SYNOPSIS OF

Exploration of novel descriptors for online writer identification

A THESIS

to be submitted by

Vivek Venugopal

for the award of the degree

of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRONICS AND ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI - 781 039, INDIA

January, 2019

Contents

1 Introduction

In the recent decade, the concerns of security have led to the necessity of employing technology to identify a person. Person identification through the medium of biometric systems refers to the recognition of individuals by their unique characteristics or traits [1]. The crux of the present dissertation is in the exploration of strategies for establishing the identity of the user with his / her handwriting. As a behavioural biometric, the goal of writer identification is to determine the authorship of a handwritten document by contrasting it against a set of enrolled samples with known authorship stored in a database. This area of research falls under the broader field of automatic handwriting recognition. However, there does exist a fine distinction between the two domains. The main focus of handwriting recognition is to rely on invariant representations of the handwritten content, typically achieved by reducing the extent of inter-writer variations. Contrast to this, works on writer identification exploit the writer-specific variations for discriminating the characteristics between the writers.

The recent trend in technology has enabled the manufacture of hand-held devices that provide a pen based input interface to capture the handwritten data. These primarily make use of an electronic stylus pen to record the data on a pressure-sensitive screen. The pen keeps track of the trajectory information of the handwriting, such as the (x, y) spatial location and time stamp. In the literature of handwriting analysis, the processing of such data is referred to as 'online'. Without loss of generality, the input to such a system is a handwritten document consisting of a sequence of strokes, which in turn contains a sequence of points. In this research, we focus our proposal towards such a text-independent online writer identification system.

Over the last decade, there has been considerable research in the area of online writer identification. The techniques proposed by researchers include the use of Gaussian Mixture Model-Universal Background Model (GMM-UBM) as in [2], information retrieval approaches [3–7], shape primitive based description [8–10], to name a few.

2 Contributions of the thesis

In this thesis, we propose a number of novel writer descriptors for text independent online writer identification systems. Our main focus is in relying on additional information as obtained from a pre-learnt codebook or sparse dictionary that can help in providing cues in deriving these descriptors. Our different proposals have been laid out in three main contributing chapters and we elaborate on the same in the following subsections.

2.1 Chapter 2: Exploration of codebook descriptors

In this part of the thesis, we derive a strategy that encodes the sequence of feature vectors extracted at sample points of the temporal trace with descriptors obtained from a codebook. These descriptors aim to capture the relative location of the feature vectors corresponding to the handwriting samples of a writer with respect to their nearest codevector in the feature space. The idea of the same comes from the intuition that the relative location of feature vectors of the same writer are more or less aligned in near-by proximity in the feature space. However, such a trend may not be prevalent when considering the feature vectors across different writers.

We begin by investigating on the merit of the Vector of Locally Aggregated Descriptor (VLAD) approach, a codebook based descriptor popularly used in the field of image retrieval [11], for online writer identification. However, we demonstrate that at times, the formulation of the VLAD is confronted with an issue, that can reduce the discrimination between writers. This can occur when the point based feature vectors corresponding to the different writers are not linearly separable in the Voronoi cell of a codevector - thereby affecting the identification rate. To alleviate this drawback, we propose a set of descriptors with a modified formulation aimed at improving writer discrimination beyond VLAD.

In our proposal, we first assign each feature vector belonging to a document to a specific codevector based on distance criterion. The proposed descriptors take into consideration, the scores of each of the attributes in a feature vector with regards of the proximity to their corresponding value in the assigned codevector. We show in this Chapter that such explicit scoring can provide useful cues for better discriminating handwritten samples of the different writers enrolled to the system, when compared to the VLAD. We formulate our strategy in a way that, for a given codebook size M, we employ the descriptors of only $M-1$ codevectors to construct the final descriptor by concatenation.

The performance of the VLAD and our proposed descriptor are evaluated on two publicly available online handwriting datasets, namely the IAM and IBM-UB1 databases. The results obtained demonstrate an improvement in writer identification rate over the system that uses the conventional VLAD formulation. Moreover, since our devised descriptor has a dimension twice that of the VLAD, we consider reducing it to half. Despite this reduction, we report improved results.

Apart from the above contribution, for constructing the codebook, and subsequently the descriptor, we derive features by incorporating a gap parameter. This aids in capturing the information from the sample points in the neighbourhood / vicinity of the point under consideration. An empirical study is also conducted on the variation of the writer identification rates for different values of the gap parameter.

2.2 Chapter 3: Exploration of sparse coded based descriptors

In the first part of the thesis presented in Chapter 2, the codebook used for formulating the writer descriptor is generated by applying the k-means algorithm. This however comes with a limitation that each feature vector gets assigned to only one prototype. As an alleviation to this issue, we consider proposing the descriptors with respect to an over-complete dictionary obtained via a sparse representation framework. Such a strategy ensures that the feature vectors are expressed as a linear combination of atoms that serve as prototypes. To the best of our knowledge, the use of sparse coding approaches for online writer identification have not been addressed till date in the literature.

The traditional pipeline typically followed in works on sparse coding involve pooling the obtained coefficients with either maximum or average operation. Though these methods do work satisfactorily, we do believe that there can still be room for improvement. In the remaining two Chapters, we demonstrate that exploring additional information from the sparse coefficients can provide us with useful cues that can model the dynamic characteristics of the writer. The incorporation of such information in turn leads to a better performance of the online writer identification system.

Keeping in line with the above idea, in Chapter 3, we move on to proposing a methodology to encode the sub-strokes of the online handwritten trace with descriptors derived from the set of dictionary atoms obtained in a sparse coding framework. The usage of sparse representation stems from the flexibility it offers over the hard clustering algorithms such as k-means in terms of representing the segmented handwritten sub-stroke of a writer as a combination of more than one prototype / atom.

Our work attempts at capturing the similarity of the attributes of each feature vector to the corresponding value in the subset of dictionary atoms that contribute with a non-zero sparse coding coefficient. Keeping this in perspective, we derive descriptors from each of the dictionary atoms, which incorporate the similarity scores of the attributes in a feature vector. The scores are in a way, indicative of the reconstruction error obtained while employing a particular dictionary atom alone for reconstruction. The idea of computing similarity scores for each attribute separately leads to a writer descriptor that captures the dynamic characteristics at a finer level there by leading to performance improvement in the IAM and IBM-UB1 databases.

In addition to the above contribution, we consider several sets of histogram based attributes / features at the sub-stroke level for constructing the dictionary atoms and subsequently the descriptors. We provide a thorough and comprehensive analysis of the histogram based attributes by addressing the issue of the selection of the bin size to be used and their relation to the writer identification rate. More specifically, we propose an entropy based strategy for determining the appropriate bin size to be selected for the generation of these histograms. Considering that the descriptors of the sparse coding framework are constructed from the resulting histogram feature sets, the selection of the bin size using our method ensures a good discrimination between the writers. This also gets reflected in the experimental results, where we show a dependence between the bin size chosen by the entropy based analysis and the corresponding writer identification performance.

2.3 Chapter 4: Exploration of saliency information in the sparse framework

In this Chapter, we once again utilize the sparse coding strategy for identifying the authorship of online handwritten documents. The contribution is on the idea of exploring additional information, that is related to quantifying the degree of importance of each of the dictionary atoms with regards to the dynamic characteristics of the writers. In this context, we define in this work, the term "saliency" [1] for an dictionary atom - the value of which is learnt from the sparse coefficients corresponding to the sub-strokes of the handwritten training data. In this Chapter, we propose two different approaches at obtaining the saliency values.

(i) In the first approach, we make use of the entropy measure computed over histograms generated from the sparse coefficients as a measure of saliency.

(ii) The saliency factor in the second approach is computed as a function of the sum pooled sparse coefficients obtained from the feature vectors corresponding to the enrolled documents of the writers in the database. Our formulation used bears resemblance to the inverse document frequency (idf) popular in the field of information retrieval.

The pre-learnt saliency values are incorporated on the traditional schemes of max and average pooling of sparse codes, thereby resulting in a modified writer descriptor. It has been shown through experiments on the IAM and IBM-UB1 databases that the modified writer descriptor obtains improved performance when compared to the traditional average or max pooled writer descriptor.

Further to the above, we also consider incorporating writer-specific adaptation of saliency values, that quantifies how important a dictionary atom is for a given writer. We employ the reconstruction error on the sub-stroke based feature vectors to derive a similarity score for each dictionary atom with regards to a writer using only his / her handwriting. The obtained scores across all the dictionary atoms are subsequently

[1] In this thesis on online writer identification, this term is not to be interpreted to that used by researchers of the image processing and computer vision community.

fused with their respective saliency values to generate adapted values for the purpose of identification. In particular, we formulate an ensemble of SVMs, wherein the descriptor to the SVM trained for a writer is based on the saliency values adapted for that writer. The final decision on the authorship is proposed as the maximum of the prediction score obtained from the SVMs.

3 Database description

In this section, we present an outline of the two databases used, together with their enrollment and evaluation protocols adopted in the experiments discussed in the thesis.

- **IAM Online Database (IAM)**: The IAM Online handwriting database [12] consists of samples of 217 writers acquired from a whiteboard. To acquire a database of handwritten sentences contained in the corpus, the texts in the corpus are split into fragments of about 50 words each. These fragments were then copied onto forms on paper and each writer was asked to write down the text contained in eight forms on the whiteboard. The online handwriting data collected contains x-coordinate, y-coordinate and time stamp information.

- **IBM_UB_1 Handwriting dataset**: The IBM_UB_1 handwriting dataset [13] contains online handwriting data from 43 writers, collected on the IBM Crosspad. For each document written by a specific writer, a summary text - a one to two page essay on a particular topic, and a corresponding 'query' text was generated. The 'query' text is a collection of approximately 25 words that attempt to distill the essence of the summary text. The basic attributes collected in this database comprise the spatial x-coordinate and y-coordinate.

In the thesis, the following protocol is used for the IAM Online Database. Out of the eight paragraphs written by each writer, four are chosen at random for enrolment and the remaining are used for performance evaluation. For the IBM-UB1 handwriting data-set, we are using the protocol as suggested in [6] - 80% of the summary document paragraphs are selected at random from each writer for enrolment and the remaining are used for judging the efficacy of the proposed method. For each database, the codevectors for the codebook in Chapter 2 and the atoms of the dictionary for Chapters 3 and 4 are learnt from the enrolled data corresponding to 25% of the number of users, that are randomly chosen. Moreover, the results of each experiment are reported both at paragraph and text-line level in the thesis.

4 Results and Discussion

In the thesis, we have elaborated a number of experiments to demonstrate the different aspects of our proposals. However, for sake of completeness of this synopsis report, we confine our discussion to only one result from each of the three Chapters.

- Recall that in Chapter 2, we formulate a writer descriptor that addresses the limitation of the VLAD. Accordingly in Figure 1, we highlight the trends of the VLAD with our proposal with regards to writer identification rates for varying sizes of codebook . The average writer identification rates for the two techniques is depicted in Figures 1 (a)-(d) for codebook sizes ranging from 5 to 100 in steps of five. Note that the blue and red curves denote the performance of VLAD and our proposal respectively. The sub-figures in the first column (a), and (c) represent the paragraph level identification rates for the IAM and IBM-UB1 databases respectively. Likewise, the text line level performance for these databases are plotted on the remaining sub-figures of (b), and (d). For simplicity, we abbreviate the average writer identification rate as IR.

  The inferences from the different plots of Figure 1 across the databases are as follows:

  (i) For the varying codebook sizes being considered, our descriptor provides a better performance than VLAD. This is owing to its ability to assign scores based on the residual / distortion values, thereby increasing the discrimination of the feature vectors of the writers.

  (ii) Taking into consideration, the trend of the performance of VLAD and the proposed descriptor, we see that the average identification rate initially increases with codebook size and becomes comparable at moderate sizes.

  (iii) Owing to coarser quantization, small-sized codebooks do not adequately capture the nuances that discriminate the features of different writers and hence the low average identification rate.

- In Chapter 3, the atoms of a dictionary (pre-learnt from the training phase) are used in conjunction with the attributes of sub-stroke feature vectors, to generate a descriptor, that encapsulates the writer specific characteristics of the document under consideration. Table 1 compares the performance of the proposal with the traditional average pooled and max pooled writer descriptor. For this experiment, the number of atoms in the dictionary was varied from 50 to 500 in steps of 50. For brevity sake, we abbreviate the average identification rate as IR and dictionary size as M. Owing to the fact that our proposed descriptor aids in capturing details of the writer at a finer level when compared to traditional sparse representation based classification strategies, we obtain improved results.

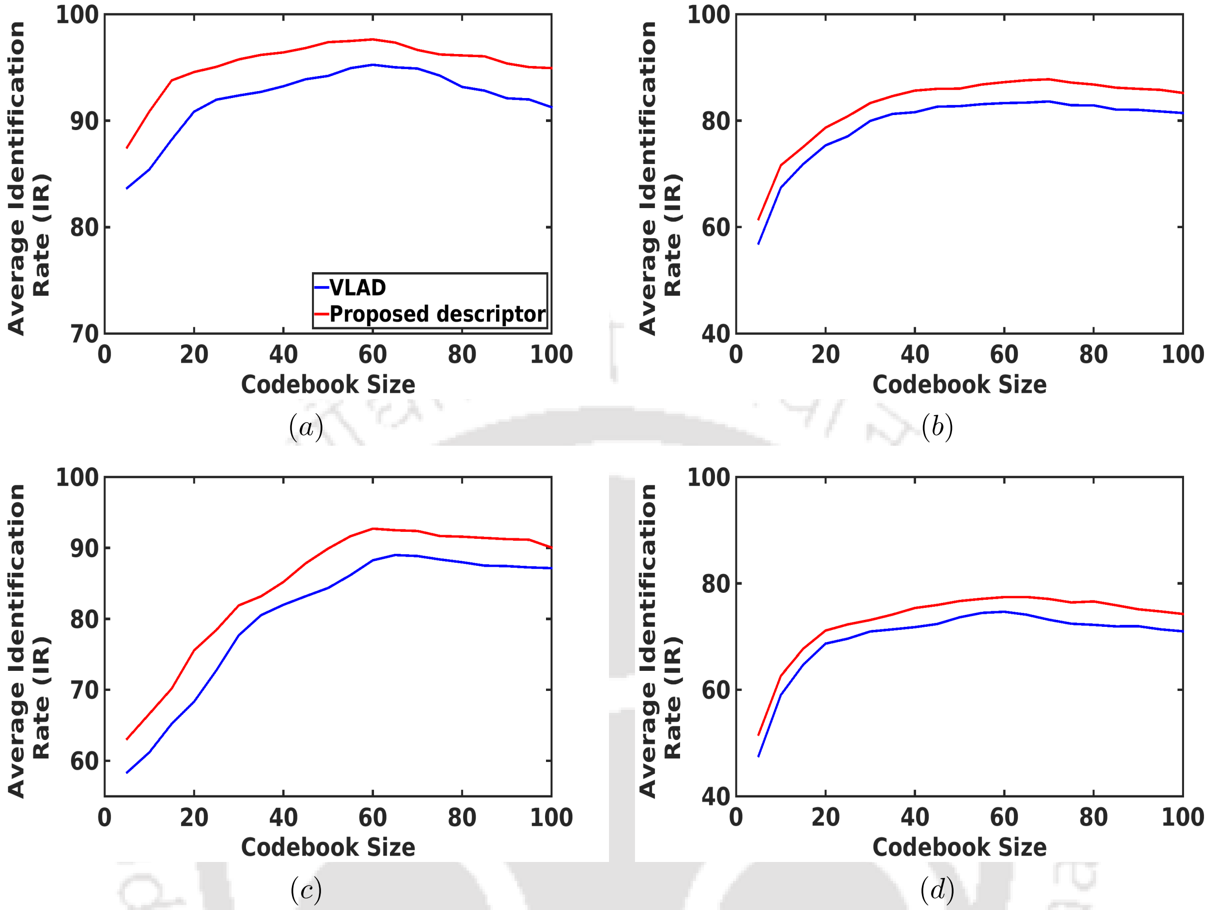

Figure 1: Trend of the average writer identification rate obtained for VLAD (shown in blue) and the proposed codebook descriptor (in red) for varying size of codebook . The sub-figures (a), and (c) represent the performance at the paragraph level for the IAM and the IBM-UB1 databases respectively, while those of (b) and (d) denote at the text-line level. From these four plots, it may be inferred that the proposal outperforms over VLAD at both paragraph and text-line levels.

Table 1: Summary of the best average writer identification rates (in %) along with dictionary size corresponding to the different sparse representation based descriptors.

Descriptor Type	IAM database				IBM UB1 database			
	Paragraph level		Text line level		Paragraph level		Text line level	
	IR	M	IR	M	IR	M	IR	M
Max pooling	98.03	400	85.89	400	94.28	400	78.57	400
Average Pooling	98.41	400	87.69	400	95.16	400	79.47	400
Proposed	**99.45**	400	**90.28**	400	**97.21**	400	**83.49**	400

- In Chapter 4, we modify the writer descriptor obtained with the traditional average / max pooling of sparse coefficients by the incorporation of saliency information of the atoms. Accordingly in Table 2, we compare the performance of our proposals with the traditional approaches. The abbreviations of the systems being presented are as follows:

Table 2: Comparison of best average writer identification rates (in %) for the descriptor obtained via the traditional average / max pooling approach (without saliency) and our proposals *EN-SL* *SP-SL*, and *SAL-ADP* (with saliency incorporation) along with dictionary size M.

System	IAM database				IBM UB1 database			
	Paragraph level		Text line level		Paragraph level		Text line level	
	IR	M	IR	M	IR	M	IR	M
SPF	98.41	400	87.69	400	95.16	400	79.47	400
SP-SL	99.07	400	88.61	400	95.78	400	80.35	400
EN-SL	99.28	400	89.54	400	96.37	400	81.61	400
SAL-ADP	**99.54**	400	**91.26**	400	**97.24**	400	**83.54**	400

- *SPF*: This system corresponds to the traditional sparse representation framework based on average / max pooling (without the incorporation of saliency).
- *EN-SL*: This system refers to the modified sparse representation framework based on average / max pooling with the incorporation of saliency obtained from the entropy approach.
- *SP-SL*: This system corresponds to the modified sparse representation framework based on average / max pooling with the incorporation of saliency obtained from the sum-pooling approach.
- *SAL-ADP*: refers to the system where the modified writer descriptor is computed by incorporating writer-specific saliency values to the writer descriptor obtained via traditional pooling strategies.

We observe that the performance of the three modified systems *EN-SL* , *SP-SL* and *SAL-ADP* (resulting from the incorporation of saliency values) attains a best performance of 99.28% , 99.07% , and 99.54% for $M = 400$ at the paragraph level on the IAM database. In contrast, the traditional descriptors of the *SPF* system (*viz.* obtained without saliency) trails behind, achieving a best average identification rate of 98.41% again at $M = 400$. Likewise, at text-line level, we see an improvement in performance accuracy from 87.69% (*SPF* system) to 89.54% (*EN-SL* system) , 88.61% (*SP-SL* system) and 91.26% (*SAL-ADP* system). Not to mention, across the different dictionary sizes considered, at both paragraph and text-line levels, the modified writer descriptors with saliency provides a better identification rate than that obtained without their incorporation. It may be noted here that a similar trend of improvement is also observed with respect to the IBM-UB1 database.

5 Publications related to thesis

Refereed Journals:

(i) Vivek Venugopal and Suresh Sundaram, “Online writer identification with sparse coding based descriptors,” IEEE Trans. Information Forensics and Security, vol. 33, no. 10, pp. 2538 - 2552, Oct 2018.

(ii) Vivek Venugopal and Suresh Sundaram, “An improved online writer identification framework using codebook descriptors,” Pattern Recognition, vol. 78, pp. 318-330, June 2018.

(iii) Vivek Venugopal and Suresh Sundaram, “A modified sparse representation classification framework for online writer identification,” IEEE Trans. Systems, Man, and Cybernetics: Systems, In press.

(iv) Vivek Venugopal, Suresh Sundaram, “An online writer identification system using regression-based feature normalization and codebook descriptors,” Expert System with Application, vol. 72, pp. 196-206, April 2017.

Under Review:

(i) Vivek Venugopal and Suresh Sundaram, “ An adaptive sparse representation framework for online writer identification”, submitted to Pattern Recognition Letters.

List of Conferences:

(i) Isht Dwivedi, Swapnil Gupta, Vivek Venugopal and Suresh Sundaram, “Online writer identification using sparse coding and histogram based descriptors ”, International Conference on Frontiers in Handwriting Recognition 2016, pp. 572-577.

6 Organization of the Thesis

(i) **Introduction**

1.1 Introduction

1.2 Overview of writer identification systems

1.3 Previous works on online writer identification

1.4 Contributions of the thesis

1.5 Conclusion

(ii) **Exploration of Codebook based descriptors**

2.1 Introduction

2.2 Preprocessing

2.3 Feature extraction

2.4 Codebook description

2.5 Writer identification

2.6 Experimental set-up

2.7 Performance evaluation

2.8 Conclusion

(iii) **Exploration of sparse coding based descriptors**

3.1 Introduction

3.2 Proposed methodology

3.3 Sub-stroke generation

3.4 Feature extraction

3.5 Sparse coding: an overview

3.6 Sparse coding based writer description

3.7 Results and discussion

3.8 Conclusion

(iv) **Exploration of saliency information in the sparse framework**

4.1 Introduction

4.2 Schematic of the proposed framework with saliency value incorporation

4.3 Entropy based saliency computation

4.4 Sum-pooling based saliency computation

4.5 Modified writer descriptor

4.6 Writer specific saliency value adaptation

4.7 Experiments and discussion

4.8 Conclusion

(v) **Summary**

5.1 List of contributions

5.2 Discussion of prior works

5.3 Possible pointers for the future

References

[1] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4–20, January 2004.

[2] A. Schlapbach, M. Liwicki, and H. Bunke, "A writer identification system for on-line whiteboard data," *Pattern Recogn.*, vol. 41, no. 7, pp. 2381–2397, 2008.

[3] G. X. Tan, C. Viard-Gaudin, and A. C. Kot, "Automatic writer identification framework for online handwritten documents using character prototypes," *Pattern Recognition*, vol. 42, no. 12, pp. 3313–3323, 2009.

[4] G.X. Tan , C. Viard-Gaudin, and A.C Kot, "Individuality of alphabet knowledge in online writer identification," *International Journal on Document Analysis and Recognition (IJDAR)*, vol. 13, no. 2, pp. 147–157, 2010.

[5] A. Chaabouni, H. Boubaker, M. Kherallah, A. M. Alimi, and H. El Abed, "Multi-fractal modeling for on-line text-independent writer identification," in *2011 International Conference on Document Analysis and Recognition*, 2011, pp. 623–627.

[6] C. Ramaiah, A. Shivram, and V. Govindaraju, "A hierarchical bayesian approach to online writer identification," *IET Biometrics*, vol. 2, no. 4, pp. 191–198, 2013.

[7] C. Ramaiah, A. Shivram, and V. Govindaraju, "A Bayesian framework for modeling accents in handwriting," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on.* IEEE, 2013, pp. 917–921.

[8] A. M. Namboodiri and S. Gupta, "Text independent writer identification from online handwriting," in *Proc. International Workshop on Frontiers in Handwriting Recognition*, vol. 10, October 2006, pp. 287–292.

[9] B. Li and T. Tan, "Online text-independent writer identification based on temporal sequence and shape codes." in *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on.* IEEE, 2009, pp. 931–935.

[10] K. Yu, Y. Wang, and T. Tan, "Writer identification using dynamic features." in *First International Conference on Biometric Authentication*, 2004, pp. 512–518.

[11] H. Jegou, F. Perronnin, M. Douze, J. Sanchez, P. Perez, and C. Schmid, "Aggregating local image descriptors into compact codes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1704–1716, 2012.

[12] M. Liwicki and H. Bunke, "IAM-OnDB-an on-line English sentence database acquired from handwritten text on a whiteboard," in *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*, 2005, pp. 956–961.

[13] A. Shivram, C. Ramaiah, S. Setlur, and V. Govindaraju, "IBM_UB_1: A dual mode unconstrained English handwriting dataset," in *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, 2013, pp. 13–17.