inliers found by the proposed method.
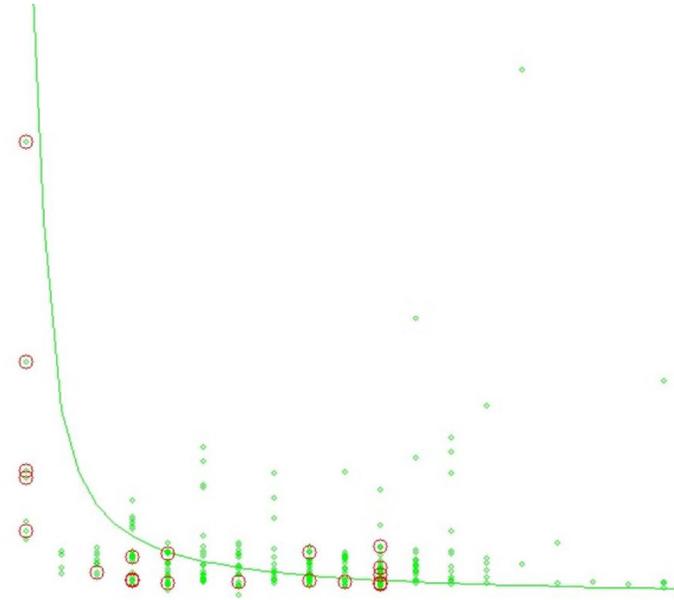
Figure 4. Decision figure of the pair of real images.
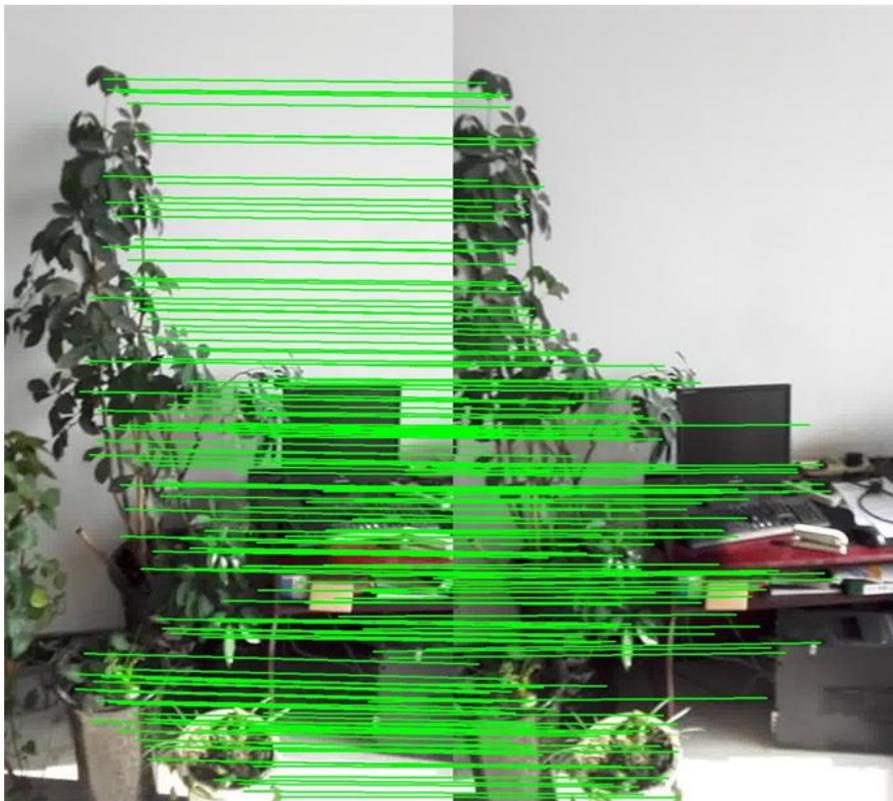
Figure 5. Matching points of the pair of real images.

As is seen in Fig. 4, most of the outliers in RANSAC (displayed by bigger circles) are under the curve, it proves the reliability of the search of density peaks.

Using different values of th, we compare the proposed method with RANSAC whose result is shown in Table 1.

Table 1. Comparison with RANSAC.

th(pixel)	Method	Time(ms)	Mean Error(pixel)
2.2	RANSAC	4.7	0.841559
	Proposed, α= 0.011	9.9	0.744619
1	RANSAC	10.3	0.438181
	Proposed, α= 0.02	13.3	0.412330
0.8	RANSAC	33.3	0.370029
	Proposed, α= 0.025	14.3	0.357731
0.5	RANSAC	126	0.241158
	Proposed, α= 0.04	17.2	0.224982

When th decreases, the time RANSAC consumes increases rapidly because Ransac picks up data points randomly and iterates until a fine F-matrix is found. However, the proposed method has a prior analysis on the characteristics of the data points using a simple algorithm. Beyond that, the mean error in pixel of the proposed method is smaller.

Comparison results with other methods are show in Table 2.

Table 2. Comparison with other methods.

Method	Time(ms)	Mean Error(pixel)
8-POINT	1.3	1.064726
7-POINT	23.6	1.706590
LMeDS	18.2	0.833023
Proposed	9.9	0.744619

3.2. Synthetic Images

To compare the methods from another point of view deeply, we use OpenGL to generate 2 images I and I' which have a known F-matrix F_0. So F_0 is the ground truth F-matrix. Then we utilize the evaluation method in section 4.1 of [8] proposed by Zhang to evaluate the F-matrix estimated by the methods described above. The evaluation method is computed following the steps (see Fig. 6):

(1) Compute the fundamental matrix F_1 the proposed method.
(2) Choose a point m_i from image I and compute the following epipolar lines:

$$l_{0i} = F_0 m_i \tag{12}$$

$$l_{1i} = F_1 m_i \tag{13}$$

(4) Choose a point m_i' on the epipolar line l_{0i} and compute the following epipolar lines:

$$l_{0i}' = F_0^T m_i' \tag{14}$$

$$l_{1i}' = F_1^T m_i' \tag{15}$$

(4) If the epipolar lines do not intersect the image plane, go back to step 2.

(5) Compute the distance d'_{1i} from the point m'_i to the line l_{1i} (c.f. Fig. 6).

(6) Compute the distance d_{1i} from the point m_i to the line l_{1i} (c.f. Fig. 6).
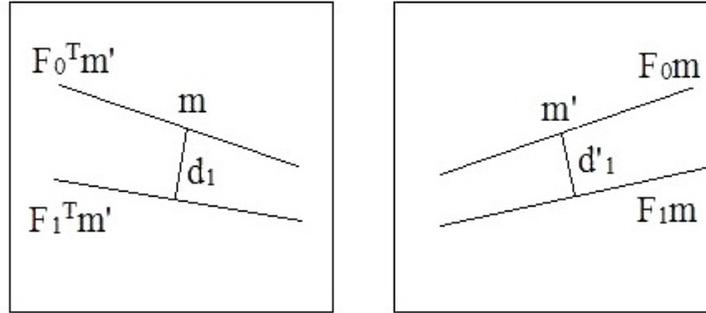
Figure 6. Illustration of the evaluation method [8].

(7) Repeat step 2 to step 6 for N times.

(8) Compute the mean value d_1 of all the d_{1i} and d'_{1i}.

Then the results that the proposed method compares with other methods are shown in Fig. 7, Fig. 8 and Table 3. The numbers in Table 3 represent the difference between the F-matrix estimated by the methods above and the ground truth F_0 while $th = 2.2, \alpha = 0.011$.

Table 3. F-matrix estimation error comparison with ground truth.

	8-POINT	7-POINT	LMeDS	RANSAC	Rroposed
Error	83.875023	42.706934	1.300233	1.330434	1.077631
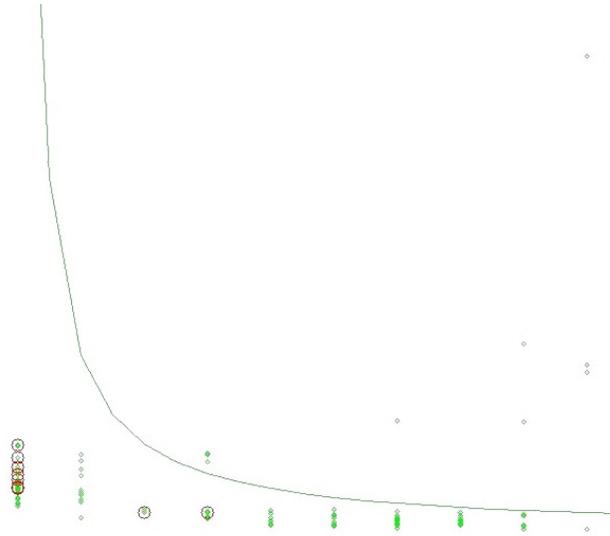
Figure 7. Decision figure of the pair of real images.
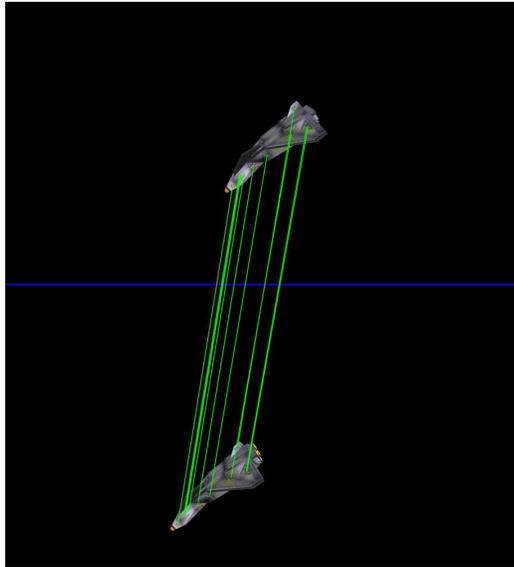
Figure 8. Matching points of the pair of real images.

4. CONCLUSION

The paper proposes the utilization of clustering to optimize the corresponding pairs of points for accurate estimation of the fundamental matrix. This approach chooses the point pairs which are likely to represent the true relationship between image pairs and suffer less from noise. As the threshold decreases, the method we use is better than the conventional RANSAC both in processing speed and accuracy. Besides, our optimization based on search of density peaks has a lower level of the error between the estimated F-matrix and the ground truth.

REFERENCES

[1]   R. Hartley and A. Zisserman, (2000) Multiple View Geometry in Computer Vision. Cambridge.

[2]   R. Hartley, (1994) Projective reconstruction and invariants from multiple images, IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(10), pp1036-1040.

[3]   H. Longuet-Higgins, (1981) A computer algorithm for reconstructing a scene from two projections, Nature 293, pp133-135.

[4]   R. Mohr, F. Veillon, and L. Quan, (1993) Relative 3d reconstruction using multiple uncalibrated images, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp543-548.

[5]   Z. Zhang, (1995) Parameter estimation techniques: A tutorial with application to conic fitting, Image and Vision Computing. In Press. Also INRIA Research Report No.2676.

[6]   P. Rousseeuw, and A. Leroy, (1987) Robust Regression and Outlier Detection, John Wiley & Sons, New York.

[7]   A. Rodriguez and A. Laio, (2014) Clustering by fast search and find of density peaks, Science 344, pp1492-1496.

[8]   Z. Zhang, (1998) Determining the Epipolar Geometry and its Uncertainty: A Review, International Journal of Computer Vision, 27(2), pp161-198.

[9]   M.A. Fischler and R.C. Bolles, (1981) Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography Comm. ACM, Vol. 24, No. 6, pp381-395.

[10] J. Lim and N. Barnes, (2010) Estimation of the Epipole Using Optical Flow at Antipodal Points, Computer Vision and Image Understanding, Vol. 114, No. 2, pp245-253.

[11] J. Oliensis, (2002) Exact Two-Image Structure from Motion, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 24, No. 12, pp1618-1633.

[12] D. Nister, (2004) An Efficient Solution to the Five-Point Relative Pose Problem, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 26, No. 6, pp756-770.

AUTHORS

Hao Wu was born on Oct. 16, 1988, in Baoji, China. He got bachelor's degree of electronic information technology and instrument in Zhejiang University in 2007. He is studying for a master's degree in Lanzhou University. His interest of research is compute vision.

Yi Wan received his Ph.D degree from the Rice University in 2002 and is currently a faculty member at the school of information science and engineering, Lanzhou University, China.

ADOPTING COLLABORATIVE WORKFLOW PATTERN: USE CASE

Antonio Capodieci, Giuseppe Del Fiore, Luca Mainetti

Department of Innovation Engineering, University of Salento, Lecce, Italy
```
{antonio.capodieci, giuseppe.delfiore,
          luca.mainetti}@unisalento.it
```

ABSTRACT

In recent years, the use of web 2.0 tools has increased in companies and organisation. This phenomenon, has modified common organisational and operative practices. This has led "knowledge workers" to change their working practices through the use of Web 2.0 communication tools. Unfortunately, these tools have not been integrated with existing enterprise information systems. This is an important problem in an organisational context because knowledge of information exchanged is needed. In previous works we demonstrate that it is possible to capture this knowledge using collaboration processes, which are processes of abstraction created in accordance with design patterns and applied to new organisational operative practices. In this article, we want to present the experience of the adoption of the methodology and the catalog of patterns from some pattern designers of a company operating in Public Administration and Finance, with the aim of shaping an information system Enterprise 2.0, which takes into account the collaborative processes.
.

KEYWORDS

Business Praticies, Collaboration Process, Business Process Patterns, Enterprise 2.0, Web 2.0, Knowledge Workers & Collaboration Tools

1. INTRODUCTION

With the rise of Enterprise 2.0 systems, not only in big companies but also in organizations new collaborative working practices have been introduced, as defined by Nial Cook [1]. These will work in addition to the traditional business processes that define the procedure of operational organization. The collaborative processes are characterized by a strong and a non-default collaboration among the participants in the process in order to achieve common goals. The collaboration develops through the combination of traditional communication tools (e-mail, telephones and direct conversations) and web 2.0 tools (chat, social networks, blogs, etc...). The actors involved in a collaborative process can choose the preferred means of cooperation (communication). In addition, the company cannot define the set of enterprise available tools in advance: the individual employee is the only one who fully understands his needs and he must be able to build and change his own virtual working environment. The collaborative processes are key components for the management of the company; however, they are executed independently from legacy information systems. Therefore, there is a clear need to integrate collaborative processes with other processes and business information. In this way, their scope and their effectiveness will be amplified, as well as they should be integrated with information systems and business processes that are under the direct management control. Only in this way, the information generated by the communication tools Web 2.0 can become part of the corporate information assets. In the book "Human Interactions the hearth and soul of business process management" [2], it is shown that the heart of any organization is always people and that their

behaviors appear to be complex and non-default. However, this complexity must be managed. To ensure the success of an organization we need to find a new model to manage the complexity and constantly changing "human-driven" processes. You cannot consider workers as if they were simply "the gears of a machine." Therefore, first, we need to find a method to formally describe "human-driven" processes and then we can try to integrate them into a software tool that supports their execution. Nevertheless, all influence strongly on classical methods of modeling both of business processes and of information systems, requiring a radical change in perspective. The authors examine in depth the nature of the "human- driven" processes and they show how this, if properly designed, can be supported by the information systems of the next generation. To formally describe "human-driven" processes (as defined in [2]) and the interaction between man and technology, the identification of workflow design patterns can be a useful approach. This is demonstrated by the authors in the article [3], where the authors propose a catalog of workflow design patterns with different granularity levels, through which the designer can get a very detailed modeling processes. A design pattern is a concept that can be defined as "a general design solution to a recurring problem". It gives him a name; abstracts and identifies the key aspects of the structure used for the solution of the problem; identifies the classes, participant instances and the distribution of responsibilities; describes when and how it can be applied. In short, it defines a problem, the typical contexts in which it is located and the optimal solution. Once you have identified the patterns of business processes, these can be used for the re-engineering process flows in different situations. In previous works, we have presented some collaborative patterns extracted from the modeling of many business processes related to a company operating in the ICT sector. These patterns have been aggregated to make a catalog of patterns. In this article, we want to present the experience of the adoption of the methodology and the catalog of patterns from some pattern designers of a company operating in Public Administration and Finance, with the aim of shaping an information system Enterprise 2.0, which takes into account the collaborative processes. The experience gained by users composing the test panel, will be assessed through an evaluation questionnaire. Finally, we will conduct an analysis of the obtained results, in order to improve the methodology and the catalog of the proposed pattern. The paper is structured as follows: the next section reports on key related work in the areas of analysis, description, identification and application of business practices, mainly to address knowledge workers' emerging needs. Section III provides readers with an overview of the methodological approach used to identify collaborative processes. Section IV describes the case studies where it was applied the methodology and pattern. Section V describes the investigation methodology used to evaluate the methodology and patterns. Section VI describes the results of the questionnaire administered to the users of the panel of the trial. Finally, Section VII summarizes our key messages and sketches future research directions.

2. RELATED WORK

Over the years, the concept of patterns has been applied in several fields. The idea was proposed by Christopher Alexander in his book A pattern language [4] where he scientifically describes an architectural system through 253 patterns that solve common problems of cities. The concept of patterns has been adopted and applied in the field of Software Engineering by the 'Gang of Four' in their famous book Design Patterns: Elements of Reusable Object-Oriented Software [5]. It has been applied in recent research in Business Process Management (www.workflowpattern.com) and other research works such as [6] that predict the proliferation of patterns for BPM. In [7] its use has described and evaluated workflow management technologies. Several methods in the identification of patterns have been proposed in the international scientific literature, such as bottom-up and top-down approaches [8], or a combination of these two [9]. Once identified, business process patterns have been used in different contexts for the re-design of the business process flow. An example is in [10] where the authors present business processes patterns in order to enhance the design of the public health care business process. Another example is in [11]

where the authors propose a methodology for business process re-design; the methodology consists of using the process context to discover the process nature and then applying the workflow patterns to the evaluating and enhancing of the current process in the given context. Another example is in [12] where the authors propose the development of a methodology for collaborative and ubiquitous learning; the methodology combines the advantages of a collaborative learning environment with the benefits of ubiquitous computing and flexibility of new digital technological devices. The previous works clarify the importance of business process patterns in order to design and re-design a business process flow made up of pre-defined activities as- signed to specific stakeholders. The patterns, and in particular the collaboration patterns, still are helpful to designers. The concept of collaboration patterns has been introduced in the definition of virtual organization. A virtual organization is 'a temporary alliance of independent enterprises that come together to share skills, core competencies and resources in order to better respond to business opportunities, supporting cooperation through computer networks' [13]. The importance of such collaboration patterns in virtual organizations has been stressed in [14] where collaboration patterns were defined as a segments of work or parts of collaboration. The authors highlighted that the reuse of collaboration patterns can be an advantage in collaborative environments, such as virtual organizations, where there is an increasing need for modelling, executing, monitoring, and supporting the dynamic nature of collaborations. In [15] there is another work that proposes some collaboration patterns related to the virtual organization. These patterns are aggregations of detailed activities into larger-scale units. In that paper, the authors present a shared workspace system in order to collect and make available observations of a virtual organization. The real value of business patterns may be appreciated when it is possible to use the patterns in the design of business processes that describe the way of operation of the companies. Graphical models can be used to represent the patterns. An example of collaboration business process patterns is in [16] where the author describes guidelines for the development of business process models using BPMN 1.2. These guidelines focus on the use of the elements in order to correctly, consistently and clearly design artifacts, but do not focus on syntax and semantics. The importance of representing collaboration business processes, in general, and collaboration patterns, in particular, is shown by the introduction of BPMN 2.0 [17]. BPMN 2.0 contains several additional elements and new types of diagrams, especially to improve the modelling of processes that span several independent organizations. In particular, BPMN 2.0 introduces the collaboration diagram and the choreography diagram. In previous work [18][19][20][21], we presented a methodology and a catalog of patterns defined through the collaborative BPMN notation. In this paper, we evaluate the adoption of the methodology and the catalog of the patterns identified in previous work by some knowledge workers.

3. METHODOLOGY AND PATTERNS CATALOG

In this section, we will briefly present the methodology and the catalog of the patterns through a concept map, referring the details to the previous works.

3.1. Methodology

The approach to be followed to identify and apply the patterns of collaborative business processes is divided into six phases:

1. The first phase is characterized by the analysis of the business environment with much attention given to identifying some of the processes and areas that are characterized by both intense collaborative activities among the workers and the need to use Web 2.0 tools.

2. The work continues in the second phase: modelling the business processes detectable within the case used, selected in the previous step through the use of BPMN.

3. In the third step, the study and comparison of the BPMN diagrams of all the modelled processes starts; to identify new patterns, it is necessary to focus on all repetitive common and atomic "segments" which are in the modelling performed in the previous phase. Particular attention should be paid to collaborative and cooperative activities, where we found a number of practices that have considerable repetitiveness.

4. Some of the "repetitive segments" detected in the previous step may already be known, so at this stage, it is necessary to verify the existence of patterns similar or identical to the segments identified. In such a case, it is better to use the known solutions that have already been applied and validated in different contexts. Otherwise, these segments can be considered new, such as new patterns.

5. During the fifth phase, the design patterns identified in the third step start to be applied to model and realize a prototype of the collaborative information system. The purpose of this step is to verify the validity of the approach adopted in the identification of the patterns and to apply those patterns in the realization of a collaborative information system.

6. Following the experimentation, in order to verify the usefulness of the use of the patterns in the context of collaborative information systems, the data of the trial (sixth phase, evaluation of design patterns) needs to be collected.

3.2. Conceptual Map

From Figure 1 we have a comprehensive view of all the patterns, through a concept map from which emerge the functional dependencies that exist among the various patterns. The identified patterns respond, therefore, to two fundamental requirements: to manage the collaboration among different actors that are called to work to accomplish a given task without a pre-defined and a pre-structured sequence; to allow you to make the most of the typical tools of Web 2.0 within the enterprise. It should be noted that the identified patterns might be used in isolation or they can be appropriately concatenated each other to model specific situations.

4. APPLICATION OF METHODOLOGY AND PATTERNS: CASE STUDY

The methodology and patterns briefly described in the previous section were introduced into a company with information systems design experience. This had the goal of modeling and designing some processes related to Public Administration and Finance. In detail:

• Public Administration: "eGovernment Administrative Acts".

• Finance: "Circular 2.0".

• Finance: "New Banking Product".

• Finance: "Conclusion of contract".

• Finance: "Proposal for Funding".
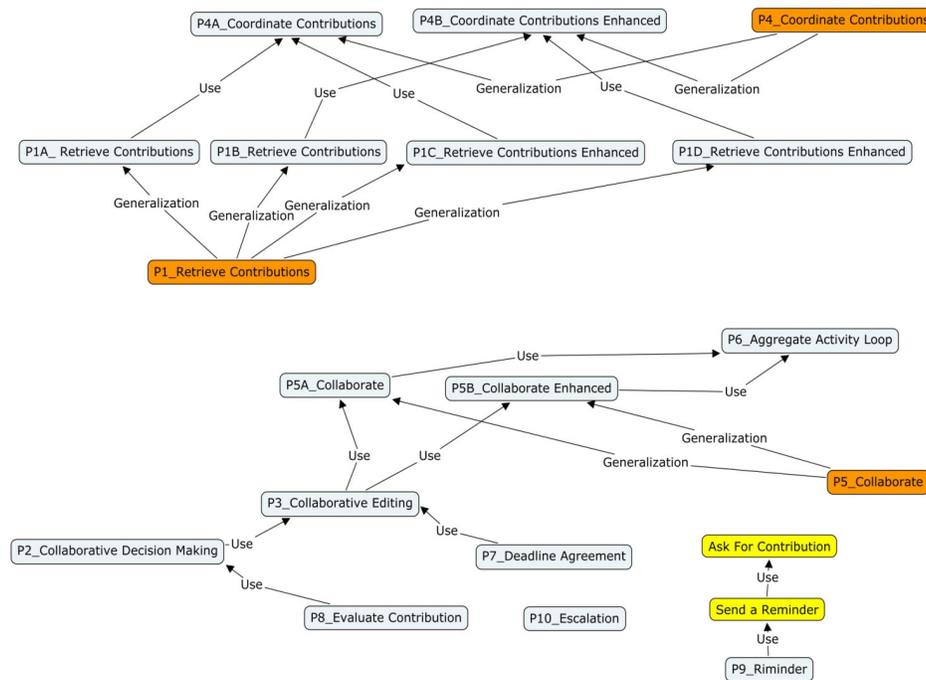
Figure 1. Conceptual Map

All these processes were modeled and subsequently re-engineered by adopting the methodology and the patterns catalog without applying to the latter substantial changes. The application of the patterns in the process of re-engineering has allowed the emergence of a series of activities that previously were not coded, but that were performed by operators.

The activity carried out has been divided into five phases:

1. Establishment of Focus Group with some employees of Links Management and Technology.

2. Discussion and study of the models resulting from the application of the methodology and of the catalog.

3. Formation of the designers on the patterns catalog, the methodology and the applicability of the latter to the contexts of reference.

4. Application of methodology and patterns to the study cases listed above from the panel of experimentation.

5. Evaluation of methodology and patterns and data collection by administering a questionnaire given to all the users of the panel of the trial.

5. INVESTIGATIVE METHODOLOGY

The criteria, which regroups the evidence and the critical issues identified, have been identified for analytical purposes. Therefore, we defined the dimensions of analysis used in the preparation of the evaluation questionnaire administered to users composing the panel of the trial. The questionnaire is divided into six distinct sections:

1. Comprehensibility: This section aims to assess the ease of identification of the operating rules that govern the methodology and logical paths used in the representation of collaborative patterns.

2. Memorability: This section aims to evaluate the simplicity and immediacy in the association of the patterns to the features that they should represent.

3. Ease of Use: This section aims to assess the ease of identification and association of the patterns in collaborative processes.

4. Extensibility / Adaptability: This section evaluates the preparation of the methodology to the introduction of new rules and new patterns and to adapting existing ones.

5. Effectiveness: This section is intended to evaluate the optimization of the modelling times of the business process obtained through the use and application of the methodology.

6. Completeness: This section evaluates the completeness in the domain of collaborative processes of the proposed patterns and of the methodology.

The questionnaire, that allowed the assessment of the following characteristics, consists of some open-ended questions, some yes / no, and many others based on a Likert scale. For each section, moreover, is present a free field called notes where users can express any idea and opinion. The choice of using the Likert scale is determined by the fact that through this scale it is possible to measure the attitudes and behaviors using a range of response options ranging from one extreme to the another (e.g. from not at all likely to extremely likely). Unlike a simple question "yes / no", a Likert scale offers the opportunity to discover the different degrees of judgment. This can be especially useful for delicate and difficult objects or subjects of investigation, as the one discussed in the following work. The number of response options helps also to more easily identify areas for improvement.

6. RESULT OF THE EXPERIMENT

In this chapter the results of the questionnaire administered to the users of the panel of the trial are presented in detail. The questionnaire was omitted for editorial limits. It was administered to four employees who hold the following business role:

1. Technical Innovation and Technology Leader Area.

2. Responsible for research.

3. Functional Analyst of the Innovation and Technology Area.

4. Functional Analyst of the Innovation and Technology Area.

As described previously, the questionnaire is divided into six distinct sections and each of them is composed of a series of questions. Almost all of the questions require an answer based on a Likert scale in four steps:

1. For Nothing (0.25 points).

2. Shortly (0.5 points).

3. Enough (0.75 points).

4. Much (1 point).

6.1. Comprehensibility

The results concerning the comprehensibility of the methodology and proposed patterns are explained in the following paragraph. The graph shown in Figure 2 illustrates in detail the results obtained by individual users for each of the sixteen questions provided in the questionnaire, in the section concerning the comprehensibility. The graph shows a good degree of intelligibility in almost all questions. An average rating of all questions is shown in the graph proposed in Figure, which shows that all users members of the panel of experimentation have found the methodology and the proposed pattern "enough" understandable.
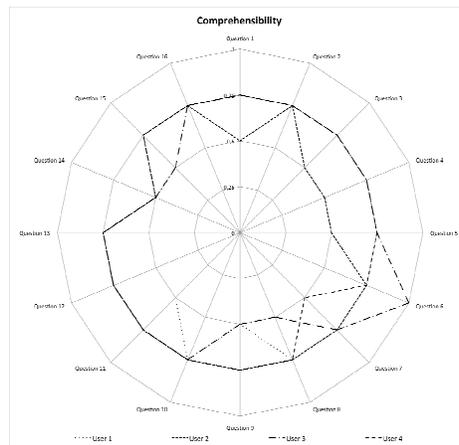
Figure 2. Comprehensibility for each user

Finally, from the observations made in the notes and the analysis of open-ended responses emerges that it is appropriate to equip all patterns with examples of applicability in order to improve the comprehensibility.
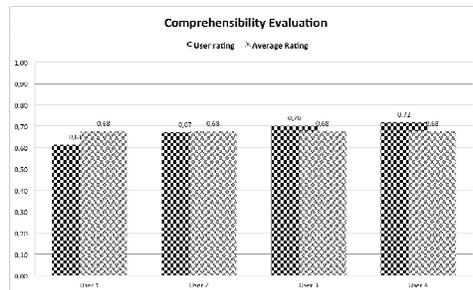
Figure 3. Comprehensibility evaluation

6.2. Storable

The following paragraph shows the results of the methodology and of the proposed pattern capability to store. The graph in Figure 4 shows a fluctuating trend with regard to the simplicity and immediacy to the respective pattern in the association of features that they must represent. This trend is justified by the different level of difficulty of storage for the various patterns.

An average rating provided by members of the panel testing is presented in the graph shown in Figure 5. This evaluation allows to state that the proposed methodology and some of the patterns are easy to memorize. Finally, it is clear from the observations made in the notes, that one of the key aspects that influenced the difficulty of storage is due to the nomenclature used in the catalog of patterns.
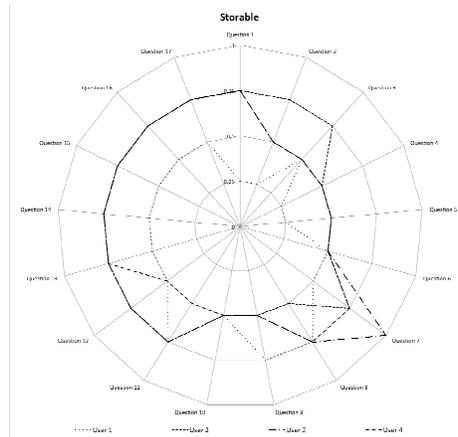
Figure 4. Storable for each user
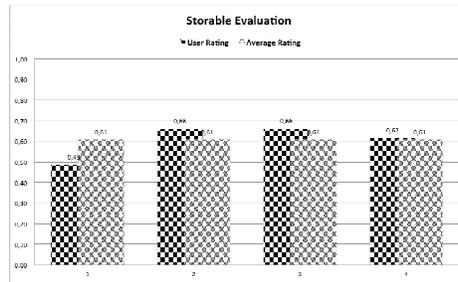
Figure 5. Storable evaluation

6.3. Ease of Use

The results regarding the ease of use of the methodology and of the proposed patterns are explained in the following paragraph. The graph in Figure 6 highlights an excellent ease of use by three out of four users.
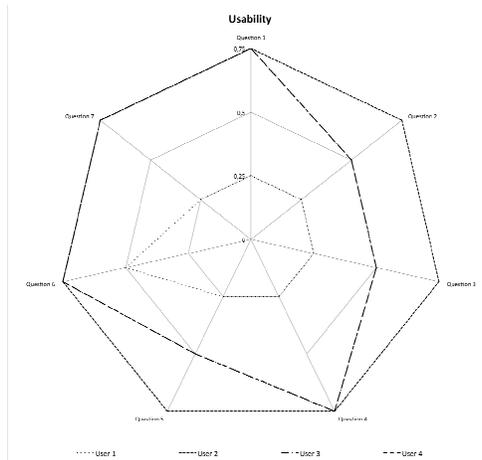
Figure 6. Usability for each user

The graph in Figure 7, puts more light that only one user (User 1) has encountered some difficulties in the use of pattern, justifying notes that this difficulty is due to the ambiguity of some of them. The rest of the users has expressed a very high degree of usability by ensuring a sufficient level average.
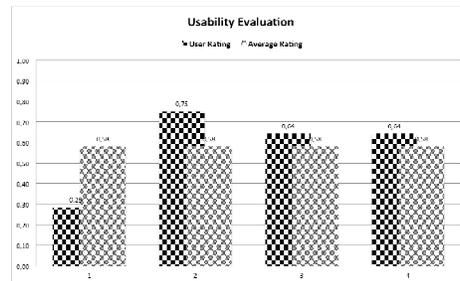
Figure 7. Usability evaluation

6.4. Extensibility/Adaptability

The following paragraph illustrates the results concerning the extensibility and adaptability of the proposed methodology and patterns. The graph in this Figure 8 shows in detail the results obtained by individual users for each of the questions given in the questionnaire, in the section on the ease of use. From the graph, it is seen easily that for the majority of users there is a good preparation of the methodology to the introduction of new rules and new patterns or to the adaption of the existing ones.
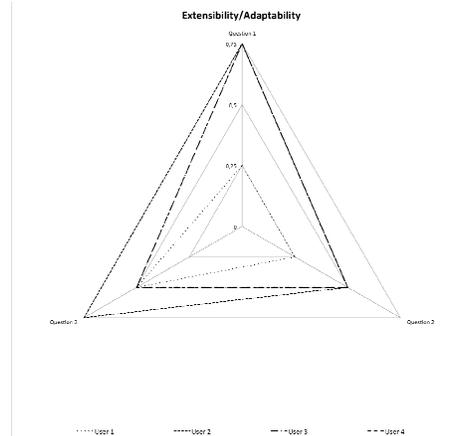
Figure 8. Extensibility / Adaptability for each user

Even in this case, the graph shown in Figure 9 puts the attention that only one user (User 1) has encountered difficulties in extending and adapting certain patterns. The rest of the users expressed a sufficient degree of adaptability-extensibility ensuring, therefore, a nearly enough average. However, from the observations made in the notes and from the analysis of the open-ended responses, it emerges that it is advisable to increase the level of generality of the pattern in order to improve the adaptability and the extensibility.
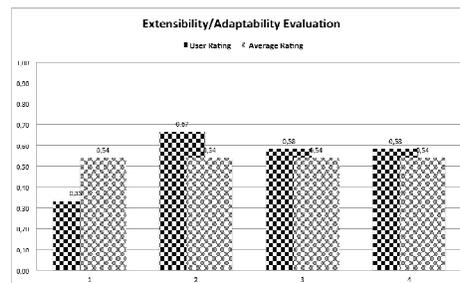
Figure 9. Extensibility / Adaptability evaluation

6.5. Effectiveness

The following section summarizes the results regarding the efficacy of the methodology and of the proposed patterns. Differently from the previous sections, in this case it was not possible to give many details about the results using diagrams, because of the very few questions with response based on Likert scale. Almost all of the questions in this section are open questions. However, it is clear from the responses received that there was a big difficulty in assessing the effectiveness due to the lack of meaningful metrics.

6.6. Completeness

The results concerning the completeness of the methodology and of the proposed patterns are explained in the following paragraph. In addition, in this case it was not possible to detail the results by using the diagram, due to the very few questions with response based on Likert scale. However, from the observations made in the notes and from the analysis of open-ended responses, it is clear that the proposed methodology is "a little" complete for the users of the panel

of experimentation. This result is justified by the fact that the methodology and patterns are still in their infancy.

7. CONCLUSIONS AND FUTURE WORKS

This document has been dealt with the analysis and testing of the methodology and of the proposed catalog of patterns concerning the modeling of processes that involve collaborative practices according to the paradigms of Enterprise 2.0. The proposed catalog has been accepted by the designers who have adopted it fully and without substantial changes.

It has been defined a survey methodology suited to identify the applicability of methodological guidelines and to report cases of corrections or improvements.

For analytical purposes, the criteria have been identified which regroups the evidence and the critical issues identified. It has been defined the dimensions of analysis used in the preparation of the evaluation questionnaire administered to members of the trial panel members. In summary, the results obtained in the six sections object of the study are:

- A good level of comprehensibility of the methodology, which can be improved by providing practical examples of application in the definition of Collaborative patterns.

- From the point of view of the ability to store, it can be enhanced by defining a nomenclature more simple and intuitive.

- Ease of use is good.

- There is a good extensibility / adaptability.

- The effectiveness is little measurable without the presence of meaningful metrics.

- The methodology has to be completed by acting on the generalization of the pattern.

As future work, we expect the adoption of the changes required, in particular on the comprehensibility and on the capability to store, on the application of the catalog in other contexts and on the implementation of the latter on different BPMS workflow systems. For example, workflow systems that work in cloud computing as described in [22].

ACKNOWLEDGEMENTS

REFERENCES

[1] N. Cook, (2008) "Enterprise 2.0: How Social Software will change the future of work," *Gower Publisher*, London.

[2] K. Harrison-Broninski, (2005) "Human Interactions: The Heart and Soul of Business Process Management," *Meghan Kiffer Pr*, 2005.

[3] S. Smirnov, M. Weidlich, J. Mendling, M. Weske, (2011) "Object-Sensitive Action Patterns in Process Model Repositories," *Business Process Management Workshop*, vol 66, pp. 251-263

[4] C. Alexander, S. Ishikawa, M. Silverstein, (1977) "A Pattern Language," Oxford Univ. Press.

[5]   E. Gamma, R. Helm, R. Johnson, R. Vlissides, (1994) "Design Patterns: Elements of Reusable Object-Oriented Software," *Addison Wesley*, Reading, MA, US.

[6]   P. Harmon, (2003) "Business Process Change: A Manager's guide to improving redesigning, and automating processes," *Morgan Kaufmann*, San Francisco, CA, US.

[7]   W. M. P. Van Der Aalst, A. H. M. ter Hofstede, M. Weske, (2003) "Business Process Management: A Survey," *In: Proceedings of the International Conference of Business Process Management*, BPM'03, Eidhoven, The Netherlands, June.

[8]   R. Baggetun, E. Rusman, C. Poggi, (2004) "Design patterns for Collaborative Learning: From practice to theory" *In: Proceedings of ED-MEDIA World conference on educational multimedia*, hypermedia and telecommunications, AACE'04, Lugano, Zwitzerland, pp. 2493–2498.

[9]   S. Retalis, P. Georgiakakis, Y. Dimitriadis, (2006) "Eliciting Design Patterns for E-learning Systems: Computer Science Education," Vol. 16, No. 2, pp. 105–118.

[10] C. Stephenson, B. Wasana, (2007) "Enhancing Best Practices in Public Health: Using Process Patterns for Business Process Management," *In: The 15th European Conference on Information Systems*, ECIS'07, St. Gallen, Switzerland.

[11] K. Bessai, B. Claudepierre, O. Saidani, S. Nurcan, (2008) "Context-Aware Business Process Evaluation and Redesign," *In: The 20th International Conference on Advanced Information Systems Engineering*, CAiSE'08, MontPellier, France,   pp. 86–95.

[12] J. Díaz Arancibia, C. Rusu, (2014) "Ubiquitous Computer-Supported Collaborative Learning: A Literature Review" *11th International Conference on Information Technology: New Generations*, pp. 593-598.

[13] H. Luczak, A. Hauser, (2005) "Knowledge management in virtual organizations," *In: Proceedings of International Conference on Services Systems and Services Management*,   ICSSSM'05, Vol. 2, Chongquing, China, pp. 898–902.

[14] Y. Verginadis, D. Apostolou, N. Papageorgiou, G. Mentzas, (2009) "Collaboration Patterns in Event-Driven Environments for Virtual Organizations," *In: AAAI   Spring Symposium Series*, SSS'09, Stanford, CA, US.

[15] R. O. Briggs, G. J. De Vendre, J. Nunamaker, (2003) "Collaboration Engineering   with ThinkLets to Pursue Sustained Success with Group Support Systems," *Journal   of Management Information Systems*, Vol. 19, No. 4, pp. 31–64.

[16] J. Kindrick, (2009) "Enterprise Architecture based on Design Primitives and Patterns business transformation agency," Available at: http://www.bta.mil/.

[17] BPMN 2.0 Specification. Available at: http://www.bpmn.org/.

[18] U. Barchetti, A. Capodieci, A. L. Guido, L. Mainetti, (2011) "Modelling Collaboration Processes Through Design Patterns," *Computing and Informatics*, 30(1), 113-135.

[19] U. Barchetti, A. Capodieci, A. L. Guido, L. Mainetti, (2012) "Collaborative Process Management for the Networked Enterprise: A Case Study," *Paper presented at the 26th International Conference on Advanced Information Networking and Applications Workshops*.

[20] A. Capodieci, G. Del Fiore, L. Mainetti, (2014) "Adopting Collaborative Business Processes to Prevent the Loss of Information in Public Administration Organisations," *World Academy of Science, Engineering and Technology, International Science Index 89, International Journal of Social, Human Science and Engineering*, 8(5), 54 - 62.

[21] A. Capodieci, G. Del Fiore, L. Mainetti, (2014) "Adopting Collaborative Business Process Patterns for an Enterprise 2.0 Banking Information System," *The Fourth International Conference on Advanced Collaborative Networks, Systems and Applications*, pp 62-71.

[22] C. Salgado, L. Baigorria, D. Riesco, G. Montejano, N. Debnath, J. Hu, (2014) "Workflow Models: Management and Quality of Process in the Cloud" *11th International Conference on Information Technology: New Generations*, 91-96.

AUTHORS

Antonio Capodieci

Antonio Capodieci graduated in Computer Engineering at Politecnico di Torino. He received a Ph.D. in Computer Science from the University of Salento, Italy, where he has gained expertise on innovations in public administration through the application of e-government (Master's Degree in Public Management and E-government).

His research interests include web design methodologies, notations and tools, and the methodology of the design of Enterprise 2.0 Information Systems. He is participating in several research projects and he has co-authored several international scientific papers.
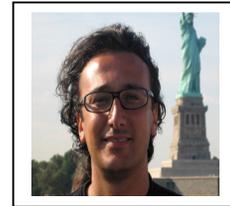
 Giuseppe Del Fiore

Giuseppe Del Fiore graduated cum laude in Computer Engineering at University of Salento, Italy, in 2013. His thesis concerned the definition and validation of Cross-layer Approach to Minimize the Energy Consumption in Wireless Sensor Networks. Since November 2013 he collaborates with IDA Lab - IDentification Automation Laboratory and GSA-Lab - Graphics & Software Architectures Lab at the Department of Innovation Engineering, University of Salento. His research activities are mainly focused on the design, development and validation of new solutions for Internet of Things. He is also involved in the creation of an information system Enterprise 2.0, which takes into account the collaborative processes which are processes of abstraction created in accordance with design patterns.

Luca Mainetti

Luca Mainetti is an associate professor of Software Engineering in the Department of Innovation Engineering at the University of Salento, Italy. His research interests include web engineering, web and services-oriented architectures, collaborative architectures, and IoT architectures. He is scientific coordinator of the GSA Lab (Graphics and Software Architectures Lab, http://www.gsalab.unisalento.it) and the IDA Lab (IDentification Automation Lab, http://www.idalab.unisalento.it). He received a Ph.D. in Computer Science from the Politecnico di Milano, Italy. He is a member of the IEEE and the ACM. He is the co-author of more than 130 international scientific papers.

INTENTIONAL BLANK

DISTRIBUTED SCRUM WHEN TURNING INTO MAINTENANCE: A SINGLE CASE STUDY

Minna Kilpala[1] and Tommi Kärkkäinen[2]

[1]Tieto Finland Oy, Mattilanniemi 6, 40100 Jyväskylä, Finland
`minna.kilpala@tieto.com`
[2]University of Jyväskylä, Department of Mathematical Information Technology,
P.O. Box 35, 40014 University of Jyväskylä, Finland
`tommi.karkkainen@jyu.fi`

ABSTRACT

Global software development using agile methods is commonplace in software industry nowadays. Scrum, as the agile development management framework, can be distributed in many ways, especially concerning how the key roles are presented in different sites. We describe here a single case study of a distributed Scrum, mainly for maintenance of the already constructed web portal. Using a qualitative method, both working well and challenging parts of the software work, as experienced by the project stakeholders, are revealed and discussed.

KEYWORDS

Scrum, software maintenance, case study, agile project management

1. INTRODUCTION

Software industry always tries to increase productivity, i.e., to produce larger amount of better software more efficiently and with reduced costs. This has not been or will not be easy; many reports and assessments continue to illustrate large portions of failed or cancelled projects. During the last decades, globalization, agility, versatility of development tools and computing platforms, domain specificity (in models, languages, and application areas), more systematic reuse in the form of product lines etc. have been affecting the ways to realize the software development efforts. But, old wisdom [1] still applies, because complexity of software entities, their conformity requirements, pressures of change, and invisibility within the distributed and hyperlinked technological environment are constantly increasing, reinforcing the unavoidable essential complexity of software development.

Globally distributed software development (i.e., Global Software Development, GSD) as a cost-effective way for software companies to minimize distances and jump the time zones for 24/7 production is commonplace. Reduced costs but higher flexibility and adaptability in this context are searched for by incorporating agile methods. However, the tension between the formal contracts and their plan-driven fulfillment, and adaptively and independently in a single-site, with direct user-interaction, working agile team, is evident.

The agile development model studied in this work is based on Scrum [2], which is more an agile project management framework than a detailed depiction of practices a la XP. Scrum's main roles are Scrum Master (SM) linking the customers and the development team and also protecting the team from external interruptions during the development cycles. Product Owner (PO) manages

the Product Backlog (PB) describing the features to be included in the software product. The software is constructed incrementally in small iterations called sprints, where the Sprint Backlog (SB) describes the subset of PBs to be realized. User stories are mainly (but not necessarily) used to describe the features in the backlogs in a semi-structured manner.

We located one literature review which summarized the used practices of Scrum in GSD projects [3]. From the review one notices that most studies in the research area are industrial reports and a knowledge gap to have sufficient amount of actual scientific information seems to exist. The actual phase of software lifecycle being studied here is related to software maintenance, whose needs are classically depicted by the famous Lehman laws [4]. More precisely, in our case the distributed team is correcting and modifying a web-based information system related to a long-lasting customer relationship with a stable PO, being involved in corrective, adaptive, and perfective maintenance [5].

Agile methods and software maintenance appears to be a scarcely studied combination. There are apparent needs for tailoring and modifications in aligning agility and maintenance: on the one side, the spirit of agility is, in principle, related to developing from the scratch without a need to comprehend an existing software system; on the other hand, small sprints and continuous integration are clearly activities suitable and valuable in maintenance as well. Moreover, both adaptive (e.g., due to changed qualitative requirements) and perfective (new user requirement from PB) forms of maintenance are incremental in nature. Actually [6] provide a pertinent characterization of the interplay between agile project and maintenance activities: "Maintenance can be seen as an endless agile project whose product backlog is changing constantly". The same authors, importantly, also point out the existence of Change request – change management type of activity resulting from a customer as PO mastering the Product Backlog. In [6] the overall analysis along the change management direction, provided recommendations on balancing the customer disturbance and agility, and being more involved in the customer's planning process on aligning business and software changes.

The empirical focus of this paper is on one organization's one individual project. Hence, we address a single case study well characterized by the definition of a case study as provided in [7, p. 13]: "an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident."

Goal of this paper is to find out if distributed Scrum can be used in software maintenance and if there are any benefits or disadvantages when doing so. Additional goal for IT Company where this study was done is to find out how this kind of work model can be improved. Research data was collected by semi-structured interviews from most of active project participants or stakeholders. Background information of case project is described by its project manager who is also another author in this study.

After this introduction we will present some studies close to our interesting area in this study as related work. From related work we will move to present our research methods and describe our study case. Study results are described after that. Eventually in discussion part our findings are compared to findings from related work with some ideas why we agree, disagree or can't say anything.

2. RELATED WORK

GSD, agile software development, and their various combinations have been studied a lot during the last decade. Without any attempt of providing a thorough coverage, we summarize next some

key findings from a set of existing research. First, the distributed agile in general and Scrum in GSD are addressed. Then, the few works that were identified on agile maintenance are recapitulated.

2.1. Globally Distributed Software Development

An experience report from My Yahoo! on *adopting agile methods in globally distributed software development* was provided by [8]. Their distribution strategy was based on a hierarchy of Scrum Masters: Chief SM in the global product team, regional SMs in the regional teams, and local SMs in the local teams. Initially the following challenges were revealed: i) cultural differences in communication (reluctance to discuss negative issues), ii) disjointed and conflicting priorities within different products of the same company, and iii) team bonding and distributed collaboration with few or no face-to-face interaction. As a whole, using a refined form of distributed Scrum, roughly 1.5 increase in productivity was obtained (FPs/dev months) with the reduction of 64% of significant bugs.

Successful modifications of Scrum practices in GSD have been described in [9], [10], [11], [12] and [2]. In [2] Paasivaara et al. provide the following list: i) tool support for daily Scrums and Backlog management, ii) synchronized 4-week sprints (and 2-week sprints by the maintenance team to provide hot fixes), and iii) weekly Scrum-of-Scrums (SoS) joining together all POs (in one site) and SMs (in two sites). In addition, i) needs for multifaceted communication, ii) joint physical presence in the critical project phases, and iii) support of POs using domain expert network were nominated as the main external practices supporting global development actions. Challenges faced were related to bad network connections, cultural differences in daily scrums, and misunderstandings between POs and developers. However, the overall experiences turned positive, because improvements in communication, trust, motivation, and product quality were documented. In a follow-up, multiple case study, the same authors [13] concluded the following further lessons (cf. [14]): i) distribute for benefits not for fashion, ii) focus on Scrum training and open communication, iii) arrange frequent visits between the sites, and iv) provide and utilize multiple communication tools especially for informal communication.

In a longitude case study of agile distributed project in Denmark and India [9], reasons why Scrum works were concluded with nine answers (summarized here in four principles): i) it can build relations, networks, trust, common language and shared target for the team, has good meeting structure for team communication and gives energy and motivation for the team, ii) it is useful in work coordination and progress tracking, iii) it creates boundary objects and boundary spanner roles, and iv) it has a simple quality assurance mechanism.

In [10] and [11] Sutherland et al. concluded that it is possible for distributed Scrum to be as productive or achieve the same velocity and quality as with a collocated team. In [12] Sutherland et al. list advantages: i) it can reproduce performance of collocated Scrum, ii) clear communication via meetings that are facilitated by Scrum, iii) high quality by applying XP practices, and iv) accurate and transparent estimation. Challenged faced in their study were: i) cultural differences, ii) sharing context and priorities, iii) managing customers with no agile experience, and iv) special local requirements like documentation on a customer language.

Concerning the positive and measurable improvements in the projects when applying and tuning distributed agile as mentioned above, very different conclusions were provided in [14] by analyzing 66 European software development projects. Their results concluded "no statistically significant correlations between variables measuring outcomes of distributed processes and the type of development project followed—agile or structured". However, identification of effective practices (from both worlds) remained as further study, because many process clusters with mixed practices were established as part of the quantitative analysis.

2.2. Agile Maintenance

In [15] the tailoring of Scrum in a single site – single organization *maintenance* situation was studied using the action research approach. The authors point out many differences between (forward) development and maintenance, noting, for example, that maintenance sprints can be subject to interruptions by urgent customer demands. Nine heuristics were suggested for agile maintenance (summarized here in five principles): i) balance sprints for multiple customer portfolio, ii) allow unexpected sprint changes, iii) emphasize knowledge sharing in the team, iv) balance face-to-face communication and documentation, and v) find other team motivators to compensate the lack of "now the software product is ready" in maintenance.

A Case study on a medium-sized Norwegian software company having adopted Evo ('Evolutionary Delivery' by Tom Gilb, close to Scrum) in maintenance was provided by [16]. Their findings concerning the experienced challenges were related to i) comprehension of the existing system, ii) effects of small change to many components (i.e., low cohesion), iii) unstable testing and lack of test coverage, and iv) need of a guru in the team, being both technical and domain knowledge champion. In conclusion, based on [16] and [15] on agility and maintenance, one notices that many original values of Scrum were abandoned when doing maintenance, e.g., related to self-organizing teams and frozen sprint Backlog without external interruptions.

3. RESEARCH METHOD

We use case study [7] as the research method. Suitable case was decided on two principles: We wanted to have globally distributed project which was already using Scrum practices and we wanted to collect experiences which could be later used to improve performance of the corresponding projects. The selected project under study is a maintenance project including also new development (i.e., perfective maintenance). Project maintains and develops company's external web portal. The web portal is composed of main sites with several, different level country specific pages, and it is integrated to several different services which bring content to the pages. Some of integrations have legal requirements behind them.

The data was collected by semi-structured interviews of the project participants and stakeholders. The interviews were held via Lync and each interview was recorded. During each interview, notes were created on the fly and attendants were able to comment and correct them immediately. There were totally 12 interviews including all project participants and most participating stakeholders. The studied case project had also project manager but she was not interviewed because she was the one doing the study.

Interviews contained nine questions of which six were background questions, two larger, main questions, and one extra question. In the main questions, issues working well and needs of improvement were gathered and, in both questions, the interviewer encouraged participants to think all possible aspects related to the theme.

Interview data was then qualitatively analyzed. First each answer in the notes was coded and different answers were collected below the same main topic. For background and extra questions, only summarizing averages were calculated. When the open coding of cites from the interview notes was transformed into different levels of abstraction in axial coding, the categorizations of the two main themes were iterated around five times between the two authors commenting and proposing changes to one another's category labels. Finally, the axial coding was stabilized to three different levels.

In the qualitative analysis of works well and improvement issues, it was counted how many times each issue was uniquely mentioned by an interviewee. When the preliminary results were obtained, a meeting was arranged with the team where the participants were able to comment the raised issues. In the second part of the meeting, three improvement needs were voted for further study and actions how to improve the activities in them were planned. Here, however, we only report the main findings form the two main issues leaving analysis of the planned changes for further studies.

3.1. Research question

Research question for this study was: Is it possible to use Scrum in globally distributed maintenance project and, if yes, how it is implemented and what benefits and disadvantages it brings? Research question will be answered with interviews. Regarding benefits and disadvantages we will also compare our results to other studies.

Interview questions were:

1. Your role and location in project?
2. When you joined team?
3. Your previous experience of Scrum?
4. Your previous experience of globally distributed projects?
5. Your previous experience of Drupal?
6. Your previous experience of other web development?
7. What is working well in project generally / personally?
8. What you'd like to improve in project generally / personally?
9. If you have questions or problems related to your responsibilities, who do you ask for help?

3.2. Case Description

This case study is done on web maintenance and development project in global IT Company. Company's main area is North Europe, but it has offsite activities in several locations like in India. Both agile methods and global delivery model (GDM, term used for GSD in company) are in constant use in company and it is common to have both in same project. Same trend was also visible in the background questions; almost all participants had more Scrum or GSD experience than just this project. In table below average experience in years from background questions are presented.

Table 1. Average experiences in background questions

Question	Average (years)
Time in team	1.5
Scrum experience	2.6
Experience in globally distributed projects	3.7
Drupal experience	2.3
Web development experience	9.4

Web development and maintenance project had been on going about a year when this study was done and, before that, there was a development project which originally created the web portal. Between the development project and the project in a current form, there was a clear change of responsibilities from development to maintenance where also most of the project team members were changed. Average time in the project was 1.5 years while shortest time was 0.8 and longest time was 2.3 years.

The Web site was developed with Drupal which is open source Web Content Management System software. Half of participants had no earlier experience of Drupal but everyone had at least two years of experience at the web development.

The Scrum team was located in the three different offices: PO and one team member in Stockholm, Sweden, SM and two team members in Helsinki, Finland, and four team members in India, Pune. Other stakeholders were located in Helsinki and Stockholm and Project Manager in Jyväskylä, Finland.

When asking project members about their role in the project, there were actually ten different roles mentioned of which two were for the other stakeholders. Scrum team had an architect, test engineer, senior developer, Front End developer, and UX designer, in addition to usual Scrum roles of SM, PO, and a team member. Each team member had one role.

All scrum roles took care of their duties according to Scrum principles with exception that team members focused on own strength areas instead of everyone doing everything inside team. In addition to those project manager took care of additional management issues like handling project in company's ERP system, following project budget, communicating with project stakeholders and searching for new team members when needed.

The project applied sprints of two weeks which started with time limited sprint planning and ended in time limited sprint reviews. There were two grooming meetings each week which were used for presenting new issues in PB and providing story points for them. They weren't always needed, but they were reserved in beforehand anyway. Retrospectives were held regularly after each sprint. After the sprint reviews, there were bug hunts of which the whole team participated. On every second Monday, after last Wednesday review, all changes were deployed to production. All above team meetings including daily scrums were held via video meetings with Hangout and each team member had a web cam. All meetings were distributed and everyone was participating from own computer in the office, some Finnish and Swedish participants occasionally also from other places like home. Hangout was also used for smaller meetings and phone calls, whenever there was need to discuss or share screen to show something and threshold to start Hangout meeting was quite low.

All participants were encouraged to discuss and ask questions in the meetings and Flowdock as chat tool was in very active use in project. Flowdock was used for daily communication, general information sharing and for asking more details of user stories currently in development. It is remarkable how well communication succeed in project when there haven't been any visits to meet other participants face-to-face.

Project didn't have any special project room or corresponding even if most of Pune and Helsinki participants were located close to each other. Project status (PB and SB) was updated continuously in Jira task tool and its Scrum Board was used to visualize sprint status and tasks. Task statuses visible in Scrum Board where: to do, in progress and done. There were also special Jira dashboards which collected generally or role based most important issues of current sprint like one for PO which included showing all issues waiting for her approval and sprint burndown chart.

Project had also shared integration test environment where all changes were committed after peer code review. Continuous flow of actions in project code base was shown in Flowdock.

4. RESULTS

Regarding to our research question we were able to confirm that it is possible to use Scrum in globally distributed maintenance project. Project we interviewed was fully functional and no lack in service level was reported. We were able to catch more improvement needs than working well issues, but none of them were so severe that it would have prevented project work. We were able to identify some improvements that need to be handled still but we were also able to identify some clear benefits of Scrum usage.

Common conclusion based on working well issues and improvements is that this model is worth of use, but it requires some more development. Process and team issues were largely identified in both sides, which gives a contradictory result, but when looking into details it becomes clear that some actions and new skills are still needed while adjusting Scrum to this environment.

The axially coded information from the interviews for the two main categories is summarized in two mind maps as provided in Pictures 1 and 2. Results are described here separately for both main themes.

4.1. Issues working well

Issues working well have two main categories where the most issues belonged to: Team and Process. Tools, mostly collaboration tools, were the next biggest top group, and Communication, especially team communication, close to it in size. Atmosphere when interviewing people of the issues working well supported these findings: participants sounded to be generally happy to use Scrum and they seemed to know each other at least on some level even if there haven't been any face-to-face meetings.

In picture below findings from well working issues are shown in mind map. Findings strongly suggest that it is possible to use Scrum in globally distributed maintenance project and that it will create some benefits. Scrum model is able to bring some of its benefits and here they are visible especially in form of advantages for team and in Scrum process generally. Each top level will be described in more details below picture.
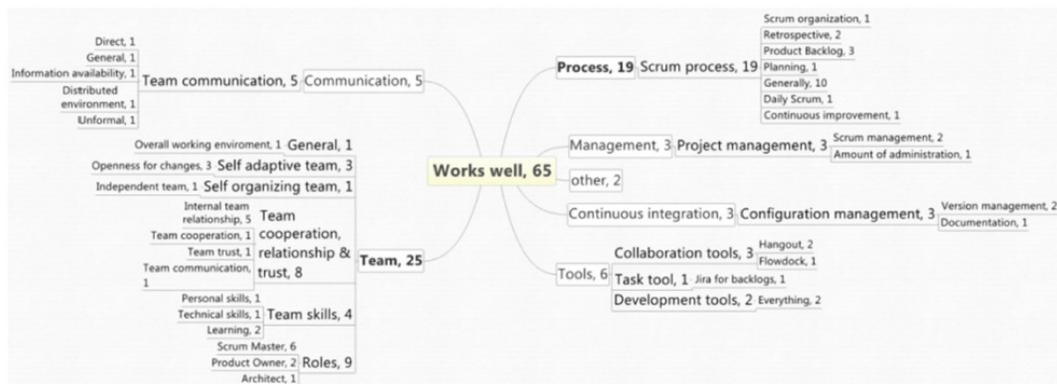
Figure 1. Works well issues mind map

4.1.1. Team

Within team, Roles got the highest amount of mentions and, especially, a good Scrum Master was found to have a central effect. The SM was good in facilitating meetings, and coordinating and supporting team. Another important issue was the combination of Team cooperation, relationship and trust, where especially internal team relationship was mentioned. Helping each other, positive attitude and asking questions are examples of things mentioned there. Also skills and Scrum related issues were mentioned: self-organized and self-adaptive team where self-adaptivity was presented as openness for changes. Therefore, we are able to say that Scrum in globally

distributed maintenance project was well-addressed by the team. It improved team spirit and trust and especially good Scrum Master was the key person in the team.

4.1.2. Process

All the good things in Scrum were linked to the top category of Process. Most comments were on the level that Scrum in general is good, but we were not able to identify exact practices behind these opinions. Interviewees generally liked to participate in a Scrum project, they saw that the use of Scrum brought many good things to the project, and it was also implemented well.

4.1.3. Other issues

Also issues linked to the categories of Tools, Communication, Management, Continuous Integration, and Other were revealed. In Tools, most mentions were for collaboration tools and, especially, Hangout as video meeting tool and Flowdock as instant messaging tool were mentioned. In Communication, there were several separate mentions but they were all related to team communication. In Management, all mentions were on Project Management and Continuous Integration was all about Configuration Management. There were two mentions in Other which we were not able to group: "service is working well" and "if no one notices your work, you are doing it well".

4.2. Needs of improvements

Improvement needs received more comments than the working well issues and they had three main top groups: Process, Team, and Communication. It is very interesting that Process and Team were also biggest top groups in the working well theme. It seemed to be easier for participants to identify improvement needs. One possible reason for this is that participants were used to have frequent retrospectives where they continuously handled improvement needs.

In the picture below findings from improvement needs are shown in mind map with three levels. Each top level will be described in more details below picture.
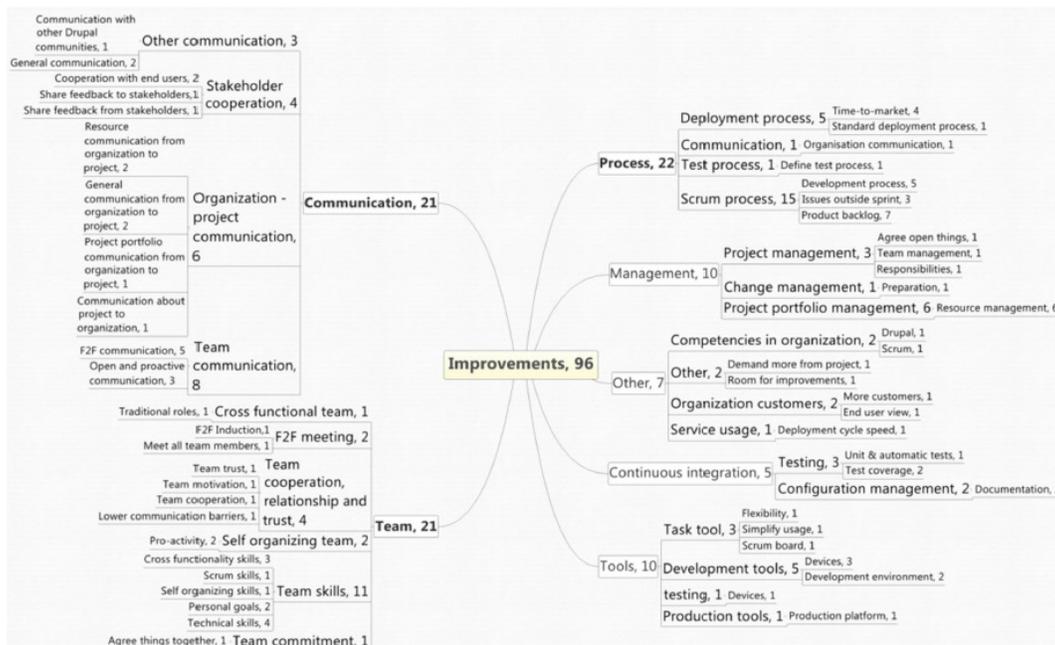

Figure 2. Improvement issues mind map

4.2.1. Process

Process was the biggest top category but only with one more mention than Team or Communication. Biggest part of Process improvements was Scrum Process, which, however, got almost as many mentions as in the Works well side. Concerning Improvements, interviewees were able to identify more specific issues.

PB was the most popular single item in Process Improvements and, there, six out of seven answers wanted more new development tasks in PB. The phase of development for the software was in a situation where the target was to maintain current service but also to create new functionalities. However, in some previous sprints the PO had prioritized mainly bug fixes in top of the PB. This would suggest that people are more interested in developing new issues than fixing bugs and Scrum is seen more as a development process model than as a maintenance process model.

In the development part related to Scrum, more flexibility and more time for real development work were the mostly raised issues. Issues outside project were mentioned too and there especially their interruptions to work. To reduce impact of these interruptions for each sprint one person was named to handle them. This is in line with suggestion above and also says that special attention must be paid on how to handle inevitable interruptions from production maintenance.

4.2.2. Team

Biggest part of team related improvements were related to skills in the team. Technical skills were the most popular issue and it is easy to see the connection to the background questions according to which the used technology was new for some team members. Another big part of team skills were the cross functionality skills. When it is compared to different roles in the team, it seems that the work inside the team was still done in a traditional (less self-adaptive) way where each participant had a certain role and corresponding tasks in the project. This suggests that more work is needed to develop team skills needed in Scrum model. This also creates need for further study: what will happen when we have fully cross-functional, self-adaptive team, will it change study results? Rest of the issues related to team improvements were mainly separated things, only mentioned once.

4.2.3. Communication

Communication issues were more evenly distributed within the different communication areas. Biggest group of communication was in the team communication where need for face-to-face communication received most mentions. The team has never actually met and it was obvious that they wanted to see each other. Another part in team communication was open and proactive communication and it is understandable that it is not so easy with persons you have never seen. Also cultural aspects need to be considered here because communication in the Nordic countries is different from India even if they weren't separately mentioned.

The next biggest issue in communication was the communication between the organization and the project. It was also the area where frustrations were visible in the interviews. It seems that fitting Scrum project in the organization and in response of maintenance was not well enough supported from the organization's side. There were some clear communication gaps between organization and project, like communication of other projects, where the same persons participated. The analyzed case is based on a project organization (meaning here that almost all work in this unit of company was made in projects) but on the maintenance side in company, ITIL (Information Technology Infrastructure Library for IT service management, see e.g. http://www.itil.org.uk/) practices were followed too. This suggests that more work is needed in how to implement maintenance Scrum project in company's processes to ensure communication between project and organization.

Another issue in communication was related to communicating with stakeholders. In stakeholder communication, there were two aspects with the same need of improvement present: stakeholders or team members didn't know how to get content in project's service in case they wanted to publish something in web portal maintained here. This refers to the phase before PB and to usage of the service.

4.2.4. Management

In the Management category, Resource Management in the Project Portfolio Management was clearly the biggest issue. There was concern that team wasn't stable, because constantly other projects were ongoing with the same members and new people joined the team. Project's role as internal development project, in a company whose main business area is to sell IT services to other companies, explains this partly. It is likely that this will also create a threat for achieving same performance as in co-located Scrum teams, but this question is left for further studies. Project and Change management got a few mentions here too.

4.2.5. Tools

In the Tools category, Development Tools created half of the mentions. Virtual servers, shared configuration management, connections and other technical issues have higher requirements in a globally distributed environment. It seemed that there were some differences in devices between locations. The project used Jira's Scrum Board as task tool and this also got some separate mentions.

4.2.6. Continuous Integration and Other

Testing and Configuration management were mentioned in Continuous integration. Especially testing coverage and documentation were found to need improvement. The separate, individual mentions without clear main context where gathered to the Other category.

4.3. Limitations

There are some identified limitations in this study. First, same person was participating in project as Project Manager and in study as Researcher. It has some limitations to results: we weren't able to interview project manager and we had to ignore all her opinions from project in order to be objective. Second, when interviewer already knew everyone to be interviewed, it might have had some effects in results. This was handled by keeping open atmosphere during interviews, encouraging participants to speak freely and letting them check notes on the fly. End result, this is now a personal opinion based on amount of results by participants, is that we were able to collect much more information than if interviewer would have been some unknown person. Third, we must consider that there were participants from three different countries (Finland, India and Sweden) with different cultures.

5. DISCUSSION

We were able to confirm that Scrum can be used in globally distributed maintenance project. Anyway, there are some aspects we found that must be addressed: i) interruptions from production, ii) ensure that team has required skills, and iii) communication between project and organization. These are partly in line with related work and in following chapters we will compare our results to them. Generally we were able to identify some same issues, some issues not, and we agree that it is possible and it might bring some benefits to use Scrum GSD.

In [3] were identified seven challenging factors. Synchronous Communication was not an issue here because biggest time difference was 4.5 hours between Sweden and India. The next biggest challenging factor in their literature review was *collaboration difficulties* and those were not visible in our case. Cooperation got only one mention and when we combined all closely related issues we still have a very small group. This is especially interesting result due to fact that there haven't been any visits or face-to-face happenings in the maintenance project and it is commonly

seen as important factor in global implementations of Scrum [13] and [17]. Communication bandwidth challenge wasn't either visible in our case but it was likely due to heavily used video conferences and chat, and the same situation was with the tools, because tools supporting distributed environments were ready and offered by the company. The same tools are also in use in other projects of the company.

Our improvement results match only somehow for challenges at [8]. We couldn't identify any of communication improvements into *cultural differences*. This is interesting because we had two different cultures involved: Nordic (Finnish, Swedish) and Indian. Conflicting priorities were found in two places: organization – project communication and project portfolio management. If combined, they would create a middle size top group in our study. Team collaboration issues are not common in our results, and internal team relationship got several mentions in the Works well side. That would suggest that lack of face-to-face can be handled with correct use of communication tools, but people will still miss it.

We were not able to identify all corresponding improvement needs for the challenges proposed in [2]. Network connections were not mentioned and cultural differences were not much visible. *Team communication* was one major category in improvements, but there were not any special comments of misunderstandings. We are able to confirm some positive experiences anyway. Team in communication and relationship was well working issue and Team and Process were biggest top groups in here. Possible reason for this might be that we were already following some of the lessons and recommendations as described in Section 2: i) shared tool supporting Backlog management, ii) distribution to achieve certain benefits like competence availability, iii) availability of several experienced Scrum persons, and iv) presence of many open communication channels whose use was encouraged. Let us also note that we had only one Scrum team under analysis.

For [9] we can confirm that i) team relation related issues were on working side, ii) for coordination and progress tracking we can agree that we didn't get that kind of issues in improvement side, and iii) it creates clear boundary objects like PB and it has *clear boundary spanner roles* like SM. Especially SM was seen as well working and important role in the project. For [10], [11] and [12] we can't estimate if velocity is close to collocated team, and for [12] we can only say that we had small group of team communication in both sides and our improvement needs are not similar to their challenges.

Based on our study, we cannot conclude rigorously whether Scrum in a distributed maintenance project is either good or bad solution [14]. We can conclude that we found more improvement needs than working well issues but, on the other hand, it seemed to be easier for participants to find detailed improvement needs. Scrum process was among the biggest groups in both sides' even if different details and levels of detail were mentioned.

Some of the nine heuristics as concluded in [15] were visible here too. It is good to remind that our project was slightly different than theirs; we had only one (internal) customer. Scrum in generally was seen as a working solution and work coordination issues were seldom mentioned as improvements, so Scrum with its sprints was a working approach to organize the maintenance work. Issues outside sprint were mentioned in Improvements and, for that, Heeager et al. just proposes to accept them. This same has been found in other studies [16] and, clearly, we can't follow Scrum strictly. In our study, team skills had a strong part in the improvements side which at quick glance seems opposite to their study, but by noting that we had, for example, several different roles in the team, this likely means that in our projects there was still some knowledge sharing to do. In our case, it looks that balance was much more on documentation than in face-to-face communication and, for project collaboration that was a good solution. Anyway, people still missed personal communication possibilities.

Compared to [16] we can find some correspondences. Comprehension, change effects and need of a guru can be linked to our findings for the need to improve the team skills. Test coverage and

testing was mentioned few times, but in our study it was not among the biggest improvement needs.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Brooks, F. P. (1987) "No Silver Bullet—Essence and Accidents of Software Engineering", *Computer 20*, pp1-16.

[2]   Paasivaara, M., Durasiewicz, S., Lassenius, C. (2008) "Using Scrum in a Globally Distributed Project: A Case Study" *Softw. Process Improve. Pract., 13,* Wiley InterScience, pp 527-544.

[3]   Hossain, E., Babar, M., Paik, H. (2009) "Using Scrum in Global Software Development: A Systematic Literature Review" *Fourth IEEE International Conference on Global Software Engineering*, pp175-184.

[4]   Lehman, M. M. (1980) "Programs, Life Cycles, and Laws of Software Evolution" *Proc. IEEE 68*, pp1060–1076.

[5]   Lientz, B. P., Swanson, E. B. (1980) *Software Maintenance Management* Addison-Wesley Longman Publishing Co., Inc., Boston, MA.

[6]   Omanovic, S., Buza, E. (2013) "Importance of Stable Velocity in Agile Maintenance", XXIV *International Conference on Information, Communication and Automation Technologies (ICAT),* IEEE.

[7]   Yin, R. (1994) *Case Study Research: Design and Methods,* Sage Publications, Unites States of America.

[8]   Lee, S., Yong, H.-S. (2010) "Distributed agile: project management in a global environment", *Empir. Software Eng. 15,* pp204-217.

[9]   Pries-Heje, L., Pries-Heje, J. (2011) "Why Scrum works: A case study from an agile distributed project in Denmark and India", *Agile Conference*, pp20-28.

[10] Sutherland, J., Viktorov, A., Blount, J., Puntikov, N. (2007) "Distributed Scrum: Agile Project Management with Outsourced Development Teams", *Proceedings of the 40th Hawaii International Conference on System Sciences,* IEEE.

[11] Sutherland, J., Schoonheim, G., Rijk, M. (2009) "Fully Distributed Scrum: Replicating Local Productivity and Quality with Offshore Teams", *Proceedings of the 42nd Hawaii International Conference on System Sciences,* IEEE.

[12] Sutherland, J., Schoonheim, G., Rustenburg, E., Rijk, M. (2008) "Fully Distributed Scrum: The Secret Sauce for Hyperproductive Offshored Development Teams", *Agile Conference,* pp339-344.

[13] Paasivaara, M., Durasiewicz, S., Lassenius, C. (2009) "Using Scrum in Distributed Agile Development: A Multiple Case Study", *Fourth IEEE International Conference on Global Software Engineering*, IEEE, pp195-204.

[14] Estler, H.-C., Nordio, M., Furia, C. A., Maye, B., Schneider, J. (2014) "Agile vs. structured distributed software development: A case study", *Empir. Software Eng, 19*, pp1197-1224.

[15] Heeager, L., Rose, J. (2014) "Optimising agile development practices for the maintenance operation: nine heuristics", *Empir. Software Eng., Springer Science+Business Media*, New York

[16] Hanssen, G., Yamashita, A. F., Conradi, R., Moonen, L. (2009) "Maintenance and Agile Development: Challenges, Opportunities and Future Directions", *Proc. ICSM 2009,* Edmonton, Canada, IEEE, pp487-490.

[17] Bannerman, P., Hossain, E., Jeffery, R. (2012) "Scrum Practice Mitigation of Global Software Development Coordination Challenges - A Distinctive Advantage?", *45th Hawaii International Conference on System Sciences*, IEEE; pp5309-5318.

AUTHORS

Minna Kilpala has ten year experience of project management and she works as senior project manager in Tieto. She has IPMA-C project management certificate and she is certificated Scrum Master and Scrum Product Owner. She has been working last 5 years with globally distributed agile projects. She has master degree in Computer Science and she is doing postgraduate studies at Jyväskylä University.

Tommi Kärkkäinen serves as a professor in the Department of Mathematical Information Technology, University of Jyväskylä. He is vice head and head of a PhD study line in COMAS (Graduate School in Computing and Mathematical Sciences) and principal professor of Software Engineering and Educational Technology MSc study lines. His research interests are related (but not limited) to computational sciences, software engineering, and computing education research. He has supervised more than one hundred MSc theses, over 20 PhD theses, and co-authored more than one hundred scientific publications.

INTENTIONAL BLANK

SUBSET SELECTION FOR LANDMARK MODERN AND HISTORIC IMAGES

Heider K. Ali[1] and Anthony Whitehead[2]

[1]Carleton University, Systems & Computer Engineering Department ,
Ottawa, ON, K1S 5B8, Canada
`heider@sce.carleton.ca`
[2]Carleton University, School of Information Technology,
Ottawa , ON, K1S 5B8, Canada
`Anthony.Whitehead@carleton.ca`

ABSTRACT

An automatic mechanism for the selection of image subset of modern and historic images out of a landmark large image set collected from the internet is designed in this paper. This selection depends on the extraction of dominant features using Gabor filtering. These features are selected carefully from a preliminary image set and fed into a neural network as a training set. The mechanism collects a raw large set of landmark images containing modern and historic images and non-landmark images as well, process these images, and finally classify them as landmark and non-landmark images. The classification performance highly depends on the number of candidate features of the landmark.

KEYWORDS

Feature Extraction, Neural Networks, Gabor Filters, Subset Selection, Image Categorization

1. INTRODUCTION

Classification of a large set of images containing both modern and historic images which leads to a selection of a subset of these images which fulfills certain technical standards is based completely on the detection and extraction of image features. Feature extraction has been received significant effort of researchers because of its great role in computer vision, image processing, and robotics fields [1]. Many standard feature extraction techniques have been applied to images to achieve optimal extraction performance and to maintain robustness by addressing the varying spatial resolutions, illumination, observer viewpoint, and rotation [2, 3, 4, 5]. Some success has been found by using a hybrid of standard techniques [6]. These techniques have achieved partial success in the feature matching in a mixed set of modern and historic images.

Object recognition techniques based on Gabor filters for feature extraction shown moderate success in extracting fundamental frequencies which represents the shape of an object [7]. Because the Gabor filters act as edge, shape and line detectors, as well as the tuning flexibility with different orientations and frequencies [8], these filters are often applied in a wide range of applications such as texture segmentation and classification [9], face recognition [10], fingerprint matching [11], and motion tracking [12].

In our case of mixed modern and historic images captured with cameras of different technologies under different environmental conditions, it is necessary to select the most appropriate image

subset from a huge image set with significant redundancy. This task can be successfully achieved by using Gabor filters for feature extraction which offer promising prospects in object recognition where the scale, rotation, translation and illumination invariant recognition can be realized within a reasonable computational time limit [8]. The selected features are trained on a neural network to recognize and extract the true features and to select the images of highest similarity.

This paper presents an automatic matching mechanism to select an image subset of landmarks out from a huge image set downloaded or collected from image websites like Google images or Flickr.

2. RELATED WORK

Selection of image subset from a large set depends widely on the image type and contents. This image nature reflected widely on the criteria of subset selection, and hence, on the efforts of researchers in image categorization and classification. Thorough analysis of image contents has to be done to extract the salient features and characteristics that characterize an image set. This extraction relies on the conditions of the image were taken in [1] .Traditional image set selection and classification is mainly relying on the analysis of the low-level features of the image to get high-level content semantics [14]. To optimize the feature detection and extraction process, the researchers worked to support their efforts by non-traditional techniques. Machine learning techniques like artificial neural networks (ANN) has obtained popularity in image subset selection [15].

Xiong et. Al. [16] proposed a BP neural network to improve the performance of image selection by segmenting and clustering the image into several visual objects and building the total feature vector, while Xu and Qu [17] developed a method based on feature matching similarity and frequency–inverse document frequency (TF-IDF) value of region is proposed. Multi-scale features used by Li and Zhao[18] in medical image classification and the classifiers use a set of complementary image features in various scales to compare the results of classification process. To optimize the feature selection process for a robust image classification, Al-Sahaf et. al. [19] introduced a Two-Tier Genetic Programming (GP) based image classification method which works on raw pixels rather than high-level features.

3. GABOR FILTERS

Gabor filters are based on the Gabor wavelets which are formed from two components, a complex sinusoidal carrier placed under a Gaussian envelope. These wavelets are based on the Gabor elementary function presented by D. Gabor in 1946 [13]. Many forms of the 2-D Gabor filter have been given. The 2-D Gabor filter $G(x, y)$ can be defined as [7]:

$$G(x, y = e^{-(\alpha^2 x_p^2 + \beta^2 y_p^2)} e^{j2\pi f_0 x_p} \tag{1}$$

Where α is the time duration of the Gaussian envelope and the plane wave, f_0 is the frequency of the carrier , $x_p = x\cos\theta + y\sin\theta$, $y_p = -x\sin\theta + y\cos\theta$ and Θ and β are the sharpness values of the major and minor axes of the Gaussian envelope. Gabor filters can be used effectively to make the classification with varying capture conditions [8]. In this work we examine the effectiveness of Gabor filters in a situation where other techniques have failed significantly [6]. Specifically, by using an image set of a landmark containing both historic and modern images with different objects captured by different technologies widely spread throughout time.

4. NEURAL NETWORKS

A multi-layer feed-forward neural network has been used in training a data set on input patterns. A three layer network was used: the input layer, a hidden layer and the output layer. For the input layer, the input feature vectors comes from the Gabor filter feature extraction stage and consists of 100 neurons applied to the neural network. These vectors are processed in the hidden layer using the scaled conjugate gradient method as the training method and the mean square error with regularization as the performance function to adjust the network output to be in the range -1.0 to 1.0. The output layer size depends on the obvious candidate features in each landmark. The following figure shows a typical multi-layer feed-forward neural network.
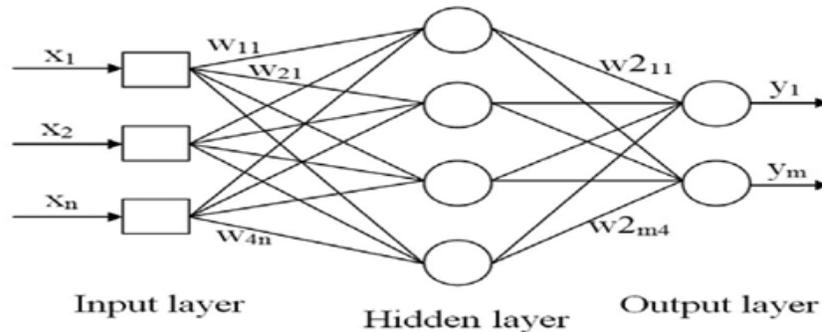
Figure 1. Multi-layer feed-forward neural network

5. DATA SUBSET SELECTION

5.1. Data sets

The data sets are divided into two categories: the preliminary data sets and the actual dataset. The preliminary data sets are set of modern and historic images for five landmarks. Every data set contains 300-350 images. This data set is primarily used as a training set to the neural network after the Gabor features has been extracted and fed into the neural network. Figure 2 shows some the images of the preliminary data set for Eiffel tower and Coliseum landmarks. Images of other three are shown in Appendix 1.

(a)   (b)
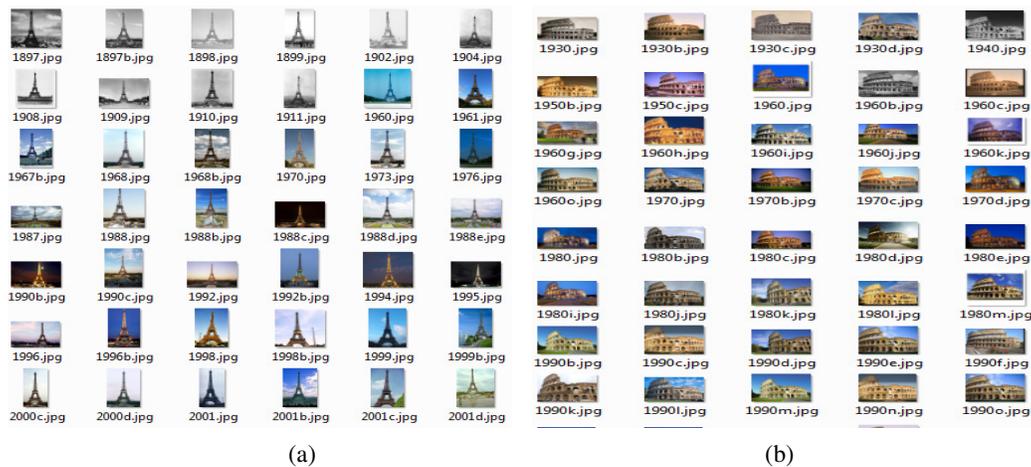
Figure 2. Examples of modern and historic images of many landmarks (a) Eiffel tower in Paris (b) Coliseum of Rome

To prepare these images for the next processing steps, these images have to be converted into gray scale images and resized to a common size. To overcome the problems of variations in brightness and contrast, as well as, the different lighting conditions, these images have to be preprocessed. To compensate for the effects of brightness and contrast, histogram equalization [20] was used. Specifically, adaptive histogram equalization (AHE) is applied to correct the problems of different lighting conditions. AHE computes many histograms; each histogram corresponds to a specific sector of the image and uses them to redistributes the image brightness values [21]. Finally, the input values are normalized to a range of [-1, 1].

The actual data set is comprised of sets of modern and historic images of five landmarks. Every set is 1000-3000 images gathered from Google Image Search. These images are placed in a folder, resized, converted to gray scale, and preprocessed to compensate for contrast and lighting problems. These images are applied to the system to classify them as matching or non-matching depending on the Gabor features of every landmark.

5.2. Feature extraction by Gabor filtering

To extract the image Gabor filters, it is required to create the Gabor wavelets. These wavelets are created using equation (1) above by using three control factors to create the Gabor kernels. Figure 3 shows the Gabor wavelets generated with three control parameters: scale, frequency and orientation.
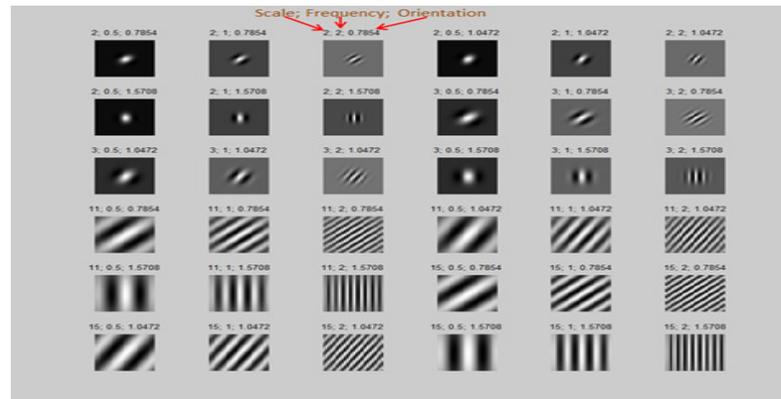

Figure 3. Gabor wavelets created using three control parameters

By applying these kernels on an image, we get the Gabor wavelets family shown in figure 4.
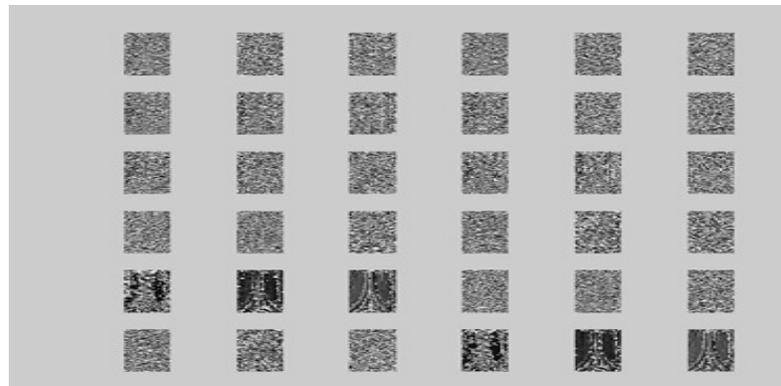

Figure 4. Gabor family

5.3. Neural network training

In the training phase of the neural network two sets of images are applied, a feature image set and non-feature image set. The feature image set consists of 100-150 images containing the feature to be tested, while the non-feature image set is 50 non feature images. Feature images completely depend on the landmark images being investigated. To facilitate the classification process, all landmark images are studied carefully to select the highest detectable features with Gabor filtering to be the candidate features. For example, the Eiffel tower images contain two regions which are assumed to have accurately detectable features. The image sets of these regions are fed into the neural network as the feature image set. For Eiffel tower landmark, the network output neurons will set to two. These outputs appear in a mutually exclusive basis, i.e. if the features exist in an image, one of the outputs will be greater than 0.8 or less than -0.5 and vice versa.

The image classification process described in algorithm 1 is shown in in figure 7. By a careful study the image set of Eiffel tower landmark; as an example, two features are nominated as candidate features that lead to a reasonable extraction and hence; a good classification. The training image sets of these two areas are fed into the neural network as training sets. Figure 5 shows a sample image of Eiffel tower, the two candidate feature regions, and the training sets fed into the network. The same process is done on the other four landmarks to select the dominant features and to prepare the training image sets of these features for next steps.
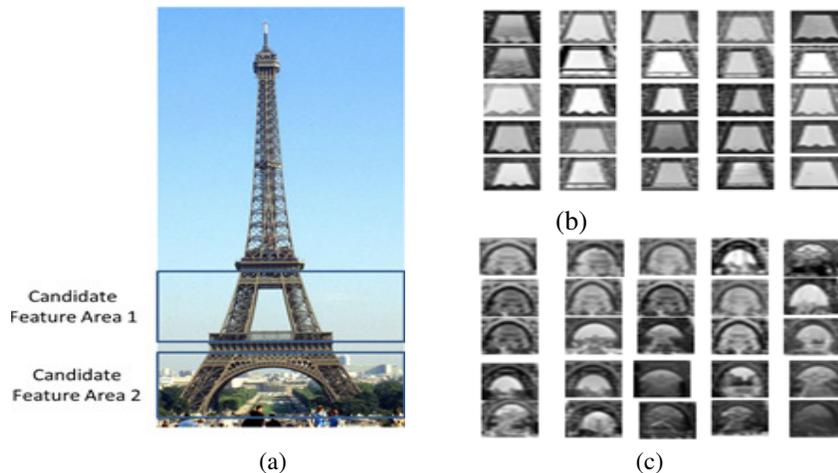
Figure 5. Eiffel tower image (a) Two candidate feature areas (b) Sample of training set for area 1 (c) Sample of training set for area 2

5.4. Performing Gabor filtering

The Gabor filtering is performed by convolving images with Gabor kernels. Generally, convolution can be done using Fast Fourier Transform (FFT). The FFT is done by each multiplying the frequency domain of Gabor kernels with the image pixel by pixel, i.e. dot product. Then, the inverse Fourier Transform (IFFT) has to be done to return the result back into spatial domain. Finally, the feature vector is built by converting the image data which convolved with the Gabor kernels to prepare these images to be trained on the neural network. Figure 6 summarizes the feature extraction process.
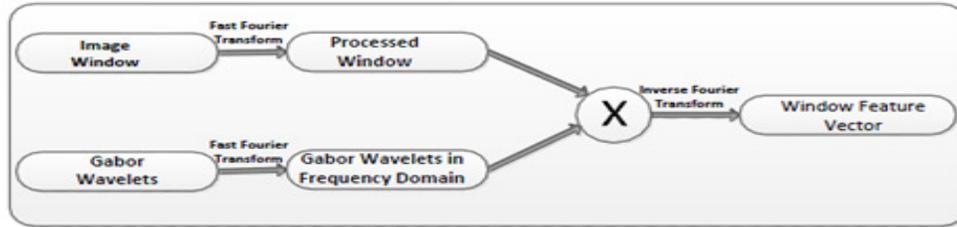
Figure 6. Feature Extraction Process

5.5. Candidate feature selection

For every landmark, the candidate feature(s) should be selected prior to training them on the neural network. The selection process depends on the nature of the landmark image. The candidate feature should be unique in the image, i.e. repeated features should not be selected. Also, the features with a larger texture similarity serve better in the selection process.

5.6. Image matching process

The image set of a landmark downloaded or collected from the internet is pre-processed by resizing them and adjusting the contrast, brightness and illumination. Then, these images are applied to the neural network recognition phase to classify them as matched images or not. The images are placed in the matched folder or un-matched folder according to algorithm 1 in figure 7.

Algorithm 1:
 Read the test feature-images.
 Read the test non-feature images.
 Set the number of candidate features.
 Set the training parameters:
   Epoch=300;
  Performance: MSE=0.3 x 10-3;
  Gradient= 10-6;
 Set neuron number to 1,
 Set all Neuron Detection factors to zero.
 Set Overall Matching factor to 1.
 While (neuron number ≤ number of candidate features)
 {
     Read neuron output;
     If (neuron output ≥ 0.8)
        Set Neuron Detection factor to 1.
     Multiply Overall Matching factor by   Neuron Detection factor.
     Increment neuron number.
 }
If (Overall Matching factor =1)
  Classify image as Matched images;
Else
  Classify image as Un-Matched image

Figure 7. Image classification algorithm

5.7. Experimental Results

The algorithm was applied to two different sets of images of five landmarks, Eiffel Tower, Coliseum of Rome, Pyramids of Giza, Dome Rock, and the Statue of Liberty. The first bunch of image data is the primary data set which comprised of 300-350 images for each of the landmarks. These images are preprocessed to improve their contrast and lighting and then cropped to maintain the most relevant parts.

The features that are most likely detected along with an example of problematic false positives are shown in figure 8. The false positive problem was partially addressed by selecting the features with highest matching scores. This leads to the selection of most unique features. The landmarks image sets were selected randomly from image websites as a mix of modern and historic images of five landmarks as well as non-landmark images. The landmark images had selected one, two, or three candidate features for evaluation. Each landmark set was applied to the system to classify them into proper folders, namely the landmark image folder and the non-landmark image folder.

(a)

(b)

(c)

(d)
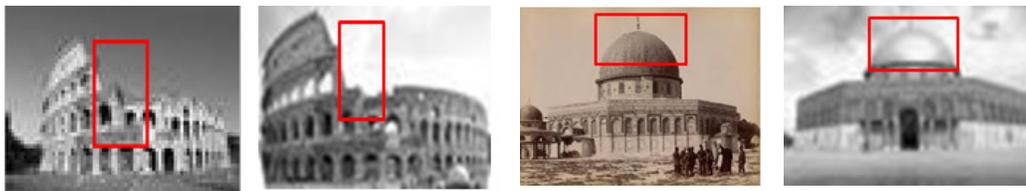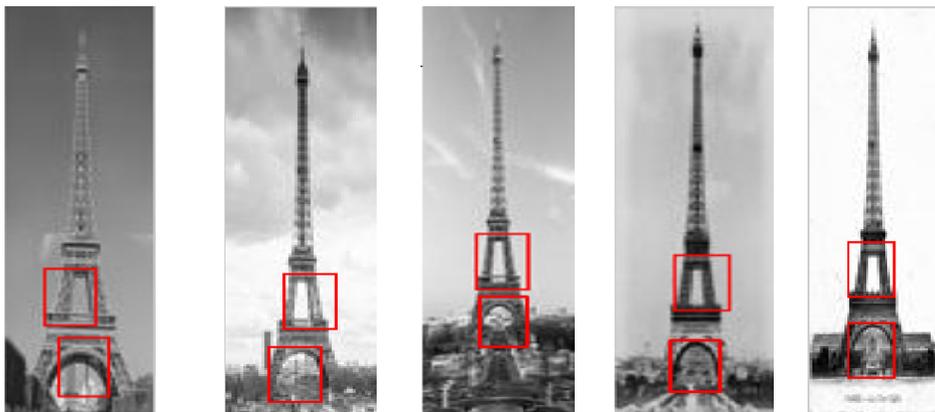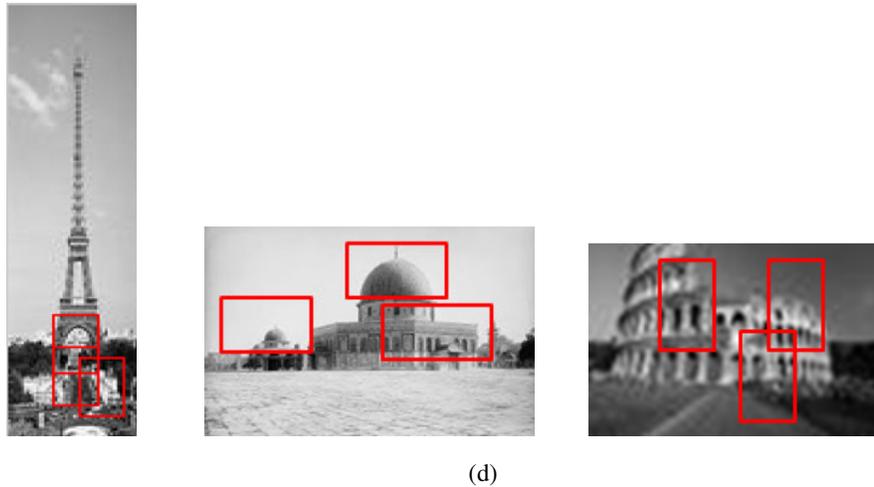
Figure 8. Feature extraction of images (a) 1- Feature (b) 2-Features (c) 3-Features (d) False positive features

The results are shown in table 1.

Table 1. Applied algorithms results on landmark images collected from the image websites

|	Total Images	True Positive	False Negative	False Positive	True Negative
One Feature Image Set	2833	1023	634	819	357
Two Features Image Set	2960	1417	780	521	242
Three Features Image Set	2219	1569	331	216	103

 The results shown in table 1 declare that this algorithm performed effectively on a mixed of modern and historic images of a large number of five landmarks collected randomly from the Google images. Despite the diverse capturing environments of these image in the sense of used camera technology, lighting conditions, and geographical variations of these images, this algorithm exploits the merging of Gabor filtering with the neural network setting to extract the dominant image features and to categorize images as landmark images or no.

To analyze the obtained results of table 1, the precision and recall metrics were applied to calculate the precision, recall, accuracy and the F-measure factor which is the harmonic mean of precision and recall [22] using the formulas given in table 2.

Table 2. Precision and recall metrics formulas

Precision	Recall	Accuracy	F1 Score
tp/(tp+fp)	tp/(tp+fn)	(tp+tn)/(tp+tn+fp+fn)	2*tp/(2*tp+fp+fn)

5.8. Subset Selection Mechanism Performance and Effectiveness

The results shown in table 3 reveals that the precision and the accuracy of this mechanism increased as the number of the candidate features of landmark images increased, and the overall performance enhanced by increasing the number of the extracted features.

The selection mechanism applied on five sets of landmarks images picked up the images with highly detectable features despite the considerable variances in images views, scales, and lighting conditions in an automated manner shows that this subset selection mechanism effectiveness is reasonably accepted and represents a promising image set categorization and selection.

Table 3. Accuracy table of the algorithm

	Precision	Recall	Accuracy	F1 Score
One Feature	0.555374593	0.617380809	0.487116131	0.826332795
Two Features	0.731166151	0.644970414	0.560472973	1.04267844
Three Features	0.878991597	0.825789474	0.753492564	1.482986767

6. CONCLUSIONS

Selection of the suitable images from a larger set of a landmark images containing both modern and historic images and captured under different lighting conditions and from different viewpoints represents a very important task when looking to match historical imaginary to modern imaginary. We show that Gabor filtering for such a task is a promising methodology when combined with a Neural Network for categorization. Dominant features from the images of many landmarks were selected as training image set to be fed to the neural network. In the categorization stage, the raw images were applied as input to the neural network. The categorization results have shown the proposed method to be a promising first step in creating historical to modern timeline image sets.

REFERENCES

[1]   M. S. Nixon & A. S. Aguado,(2012) *Feature Extraction & Image Processing for Computer Vision*, 3rd Ed. Academic Press.

[2]   M. D. Lowe,(2004) "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110.

[3]   H. Bay, et al.,(2006) "Surf: Speeded up robust features," In *European Conference on Computer Vision.*

[4]   E. Rublee, et al.,(2011) "ORB: an efficient alternative to SIFT or SURF", In *IEEE International Conference on Computer Vision*, pp. 2564-2571.

[5]   M. Calonder, et al.,(2010) "BRIEF: Binary Robust Inde-pendent Elementary Features", *Proceedings of the 11th European Conference on Computer vision*: Part IV, pp. 778-792.

[6]   H. K. Ali, & A. Whitehead,(2014)"Feature Matching for Aligning Historical and Modern Images", *International Journal of Computer Applications IJCA*, Vol. 21, No. 3, pp. 188-201.

[7]   J. Kamarainen, et al.,(2002) "Fundamental frequency Gabor filters for object recognition", In *Proceedings of the 16th International Conference on Pattern Recognition*, pp. 628-631.

[8]   S. Berisha,(2009) "Image classification using Gabor filters and machine learning", *MSc Thesis, Wake Forest University*, Winston-Salem, North Carolina.

[9]   A.C. Bovic, et al.,(1990) "Multichannel texture analysis using localized spatial filters", *IEEE Trans. Patter Analysis and Machine Intelligence*.

[10] Y. B. Jemaa & S. Khanfir,(2009) "Automatic local Gabor features extraction for face recognition", *Inter Journal of Comp. Science and Information Security*, Vo. 3, No. 1.

[11] M. U. Munir & M.. Y. Javed,(2004) "Fingerprint Matching using Gabor Filters", *In Proceedings of National Conference on Emerging Technologies.*

[12] B.S. Manjunath, et al.,(1996) "A new approach to image feature detection with applications ", *Pattern Recognition.*

[13] D. Gabor,(1946) " Theory of communications", *Journal of Inter-national Electrical Engineers*, Vo. 93,pp. 427-457.

[14] S. Chang et. al. (1998) ," Semantic Visual Templates: Linking Visual Features to Semantics" In *Proceedings of International Conference on Image Processing*. pp. 531-535.

[15] N. A. Mahmon & N. Ya'acob,(2014) "A Review on Classification of Satellite Image Using Artificial Neural Network (ANN)", In *IEEE 5th Control and System Graduate Research Colloquium*. pp. 153-157.

[16] Z. Xiong et. al., (2012) " An Algorithm of I mage Classification Based on BP Neural Network", In *IEEE International Conference of Computer Science and Automation Engineering (CSAE)* . pp. 523-526.

[17] D. Xu & Z. Qu, (2013) "An Image Classification Method Based on Matching Similarity and TF-IDF Value of Region", In *Sixth International Symposium on Computational Intelligence and Design*. pp. 112-115.

[18] B. Li et. al. (2013) " Multi-Scale Feature Based Medical Image Classification", *In 3rd International Conference on Computer Science and Network Technology*. pp. 1182-1186.

[19] H. Al-Sahaf et. al., (2012) " Extracting Image Features for Classification By Two-Tier Genetic Programming", In *IEEE World Congress on Computational Intelligence*.

[20] R. C. Gonzalez a& R. E. Woods,(2008) "Digital Image Pro-cessing" , *Prentice-Hall*, 3rd Edition.

[21]   D. J. Ketcham, et al.,(1974) "Image enhancement techniques for cockpit displays", *Tech. rep., Hughes Aircraft*.

[22] D. M. W. Powers,(2007)"Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation*", Journal of Machine Learning Technologies,* Vo. 2, No. 1, pp. 37–63.

APPENDIX 1

(a)

(b)

(c)
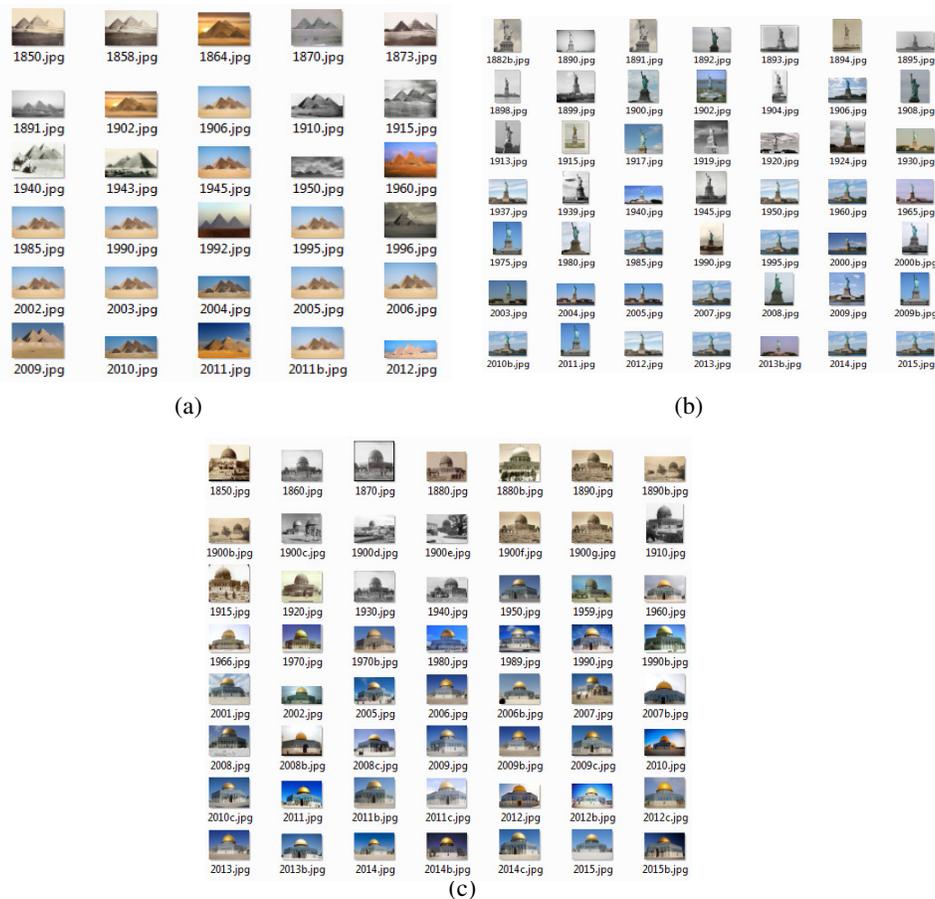
Figure A-1. Examples of modern and historic images of many landmarks   (a) Pyramids of Giza(b) Statue of Liberty (c) Dome of Rock

AUTHORS

Heider K Ali is doing his PhD in Electrical and Computer Engineering at Carleton University. He works on the feature matching of modern to historic photos of many worldwide landmarks and the investigation of the problems of dealing with historic photos. He got his MSc in computer engineering in 1991 and BSc in control and systems engineering in 1986 both from the University of Technology, Iraq. His fields of interest are computer vision, panorama building, historic to modern photos matching, software engineering and machine learning.

Anthony Whitehead is an Associate Professor and Director of the School of Information Technology at Carleton University. He is also the current Chair of the Human Computer Interaction Program at Carleton University. His research interests can be described as practical applications of machine learning, pattern matching and computer vision as they relate to interactive applications (including games), computer graphics, vision systems and information management.

INTENTIONAL BLANK

DIAGNOSING LEARNER ERRORS ON ARABIC CLOSED-CLASS ITEMS

Hayat Alrefaie and Prof. Allan Ramsay

The University of Manchester, Manchester, United Kingdom

ABSTRACT

Language learners are faced with a myriad of tasks: they have to learn how to make the sounds of the language they are learning, they have to learn its grammatical patterns, and they have to learn its vocabulary. There are numerous computational tools to help with the first two of these tasks. Providing help with the third raises a number of new challenges. The current paper describes a tool which will help learner to understand how to use 'closed-class' lexical items: this is a particularly taxing problem. To learn that the Arabic for 'office' is مكتب (mktb) is fairly straightforward: keep saying to yourself 'office'= مكتب (mktb) over and over again until it sticks. But learning the Arabic equivalent of 'on' is more difficult. The paper outlines a mechanism for providing diagnostic information about particular examples, with the aim of helping the learner to understand why a particular translation of a given closed-class item is appropriate in one situation but not in another.

KEYWORDS

CALL tool, Diagnostic, Meaning postulates, logical form & Arabic Preposition.

1. INTRODUCTION

The goal of the work here is to help people trying to learn Arabic to understand how to use closed- class items, particularly prepositions whose primary meaning is spatial, correctly. In general, learning vocabulary takes a lot of work; because in order to be fluent in a language you need to know a very large number of words (a typical adult speaker will have active knowledge of 10-40K words). Acquiring this set takes time and effort, but most of the time you just have to learn that the equivalent of '*office*' is مكتب (*mktb*), the equivalent of '*man*' is رجل (*rgl*), the equivalent of '*boy*' is ولد (*wld*), . . . Lots and lots of pairs, but mostly not too difficult.

Closed class items, however, pose a different kind of problem. There are, by definition, not that many of them. There are a few dozen English prepositions, and a similar set of Arabic ones. Their meanings, however, are very slippery, and the correct choice of preposition for locating one item with respect to another is hard to predict.

When you want to know the meaning of a word, it is common practice to look it up in a dictionary. With a word like '*in*', however, this turns out not to be a very effective strategy. A typical dictionary definition of '*in*' says 'used to indicate location or position within something'. When you look up '*within*' in the same dictionary, you find 'inside something', and when you look up '*inside*' you find 'an interior or internal part or place: the part within'. Dictionary definitions of prepositions generally take you in circles in this way, without bottoming out in anything which will tell you whether you should use '*in*' or '*on*' in any given case (why do

people get in a car but on a bus? Why is my office in the computer science building but on the second floor? . . .).

The key to understanding how prepositions work is the observation that they express a relationship between some parts of some facets of the entities being related. The second part of this statement involves ideas related to Pustejovsky's [30] notion of 'qualia structure' or Dowty's [15] more general notion of 'logical space':

There will be as many axes of logical space as there are kinds of measurement. [. . .] *Each axis might have a different mathematical structure according to the discriminations that can be appropriately made in each case* [p. 126].

Consider the examples in (1):

(1)   a. I read it in the park.
         b. I read it in the Guardian.
         c. I read it in January.

(1a) expresses a relationship between a reading event and a place, (1b) expresses a relationship between a reading event and a source of information' (1c) expresses a relationship between a reading event and an interval. It is hard to see how a single entity can be in the same relationship to a set of objects of such different kinds. To allow for this, we assume that objects can be viewed in different ways, and that the first task for a preposition is to pick out compatible views of the figure and ground–to pick out the physical view of reading in (1a), as a process of directing your gaze so that light from the object being read is captured on your retina; the abstract view, in (1b), as a process of transferring information from an external source into your mind; and the temporal view, in (1c), as something which starts at some moment and ends some time later.

Unlike Pustejovsky, we make no particular claims about the number of possible kinds of view that underlie natural language. We implement the notion by using a simple semi-partitioned type lattice, since this way of representing types can be managed extremely efficiently, and suffices for the task at hand. We will write $T_1 \sim T_2$ to say that T_1 and T_2 are compatible with respect to the type lattice.

Having found compatible views of the figure and ground, the preposition then states some relation-ship between appropriate views of the two entities. '*in*', for instance, says that the (selected view of the) ground has an interior which encloses some region in the relevant space, and that every part of the (compatible view of the) figure is contained in this region. '*on*', on the other hand, says that the (selected view of the) ground has an orientable surface, and that the (compatible view of the) figure has a set of contiguous points which lie an arbitrarily small distance from the surface.

Reducing '*F in G*' to $(S \sim S')\&(view(F, S) \subseteq interior(view(G, S')))$ avoids the circularity of saying that '*in*' means '*inside*' because we can give precise definitions of the interior of some entity in a given space, e.g.

- If *X* has an embedding *E* into R^n, then interior*(X)* is the set of members p of *X* such that *E(p)* is not a member of *X* but there are members p_1, p_2 of X such that *E(p)* lies on the line from $E(p_1)$ to $E(p_2)$. Insofar as our everyday notion of physical space can be captured by mapping points in space to R^3, this enables us to talk about the interiors of physical objects.

- If *X* is a dense partially ordered set of points, an interval has a set of lower bounds (points which are not preceded by any other point in the set) and upper bounds (defined likewise). The interior of an interval is then the interior of the set consisting of its lower and upper bounds. Most of the commonsense approaches to time discussed by authors such as [32]; [33]; [4] and linguistic approaches discussed by authors such as [35]; [23] involve at least dense partial orderings, and hence can be covered by this definition.

- The interior of a simple unstructured set is the set itself. To say that someone is in my class is simply to say they are a member of the set of people who I am teaching.

In other words, if we have two items placed in some abstract space with an appropriate structure (distance if we want to use '*at*', regionhood if we want to use '*in*', vertical direction if we want to use '*on*', . . .) then we can locate them with respect to each other in that space, using clear definitions that exploit the structure of the space. Of course, the two items have to exist in the same space–it is not possible to say that a physical object is contained in a temporal space (* '*London is in January*'), or that a temporal object is contained in an abstract space (* '*January is in my best paper on syntax*'). Hence the first task when interpreting a preposition is to find views of the entities in question which are in the same space.

 Not every item in a given space, however, has the properties required by a given preposition. It is perfectly possible to have a physical object which does not have a non-empty interior (* '*my office is in the second floor*'), or an interval that does not have upper and lower bounds (* '*I have been knowing how to play tennis since I was 12*' is problematic because once you know something, you don't stop knowing it, so the interval denoted by the verb '*know*' has no (natural) upper bound). For a preposition to locate one entity with respect to another, the entities in question have to have the right properties as well as being of the right general kinds. We will return to this in Section 2.

1.1. English and Arabic prepositions

The main aim of the current paper is to provide English learners of Arabic with diagnostic information about their use of essentially spatial prepositions by exploiting the notions described above. The problem is that while a given English preposition may have a 'typical' translation into Arabic, there will be numerous exceptions to this standard translation, and the causes of these exceptions are hard to discern.

Consider (2), where في(*fy*) and على(*ElY*) are normally taken mean '*in*' and '*on*' respectively.

(2)   a.   My office is on the second floor.
      b.   * مكتبي على الطابق الثاني. (*mktby Ely AlTAbq AlvAny.*)
      c.   مكتبي في الطابق الثاني (*mktby fy AlTAbq AlvAny.*)

The natural way for an English person to think of a floor is as a flat surface, with no interior, so that if you want to locate a room with respect to a floor you have to say that the room is '*on*' the floor. Arabic speakers, on the other hand, conceptualize a طابق *(TAbq)* as a container, like a very large room, and hence want to say the مكتب (*mktb*) is inside the طابق *(TAbq)* rather than on its surface. Why speakers of one language have different ideas of what various everyday objects are like is beyond the scope of the work reported here. It just seems to be the case. What matters is that it does seem to be the case, and it is this underlying conceptual framework that drives the choice of particular prepositions for linking different kinds of figures and grounds.

2. MEANINGS, MODELS AND INFERENCE

To make any computational sense of this idea, we have to give precise characterisations of how the various prepositions work, and we have to describe the entities that they can link. We use a framework in which natural language texts are translated, through compositional rules, into some suitable logic; we then write meaning postulates to capture the consequences of using particular combinations of words; and finally we carry out abductive inference to investigate the use of different prepositions for linking the same figure and ground in two languages.

2.1 Logical forms

The framework used here follows the tradition started by Montague [26]; [16] of annotating grammatical rules with semantic interpretations which can be used to 'construct an interpretation by combining the interpretations of the parts'. This process lets you construct formal paraphrases ('logical forms', LFs) in some suitable logic of any sentences that you can parse. There have been many developments in this area since Montague's original presentation of this idea: in particular, the choice of a suitable logic has been extended to cover the notion that natural language is situated, i.e. that utterances have to be interpreted with respect to the context in which they are produced [21]; [6]; [19]; [18] and to take account of the fact that natural languages make extensive use of constructions such as defaults [34]; [5] and propositional attitudes which require highly expressive logics [37]; [1]; [2]. The general principle remains the same: if you have a structural analysis of a text which is obtained by applying the rules of a grammar, you can annotate the rules of that grammar with semantic composition rules which will let you produce an LF. And if you have an LF in some suitable logic, and you have an inference engine for that logic, then you can reason about what has been said.

There are, clearly, two limiting factors to be considered here:

The structural analysis has to be obtained by using the rules of a grammar Modern data driven parsers often have no explicit grammar [25]; [28]; [27], or have a very (very) large context-free grammar induced from a treebank [11]; [9]. Clearly, it is not possible to annotate the rules of the grammar being used if there is in fact no grammar; and it is infeasible to do so in cases where the induced grammar contains tens of thousands of rules. Bos?Bos:08a describes a system which produces semantic interpretations using a fairly wide-coverage grammar, but grammar-based parsers can become very slow when faced with long (40+ words) sentences. In the present context, however, we are working with comparatively short sentences, since the task is to compare interpretations of English sentences set as translation exercises with interpretations of putative translations into Arabic. As such we can control the length of the sentences under consideration, thus making it possible to use an HPSG-like grammar and a fairly orthodox chart parser. The grammar we are using has an underlying set of principles, together with a small set of language specific constraints. We can thus use essentially the same grammar for parsing the English and Arabic texts. This has the enormous advantage that exactly the same machinery can be used for constructing LFs for both languages, which makes it comparatively easy to compare them. Fig 1 shows the LFs for (2a) and (2b).

These LFs are fairly orthodox. The main points to note are

- The LFs contain the surface utterance type–*claim, query*, . . . There are two reasons for this: (i) the utterance type is part of the meaning of the utterance, and hence it naturally belongs in the LF; and (ii) in some cases, notably the English determiner *'any',* there are quantifiers which are naturally seen as out scoping the utterance type: these simply cannot be accommodated without including the utterance type in the LF.

- Referring terms in the original text are represented as reference terms of the form *ref(λX (. . . X . . .))*. Such terms have to be anchored with respect to the context in order to determine the significance of the utterance[6].
- The terms in the Arabic LF make use of predicates with names like كتب *office*, where we use the English gloss to distinguish between distinct terms with the same root (the stem مكتب(*mktb*) is derived from the root كتب(*ktb*), so since we use roots rather than stems in our lexicon it is the root that appears in the LF).

utt(*claim*,
        on(*ref(λEown(ref(λ F(speaker(F))), E)&office(E)),*
            ref(λGfloor(G)&second(G, λH(floor(H))))))

utt(*claim*,
        *ElY(ref(λDowner(ref(λ E(speaker(E)))),D)&*كتب *office(D)),*
            *ref(λF*طابق *floor(F)* ثاني *& second(F))))*

Figure 1. Logical forms for (2)

You have to have an inference engine for the chosen logic There is no point in constructing any kind of meaning representation unless you are going to **do** something with it. In general, what you want to do when confronted with an utterance is to reason about it: to see what follows from it, to think about why the speaker produced, to try to decide what you should do in response to it, . . .

This means that you have to have some kind of engine that is capable of carrying out inference over formulae in the formalism that you have chosen as your representation. This is true no matter whether the meaning representation is just the original string of words and the inference engine is string-edit distance between sequences of words; or the meaning representation is the dependency tree for the text and the inference engine is some form of tree-edit distance [39]; [3], as is commonly employed for textual entailment (Dagan et al., 2005); or whether it is a translation into first-order logic and the inference engine is a theorem prover for first order logic [7]. Simple string/tree-based representations have two advantages: (i) they are robust, since to construct the representation you simply invoke a morphological analyser (for string-based representation) or some kind of dependency parser (for tree-based ones). It doesn't even matter that the analyser may not be particularly accurate, since you can get some information out of a parse tree even if contains some mistakes. (ii) the inference engines are fairly easy to implement–string edit distance algorithms which exploit lexical relations when decided the cost of substituting one term for another are straightforward and the more complex ones (such as [39]) are widely downloadable. They do not, however, support detailed chains of reasoning, and they do not generally pay much attention to whether the reasoning that is carried out is sound. Indeed, typical measures of success for shallow inference report on precision and recall values, where precision means 'How often are the entailments discovered by the system valid?' In other words, such systems recognise from the outset that the inference that they are carrying out may not be sound. Approaches based on translation into logic have the obverse advantages: (i) they are sound, since the target logics have well-defined proof and model theories, so the steps that can be carried out are guaranteed to lead to true conclusions if the premises are true. (ii) They are efficient, and hence can be used to carry out long chains of inference. They also have the obverse problems. In particular, it is difficult to construct them, since the standard approach involves parsing the input text using a grammar which has been annotated with semantic interpretations, and as noted above most systems for parsing free text with long sentences use either a grammar which has been inferred from a treebank, and which is very hard to annotate, or no grammar at all. For the current task, the need to carry out detailed chains of inference outweighs the difficulties that occur when you try to parse long natural texts. For our particular task we have a great deal of control over

what the learner will write, because their task is to choose the right prepositions to translate a set of target texts. If the target texts are fairly simple (say up to ten or twelve words) then the texts that users generate will also be fairly simple–you would hardly translate '*I am on the train*' by a sentence with 40 words in it. We therefore use translations into property theory[37]; [10] ,which allows properties and propositions to be treated as arguments and hence allows for finer-grained distinctions than are permitted in first-order logic, as our target representations; and we use a version of the theorem prover described in [31] as our inference engine.

2.2 Meaning postulates

In order to reason about the consequences of some utterance, you have to have some knowledge to reason with. If you have a perfect translation from English or Arabic into some formalism and you have a theorem prover for that formalism, you will still not be able to do anything unless you also have an appropriate body of background knowledge.

The knowledge that we require involves describing the circumstances under which some preposition can use for locating one entity with respect to another. We write these down as meaning postulates in property theory–typical rules for describing the use of the preposition *'in'* and *'on'* are given in Fig 2.

$\forall B \forall C$: {*in(B, C)*}
$\quad\quad \forall D \forall E \forall F$: {*view(B, F) & type(F,D) & dim(F, E)*}
$\quad\quad\quad\quad \forall G \forall H$: {*view(C,H)*
$\quad\quad\quad\quad\quad\quad$ & *type (H,G)&(D ~ G) & dim(H, E)*}
$\quad\quad\quad\quad\quad \forall I$: {*interior(H, I)*}
$\quad\quad\quad\quad\quad\quad F \subset I$ & *located(B, C)*

$\forall B \forall C$: {*on(B, C)*}
$\quad\quad \forall D \forall E$: {*view(B, E) & type(E,D)*}
$\quad\quad\quad\quad \forall F \forall C$: {*view(C,G) & type(G, F) & D ~ F*}
$\quad\quad\quad\quad\quad \forall H$: {*bottom(E,H)*}
$\quad\quad\quad\quad\quad\quad \forall I$: {*surface(G, I)*}
$\quad\quad\quad\quad\quad\quad\quad$ *touching(H, I)&located(B, C)*

Figure 2: Meaning postulates for *'in'* and *'on'*

The first rule in Fig 2 says that if the preposition *'in'* has been used to link to entities *B* and *C* then B will be successfully located with respect to *C* if there are views *F* and *H* of *B* and *C* where *F* and *H* have compatible types *D* and *G* and *F* has an interior *I* with the same dimensionality *E* as *H*. Under those circumstances *F* is a subset of *I,* and we have located *B* with respect to *C*. In other words *'in'* works for locating a figure entity with respect to a ground if the ground has a suitable interior. Similarly, the second rule in Fig 2 says that the preposition *'on'* may be used to locate a figure with respect to a ground if the ground has a view which has a surface, and the figure has to have a compatible view which has a bottom.

Given that Arabic has prepositions, في(*fy*) and على (*ElY*), which seem to carry the same meanings as 'in' and 'on', we can write parallel meaning postulates for them. The Arabic spatial prepositions seem to mean almost exactly the same as the English ones, and hence we use exactly the same rules for them. We also need meaning postulates to tell us what views, with what properties, particular words induce. The two languages do differ at this point: to continue with our running example, 'building' andمبنى (*mbnY*) both denote entities with interiors; but although *'floor'* and طابق (*TAbq*) are roughly equivalent, a floor is a 2D surface, with no interior, whereas a

طابق is a 3D space with an interior and no surface. To capture these observations, we need some general rules about what kinds of entities have interiors, and surfaces, and so on

$\forall B \forall C \forall D: \{embedding(C, B, R^D)\}$
　　　　$\exists E interior(B, E) \& \exists F(embedding(F, E, R^D))$

$\forall B \forall C: \{embedding(C, B, R^3)\}$
　　　　$\forall D : \{type(B,D)\} \& (\exists E top(B,E) \& type(E,D)) \& \exists F bottom(B, F) \& type(F,D)$

$\forall B \forall C \forall D: \{embedding(C, B, R^2) \& orientable(C)\}$
　　　　$\exists E(surface(B, E))$

. . .

Figure 3: Rules about R^N

The first rule in Fig. 3 says that if there is an embedding C of B into R^N then B has an interior which also has an embedding into this set. This is not the only kind of entity that can have an interior, but it is a very common one. The second rule says that if there is an embedding of B into R^N then B has members which are top and bottom points. Again, there are other kinds of entity that can have top and bottom points, but this is a common one. The third says that if there is an orientable embedding of B in to R^2 then B has a surface. The rules in Fig. 3 are themselves derivable from basic properties of RN and the definitions of *interior*, *top* and *bottom* for these spaces. We are not, in this paper, concerned with deriving such rules from first principles (see [29] for work along these lines). What we need is to be able to say what kinds of things have interiors and so on. We then need to say things about particular entities. It is at this point that the nuances of different languages emerge.

$\forall B :\{office(B)\}$
　　$type(B,physical) \& \exists C(embedding(C, B, R^3))$

$\forall B :\{$ كتب $office(B)\}$
　　$type(B, physical) \& \exists C(embedding(C, B, R^3))$

$\forall B :\{building(B)\}$
　　$type(B, physical) \& \exists C(embedding(C, B, R^3))$

$\forall B :\{$ مبنى $building(B)\}$
　　$type(B, physical) \& \exists C(embedding(C, B, R^3))$

$\forall B :\{floor(B)\}$
　　$type(B, physical)$
$\& \exists C embedding(C, B, R^2) \& orientable(C)$

$\forall B :\{$ طابق $floor(B)\}$
　　$type(B, physical) \& \exists C(embedding(C, B, R^3))$

Figure 4: Buildings, offices & floors

Fig. 4 says that the English and Arabic words for building and office denote entities with the same spatial properties, whereas the words 'floor' and طابق(*TAbq*), despite being rough translations of one another, differ when you look at them in detail–a 'floor' is an oriented 2D surface, a طابق *(TAbq)* is a 3D container.

Given these descriptions, it is easy to see why English and Arabic speakers use a preposition (*'in'*, في *(fy))* which talks about interiors for locating an office with respect to a building, but English speakers use one (*'on'*) which talks about surfaces for locating an office with respect to a floor while Arabic speakers continue to use في*(fy)* for these kinds of objects. *'in'* just won't work for locating the office with respect to the floor, because a floor doesn't have an interior.

3. MODELS VS ABDUCTION

Turning this observation into a CALL tool is not straightforward. There are two plausible ways to proceed: we can either use the theorem prover as a model builder, and compare the models that are generated from an English utterance and the English meaning postulate and an Arabic utterance and the Arabic meaning postulates; or we can adapt the theorem prover to work abductively, trying to fill in the gaps in the use of the prepositions.

3.1 Model building

Suppose that our learner has been asked to translate 'My office is on the second floor.', and they have done so by producing the sentence. مكتبي على الطابق الثاني *(mktby Ely AlTAbq AlvAny.)*

This is not an unreasonable translation, since على (*Ely*) is often taken as the translation of 'on'. Unfortunately, as we have just seen, it is wrong, because the طابق (*TAbq*) is conceptualised as a container rather than a surface. The models for these two sentences are shown in Fig. 5. The two models in Fig. 5 contain a substantial amount of information that is not present in the original LFs for the two sentences. That is as it should be–a model contains information that arises when you combine an interpretation with a body of background knowledge. If you look carefully, it is possible to see that the English model contains the fact that the office, #3228, has been located with respect to the floor, #3229, and that this has happened because the floor has an oriented embedding #3274 into R^2 and hence has a surface, #3265, whereas the طابق floor, #3655, in the Arabic model has an embedding into R^3 rather than R^2, and hence has no surface for the كتب office, # 3654, to rest on. The problem is that you have to look very carefully in order to spot the relevant differences between the two models. Given that we are concerned with the use of prepositions for locating one entity with respect to another, it makes sense to look for the entries in the models that correspond directly to the prepositions, and to then look to see whether the entities referred to in these entries are located with respect to one another. In order to use this as a CALL tool, however, you have to be able to explain why the entities linked by على are not located, and that is very hard to do just using the models.

3.2 Abductive reasoning

As noted above, the CALL tool is to be used in a setting where the learner has been invited to translate a sentence which uses a preposition to locate one object with respect to another from English into Arabic (or Arabic into English: the machinery will work in either direction). In order to diagnose any errors.

'my office is on the second floor'
[*office*(#3228),
bottom(#3228, #3272),
dim(#3228, 3),
interior(#3228, #3266),
located(#3228, #3229),
on(#3228, #3229),
type(#3228, *physical*),
floor(#3229),
dim(#3229, 2),
interior(#3229, #3268),
second(#3229, λB(floor(B))),
surface(#3229, #3265),
type(#3229, *physical*),
dim(#3266, 3),
interior(#3266, #3267),
dim(#3268, 2),
interior(#3268, #3269),
embedding(#3270, #3266, $\mathcal{R}^3$),
embedding(#3271, #3268, $\mathcal{R}^2$),
member(#3272, #3228),
touching(#3272, #3265),
type(#3272, *physical*),
embedding(#3273, #3228, $\mathcal{R}^3$),
orientable(#3274),
embedding(#3274, #3229, $\mathcal{R}^2$),
own(#*user*, #3228)]

مكتبي في الطابق الثاني.(*mktby fy alṭabq altāny*.)
[*office*مكتب(#3654),
ElY(#3654, #3655),
bottom(#3654, #3691),
interior(#3654, #3685),
type(#3654, *physical*),
*floor*طابق(#3655),
*second*في(#3655),
bottom(#3655, #3692),
interior(#3655, #3687),
type(#3655, *physical*),
interior(#3685, #3686),
interior(#3687, #3688),
embedding(#3689, #3685, $\mathcal{R}^3$),
embedding(#3690, #3687, $\mathcal{R}^3$),
type(#3691, *physical*),
type(#3692, *physical*),
embedding(#3693, #3654, $\mathcal{R}^3$),
embedding(#3694, #3655, $\mathcal{R}^3$),
owner(#*user*, #3654),
type(#*user*, *human*)]

Figure 5: Models for (2a) and (2b)

They may have made and provide them with appropriate feedback, then, we want to look for places where they have used a preposition to link two items without successfully locating them. We have a collection of rules that describe the conditions under which some preposition **will** successfully locate the figure with respect to the ground, as in Fig. 2. Given a model, we can look for places where someone has used a preposition without locating the relevant entities by looking at each rule that has *located(B,C)* has its consequent where the **first** element of the antecedent (which will always specify that some preposition has been used) is satisfied but the consequent is not. We can then use the theorem prover abductively, by allowing proofs with missing steps, where we can prespecify the number of steps that can be omitted. In order to find the minimal set of missing steps, we call the theorem prover once allowing one missing step, then allowing two, then . . . until a proof is found. This does increase the time taken for finding a suitable proof, but if we simply allow proofs with unbounded numbers of missing steps then we are likely to be led to incorrect diagnoses. Thus if we try to carry out an abductive proof that two items mentioned in the model of (2b) are located, we will find that the rule using على (*Ely*) would have worked if we had been able to prove *surface(#3655, B)* for some *B*. So the primary reason why (2b) failed to locate the مكتب (*mktb*) is that the طابق (*TAbq*) does not have a surface. We could, therefore, provide a message to that effect to the learner as the first round of feedback–'You tried to use على (*Ely*) as the translation of *'on'*, but it doesn't work in this case because although طابق (*TAbq*) is the correct translation of *'floor'*, a طابق (*TAbq*) does not have a surface'. We can follow this up with more detailed feedback. If the learner is not satisfied with this as a comment, he or she can ask why a طابق (*TAbq*) does not have a surface. An abductive proof of *surface (#3655, B)* produces the response that this would have worked if we had been able to show *embedding (B,#3655,R2)&orientable(B)*. It is harder to see how this could be turned into a comprehensible diagnostic message for the learner, but the general principle is simple. The reason why some attempted translation fails to produce a target consequence can be found by carrying an abductive proof of the target and recording the missing steps. In many cases, these missing steps can be converted to meaningful diagnostic messages; and if the learner wants to know more about what went wrong, then abductive proofs of the missing steps themselves can be derived, though as this process gets deeper the rules that are being used become more and more abstract, to the point where they will not in fact produce output will help the learner to understand what they have done wrong.

4. FUTURE WORKS

In order to turn the work described here into a functioning CALL tool, we have to carry out two major tasks:

1. We have to write meaning postulates for a much larger set of open-class words. We have rules like the ones in Fig. 2 for the major spatial prepositions, and we have general rules of the kind in Fig. 3 for things like the interior of an object, its surface, its ceiling, its start and end. For the tool to be useful in a classroom setting, however, we need a wide set of entities that can enter into these relationship, which means writing rules like the ones in Fig. 4 for a large number of words.

2. We need to convert the output of the abductive reasoner, which takes the form of a set of missing steps, into meaningful diagnostic messages. We anticipate using canned phrases with slots in for the purpose, but even doing this requires some care because the missing steps contain Skolem constants, which will be entirely opaque to the user, and deriving appropriate referring expressions from these is not a trivial matter [13]; [38].

The first of these tasks is fairly straightforward, if tedious. The second requires more work, but we believe that the outcome will be a tool which provides deeper information about how to use a given preposition than can be obtained using statistical methods such as those proposed by [17]; [14]; [36].

REFERENCES

[1]   Aczel, P. (1980). Frege structures and the notions of proposition, truth and set. In J. Barwise (Ed.), The Kleene symposium, Amsterdam, pp. 31–39. North Holland.
[2]   Aczel, P. (1988). Non-Well-Founded-Sets. Stanford: CSLI Publications.
[3]   Alabbas, M. and A. M. Ramsay (2013). Natural language inference for Arabic using extended tree edit distance with subtrees. Journal of Artificial Intelligence Research 48, 1–22.
[4]   Allen, J. F. (1984). Towards a general theory of time and action. Artificial Intelligence 23(2), 123–154.
[5]   Asher, N. (1992). A default, truth conditional semantics for the progressive. Linguistics and Philosophy 15, 463–508.
[6]   Barwise, J. and J. Perry (1983). Situations and Attitudes. Cambridge, MA: Bradford Books.
[7]   Bos, J. (2008). Wide-coverage semantic analysis with Boxer. In J. Bos and R. Delmonte (Eds.), Semantics in Text Processing (STEP 2008), Venice, pp. 277–286. College Publications.
[8]   Carnap, R. (1936). Testability and meaning. Philosophy of Science 3, 401–467.
[9]   Charniak, E. (2000). A maximum-entropy-inspired parser. In Proceedings of the 1st North American Chapter of the Association for Computational Linguistics Conference, NAACL 2000, Stroudsburg, PA, USA, pp. 132–139. Association for Computational Linguistics.
[10]   Chierchia, G. and R. Turner (1987). Semantics and property theory. Linguistics and Philosophy 11(3).
[11]   Collins, M. (1997). Three generative, lexicalised models for statistical parsing. In Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics, Madrid, pp. 16–23.
[12]   Dagan, I., B. Magnini, and O. Glickman (2005). The PASCAL recognising textual entailment challenge. In Proceedings of Pascal Challenge Workshop on Recognizing Textual Entailment.
[13]   Dale, R. and E. Reiter (1995). Computational interpretations of the Gricean maxims in the generation of referring expressions. Cognitive Science 19(2), 233–263.
[14]   De Felice, R. and S. G. Pulman (2008). A classifier-based approach to preposition and determiner error correction in l2 English. In Proceedings of the 22Nd International Conference on Computational Linguistics - Volume 1, COLING '08, Stroudsburg, PA, USA, pp. 169–176. Association for Computational Linguistics.
[15]   Dowty, D. R. (1979). Word Meaning and Montague Grammar. Dordrecht: D Reidel.
[16]   Dowty, D. R., R. E. Wall, and S. Peters (1981). Introduction to Montague Semantics. Dordrecht: D Reidel.

[17] Felice, R. D. and S. G. Pulman (2007). Automatically aquiring models of preposition use. In Proceedings of the Fourth ACL-SIGSEM Workshop on Prepositions.

[18] Groenendijk, J. and M. Stokhof (1991). Dynamic predicate logic. Linguistics and Philosophy 14, 39–100.

[19] Heim, I. (1983). File change semantics and the familiarity theory of definiteness. In R. Ba¨uerle, C. [20] Schwarze, and A. von Stechow (Eds.) . Meaning, Use and Interpretation of Language, Berlin, pp. 164–189. Walter de Gruyter.

[21] Kamp, H. (1984). A theory of truth and semantic representation. In J. A. G. Groenendijk, T. M. V. [22] Janssen, and M. B. J. Stokhof (Eds.), Formal Methods in the Study of Language, Dordrecht, pp. 277–322. Foris Publications.

[23] Kamp, H. and U. Reyle (1993). From discourse to logic: introduction to model theoretic semantics of natural language. Dordrecht: Kluwer Academic Press.

[24] Loveland, D. W., D. W. Reed, and D. S. Wilson (1995). Satchmore: Satchmo with relevancy. Journal of Automated Reasoning 14(2), 325–351.

[25] McDonald, R., K. Lerman, and F. Pereira (2006). Multilingual dependency parsing with a two-stage discriminative parser. In Tenth Conference on Computational Natural Language Learning (CoNLL-X), New York.

[26] Montague, R. (1974). The proper treatment of quantification in ordinary English. In R. Thomason (Ed.), Formal Philosophy: Selected Papers of Richard Montague, New Haven. Yale University Press.

[27] Nivre, J. (2010). Dependency parsing. Language and Linguistics Compass 4(3), 138–152.

[28] Nivre, J., J. Hall, J. Nilsson, A. Chanev, G. Eryigit, S. K¨ubler, S. Marinov, and E. Marsi (2007). MaltParser: A language-independent system for data-driven dependency parsing. Natural Language Engineering 13(02), 95–135.

[29] Pratt-Hartmann, I. (2005). Temporal prepositions and their logic. Artif. Intell. 166(1-2), 1–36.

[30] Pustejovsky, J. (1991). The generative lexicon. Computational Linguistics 17(4), 409–441.

[31] Ramsay, A. M. (2001). Theorem proving for untyped constructive _-calculus: implementation and application. Logic Journal of the Interest Group in Pure and Applied Logics 9(1), 89–106.

[32] Reichenbach, H. (1956). The Direction of Time. Berkeley: University of California Press.

[33] Reichenbach, H. (1958). The Philosophy of Space and Time. New York: Dover Books.

[34] Reiter, R. (1980). A logic for default reasoning. Artificial Intelligence 13(1), 81–132.

[35] Smith, C. S. (1991). The Parameter of Aspect. Dordrecht: Kluwer Academic Publishers.

[36] Tetreault, J. R. and M. Chodorow (2008). The ups and downs of preposition error detection in esl writing. In Proceedings of the 22Nd International Conference on Computational Linguistics - Volume 1, COLING '08, Stroudsburg, PA, USA, pp. 865–872. Association for Computational Linguistics.

[37] Turner, R. (1987). A theory of properties. Journal of Symbolic Logic 52(2), 455–472.

[38] Van Deemter, K. (2000). Generating referring expressions: beyond the incremental algorithm. In H. C. Bunt (Ed.), 4th International Workshop on Computational Semantics, University of Tilburg, pp. 50–66.

[39] Zhang, K. and D. Shasha (1989). Simple fast algorithms for the editing distance   between trees and related problems. SIAM J. Comput. 18(6), 1245–1262.

AUTHORS

1. Hayat M A. Alrefaie
Postgraduate Student
School of Computer Science
University of Manchester

2. Prof Allan Ramsay

Professor of Formal Linguistics
School of Computer Science
University of Manchester

AUTHOR INDEX