ISSN: 2322-5211

Journal of AI and Data Mining

Volume No. 12
Issue No. 1
January - April 2024

ENRICHED PUBLICATIONS PVT. LTD

JE-18, Gupta Colony, Khirki, Extn, Malviya
Nagar, New Delhi-110017
PHONE : - + 91-8877340707
E-Mail : info@enrichedpublications.com

ISSN: 2322-5211

Journal of AI and Data Mining

Aims and Scope

The Journal of Artificial Intelligence & Data Mining (JAIDM) is a scientific journal that aims to develop the international exchange of scientific and technical information in all areas of Artificial Intelligence and Data Mining.

All manuscripts with significant research results in the scope of the journal are welcome if they are not published or not being considered for publication elsewhere.

Journal of Artificial Intelligence & Data Mining appears quarterly considering the increasing importance of rapid, effective, international communication. JAIDM offers:

- Publication within a short period after acceptance
- On-line publication in advance of the printed journal
- One journal copy will be sent to the corresponding author
- Printing, Processing and Delivery without any charges

Topics of interest include, but are not limited to, the following:
Artificial Intelligence Algorithms, Tools & Applications
Data Mining and Machine Learning Tools
Semantic Web Techniques and Technologies
Soft computing theory and applications
Web Intelligence Applications & Search
Bioinformatics
Natural Language Processing
Computer Vision and Image Processing
Speech Understanding
Fuzzy Logic
Information Retrieval
Intelligent System Architectures
Knowledge-based/ Expert Systems
Automatic Control
Neural Networks
Parallel Processing
Pattern Recognition
Software & Hardware Architectures

ISSN: 2322-5211

Journal of AI and Data Mining

ISSN: 2322-5211

Journal of AI and Data Mining

(Volume No. 12, Issue No. 1, January - April 2024)

Contents

Data mining for decision making in engineering optimal design

A. Mosavi
University of Debrecen, Faculty of Informatics, Hungary.

ABSTRACT

Often in modeling the engineering optimization design problems, the value of objective function(s) is not clearly defined in terms of design variables. Instead it is obtained by some numerical analysis such as finite element structural analysis, fluid mechanics analysis, and thermodynamic analyses. Yet, the numerical analyses are considerably time consuming to obtain the final value of objective function(s). For the reason of reducing the number of analyses as few as possible, our methodology works as a supporting tool to the metamodels. The research in meta-modeling for multi-objective optimization are relatively young and there is still much research capacity to further explore. Here is shown that visualizing the problem on the basis of the randomly sampled geometrical big-data of computer aided design (CAD) and computer aided engineering (CAE) simulation results, combined with utilizing classification tool of data mining could be effective as a supporting system to the available meta-modeling approaches. To evaluate the effectiveness of the proposed method, a case study in 3D wing optimal design is proposed. Discussion focusing on how effective the proposed methodology could be in further practical engineering design problems is presented.

Keywords: Data Mining, classification, Multi-objective Optimization, Engineering Optimization, Meta Modeling.

1. Introduction

The research field of considering decision problems with multiple conflicting objectives is known as multiple criteria decision making (MCDM) [1]. Solving a multi-objective optimization problem has been characterized as supporting the decision maker (DM) in finding the best solution for the DM's problem. DM and optimization typically create an interactive procedure for finding the most preferred solutions. Yet, despite the increasing level of complexity, it has been often tried to pay attention to improving all the defined objective functions instead of reducing or ignoring some of them. Although due to the increased complexity, this would apply complications where objective functions are visualized by trade-off analysis methods as well studied in [9, 10, 25, 26, 35, 37]. According to [1], the general form of the multiobjective optimization problems can be stated as; Minimize $\mathbf{f}(x) = \{f_1(x), \dots, f_m(x)\}$, Subjected to $x \in \Omega$, where $\mathbf{x} \in \mathbb{R}^n$ is a vector of n decision variables; $\mathbf{X} \subset \mathbb{R}^n$ is the feasible region and is specified as a set of constraints on the decision variables; $\mathbf{f} : \Omega \rightarrow \mathbb{R}^m$ is made of objective functions subjected to be minimization. Objective vectors are images of decision vectors written as $\mathbf{z} = \mathbf{f}(x) = \{f_1(x), \dots, f_m(x)\}$. Yet an objective vector is considered optimal if none of its components can be improved without worsening at least one of the others. An objective vector z is said to dominate z denoted as $\mathbf{z} \prec \mathbf{z}'$, if $z_k \leq z'_k$ for all k and there exists at least one h that $z_h \leq z'_h$. point z is Pareto optimal if there is no other $x \in \Omega$ such that f(x) dominates The set of Pareto optimal points is called Pareto set (PS). And the corresponding set of Pareto optimal objective vectors is called Pareto front (PF).

Solving a multi-objective optimization problem would be done by providing the DM with the optimal solution according to some certain utility criteria allowing to choose among competing PF. Such utility criteria are often inconsistent, difficult to formalize and subjected to revision. The complete process of

MCDM has two parts (1) multi-objective optimization process which tries to find the PF solutions (2) decision making process which tries to make the best decision out of the possible choices. In dealing with increased complexity, this paper focuses on the first part which mostly deals with variables, constraints and objective functions.

1.1. Computational intelligence and multi objective optimization

Developing the methods for multi-objective optimization using computational intelligence along with real applications appeared to be quite young. However it has been observed that techniques of computational intelligence are indeed effective [3, 7, 15, 27]. On the other hand, the techniques of multi-objective optimization by themselves can also be applied to develop and to improve the effective methods in computational intelligence [2]. Currently there are many computational intelligence-based algorithms available to generate PF [1, 16, 17, 30]. However, it is still difficult to generate and visualize the PF in the cases with more than three objectives. In this situation, methods of sequential approximate optimization of computational intelligence with meta-modeling are recognized to be very effective in a series of practical problems [1, 4].

1.2 Meta-modeling and multi-objective optimization in shape optimization

Meta-modeling is a method for building simple and computationally inexpensive models, which replicate the complex relationships. However the research in meta-modeling for multi-objective optimization is relatively young and there is still much to do. So far there existed only a few standards for comparisons of methods, and little is yet known about the relative performance and effectiveness of different approaches [4, 15]. The most famous methods of Meta-modeling are known as response surface methods (RSM) and design of experiments (DOE). Although it is concluded in previous studies [16, 18, 19, 20], in the future research, scalability of MCDM models in terms of variables' dimension and objective space's dimension will become more demanding.

This is because the models have to be capable of dealing with higher computation cost, noise and uncertainties.

According to [18], the application of meta modeling optimization methods in industrial optimization problems is discussed. Some of the major difficulties in real-life engineering design problems counted: (1) there are numerous objective functions to be involved, (2) the function form of criteria is a black box, which cannot be explicitly given in terms of design variables, and (3) there are a huge number of unranked and non-organized input variables to be considered. Additionally in engineering design problems, often the value of objective functions is not clearly defined in terms of design variables. Instead it is obtained by some numerical analyses, such as FE structural analysis [34, 37], fluid mechanics analysis [7, 16, 17, 32], thermodynamic analysis [30], chemical reactions [3]. These analyses for obtaining a single value for an objective function are often time consuming.

Considering the high computation costs, the number of CAE evaluations/calculations are subjected to minimization with the aid of meta models [18]. In order to make the number of analyses as few as possible, sequential approximate optimization is one of the possible methods, utilizing machine learning techniques for identifying the form of objective functions and optimizing the predicted objective function.

Machine learning techniques have been applied for approximating the black-box of CAE function in many practical projects [1, 9, 10 ,25 ,37]. Although the major problems in these realms would be (1) how to approach an ideal approximation of the objective function based on as few sample data as possible (2) how to choose additional data effectively. The objective functions are modeled by fitting a function through the evaluated points. This model is then used to help the prediction value of future search points.

Therefore, those high performance regions of design space can be identified more rapidly. Moreover the aspects of dimensionality, noise and expensiveness of evaluations are related to method selection [32]. However, according to Bruyneel et al. [18] for the multi-objective capable version of meta-modeling algorithms further aspects such as the improvement in a Pareto approximation set and modeling the objective function must be considered.

Today, numerical methods make it possible to obtain models or simulations of quite complex and large scale systems [7, 8, 20, 22]. But there are still difficulties when the system is being modeled numerically. In this situation, modeling the simplified models is an effective method, generating a simple model that captures only the relevant input and output variables instead of modeling the whole design space [3, 20, 22].

The increasing desire to apply optimization methods in expensive CAE domains is driving forward the research in meta-modeling. The RSM is probably the most widely applied to metamodeling. The process of a meta-model from big data is related to classical regression methods and also to machine learning [4, 37]. When the model is updated using new samples, classical DOE principles are not effective. In meta-modeling, the training data sets are often highly correlated, which can affect the estimation of goodness of fit and generalization performance. Yet Meta-modeling brings together a number of different fields to tackle the problem on optimizing the expensive functions. On the other hand the classical DOE methods with employing evolutionary algorithms have delivered more advantages in this realm. Figure 1 describes the common arrangement of meta-modeling tools in multi-objective optimization processes of engineering design. It is worth mentioning that the other well-known CADOptimization integrations for shape optimization e.g. [24, 29, 31] would also follow the described scheme.

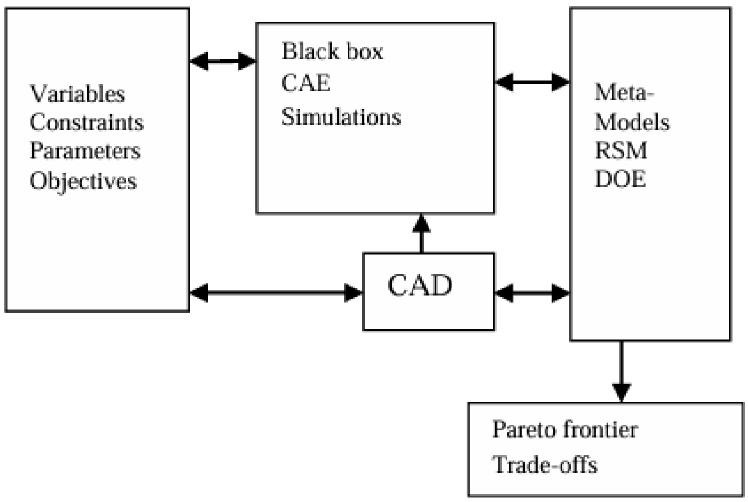


Figure 1. Meta-modeling tools in multi-objective optimization process.

2. Data mining classification in engineering design applications

The particular advantage of evolutionary algorithms (EAs) [11] in the multi-objective optimization (EMO) applications [19] is that they work with a population of solutions. Therefore, they can search for several Pareto optimal solutions providing the DM with a set of alternatives to choose from [14]. EMO-based techniques have an application where mathematical-based methods have difficulties with. EMO are also helpful in knowledge discovery related tasks in particular for mining the data samples achieved from CAE and CAD systems [29, 31]. Useful mined information from the obtained EMO trade-off solutions have been considered in many real-life engineering design problems.

2.1. Classifications

Finding useful information in large volumes of data drives the development of data mining procedure forward. Data mining classification process refers to the induction of rules that discriminate between organized data in several classes so as to gain predictive power [5]. There are some example applications of data mining classification in evolutionary multi-objective optimization available in the literature of [1, 6, 12, 19] where the goal of the classification algorithms is to discover rules by accessing the training sets. Then the discovered rules are evaluated using the test sets, which could not be seen during the training tasks [5].

In the classification procedures, the main goal is to use observed data to build a model, which is able to predict the categorical or nominal class of a dependent variable given the value of the independent variables [5]. Obayashi [12] for the reason of mining the engineering multi-objective optimization and visualization data applied selforganizing maps (SOM) along with a data clustering method. Moreover Witkowski et al. [13] and Mosavi [7, 20, 22] used classification tools of data mining for decision making supporting process to multi-objective optimization.

2.2. Modeling the problem

According to [1], before any optimization takes place, the problem must first be accurately modeled. In this case, identifying all the dimensions of the problem, such as formulation of the optimization problem with specifying decision variables, objectives, constraints, and variable bounds is an important task. Here the methodology proposes that mining the available sample data before actual modeling will indeed help to better model the problem as it delivers more information about the importance of input variables and could in fact rank the input variables. The proposed method of classification, also earlier utilized in [7, 20, 22], presented in Figure 2, is set to mine the input variables which are in fact associated with the final CAE data.

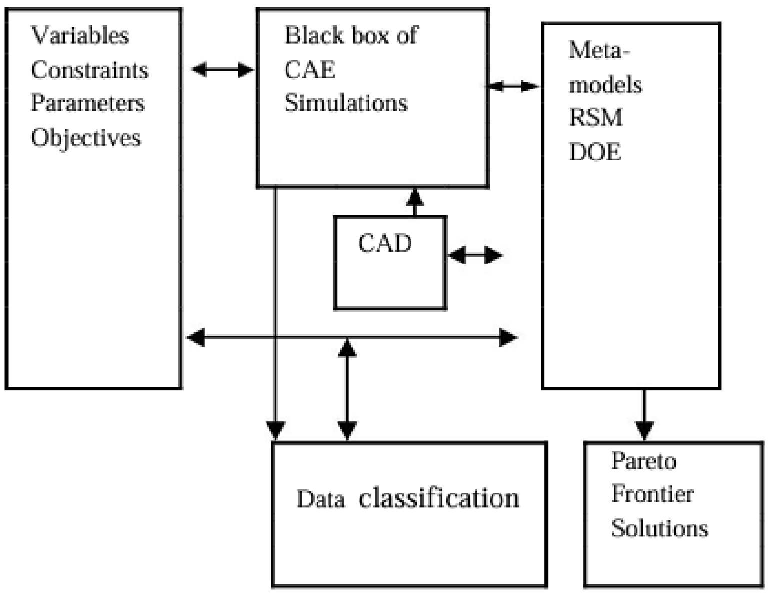


Figure 2. Supporting the meta-modeling process by mining the data

3. Study case; three-objective and 42-variale optimization problem

The applications in engineering optimal design have numerous disciplines to bring into the consideration. In mechanical engineering, the structural simulation is tightly integrated more than one discipline [18, 21, 22, 23, 24, 36]. Meanwhile, the trend nowadays is to utilize independent computational codes for each discipline [32]. In this situation, the aim of MCDM tools is to develop methods in order to guarantee that all physical variables are involved in the model. Bo et al. [28] in aerodynamic optimization of a 3D wing has tried to utilize the multi-objective optimization techniques in a multidisciplinary environment. In the similar cases [20, 24, 29, 32] in order to approach the optimal shape in an aerospace engineering optimization problem, the multiobjective optimization techniques are necessary to deal with all important objectives and variables efficiently. Here the optimization challenge is to identify as many optimal designs as possible to provide a choice of better decision. However with an increased number of design variables the modeling task, in a multidisciplinary environment, is getting even ever complicated. Therefore the multi-objective optimization tasks become more difficult with the increasing number of variables [20, 35]. Although the recent advances in parametric CAD/CAE integrations [24, 29, 31] have reduced the complexity of the approach in some levels.

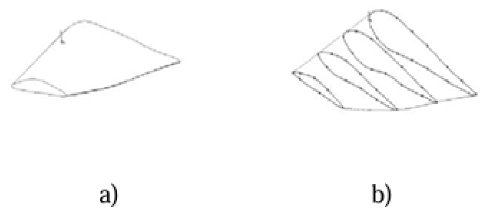
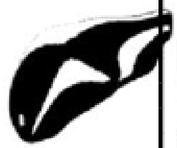
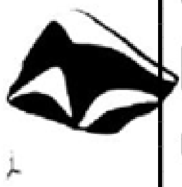
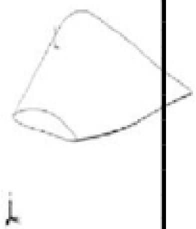
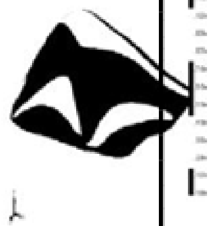
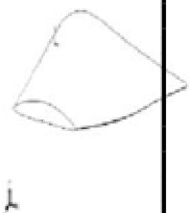
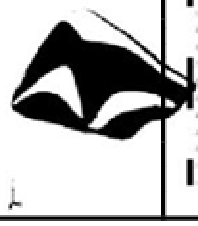
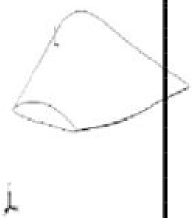
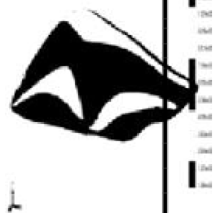
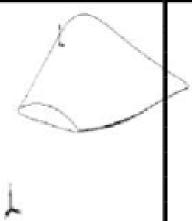
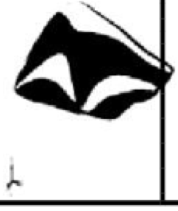
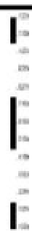


Figure 3. Airfoil geometry, modeled by S-plines [12, 14, 33, 34]

The airfoil of Figure 3 part (a) is subjected for shape improvement. The shape needs to be optimized in order to deliver minimum displacement distribution in terms of applied pressure on the surface. Figure 3, part (b) shows the basic curves of the surface modeled by S-plines. Here the proposed S-pline geometrical modeling methodology of Albers et al. [36] is successfully adapted and utilized. In the study case for modeling the 3D wing surface, four curve profiles have been with 42 points utilized. The coordinates of all points are supplied by a digitizer in which each point includes three dimensions of X, Y, and Z. Consequently the case, by adding the variable constraints, would include 126 columns plus three objectives which are going to highly increase the complexity. In fact, an optimal configuration of 42 variables supposed to satisfy the following three described objectives. The objectives are listed as follow: Objective 1: Minimizing the displacement distribution in the airfoil for constant pressure value of α. Objective 2: Minimizing the displacement distribution in the airfoil for constant pressure value of 2α. Objective 3: Minimizing the displacement distribution in the airfoil in constant pressure value of 4α.

In the described multi-objective optimization problem the number of variables is subjected to minimization before the multi-objective optimization modeling process takes place in order to evolve a large scale design space to the smaller and much more handy design space. Here the proposed and utilized model reduction methodology differs from the previous study Filomeno et al. [35] in terms of applicability and ease of use in general multi-objective optimization design applications.

Table 1. Training dataset including five CAE calculations' results

	Variables Configuration : V1-V42	CAD Model	Displacement Distribution	Objective Results
1	0,1,1.2,1,0.8,0.4,0.2,0,-0.4,-0.48, 0.6,-0.8,-0.72, 0,0.84,0.99,0.84,0.62,0.26,0,-0.20,-0.40,-0.36,-0.70,-0.58, 0,0.59,0.78,0.56,0.30,0,-0.21,-0.24,-0.38,-0.38 0,0.26,0.50,0.39,-0.03,-0.10,-0.12,			 O1=c O2=c O3=c
2	0,1.1,1.21,.9,0.82,0.42,0.18,.1,-0.41,-0.46,-0.62,-0.81,-0.70, 0,0.86,0.1,0.82,0.60,0.25,0.01,-0.20,-0.39,-0.39,-0.70,-0.58, 0,0.58,0.76,0.57,0.32,0,-0.21,-0.23,-0.37,-0.39 0,0.26,0.54,0.40,-0.03,-0.1,-0.1,			 O1=b O2=c O3=d
3	0,1,1.2,1,0.8,0.4,0.2,0,-0.4,-0.48,-0.6,-0.8,-0.72, 0,.88,0.99,0.84,0.62,0.26,0,-0.23,-0.35,-0.37,-0.70,-0.54, 0,0.58,0.76,0.58,0.31,0,-0.23,-0.23,-0.37,-0.37 0,0.24,0.50,0.40,-0.03,-0.13,-0.10,			 O1=b O2=c O3=b
4	0,1.3,1.23,1.06,0.83,0.41,0.28,0.07,-0.41,-0.48,-0.6,-0.8,-0.78,0,0.84,.92,0.84,0.62,0.26,0,-0.23,-0.39,-0.37,-0.70,-0.54,0,0.58,0.76,0.58,0.31,0,-0.24,-0.22,-0.36,-0.38, 0,0.24,0.52,0.38,-0.02,-0.12,-0.12,			 O1=d O2=c O3=b
4	0,1.3,1.23,1.06,0.83,0.41,0.28,0.07,-0.41,-0.48,-0.6,-0.8,-0.78,0,0.84,.92,0.84,0.62,0.26,0,-0.23,-0.39,-0.37,-0.70,-0.54,0,0.58,0.76,0.58,0.31,0,-0.24,-0.22,-0.36,-0.38, 0,0.24,0.52,0.38,-0.02,-0.12,-0.12,			 O1=d O2=c O3=b
5	0,1.01,1.21,1,0.8,0.4,0.21,0-0.41,-0.47,-0.59,-0.79,-0.69, 0,0.80,1.01,0.86,0.64,0.26,-0.01,-0.20,-0.40,-0.40,-0.72,-0.56, 0,0.58,0.76,0.58,0.31,0,-0.23,-0.23,-0.37,-0.37 0,0.24,0.52,0.38,-0.06,-0.10,-0.10,			 O1=c O2=d O3=e

The dataset of big data for data mining is supplied from the Table I. The table has gathered a collection of initial dataset including shapes' geometries and simulation results from five CAE calculations, based on random initial values of variables, which in the proposed method will be mined. In the next section, the discussion of how the dataset of five random CAE calculations are being utilized for creating the smaller design space for a multi-objective optimization model is made.

4. Methodology and experimental results

The effectiveness of data mining tools in multiobjective optimization problems presented by Coello et al. [2] and earlier in [5] the classification rules for evolutionary multi-objective algorithms were well implemented, in which along with the research work of Witkowski et al. [13] forms the proposed methodology working via a novel workflow. The workflow of data mining procedure methodology is described in Figure 4. In this method, the classification task is utilized to create several classifiers or decision trees. In the next steps, the most important variables, which have more effects on the objectives, are detected.

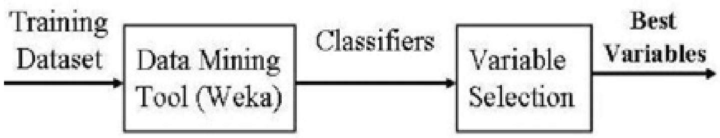


Figure 4. Proposed methodology workflow

Regressions and model trees are constructed by a decision tree building an initial tree. However, most decision tree algorithms choose the splitting attribute to maximize the information gain. It is appropriate for numeric prediction to minimize the intra subset variation in the class values under each branch.

The splitting criterion is used to determine which variable is better to split the portion T of the training data. Based on the treating the standard deviation of the objective values in T as a measure of the error and calculation the expected reduction in error as a result of testing each variable is calculated. Meanwhile the variables, which maximize the expected error reduction, are chosen for splitting. The splitting process terminates when the objective values of the instances vary very slightly, that is, when their standard deviation has only a small fraction of the standard deviation of the original instance set. Splitting also terminates when just a few instances remain. Experiments show that the obtained results are not very sensitive to the exact choice of these thresholds. Data mining classifier package of Weka provides implementations of learning algorithms and dataset which could be preprocessed and fed into a learning scheme, and analyze the resulting classifier and its performance. The workbench includes methods for all the standard data mining problems such as regression, classification, clustering, association rule mining, and attribute selection. Weka also includes many data visualization facilities and data preprocessing tools. Here three different data mining classification algorithms i.e. J48, BFTree, LADTree are applied and their performance is compared to choose attribute importance. The mean absolute error (MAE) and root mean squared error (RMSE) of the class probability is estimated and assigned by the algorithm output. The RMSE is the square root of the average quadratic loss and the MAE is calculated in a similar way using the absolute instead of the squared difference. The comparison between importance ranking results is obtained by our experiments listed in Table II. It is concluded that in the worst case, more than 55% variable reduction is achieved. As one can see, BFTree and J48 algorithms have classified the datasets with less number of variables. While in LADTree algorithms, at least seven variables have utilized to classify dataset. The variables number 15 and 24 play much more important role in effecting the first objective (O1). Variables number 41 and 35 also have the more effects on third objective (O3) as well. According to the experimental results, it is possible to optimize the model by reducing the 45% number of variables. In Table II, two types of classification error (MAE, RMSE) are shown for all algorithms corresponding to different class of objectives.

Table 2. Variables importance ranking for three classification methods

Classification Method	MAE	RMSE	Effective Variables	Objectives
BFTree	0.370 0.412 0.418	0.517 0.519 0.555	15 23 41	O_1 O_2 O3
J48	0.309 0.482 0.378	0.514 0.642 0.590	15,24 13 35,41	O_1 O2 O_3
LAD Tree	0.277 0.604 0.365	0.500 0.769 0.584	15,24,2,32,41,39,3 23,22,18,15,42,2,17, 20 41,35,9,17,11,38,37, 16	O1 O_2 O_3

5. Conclusions

In order to extract more information from the optimization variables in a reasonable way, the classification task of data mining has been applied. Variables were ranked and organized utilizing three different classification algorithms. The results show the reduced number of variables speeds up and scales up the process of optimization within a preprocessing step. The utilized data mining tool has found to be effective in this regard. Additionally, it is evidenced that the growing complexity can be handled by a preprocessing step utilizing data mining classification tools. The modified methodology is demonstrated successfully in the framework and the author believes that the process is simple and fast. Future research should focus on the effectiveness of the proposed data reduction process. Also, trying other data mining tasks such as clustering, association rules, and comparison could be beneficial. Although in real-life applications where the optimal design problem has to be considered by inclusion of multiple criteria, a combination of the proposed method with the other developed MCDM tools [38-46] would be effective.

References

[1] J. Branke, K. Deb, K. Miettinen, R. Słowinski, Multiobjective, Optimization Heidelberg New York, 2008) (Springer, Berlin,

[2] C. Coello, S. Dehuri, S. Ghosh, Swarm Intelligence for Multi-objective Problems in Data Mining (Springer, Berlin, Heidelberg New York, 2009)

[3] A.Mosavi, "Applications of Interactive Methods of MOO in Chemical Engineering Problems," Global Journal of Researches in Engineering, Vol. 10, No. 3, p.8. (2010).

[4] J. Knowles, H. Nakayama, Meta-Modeling in Multiobjective Optimization Heidelberg New York, 2008) (Springer, Berlin, Heidelberg New York, 2008)

[5] K. K. Kshetrapalapuram, M. Kirley, Mining Classificationm Rules Using Evolutionary Multiobjective Algorithms, Knowledge-Based Intelligent, Information and Engineering Systems. 3683 (Springer Berlin, Heidelberg, 2005).
[6] Alex A. Freitas, On Objective Measures of Rule Surprisingness, Principles of Data Mining and Knowledge Discovery. 1510, (2008).
[7] A. Mosavi, "The Large Scale System of Multiple Criteria Decision Making; Pre-Processing," Large Scale Complex Systems Theory and Applications, Vol. 9, 2010, pp. 354- 359.
[8] A. Adejuwon and A. Mosavi, "Domain Driven Data Mining; Application to Business," International Journal of Computer Science Issues, Vol. 7, No 2, 2010, pp. 4144.
[9] Mosavi, M.Azodinia Abbas S. Milani, Kasun N. Hewage and M.Yeheyis, "Reconsidering the Multiple Criteria Decision Making Problems of Construction Workers with the Aid of Grapheur," International ANSYS and EnginSoft Conference, Verona, Italy, 2011.
[10] E. Foldi, A. Delavar, A. Mosavi, K. N. Hewage, A. S. Milani, A. A. Moussavi and M. Yeheyis, "Reconsidering the Multiple Criteria Decision Making Problems of Construction Projects; Using Advanced Visualization and Data Mining Tools," Conference of PhD Students in Computer Science, Szeged, Hungary, 28-30 June 2012.
[11] X. Llor, D.E. Goldberg, I. Traus, E.Bernad, Generalisation and Model Selection in Supervised Learning with Evolutionary Computation, Applications of Evolutionary Computing. 2611, (2003).
[12] S.Obayashi, evolutionary multiobjective optimization and Visualization, new developments in computational fluid dynamics (Springer, 2005)
[13] K. Witkowski, M. Tushar, Decision making in multiobjective optimization for industrial applicationData mining and visualization of Pareto, In Proceedings of 7th European LS-DYNA Conference, 416-423 (2009).
[14] K. Deb, Current Trends in Evolutionary MultiObjective Optimization, International Journal for Simulation and Multidisciplinary Design Optimization. 2, 1–8 (2007).
[15] A.Mosavi, "The Multiobjective Optimization Package of IOSO; Applications and Future Trends," In: CSCS, Conference of PhD Students in Computer Science, University of Szeged, Szeged, Hungary, 2010, p. 55.
[16] A.Mosavi, "Computer Design and Simulation of Built Environment; Application to Forest," Proceeding of ICECS'09, The Second IEEE International Conference on Environmental and Computer Science, Dubai, 28-30 December 2009, pp. 81-85.
[17] A.Mosavi, "Parametric Modeling of Trees and Using Integrated CAD/CFD and Optimization Tools: Application to Creating the Optimal Planting Patterns for New Forests," Proceedings of 2nd International Conference Wind Effects on Trees, Albert-LudwigsUniversity of Freiburg, Freiburg, 2009.
[18] M. Bruyneel, B. Colson, P. Jetteur, C.Raick, A. Remouchamps, S. Grihon, Recent progress in the optimal design of composite structures: industrial solution procedures on case studies, Int. J. Simul. Multidisci. Des. Optim. 2, 283-288 (2008).
[19] S. E. Bedingfield, K. A. Smith, Evolutionary Rule Generation Classification and its Application to Multiclass Data, Computational Science, Springer. 2660, (2003).
[20] Mosavi, Multiple Criteria Decision-Making Preprocessing Using Data Mining Tools, International Journal of Computer Science Issues. 7, 26-34 (2010).
[21] V. Arularasan, Modeling and simulation of a parallel plate heat sink using computational fluid dynamics, Int J Adv Manuf Technol. 5, 172-183(2008).
[22] M. Esmaeili and A. Mosavi, "Variable Reduction for Multi-Objective Optimization Using Data Mining Tech-niques; Application to Aerospace Structures," Proceeding of ICCET, the 2nd IEEE International Conference on Computer Engineering and Technology, Vol. 5, Chengdu, 16-18 April

2010, pp. 333- 337.
[23] I. Olcer, a hybrid approach for multi-objective combinatorial optimization problems in ship design and shipping, Computers & Operations Research. 35, 2760 – 277(2007).
[24] L. Toussaint, N. Lebaal, D. Schlegel, S. Gomes, Automatic Optimization of Air Conduct Design Using Experimental Data and Numerical Results, Int. J. Simul. Multidisci. Des. Optim. 4, 77-83 (2010).
[25] A.Mosavi, M. Hoffmann and A. S. Milani, "Optimal Design of the NURBS Curves and Surfaces Utilizing Multiobjective Optimization and Decision Making Algorithms of RSO," Conference of PhD Students in Mathematics, Szeged, Hungary, Jnue 2012.
[26] Mosavi, M. Hoffmann and A. S. Milani, "Adapting the Reactive Search Optimization and Visualization Algorithms for Multiobjective Optimization Problems; Application to Geometry," Conference of PhD Students in Computer Science, Szeged, Hungary, June 2012.
[27] Mosavi, A. S. Milani, M. Hoffmann and M. Komeili, "Multiple Criteria Decision Making Integrated with Mechanical Modeling of Draping for Material Selection of Textile Composites," ECCM15, 15th Eeuropean Conference on Composite Materials, Venice, Italy, 24-28 June, 2012.
[28] Y. Bo, X. ANY, Aerodynamic optimization of 3D wing based on iSIGHT, Appl. Math. Mech. -Engl. Ed. 5, 603–610 (2008).
[29] J. B. Bluntzer, S. Gomes, D.h. Bassir, A. Varret and J.c. Sagot, Direct multi-objective optimization of parametric geometrical models stored in PLM systems to improve functional product design, Int. J. Simul. Multidisci. Des. Optim. 2, 83-90(2008).
[30] A.Mosavi, "Application of Multiobjective Optimization Packages in Design of an Evaporator Coil," World Academy of Science, Engineering and Technology, Vol. 61, 2010, pp. 25-29.
[31] P. Vik, D. Luís, P. Guilherme, J. Oliveira, Automatic Generation of Computer Models through the Integration of Production Systems Design Software Tools, Int. J. Simul. Multidisci. Des. Optim. 4, 141-148 (2010).
[32] Mosavi, "Hydrodynamic Design and Optimization: Application to Design a General Case for Extra Equipments on the Submarine's Hull," Proceeding on IEEE International Conference on Computer Technology and Development, ICCTD'09, Vol. 2, Kota Kinabalu, 13-15 November 2009, pp. 139-143.
[33] Mosavi, "Multiobjective Optimization of Spline Curves Using modeFrontier," Proceedings of 13 Internaional Conference on Engineering Optimization and In-ternational Mode Frontier users' Meeting, Trieste, 2010.
[34] Mosavi, "On Engineering Optimization the Splined Profiles," Proceedings of International Conference on Engineering Optimization and International modeFrontier Users' Meeting, Trieste, 2010.
[35] R. Filomeno, C. Coelho, P. Breitkopf, C. KnopfLenoir, Model reduction for multidisciplinary optimization - application to a 2D wing, Struct Multidisc Optim, 7, 29–48 (2008).
[36] Albers, N. Leon-Rovira, Development of an engine crankshaft in a framework of computer-aided innovation, Computers in Industry, 60, 604– 612 (2009).
[37] Mosavi, Optimization; A.Vaezipour, "Reactive Search Application to Multiobjective Optimization Problems." Applied Mathematics, 3, no.30 (2012): 1572-1582.
[38] Mosavi, "A MCDM Software Tool for the Automated Design Environments," 26th Europian Conference on Operational Research, Rome 2013, EURO - INFORMS XXVI.
[39] A.Vaezipour, A.Mosavi, U.Seigeroth, "Machine learning integrated optimization for decision making," 26th Europian Conference on Operational Research, Rome 2013, EURO - INFORMS XXVI.
[40] Mosavi, "Multiobjective Optimization package of IOSO," Mini EURO Conference EUROPT, Izmir, 2010.

[41] Vaezipour, A.Mosavi, "Managing Decision Making Within Enterprise," International CAE Conference, Verona, Italy, 22-23 Oct 2012.
[42] Vaezipour, A.Mosavi, "Enterprise Decision Management With the Aid of Advanced Business Intelligence and Interactive Visualization Tools," International CAE Conference, Verona, Italy, 22-23 Oct 2012
[43] Mosavi, "Data Mining for Business Applications," OGIK_ISBIS, Hungary, 2010.
[44] Mosavi, "Data Mining for Business Applications and Business Decision-Making: Challenges and Future Trends," OGIK_ISBIS, Hungary, 2010.
[45] Mosavi, Nage Peter, Miklos Hoffmann, "Automatic multi-objective surface design optimisation using modeFRONTIER's CAD/CAE integrated system: Application to military submarine sail," EnginSoft International Conference, CAE Technologies for Industry, Italy, 2009.
[46] Mosavi, "Recent developments and innovations in engineering optimization," Spring Wind Conference, Pecs, Hungary, 2010.

Prioritizing the ordering of URL queue in focused crawler

D. Koundal
University Institute of Engineering and Technology, Panjab University, Chandigarh, India

ABSTRACT

The enormous growth of the World Wide Web in recent years has made it necessary to perform resource discovery efficiently. For a crawler, it is not a simple task to download the domain specific web pages. This unfocused approach often shows undesired results. Therefore, several new ideas have been proposed, and crawling is a key technique, which is able to crawl particular topical portions of the World Wide Web quickly without having to explore all web pages. Focused crawling is a technique, which is able to crawl particular topics quickly and efficiently without exploring all WebPages. The proposed approach does not only use keywords for the crawl, but also rely on high-level background knowledge with concepts and relations, which are compared with the texts of the searched page. In this paper, a combined crawling strategy is proposed that integrates the link analysis algorithm with association metric. An approach is followed to find out the relevant pages before the process of crawling and to prioritize the URL queue from downloading higher relevant pages to an optimal level based on domain dependent ontology. This strategy makes use of ontology to estimate the semantic contents of the URL without exploring which in turn strengthen the ordering metric for URL queue and leads to the retrieval of most relevant pages.

Keywords: WebCrawler, Importance-metrics, Association - metric, Ontology.

1. Introduction

A crawler is a constituent of search engine that retrieves Web pages by strolling around the Internet following one link to another. A focused crawling algorithm weights a page and extracts the URLs. By rating the URLs, the crawler decides which page to retrieve next. A focused crawler fetches the page that locates on the head of its queue, examines the page and assigns a score to each URL. According to the scores inserted into the queue, the queue will organize itself in order to place URLs with higher scores in the queue head so that they first will be processed. Again, the crawler will fetch the URL on the head of the queue for new processing [1].

Intuitively, the term in-links refers to the hyperlinks pointing to a page. Usually, the larger the number of in-links, the higher a page will be rated. The assumption is made that if two pages are linked to each other, they are likely to be on the same topic. Anchor text can provide a good source of information about a target page, because it signifies how people linking to the page actually describe it. Several studies have tried to use either the anchor text or the text close to it to predict a target page's content. Researchers have developed several link-analysis algorithms over the past few years [2-11]. The most popular link-based Web analysis algorithm includes Page Rank. A major problem of a focused crawler is to effectively order the links at the crawl frontier so that a maximum number of relevant pages are loaded, while loading only a minimum number of irrelevant pages. This is a challenging task because most of the existing focused crawlers use local search algorithms in Web searching. This may miss a relevant page if there does not exist a chain of hyperlinks that connects one of the seed pages to that relevant page. The whole paper divides into the following sections: The section 2 discusses the related work done so far on this challenge. Section 3 gives various prioritizing algorithms. Section 4 tells about association metric based on ontology. Section 5 deals with proposed work on this challenge. The results of experimental evaluation presented in section 6. The implementation details are given in section 7. The section 8 covers conclusion.

2.Related work

Most of the focused crawling techniques use link structures of the web to improve ordering of URLs in priority queue. A recurring problem in a focused crawling is finding relevant page that is surrounded by non-relevant pages. One remedy presented in [12] by Aggarwal et al. uses the characteristics of the linkage structure of the web while performing the crawl by introducing a concept of “intelligent crawling” where the user can specify an arbitrary predicate (e.g. keywords, document similarity, anything that can be implemented as a function which determines documents relevance to the crawl based on URL and page content) and the system adapts itself in order to maximize the harvest rate. Ehrig et al. in [13] in another approach named as CATYRPEL consider an ontology-based algorithm for page relevance computation. After preprocessing, entities (words occurring in the ontology) are extracted from the page and counted. Relevance of the page with regard to user selected entities of interest is then computed by using several measures on ontology graph (e.g. direct match, taxonomic and more complex relationships). The evaluation of the importance of the page P as I (P) uses some metrics [14]. Cho et al. proposed an approach calculating the PageRank score on the graph induced by pages downloaded and then using this score as a priority of URLs extracted from a page. This may be due to the fact that the PageRank score is calculated on a very small, non-random subset of the web and also that the PageRank algorithm is too general for use in topic-driven tasks. L. page et al. in [15] proposed an approach for calculating the PageRank score on the graph induced by pages downloaded so far and then using this score as a priority of URLs extracted from a page. They show some improvement over the standard Breadth-first algorithm. Ontology based web crawler [16] estimates the semantic content of the link of the URL in a given set of documents based on the domain dependent ontology, which in turn reinforces the metric that is used for prioritizing the URL queue. The link representing concepts in the ontology knowledge path is given higher priority. However in this work, the content of the page based on the concepts is also used for determining the relevancy of the page. An approach presented by [17] is used to prioritize the ordering of URLs through using association metric along with other importance metric. The rank or relevancy score of the URL is calculated based on the division score with respect to topic keywords available in a division i.e., finding out how many topic keywords there are in a division in which this particular URL exists and calculates the total relevancy of parent page of the relevancy score of the URL page [18]. The maximal set of relevant and quality page is to be retrieved [19]. In this proposed approach, a combination of importance metric and association metric are presented in order to obtain ordering metric for prioritizing the URLs in queue on the basis of syntactic as well as semantic nature of URL.

3. Importance Metric

For a given Webpage p, there are different types of importance metrics, which are as follow: Back link Count I(P) is the number of links to page p that seem over the entire Web. Intuitively, a page p that is linked by many pages is more important than one that is rarely referenced. This type of “citation count” has been used widely to evaluate the impact of published papers.

Page Rank

Page Rank is the connectivity-based page quality metric suggested by Page et al. [15]. It is a static measure to rank pages in the absence of any queries. That is, PageRank computes the “global worth” of each page. Intuitively, the Page Rank measure of a page is similar to its in-degree, which is a possible measure of the significance of a page. The PageRank of a page will be high, if many pages with a high PageRank have links to it, and a page having few outgoing links contributes more weight to the pages, it links to a page containing many outgoing links. Thus, a link from the Yahoo home page counts the same as a link from some individual's home page. However, since the Yahoo home page is more important (it has a much higher IB count) it would make sense to value that link more highly. The weighted back link count of page p is given by

$$IR(p) = (1-d) + d(IR(t1)/c1 + ... + IR(tn)/cn)$$

4.Association metric with Ontology

Ontology serves as metadata schemas, providing a controlled vocabulary of concepts, each with unambiguously defined and machine-process able domain theories, ontologies help people and machines to communicate succinctly - supporting semantics exchange, not merely syntax.

Ontology is a description (like a formal specification of a program) of the concepts and relationships that can be for an agent or a community of agents. The essential of an ontology is " ***is a***"hierarchy. The Reference Ontology thus created would have the following associations like "is a", "part of", "has" relationships. The association metric for the URL u is estimated based on its relevancy with the reference ontology using proper text classification algorithms. Once the page p of the URL u is downloaded, the association metric for this page p is also calculated and preserved, as it will be a parent page for many links to be crawled. AS(p) is the same as all links from that page p but it utilizes the Web's hyperlink structure to retrieve new pages by traversing links from previously retrieve ones. Here an ontology-based strategy is taken into account for page relevance computation. After preprocessing, entities (words occurring in the ontology) are extracted from the page and counted and weight of the page is then calculated. With this, a candidate list of Web pages in order of increasing a priority is maintained. In next section, the core elements of proposed work are discussed in detail.

5. Proposed Work

A. System Overview

The focused crawling method consists of two interconnected cycles. The first cycle is ontology cycle that defines the crawling target in the form of instantiated ontology. This cycle also presents the output of the crawling process to the user in the form of a document list and proposals for enhancement of the already existing ontology to the user. The second cycle comprises the Internet crawler. It intermingles automatically with the data contained on the Web and retrieves them then it connects to the ontology to determine relevance. The relevance computation is used to select relevant documents for the user and to focus on links for the further search for relevant documents and metadata available on the Web. Our proposed focused crawler is based on domain dependent ontology has following components: All_URLs queue is employed for storing the list of URLs to download. Metric Module persistently scans through All_URLs to make the refinement decision. It

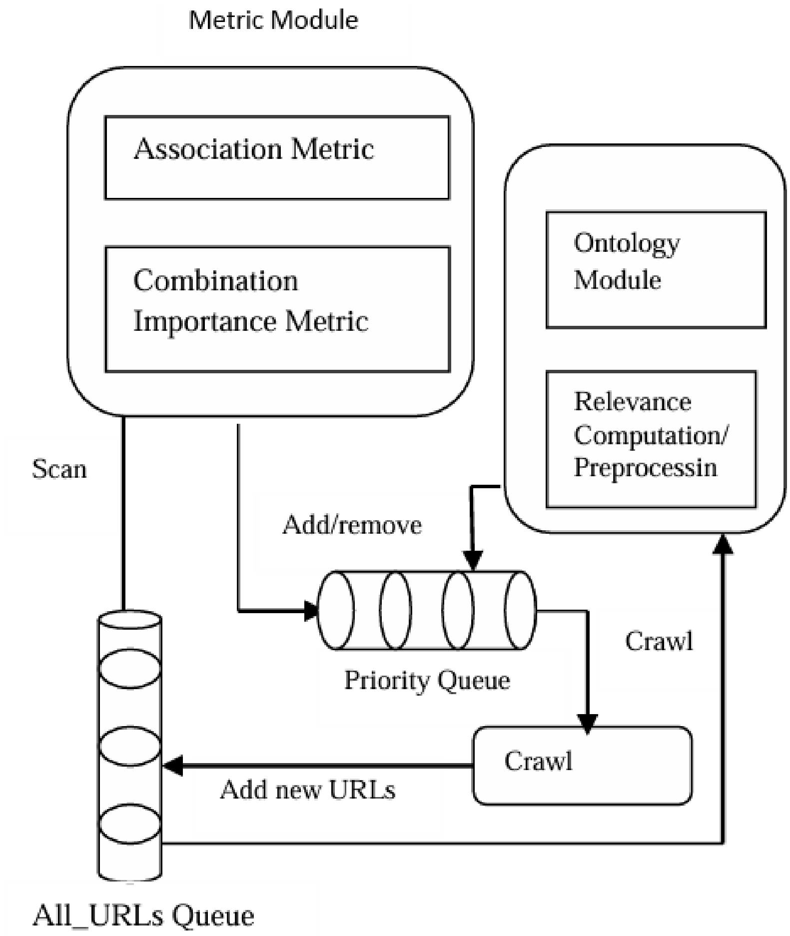


Figure 2. Prototype architecture of ontology based focused crawler

schedules for replacement of the less-important pages in priority queue with the more important page. Metric Module is a collection of Association metric and Combination Metric. Ontology module works as background knowledge for a crawler to search in the web. It has been widely accepted that ontology is the core ingredient for the Semantic Web. This will have to be extended for the relevance measure of focused crawler. For this purpose, it is a formal and declarative representation, which includes the vocabulary (or names) for referring to the terms in that subject area and the logical statements that describe what the terms are, how they are related to each other, and how they can or cannot be related to each other. Ontology therefore provides a vocabulary for representing and communicating knowledge about some topics and a set of relationships that hold among the terms in the vocabulary. After preprocessing like HTML tag removal, stemming, lexical entries of the ontology are matched against the URLs and a relevance score is computed. Relevance computation is a function, which tries to map the content (e.g. natural language text, hyperlinks) of a Web document against the accessible ontology to gain an overall relevancescore. Crawl Module is started with a given set of links. The links are retrieved according to their rank.

Priority queue is used for placing the URLs to be crawled in the front. The URLs in priority queue is chosen by metric module. The processed web resources are indexed and stored in a database and then stored resources are being semantically analyzed and rated in the context of a given ontology. The crawl frontier is implemented by a standard DBMS system.

All crawling modules share the data structures needed for the interaction with the crawler. The prototype maintains a list of unvisited URLs called the frontier. This is initialized with the seed URLs specified at the configuration file. Besides the frontier, the simulator contains a queue. The scheduling algorithm fills it with the first k URLs of the frontier, where k is the size of the queue mentioned above, once the scheduling algorithm has been applied to the frontier. Each crawling loop involves picking the next URL from the queue, fetching the page corresponding to the URL from the local database that simulates the Web and determining whether the page is relevant or not. If the page is not in the database, the simulation tool can fetch this page from the real Web and store it into the local repository. If the page is relevant, the outgoing links of this page are extracted and added to the frontier, as long as they are not already in it. The crawling process stops once a certain end condition is fulfilled, usually when a certain number of pages have been crawled or when the simulator is ready to crawl another page and the frontier is empty. If the queue is empty, the scheduling algorithm is applied and fills the queue with the first k URLs of the frontier, as long as the frontier contains k URLs. If the frontier doesn't contain k URLs, the queue is filled with all the URLs of the frontier.

B. Proposed Prioritizing Algorithm:

The proposed crawler will work according to the following segment of code. Input: seed URLs: start_urls Assumption: Initially form beginning assumes Priority queue is full. Output: Replacing "less important" pages with "more important pages" in a priority queue based on domain specific ontology. enqueue (url_queue, start_urls); While (not empty (url_queue) and not termination) { url = dequeue (url_queue); page = crawl_page (url); enqueue (crawled_pages, (url, page)); url_list = extract_urls (page); For each page p in crawled_pages

Association_weight_page = AS(p); // compute
association weight (metric) of page
End loop
For each u in url_list
enqueue (links, (url,));
If [not in url_queue] and [(,-) not in
crawled_pages]
enqueue (url_queue, u);
Association_Weight_URL = AS(u); //compute
association weight of URL
Combination_Importance = CI(u); //CI(u)=
pagerank[u]+ backlink[p]
End loop
Ordering_metric = O (u);
//

$$O[u] = b_1 CI(u) + b_2 AS(u) + b_3[AS(p_1) + AS(p_2) + \ldots\ldots + AS(p_n)] + b_4 TD[u]$$

where p1, p2 ...pn are the parent pages to this url u

reorder_queue (url_queue); //*based on O[u]*

C. Ordering Metric O (u)

The ordering metric O is used by the crawler for this selection, i.e., it selects the URL u such that O(u) has the highest value among all URLs in the queue. In our experiments, we explore the types of ordering metrics that are best suited for either IB (p) or IR (p). The Ordering Metric O(u) used for reordering the URL queue in our crawler is a composite metric defined as follows: CI (u) = Page Rank[u]

$$CI(u) = Page\ Rank[u]$$

$$O[u] = b_1 CI(u) + b_2 AS(u) + b_3[AS(p_1) + AS(p_2) + \ldots\ldots$$
$$+ AS(p_n)] + b_4 TD[u]$$

Where, pi is the ith Parent page of URL u to be crawled and b1, b2, b3, b4 are real constants to be evaluated from the results of our crawl. The proposed new ordering metric will solve the major problem of finding the relevancy of the pages before the process of crawling, as well as plays an important role in estimating the relevancy of the links in the page to an optimal level.

6. Implementation details

The implementation of our ontology embedded crawler is an application with in the KAON, the Karlsruhe Ontology and Semantic Web tool suite. The underlying data structure is provided by KAON-API. The crawler is designed with the TextToOnto tool i.e. KAON Workbench. The tight integration of the crawler with the ontology and metadata management component is also important to allow for quick adaption and extension of the structures. The proposed framework for focused crawling has been implemented in KAON framework and is written in Java.

7. Experimental Results

The results of this paper are the relevant web pages obtained from crawled pages for the different three seed URLs. The resulting comparison charts are drawn using Microsoft Excel software. Graphical interpretations of these results are also shown here. Performance Metrics In order to evaluate the performance of a given scheduling algorithm, the metric used is: Harvest rate Harvest rate is a common measure on how well a focused crawler performs. It is expressed
as HR= r/p,
Where, HR is the harvest rate, r is the number of relevant pages found and p is the number of pages downloaded.
Seed URLs
For the crawler to start crawling we provide some seed URLs.
http://www.puchd.ac.in
(Panjab
University),
http://www.du.ac.in (Delhi university),
http://www.ignou.ac.in/ (Indra Gandhi National
Open University).
Scenario
1. http://www.puchd.ac.in/
In first experimental run, total 1000 pages were crawled from which 478 relevant pages were obtained. Therefore, the harvest ratio obtained for this crawler run is 48%. The harvest ratio for seed URL http://www.puchd.ac.in:80/ is shown in Figure 4.
From first crawler run, the sample of top ten URLs of obtained results set is shown in Table 1 as:
2. http://www.du.ac.in/
In second experimental run, 464 relevant pages were obtained from total crawled pages i.e. 1,000.

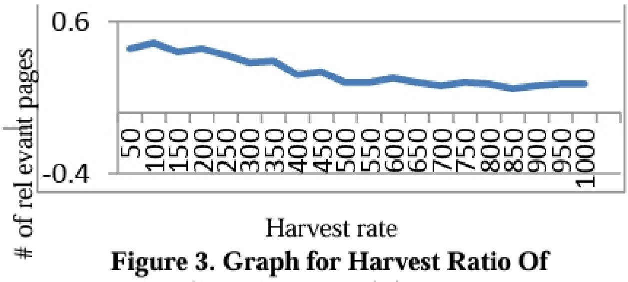


Figure 3. Graph for Harvest Ratio Of http://www.puchd.ac.in/

Table 1. Top 10 results for Panjab University

rank	Web Page
1	http://directory.puchd.ac.in:80/
2	http://exams.puchd.ac.in:80/
3	http://uiet.puchd.ac.in:80/
4	http://puchd.ac.in:80/prospectus.php
5	http://punet.puchd.ac.in:80/
6	http://forms.puchd.ac.in:80/
7	http://admissions.puchd.ac.in:80/
8	http://results.puchd.ac.in:80/
9	http://tenders.puchd.ac.in:80/
10	http://alumni.puchd.ac.in:80/

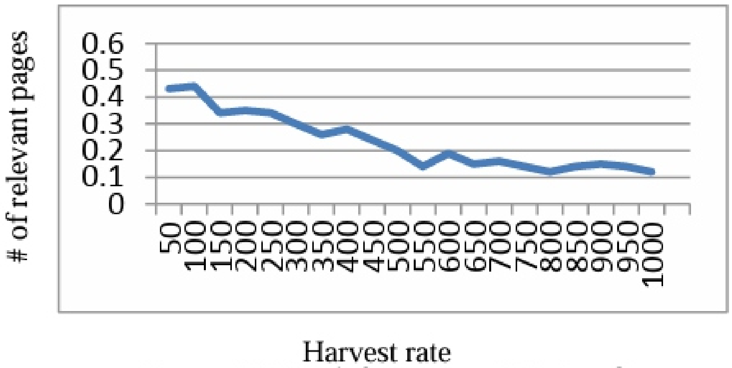


Figure 5. Graph for Harvest Ratio of http://www.du.ac.in/

Therefore, the harvest ratio obtained in this second run is 46% which is shown in Figure 5.

1. **http://www.ignou.ac.in/**

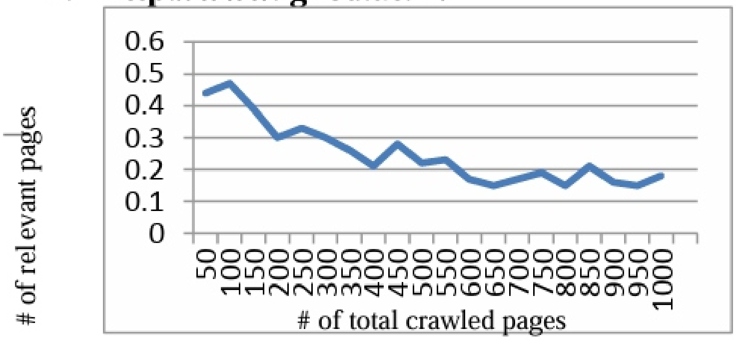


Figure 6. Graph of harvest ratio of http://www.ignou.ac.in/

In the third experimental run, 496 relevant pages are obtained from 1000 crawled pages. Therefore, the harvest ratio obtained in this third run is .49% as shown in Figure 6.

A. Average Harvest Rate Of Three Experimental Run

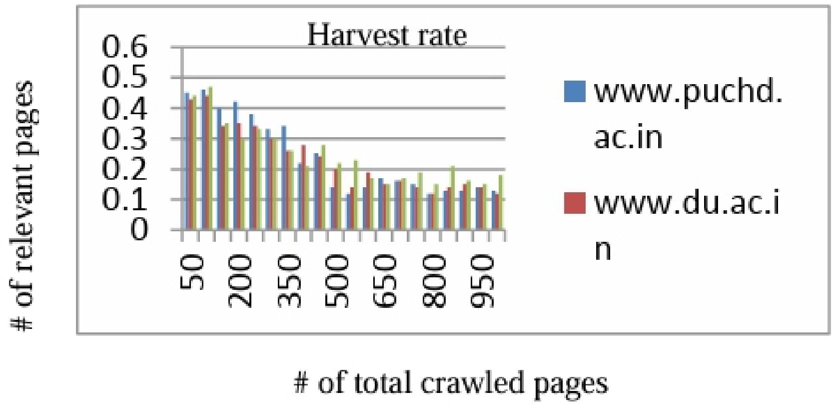


Figure 7. Average Harvest ratio of above three URLs

In above three experimental runs, total 3,000 webpages were crawled from which total of 1,434 pages were obtained. The above results of these three seed URLs i.e www.puchd.ac.in/, www.du.ac.in/, www.ignou.ac.in/ show that our ontology-based focused crawler is better than standard crawler and having average harvest ratio of 48%.

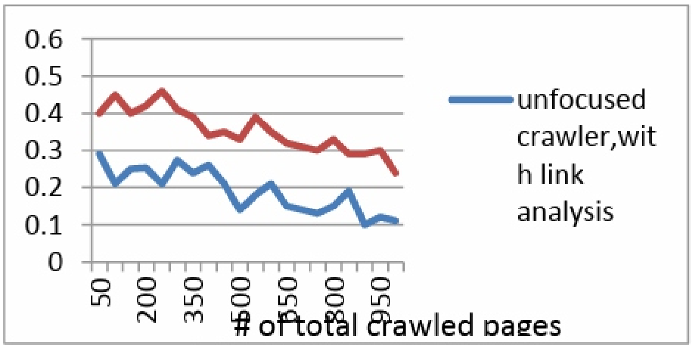
B.Comparison Of Unfocused Crawler And Ontology-Based Crawler:

The literature analysis shows that unfocused crawler with link analysis algorithm crawled 350 pages out of 1000 pages i.e. the obtained harvest ratio is 35% as shown in Table 2.

Table 2. Simulation results of different algorithm

Strategy	# of pages visited	# of relevant pages visited	Harvest Ratio
Breadth First	1,000	287	28%
PageRank	1,000	350	35%
Ontology based crawler	1,000	478	48%

Another evaluation run shows that more relevant pages were` obtained using ontology-based crawler rather than unfocused crawler is given in Figure 8. With the help of ontology-based crawler using analysis algorithm, the harvest ratio obtained is 48%, while with unfocused crawler having link analysis algorithm, the harvest ratio obtained is 35%. This shows that more relevant pages can be retrieved by using ontology with our proposed combined strategy.



# of total crawled pages

Figure 8. Comparison of unfocused crawler, with link analysis and ontology-based crawler, with link analysis algorithm

In this paper, a combined strategy of link analysis algorithm guided by topic ontology is proposed in order to efficiently discover pages relevant to the domain of interest. The prototype uses the structured information in the ontology to guide the crawler in its search for web pages that are relevant to the topic specified in the ontology. The test results show that the use of link analysis in our prototype gives a slight increase in the harvest rate.

Our crawler depends on rating the links which in turn enhance the discovery mechanism, with the introduction of combination of importance metric, this distinguishes our approach from existing approaches as the link with the higher calculated rank will be visited next. A final conclusion of this work is the realization that it is definitely worth using advanced knowledge structures when searching a specific domain on the Internet and it is possible to extract much more information from the large distributed database Internet as today's applications allow. This makes it an effective tool for the Semantic Web environment. This may result in improving the performance in the area of focused crawling and overcomes the various drawbacks of the current approaches.

References

[1] Blaz Novak, "A Survey of Focused Web Crawling Algorithms" SIKDD 2004 multi conference IS2004, 1215 Oct 2004.

[2] Hiep Phuc Luong, Susan Gauch, Qiang Wang, 2009. Ontology-based Focused Crawling, International Conference on Information, Process, and Knowledge Management, pp. 123-128.

[3] Li, H., Peng, Q. Q., Du, Y. J., Zhao, Y., Chen, S. M., Gao, Z. Q. (2009). Focused web crawling strategy based on web semantics analysis and web links analysis. Journal of Computational Information Systems, 5(6), 1793-1800

[4] Batsakisa S, Petrakisa EGM, Milios E. Improving the performance of focused web crawlers. Data Knowl Eng 2009;68(10):1001–13

[5] Chakrabarti S, van den Berg M, Dom B. Focused crawling: a new approach to topic-specific web resource discovery. Comput Netw 1999;31(11–16):1623–40.

[6] Chakrabarti, S., M. Berg, B. Dom, Fo cused crawling: A new approach to topic-sp ecific web resource discovery, Computer Networks and ISDN Systems, 31 (11-16), 1999, 1623-1640

[7] S´anchez, D., M. Batet, D. Isern. Ontology-based information content computation. Knowledge-Based Systems, 24 (2011), 297-303

[8] Mohen Jamali, Hassan Sayyadi, Babak Bagheri, Hariri and Hassan Abolhassani, 2006. A method of focused crawling using combination of link structure and content similarity, Proceedings of the International Conference on Web Intelligence.

[9] Kozanidis L. An ontology-based focused crawler. In: LNCS 5039. Springer; 2008. p. 376–9.

[10] Liu Z, Du Y, Zhao Y. Focused crawler based on domain ontology and FCA. J Inform Comput Sci 2011;8(10):1909–17

[11] Ester M., Gro M. and Kriegel H.-P.: 2001, Focused Web crawling: A generic framework for specifying the user interest and for adaptive crawling strategies,
Technical report, Institute for Computer Science, University of Munich.

[12] Aggarwal, C., F. Al-Garawi and P. Yu. "Intelligent Crawling on the World Wide Web with Arbitrary Predicates", In Proceedings of the 10th International WWW Conference, Hong Kong, May 2001.

[13] Ehrig M. and A. Maedche "Ontology-Focused Crawling of Web Documents" Proc. the 2003 ACM symposium on applied computing.

[14] Cho, J., H.Garcia - Molina, and L. Page. Efficient crawling through URL ordering. Computer Networks, 30(17):161172, 1998.

[15] Page, L., S. Brin, R. Motwani, T. Winograd. "The PageRank Citation Ranking: Bringing Order to the Web", Stanford Digital Library Technologies Project.

[16] Ganesh, S., M. Jayaraj, V. Kalyan, and G. Aghila, "Ontology-based Web Crawler," Proc. of the International Conference on Information Technology: Coding and Computing, Las Vegas, NV, USA, pp.337341, 2004.

[17] Deepika Koundal, Mukesh Kumar, Renu Vig, "Prioritizing the URLs in Ontology based Crawler" published and presented at International Conference of IEEE- AICC '2009 at Thapar University, Patiala.

[18] Debashis Hati, Amritesh kumar, 2010. An approach for identifying URLs based on Division score and link score in focused crawler, International journal of computer applications, Volume 2 – No.3.

[19] Debashis Hati, Amritesh Kumar, Lizashree Mishra, 2010. Unvisited URL Relevancy Calculation in Focused Crawling Based on Naïve Bayesian Classification, International Journal of Computer Applications, Volume 3- No.9.

A No-Reference Blur Metric based on Second-Order Gradients of Image

T. Askari Javaran1*, A. Alidadi2 and S. R. Arab2
1. Faculty of Computer science, Higher Education Complex of Bam, Bam, Iran.
2. Faculty of Computer Engineering, Higher Education Complex of Bam, Bam, Iran.

ABSTRACT

Estimation of the blurriness value in an image is an important issue in the image processing applications such as image deblurring. In this paper, a no-reference blur metric with a low computational cost is proposed, which is based on the difference between the second-order gradients of a sharp image and the one associated with its blurred version.
The experiments, in this work, are performed on four databases including CSIQ, TID2008, IVC, and LIVE. The experimental results obtained indicate the capability of the proposed blur metric in measuring image blurriness and also the low computational cost compared with the other existing approaches.

Keywords: No-reference Blur Metric, Blur Estimation, Second-order Gradients.

1. Introduction

Blur is a phenomenon that makes the details of an image not clearly visible, and its edges are weakened. In the process of reconstructing a sharp image of its blurry version, an accurate estimation of blur is the first step. Therefore, it is required to define a metric that measures the blurriness value of an image. Several metrics have been introduced for estimation of the blurriness value in an image. However, most of these metrics are based on sophisticated algorithms, and consequently, are time-consuming.
According to the research works, we can say that there are five standard categories of blur metrics. The energy of image can be used in order to estimate the amount of blurriness, because the blur smoothens the image and reduces its energy. This phenomenon is used for image blur estimation in the first category of blur metrics [1]. For blur estimation, in [2], the counted number of high frequency DCT coefficients above a threshold is used. The energy ratio of the high frequency coefficients to the low ones has been used for the estimation of blurriness value in an image in [3]. A blind image blur evaluation has been presented in [4] based on discrete Tchebichef moments. First, the gradient of a blurred image is computed in order to account for the shape. Then the gradient image is divided into equal-size blocks and the Tchebichef moments are calculated to characterize the image shape. The energy of a block is computed as the sum of squared non-DC moment values. Finally, the proposed image blur score is defined as the variance-normalized moment energy. The edges of an image have been considered in the second category of blur metrics. In [5], the edges and their width are extracted by vertical and horizontal gradients. In [6], the edges have been extracted by local gradients. The concept of Just Noticeable Blur (JNB) has been employed with the edge detection in [7]. JNB is a perceptual model that specifies the probability of blur detection by the human eye. JNB has been improved by the Cumulative Probability of Blur Detection (CPBD) in [8]. CPBD is based on a probability framework on blur perception sensed by the human eye in different illumination conditions [8]. In [9], the edge information has been extracted by a Toggle operator and used as weight of the local patterns. A support vector regression method is used to

train a predictive model for blur estimation. In [10], the second derivative values along two directions have been combined in order to get the amplitude of the second derivative at the edge point. The ratio of the edge points that have these second derivative amplitudes greater than a particular threshold is calculated as the blurriness value. The blur metrics in the third category are the statistical methods based on the distribution of the pixel intensities or transform coefficients. The methods proposed in [11,12] use the fact that the sharper images have a greater variance or entropy in their pixel intensities. In [13], the stretch of DCT coefficients distribution has been used as a measure for the estimation of blur. The local phase has coherence in the image discriminating features and therefore the Local Phase Coherence (LPC) has been used to estimate the amount of blur in a given image in [14]. LPC can be extracted from the complex wavelet transform domain. In order to estimate the blurriness value, in [15], the differences between local histograms in a given test image and the blurred version have been used. In [16], a blur metric has been proposed that is based on the difference between discrete cosine transform (DCT) of a sharp image and that of the blurred version. In [17], the shape information has been acquired by computing the gradient map. Then the grayscale image, gradient map, and saliency map are divided into blocks of the same size. The blocks of the gradient map are converted into DCT coefficients, from which the response function of singular values (RFSV) are generated. The sum of RFSV is then utilized to characterize the image blur. In [18], the qualityaware features have been extracted as the gradient of log-likelihood on the natural scene statistics model in order to account for the across space and orientation correlation simultaneously by means of multivariate Gaussian mixture model (GMM). In [19], the spatial and temporal features of image sequences, extracted by convolutional neural networks and long short term memory (LSTM), respectively, have been used to evaluate the degree of image distortion. Then the proposed model is learned to predict the scores of image patches. Finally, a pooling strategy is designed in order to evaluate the quality score of the whole image.

The fourth category of blur metrics are the ones that use the local gradient measures. The Singular Value Decomposition (SVD) has been used to estimate the blurriness value in [20]. In another study, the sharpness value of a given image has been estimated using the relative gradient intensity corresponding to the two greatest singular values. In [21], a measure has been presented based on a statistical analysis of local edge gradients.

The fifth category of blur metrics are the ones that are provided from a combination of the other measures in four categories. The authors of [22] have proposed a measure based on the total variation in the spatial space (sum of the absolute difference between an image and a spatially shifted version of the image) and the slope of the magnitude spectrum in the frequency space. The total variation represents the gradient of image in the vertical or horizontal direction. Therefore, the total variation is a feature of the forth category. Also, the slope of the magnitude spectrum in the frequency space is a statistical measure. This statistical measure is based on the distribution of the image transform from the frequency domain. Hence, this feature is in the third category of blur metrics. Indeed, the blur metric proposed in [22] is a combination of the third and forth categories. The method proposed in [23] is based on both the multi-scale gradients and the wavelet decomposition of the images. Therefore, this blurriness metric is a combination of the third and forth categories.

In [24], the fuzzy membership of pixels have been obtained via the MC-FCM (Markov Constraints to the Fuzzy-C-Means (MC-FCM)) clustering algorithm, and then, to leverage fuzzy membership from MC-FCM, the blur assessment toward pixels in the edge zone has been provided by modifying the Shannon's entropy. The correlations between the degradations of image qualities and their corresponding hierarchical feature sets have been used for image blur assessment in [25]. The deep residual network, which possesses multiple levels for feature integration, is employed to extract the deep semantics for a high-level visual content representation. By fusing the local structure and the deep semantics, a

for a high-level visual content representation. By fusing the local structure and the deep semantics, a hierarchical feature set is acquired.
In this paper, an evaluation metric for estimating blurriness in a given image is proposed. The proposed metric is a no reference one, i.e., from an image (without a reference image) estimates the amount of blurriness. This metric is based on a simple feature: if we blur a sharp image with a blur filter, there is a significant difference between the edges of the sharp and blurred versions. The proposed blur metric is based on this difference.
The experimental results show that the proposed blur metric can well-estimate the amount of blurriness for various types of blur and images with different complexities. In addition, it can measure the amount of blurriness with a low time complexity.
The rest of this paper is organized as follows. The proposed blur metric is presented in Section 2. In Section 3, the efficiency of the proposed blur metric is compared with some other existing blur metrics. Finally, we discuss and conclude the proposed method in Section 4.

2. Proposed Method

The blurring process makes the details of image not clearly visible and weakens its edges. Suppose that we have a sharp image. If this image is blurred (via a blurring filter), the amount of weakening in its edges is visible and significant. In other words, the difference between the edges of the sharp image and the ones of the blurred version is noticeable and significant. Now suppose that we blur the same blurry image (via the same blurring filter). The amount of damage on the edges of the blurry image is not very noticeable. In other words, there is not much difference between the edges of a blurred image and its re-blurred version. A sharp image, its blurred version using a low-pass filter, and the reblurred image using the same filter are presented in figure 1. The edges of these images are also presented. As shown, the difference between the edges of the original and the blurred images is very significant. However, the edges of the blurred and the re-blurred images are not significantly different (at least visually). The mathematical derivations and the concept of blurring and re-blurring an image can be found in [26,27] and [16]. Suppose that we refer to the sharp image as f the blurred version (via a low-pass filter) as g1, and, the re-blurred version (via the same filter) as g2. In order to better clarify, we obtained the summation of square difference error (SSDE) between the edges of f and the edges of g1, and between the edges of g1 and the ones of g2 for five images chosen from the CSIQ database [28]. The results obtained are plotted in figure 2 in the bar form. SSDE between the edges of the sharp image and the ones of the blurred one is given in blue, and SSDE between the edges of the blurred and reblurred versions is given in green. As seen, for all the five images, SSDE between the edges of the sharp image and the ones of the blurred version is significantly larger than SSDE between the edges of the blurred and the re-blurred versions.

(a) (b)

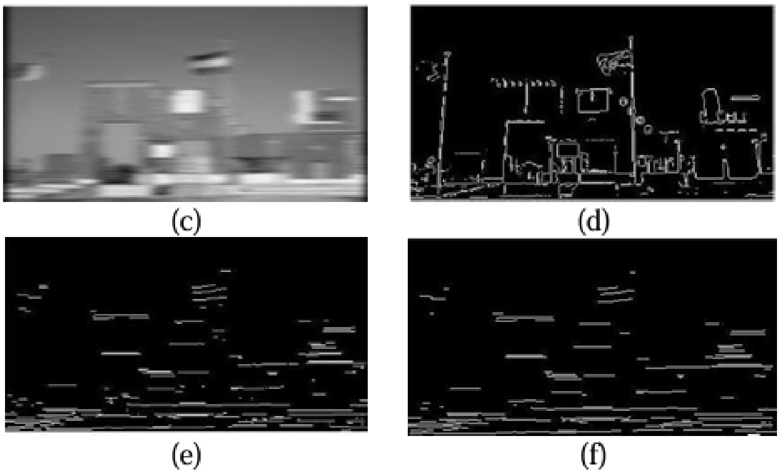


Figure 1. Difference between the original image and the blurred one: (a) original sharp image; (b) blurred version using a low-pass filter (an average filter with a 1×15 window size); (c) re-blurred image using the same filter; (d), (e), and (f) edges of images shown in (a), (b), and (c), respectively.

This phenomenon is better happened for the second-order edges (gradients) of the images. In figure 3, the second-order edges of the images shown in figure 1(a), (b), and (c) are presented. In figure 4, we plotted SSDE between the secondorder gradients of f and the second-order gradients of g1, and between the second-order gradients of g1 and the ones of g2 in the bar form for five images (chosen in Figure 2). As it can be seen, these differences are well-shown for the second gradients. As it can be seen in figure 3, the number of nonzero second-order edges for the original image is more than those for the blurred one. Due to losing details and weakening edges, a large part of the second-order edges in the blurred image is lost. However, the number of non-zero second-order edges in the blurred and re-blurred images is nearly the same, because only a small part of the other remaining second-order edges is lost in the re-blurred image. As a consequence, we can use the difference between the second-order edges in a given image and the blurred version to introduce a blur metric.

Before introducing the blur metric, the process of calculating the second-order gradients is explained. In order to detect the edge pixels in the given image, first, Canny edge detector is applied to that image. The second derivatives along the horizontal and vertical directions are calculated for all locations in the image that are classified as the edge points.

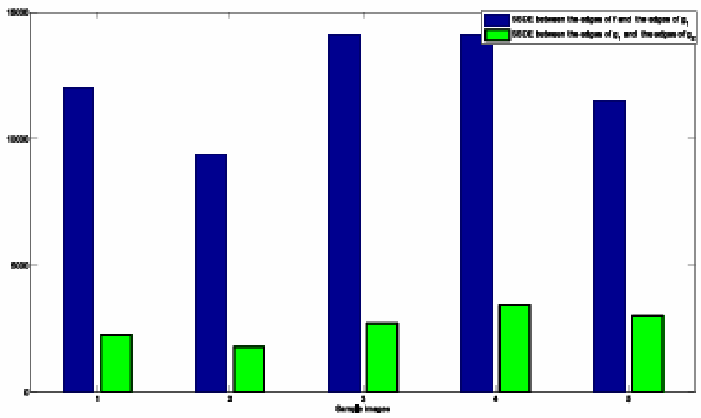
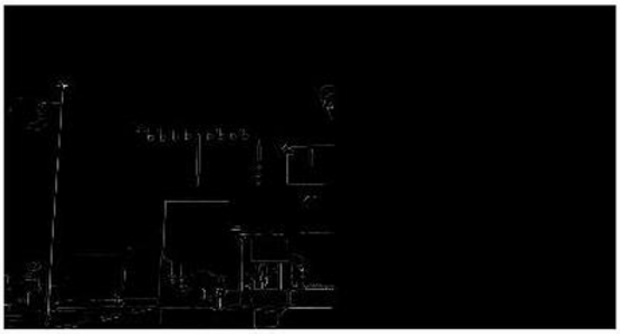


Figure 2. Comparison of SSDE between the edges of the sharp and blurred images and SSDE between the edges of the blurred and re-blurred images for five chosen images (in size of 512 × 512) from CSIQ.



(a)

(b)

(c)

Figure 3. Second-order edges of the images shown in Figure 1(a), (b), and (c).

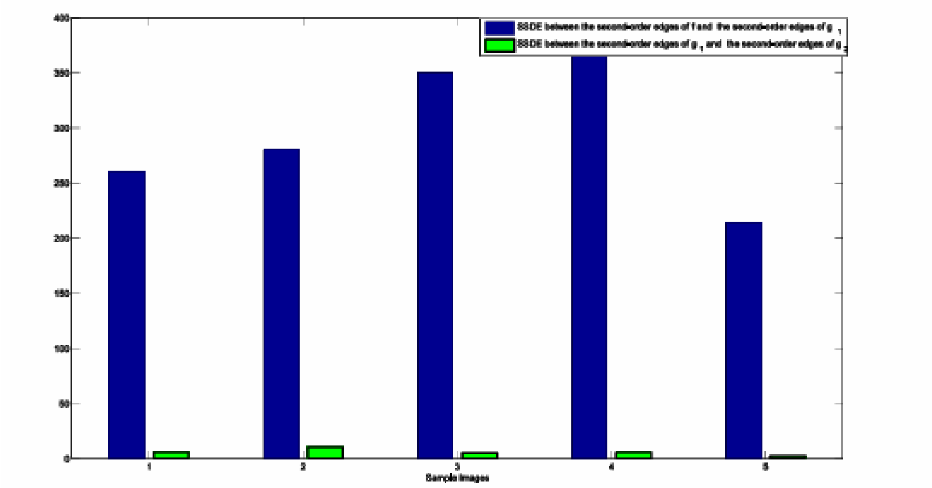


Figure 4. Comparison of SSDE between the second-order edges of the sharp and blurred images and SSDE between the second-order edges of the blurred and re-blurred images for the same five chosen images in Figure 2.

2.1. Blur Metric

For the horizontal case, denoted as the index x, the second derivative is calculated as:

$$G_{xx}(x,y) = I(x,y) - 2I(x-1,y) + I(x-2,y), \quad (1)$$

For the vertical case, denoted as index y, it is calculated as:

$$G_{yy}(x,y) = I(x,y) - 2I(x,y-1) + I(x,y-2), \quad (2)$$

where *I (x, y)* represents the pixel value intensity at the edge point location *I (x, y).*
derivative at the edge point, the second derivative values along two directions are combined as follows (for all edge points):

$$G(x,y) = G_{xx}(x,y)^2 + G_{yy}(x,y)^2 \quad (3)$$

In order to take the effect of the above-mentioned difference into account (the difference between the second-order edges in a given image and the blurred version), the l_1 norm ratio between the second-order gradients of the given image and those of the blurred version, is suggested as a metric. This blur metric is defined as follows:

$$\boldsymbol{\beta}(I) = \frac{\| G_b \|}{\| G \|}, \quad (4)$$

where G_b represents the second-order edges of the blurred version of *I*, which is obtained by applying a low pass filter on I. G ‖ ‖ is l1 norm, and is defined as follows:

$$| \; G\| = \sum_{x,y} G(x,y) \quad (5)$$

The blurriness estimated using Eq. 4 is a value within [0,1], in which a closer value to 1 indicates that the image is more blurry, and consequently, a closer value to 0 indicates that the image is more sharp. The following is an explanation of this phenomenon.

As mentioned earlier, for a very sharp image, the difference between the second-order gradients of the given image and that of the blurred version is large. Thus the amount of the fraction denominator is much greater than the fraction face. Therefore, the amount of fraction will be near to 0. As a result, the value of βfor a sharp image is close to zero for a sharp image. On the other hand, for a very blurry image, the difference between its second-order gradients and that of the blurred version is small, so the amount of the fraction denominator is not much greater than the fraction face. Hence, the amount of fraction will be near to 1. Consequently, the value of βfor a blurry image is close to one.

The blurriness value (( ) I β) of the sharp image shown in figure 1 is about 0.37, whereas, it is about 0.71 for the blurred and 0.87 for the reblurred ones, shown in figures 1(b) and (c), respectively.

3. Experimental Results

We evaluated the performance of the proposed blur metric by applying it to estimate the blurriness of the selected images. We selected four popular databases that contained blurry images: CSIQ [28], TID2008 [29], LIVE [30], and IVC [31]. Each database consists of the original images, distorted versions using Gaussian blurring at different levels. There are 150, 100, 145, and 20 blur images in the CSIQ, TID2008, LIVE, and IVC databases, respectively. The mean opinion scores (MOSs) scores for all the distorted images in all the four databases are presented.

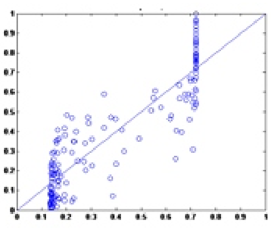
The VQEG report [32] has proposed the suggestions to measure how well the metric values correlate with the provided MOS values and to objectively evaluate its performance. As several researchers have done, we followed these suggestions. In the VQEG report, four indicators are suggested to compute: Spearman's RankOrder Correlation Coefficient (SRCC), Kendall's Rank-Order Correlation Coefficient (KRCC), Pearson's Linear Correlation Coefficient (PLCC), and Root Mean Squared Error (RMSE). Both SRCC and KRCC are used to validate prediction monotonicity [33].

In order to evaluate the prediction accuracy, PLCC and RMSE are used [33]. If a given measure yields high values in PLCC, SRCC, and KRCC; and low values in RMSE, it will be a good objective quality measure [33]. In [34,35], the definition of these indicators and more details can be found. The results of the proposed method were compared with those obtained using the most cited no-reference blur metrics: JNBM [7], CPBD [8], LPC-SI [14], BLIINDS-II [36], and NI-DCT [16].

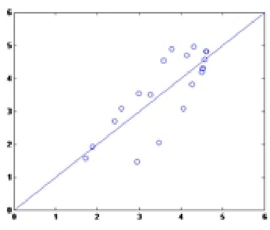
For these six metrics (five above mentioned blur metrics along with the proposed metric), the four indicators introduced earlier were computed. The results obtained are shown in table 1. As it can be concluded, the performance of the proposed blur metric is comparable to the other blur metrics; in some cases, it is the best. As suggested in the VQEG report, and as done by other researchers, we showed the scatter plots of the MOSs versus the blurriness values estimated by the six blur metrics. This is done for the visual inspection of the correlation between the estimated blurriness values and MOSs. The results obtained are shown in figure 5. In this figure, each sample point represents one test image. As suggested in the VQEG report, a logistic fitting function was used to provide a nonlinear mapping between the scores to accommodate for the quality rating compression at the extremes of the test. As it can be seen, under comparison, the sample points for the proposed blur metric generally tend to be clustered closer to the diagonal lines than the other five blur metrics.

Table 1. Performance evaluation of the six blur metrics on four databases.

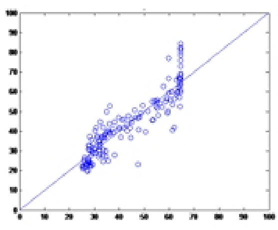
	LIVE (145 blurred images) [30]			
	SRCC	**KRCC**	**PLCC**	**RMSE**
JNBM [7]	0.7876	0.6069	0.8161	9.0857
CPBD [8]	0.9194	0.7653	0.8955	6.9971
BLINDS-II [36]	0.8242	0.6404	0.8623	7.9629
LPC-SI [14]	**0.9394**	**0.7785**	0.9182	6.2288
NI-DCT [16]	0.9282	0.7701	0.9408	5.3289
The proposed	0.9322	0.7782	**0.9456**	**5.0234**
	CSIQ150 (145 blurred images) [28]			
	SRCC	**SRCC**	**SRCC**	**SRCC**
JNBM [7]	0.7624	0.5976	0.8061	0.1669
CPBD [8]	0.8853	0.6646	0.8822	0.1349
BLINDS-II [36]	0.8396	0.709	0.876	0.1382
LPC-SI [14]	**0.9071**	0.7205	0.9158	0.1151
NI-DCT [16]	0.8888	0.7162	0.9224	0.1107
The proposed	0.8991	**0.7214**	**0.9257**	**0.1102**
	IVC (20 blurred images) [31]			
	SRCC	**SRCC**	**SRCC**	**SRCC**
JNBM [7]	0.6659	0.4974	0.6983	0.8172
CPBD [8]	0.769	0.6138	0.8012	0.6832
BLINDS-II [36]	0.8397	0.6667	0.8983	0.5016
LPC-SI [14]	0.9398	0.8042	0.9726	0.2653
NI-DCT [16]	**0.9782**	**0.9101**	**0.9905**	**0.1567**
The proposed	0.9723	0.9087	0.9889	0.1619
	TID2008 (100 blurred images) [29]			
	SRCC	**SRCC**	**SRCC**	**SRCC**
JNBM [7]	0.6667	0.4951	0.6931	0.8459
CPBD [8]	0.8414	0.6301	0.8237	0.6655
BLINDS-II [36]	0.6972	0.4793	0.6952	0.8435
LPC-SI [14]	0.8561	0.6362	0.8574	**0.604**
NI-DCT [16]	0.833	0.6107	0.841	0.6349
The proposed	**0.8565**	**0.6373**	**0.8593**	0.6076



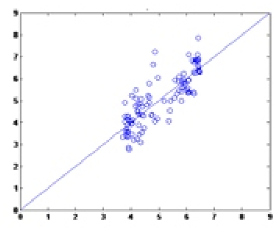
CPBD-CSIQ, Cr=0.8822



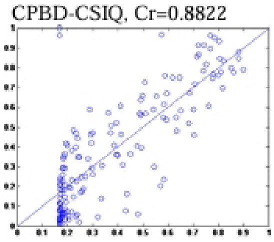
CPBD- IVC, Cr=0.8012



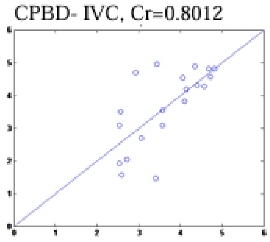
CPBD- LIVE, Cr=0.8955



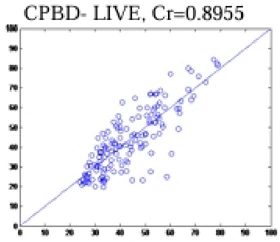
CPBD-TID2008, Cr=0.8237



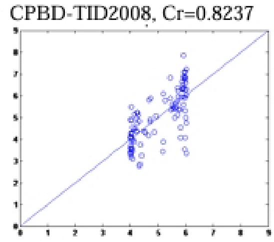
JNBM- -CSIQ, Cr=0.8061



JNBM- IVC, Cr=0.6983



JNBM- LIVE, Cr=0.8161



JNBM- TID2008, Cr=0.6931

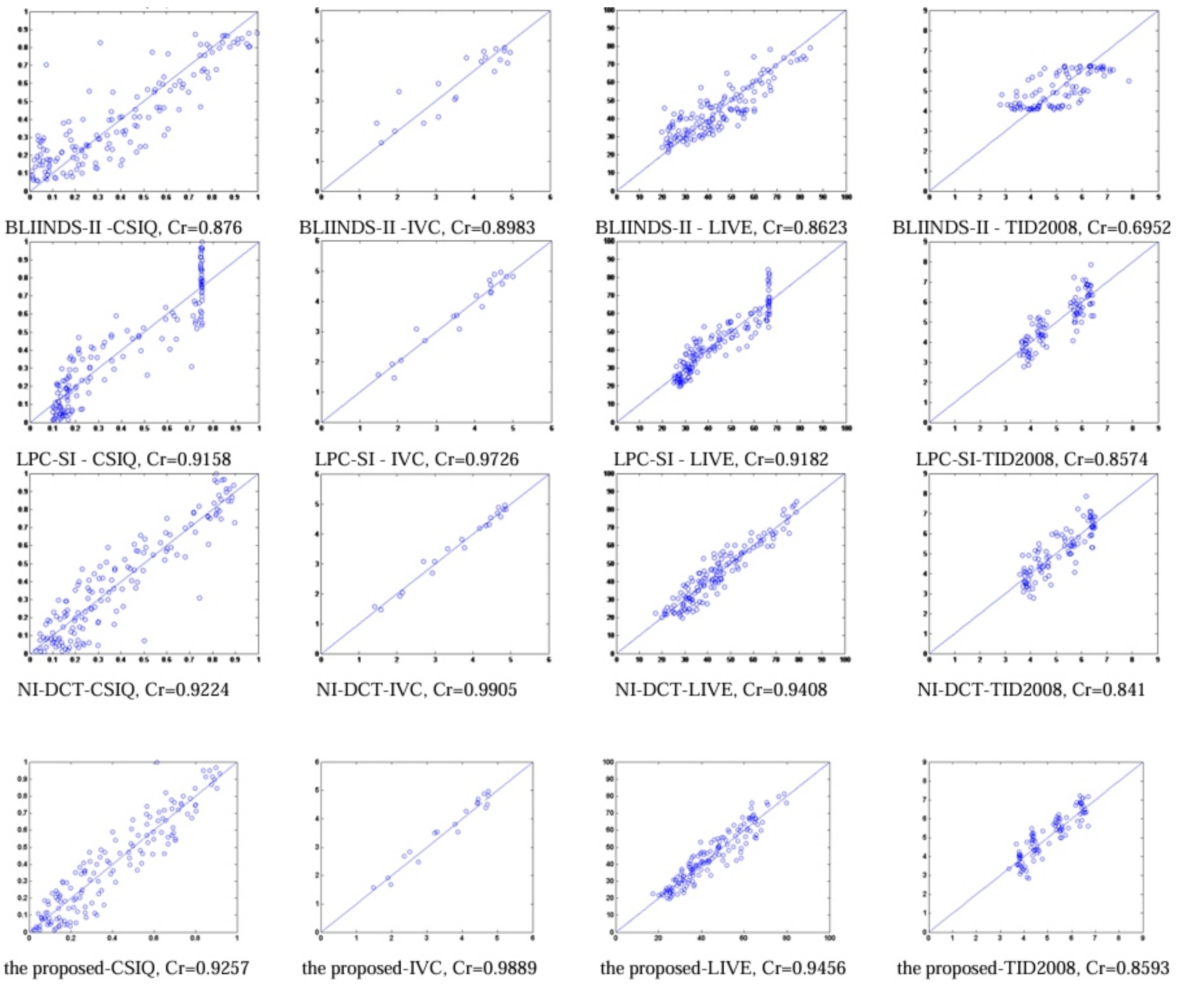


Figure 5. Scatter plots between MOSs and the values estimated (after nonlinear mapping) by the six blur metrics over the four blur image databases, with the correlation coefficient (Cr). Top to bottom rows: CPBD [8], JNBM [7], BLIINDS-II [36], LPC-SI [14], NI-DCT [16] and the proposed blur metric; Left to right columns: CSIQ, IVC, LIVE, and TID2008 databases.

In order to compare the runtime of the six blur metrics, another experiment was applied on 150 images with 512 × 512 resolutions from the CSIQ database. This test was performed on a computer configured with Intel Core i3 CPU 3.60 GHz, 4 GB RAM, Windows 7 64-bit, and MATLAB 8.3. Table 2 shows the runtime of the six blur metrics. Although the BLIINDS-II algorithm is a fast algorithm, it requires a long training process [36]. However, the proposed blur metric is the fastest algorithm.

4. Conclusion and Discussion

In this paper, a metric was proposed for estimating blur in the image. This metric was defined on the basis of the difference between the second-order derivations of the original image and the second-order derivations of the blurred version. If the given image is sharp, there is a significant difference between its second-order derivations and the ones of the blurred version. However, if the image is blurry, this difference is less. This phenomenon is taken to help define a measure of blur. This metric performs better than the other methods. The high speed of this metric is another feature.

Table 2. Runtime comparisons of blur metrics for images of 512×512 resolution.

Blur metric	Runtime (second)
JNBM [7]	0.8537
CPBD [8]	0.969
BLINDS-II [36]	0.0853
LPC-SI [14]	2.2763
NI-DCT [16]	0.2027
The proposed	**0.0652**

References

[1] J. Antkowiak, T. Jamal Baina, F.V. Baroncini, N. Chateau, F. FranceTelecom, A.C.F. Pessoa, F. Philips, "Final report from the video quality experts group on the validation of objective models of video quality assessment," In: VQEG meeting, Ottawa, Canada, March, 2000.

[2] D. B. L. Bong and B. E. Khoo, "An efficient and training-free blind image blur assessment in the spatial domain," IEICE transactions on Information and Systems, vol. 97, no. 7, pp. 1864–1871, 2000.

[3] D. B. L. Bong and B. E. Khoo, "Blind image blur assessment by using valid reblur range and histogram shape difference," Signal Processing: Image Communication, vol. 29, no. 6, pp. 699–710, 2014.

[4] P. Bromiley, "Products and convolutions of gaussian distributions," Medical School, Univ. Manchester, Manchester, UK, Tech. Rep, vol. 3, 2003.

[5] J. Caviedes and F. Oberti, "A new sharpness metric based on local kurtosis, edge and energy information," Signal Processing: Image Communication, vol. 19, no. 2, pp. 147–161, 2004.

[6] M.-J. Chen and A. C. Bovik, "No-reference image blur assessment using multiscale gradient," EURASIP Journal on image and video processing, vol. 2011, no. 1, p. 3, 2011.

[7] N. N. K. Chern, P. A. Neow, and M. H. Ang, "Practical issues in pixel-based autofocusing for machine vision," in Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164), vol. 3, pp. 2791–2796, IEEE, 2001.

[8] S. Chora's, A. Giełczyk, and M. Choraś, "Ten Years of Image Processing and Communications." International Conference on Image Processing and Communications. Springer, Cham, 2018.

[9] S. Erasmus and K. Smith, "An automatic focusing and astigmatism correction system for the sem and ctem," Journal of Microscopy, vol. 127, no. 2, pp. 185199, 1982.

[10] C. Feichtenhofer, H. Fassold, and P. Schallauer, "A perceptual image sharpness metric based on local edge gradient analysis," IEEE Signal Processing Letters, vol. 20, no. 4, pp. 379–382, 2013.

[11] R. Ferzli, and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb)," IEEE transactions on image processing, vol. 18, no. 4, pp. 717–728, 2009.

[12] G. Ghosh Roy, "A Simple Second Derivative Based Blur Estimation Technique," (Doctoral dissertation, The Ohio State University, 2013.

[13] R. Hassen, Z. Wang, and M. M. Salama, "Image sharpness assessment based on local phase coherence," IEEE Transactions on Image Processing, vol. 22, no. 7, pp. 2798–2810, 2013.

[14] L. He, Y. Zhong, W. Lu, and X. Gao, "A visual residual perception optimized network for blind image quality assessment," IEEE Access, vol. 7, pp. 176087176098, 2018.

[15] T. A. Javaran, H. Hassanpour, and V. Abolghasemi, "A noise-immune no-reference metric for estimating blurriness value of an image," Signal Processing: Image Communication, Elsevier, vol. 47, pp. 218-228, 2016.

[16] E. C. Larson, and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," Journal of Electronic Imaging, vol. 19, no. 1, 011006, 2010.
[17] P. Le Callet, and F. Autrusseau, "Subjective quality assessment irccyn/ivc database: Available: http://hal.univ-nantes.fr/hal-00580755/. [Accessed January 10, 2021].
[18] L. Li, W. Lin, X. Wang, G. Yang, K. Bahrami, and A. C. Kot, "No-reference image blur assessment based on discrete orthogonal moments," IEEE transactions on cybernetics, vol. 46, no. 1, pp. 39–50, 2015.
[19] Q. Li, W. Lin, K. Gu, Y. Zhang, and Y. Fang, "Blind image quality assessment based on joint logcontrast statistics," Neurocomputing, vol. 331, pp. 189198, 2019.
[20] L. Liu, J. Gong, H. Huang, and Q. Sang, "Blind image blur metric based on orientation-aware local patterns," Signal Processing: Image Communication, vol. 80, 115654, 2020.
[21] X. Marichal, W. Y. Ma, and H. Zhang, "Blur determination in the compressed domain using DCT information," [Conference Proceedings]. In Image processing, 1999. icip 99. proceedings. 1999 international conference on (Vol. 2, p. 386-390). IEEE, 1999.
[22] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "A no-reference perceptual blur metric," In Proceedings. International conference on image processing (Vol. 3, pp. III–III), 2002.
[23] N. D. Narvekar,and L. J. Karam, "A no-reference image blur metric based on the cumulative probability of blur detection (CPBD)," IEEE Transactions on Image Processing, vol. 20, no. 9, pp. 2678–2683, 2011.
[24] E. Ong, W. Lin, Z. Lu, X. Yang, S. Yao, F. Pan, L. Jiang, and F. Moschetti, "A no-reference quality metric for measuring image blur," in Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings., vol. 1, pp. 469–472, IEEE, 2003.
[25] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "Tid2008-a database for evaluation of full-reference visual quality assessment metrics," Advances of Modern Radioelectronics, vol. 10, no. 4, pp. 30–45, 2009.
[26] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the dct domain," IEEE transactions on Image Processing, vol. 21, no. 8, pp. 3339–3352, 2012.
[27] D. Shaked, and I. Tastl, "Sharpness measure: Towards automatic image enhancement," In IEEE international conference on image processing 2005 (Vol. 1, pp. I–937), 2005
[28] H. Sheikh, "Live image quality assessment database release 2," Available: http://live.ece.utexas.edu/research/quality, [Accessed January 10, 2021].
[29] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," IEEE Transactions on image processing, vol. 15, no. 11, pp. 3440–3451, 2006
[30] C. T. Vu, T. D. Phan, and D. M. Chandler, "S3: A spectral and spatial measure of local perceived sharpness in natural images," IEEE transactions on image processing, vol. 21, no. 3, pp. 934–945, 2011
[31] J. Wu, J. Zeng, W. Dong, G. Shi, and W. Lin, "Blind image quality assessment with hierarchy: Degradation from local structure to deep semantics," Journal of Visual Communication and Image Representation, vol. 58, pp. 353–362, 2019.
[32] S. Xu, S. Jiang, and W. Min, "No-reference/blind image quality assessment: a survey," IETE Technical Review, vol. 34, no. 3, pp. 223– 245, 2017.
[33] Y. Xu, W. Zheng, J. Qi, and Q. Li, "Blind image blur assessment based on markov-constrained fcm and blur entropy," In 2019 IEEE international conference on image processing (icip) (pp. 4519–4523), 2019.
[34] S. Zhang, P. Li, X. Xu, L. Li, and C. C. Chang, "No- reference image blur assessment based on

response function of singular values," Symmetry, vol. 10, no. 8, 304, 2018.
[35] E. Fadaei-Kermani, G. Barani, M. GhaeiniHessaroeyeh, "Drought Monitoring and Prediction using K-Nearest Neighbor Algorithm," Journal of AI and Data Mining, vol. 5, no. 2, pp. 319-325. doi: 10.22044/jadm.2017.881, 2017.
[36] X. Zhu, and P. Milanfar, "A no-reference sharpness metric sensitive to blur and noise," In 2009 international workshop on quality of multimedia experience (pp. 64–69), 2009.

Improved Facial Action Unit Recognition using Local and Global Face Features

Amin Rahmati Sardashti and Foad Ghaderi
Human-Computer Interaction Lab., Faculty of Electrical and Computer Engineering,
Tarbiat Modares University, Tehran, Iran.

ABSTRACT

Every facial expression involves one or more facial action units appearing on the face. Therefore, action unit recognition is commonly used to enhance the facial expression detection performance. It is important to identify subtle changes in the face when particular action units occur. In this paper, we propose an architecture that employs the local features extracted from specific regions of face while using the global features taken from the whole face. To this end, we combine the SPPNet and FPN modules to architect an end-to-end network for facial action unit recognition. First, different predefined regions of face are detected, and next, the SPPNet module capture deformations in the detected regions. The SPPNet module focuses on each region separately and cannot take into account possible changes in the other areas of the face. In parallel, the FPN module finds the global features related to each of the facial regions. By combining the two modules, the proposed architecture is able to capture both the local and global facial features, and enhance the performance of action unit recognition task. The experimental results on the DISFA dataset demonstrate the effectiveness of our method.

Keywords: Facial action recognition, Facial action units, Deep learning.

1. Introduction

People's intentions, expressions, physical or mental states usually appear in their faces, and it is believed that people's face say a lot about them. Facial behavior analysis is one of the most popular research areas in affective computing, human-computer interaction (HCI), and machine vision. Previous research works show that people can not completely prevent their intentions and internal states from being represented on their faces [1]. That means, analyzing people's facial behavior can help us understand their goals and intentions.

As mentioned in [2], facial behaviors can be described using two different approaches, i.e., facial expressions and facial action units (AU). Facial expressions are nothing except occurrence of meaningful combinations of facial AUs, which are movements of one or more facial muscles. combinations of two or more different AUs and their appearances on the face depict unique facial expressions. Mapping between the combination of AUs and the corresponding facial expressions is presented in [3]. Using such a mapping, if the AUs and their combinations are identified in a face, facial expressions can also be practically distinguished.

In an attempt for systematic analysis of human facial behavior, Ekman and Friesen developed the facial action coding system (FACS), which is a comprehensive reference system for studying facial actions based on anatomy of human face [4]. The goal of AU detection in a given facial image (or in a sequence of frames in a video) is to measure the similarity of facial muscle movements with those defined in FACS.

Action unit detection is a difficult task and no one can perform it with high performance if they don't have prior knowledge. Nevertheless, manual annotation is time-consuming and expensive, such that it takes more than 30 minutes for an expert to annotate one minute of a video clip [5]. Moreover, subtle changes in parts of face during the AU occurrence yield to variations in AU appearance, which causes more challenges in the AU recognition task. On the other hand, there are more technical challenges in automatic AU detection, namely, lack of large datasets with AU annotations, diverse subjects, and imbalanced AU datasets.

The action unit recognition methods can be divided into two group, i.e., those that use the whole face, and those that first divide the face image into parts related to the AUs and afterwards classify each part separately. In the latter, it is possible to tackle the subtle facial variations more carefully, however, global features and the relations and the dependencies among the AUs are missing. In this work, we combined the two approaches to take advantage of both local and global information simultaneously. This can help to obtain higher recognition rates and eliminate possible flaws.

The rest of the paper is organized as what follow.

An overview of the previous research works in the field of facial AU detection is listed in Section 2. Our proposed method is presented in detail in Section 3. The experimental results are discussed in Section 4. Finally, the paper is concluded in Section 5.

2. Related Works

Many efforts have been made over the previous years in the research filed of AU detection to extract useful features for enhancing detection rate. Static two-dimensional image representation is one of the famous methods of facial AU feature extraction [6]. In this approach, facial features are divide into two categories i.e., appearance and geometry. Gabor wavelets [7, 8], Haar feature [9], scale-invariant feature transform (SIFT) [10], and local binary pattern [11] are the most common handcrafted appearance-based features. On the other hand, deformations in the various components of face convey information that constitute geometric features and can be measured by optical flows [12] or dislocation of landmark points [13, 14]. Some researchers have used a combination of these two feature representation approaches to improve the overall performance [6]. The authors of [15] proposed Multiple kernel Learning. The authors of [16] used the SimpleMKL algorithm, combined the two types of features, and averaged the outcome to exploit the temporal information in sequences. The authors of [17] proposed a multi-conditional latent variable dependencies at both feature and model level into the proposed manifold learning for AU recognition by introducing topological and relational constraints.

The power of deep learning algorithms and their efficiency in various fields has led to the recent use of these techniques in the AU recognition task. The authors of [18] proposed AU R-CNN, in which by designing the AU partition rule, the images are decomposed into a bunch of AUrelated bounding boxes and different regions of face are localized. The regions are then merged to obtain the image-level prediction. The authors of [19] proposed a hybrid CNN-RNN network for human action recognition from video. Shao et al. suggested to jointly perform AU recognition and face alignment in order to use the specific AU positions provided by landmarks [20]. They further captured local AU-related characteristics via spatial attention mechanism [21]. The authors of [22] proposed Geodesic Guided Convolution (GeoConv) for AU recognition by embedding 3D manifold information into 2D convolutions in which the convolutional kernel is weighted by geodesic distances on the 3D facial surface. In an attempt to assess the effectiveness of 2D and 3D CNNs in human action recognition task, the authors of [23] evaluated these networks in hand gesture recognition task.

In a separate line of research, the encoderdecoder models have been employed in this context. [24] used graph convolutional networks (GCN) for AU relation modeling. They used autoencoders to extract latent representation of AUrelated regions to be fed to GCN for modeling AU relationships. In [25], a deep structured inference network (DSIN) for AU recognition is proposed. This structure passes information obtained from extracted image features and the structure inference between predictions straightforwardly to capture the relationship between AUs. The authors of [26] proposed the AU semantic relationship embedded representation learning (SRERL) framework that first extracts global feature maps over the whole face image. Then, they process cropped features from the global feature maps, separately. Finally, they used gated graph neural networks (GGNN) to capture correlations among AUs. A Meta Auxiliary Learning method (MAL) is proposed in [27] in which adaptive weights are used for learning facial expression.".

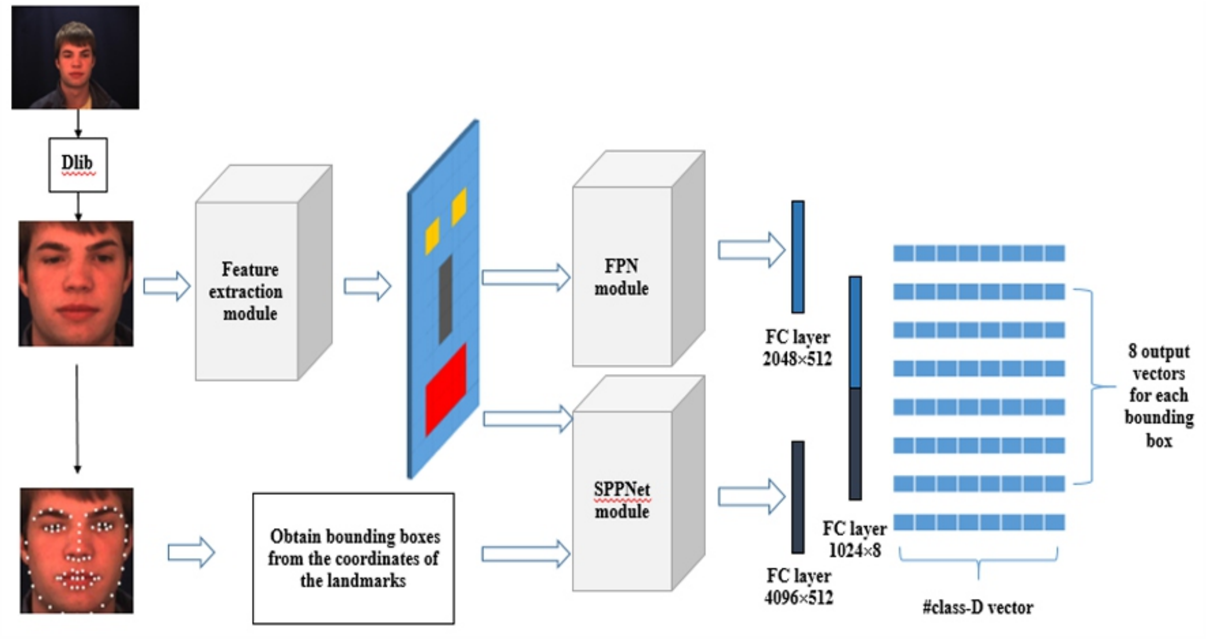


Figure 1. Outline of the proposed framework. The extracted feature maps from the initial and middle layers of ResNet 101 are fed to the FPN and SPPNet modules. Coordinates of the bounding boxes are also used by the SPPNet module. FPN and SPPNet modules work together as core modules for AU recognition.

3. Proposed Method

3.1. Overview

Most of the existing methods use only local or global features. To overcome this problem, we propose a novel architecture that covers the shortcomings of other methods. The main idea behind our proposed architecture is using both the local and global information extracted from the input data. In other words, we identify action units from locally segmented face regions while analyzing the whole face simultaneously. We combine the outcome of the two processing flows to recognize facial expressions. To this end, we use SPPNet [28] and FPN [29] in our architecture. Our proposed framework, as shown in Figure 1, consists of three parts: the initial convolutional layer, the SPPNet module, and the FPN module. In the convolutional layer, we first crop faces from the images using 68 landmark points. Then, following the [18] approach we use "expert prior knowledge" to extract the coordinates of the face regions of interest (RoIs). As shown in Figure 2, this yield to eight bounding boxes each containing specific regions of face where the desired Aus happen. After that, we employ a ResNet-101 [30] and extract the feature map from each face image.

In order to avoid overfitting of the model, we freeze conv1-res4 layer. Considering the fact that the sizes of bounding boxes are not fixed for different faces, we use RoI pooling layer to fix the sizes of the feature maps obtained from conv1-res4. For this step, we map the coordinates of the bounding boxes from the original image to the feature maps and extract them. This feature maps that belong to the bounding boxes are fed to the SPPNet, and the original feature maps that belong to the whole image face are fed to the SPPNet and the FPN modules. The outputs of the two modules are concatenated, and the final fully-connected (FC) layer's output is treated as each class probability. Finally, because the prediction was at the RoI level and belonged to each bounding box, we returned the prediction results to the image level by merging the hit of each AU or AUs in each box. The details of SPPNet and the FPN modules are explained in the sequel.

3.2. SPPNet module

One of the important properties of SPPNet is the use of multi-level spatial bins. It has been shown in [28] that this architecture is robust to object deformations. In our problem, each extracted face RoI is different in scale and level of deformation, e.g., eye regions are small and have subtle deformation compared to the other regions. Therefore, in the SPPNet module, we first import the feature map into two different branches. In the lower branch, the feature maps are given to Res5, and in the other one we configure the 4-level pyramid pooling {7×7, 3×3, 2×2, (1×1) ×2} with the total of 64 bins. In order to reduce the number of channels a 1×1 Conv is used.

Figure 2. Eight bounding boxes for each part of the face are defined. One or more AUs may occur in each bounding box.

Finally, the outcome of the two branches are concatenated and fed to a fully connected layer. Details of the SPPNet module are illustrated in Figure 3.

3.3. FPN Module

In this module, first the low-level and high resolution features (obtained from the previous layers) are transformed to high-level and lowresolution features using a Res5 block. Input and output of the Res5 block are feature maps with 1024×M×N, and 2048×M/2×N/2 respectively. Then the two set of features (before and after Res5) are combined to get more convenient representation. To this end, we first up-sample the output of Res5 by a factor of 2, because of the output of Res5 reduced by a factor of 2. Since the two feature sets have different dimensionalities, we fixed their dimensionality using convolutional layers, and pass both sets of features through 1×1 convolution layers. Then, the feature sets are added,

and a 3×3 convolution layer is applied to prevent the aliasing effect of upsampling [29]. Since the face images are divided into eight separate regions in SPP module (see Figure 2), we replicate each feature map eight times, i.e., for each bounding box we consider a feature map of the whole face.
Elements of these eight feature maps are separately element-wise multiplied by a set of learnable coefficients as follows

$$F(x_{ij}) = \sum_{k=1}^{N} \sum_{l=1}^{C} W_{ij} * x_{ij} \quad (1)$$

where all the *Wij* coefficients are set to initial value of 0.01, N is set to 8, C is the number of

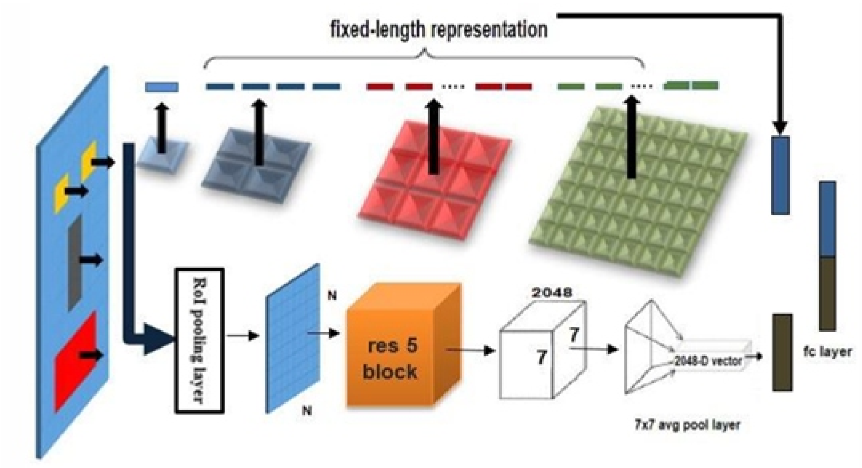


Figure 3. The SPPNet module, in which feature maps that are extracted from the initial and middle layers of ResNet101 (conv1-res4) are imported into two branches and then concatenated together and fed to a FC layer.

channels, and i,j are the coordinates of the feature map elements. The details of the FPN module are illustrated in Figure 4. This way, more attention is paid to the parts of the whole face feature maps for each bounding box that convey more informative features. At the final stage of the FPN module, these feature maps are fed into the FC layer.

4. Experiments

In this section, the dataset and the experimental setup are presented first. Details of the evaluations and comparative results are provided afterwards.

4.1. Dataset

Our proposed model is evaluated on the publicly available dataset DISFA [31]. This dataset contains 54 videos, where 27 of them were captured from the left and the rest were recorded from the right side of the subject's faces. Twentyseven young adults with diverse ethnicities participated. Each video consists of 4,485 frames, summing up to a total of about 260,000 frames. The frames are manually labeled with AU intensity on a six-point ordinal scale. Using [26, 20, 22] methods, we only considered those frames with intensities equal to or greater than 2 as positive. There are 12 AUs included in the DISFA dataset. For evaluating our framework we used 8 of them, i.e., action units 1, 2, 4, 6, 9, 12, 25, and 26.

Table 1. Results of different methods in action unit recognition task on DISFA dataset. Reported numbers are F1-frames, and bracketed and bold numbers represent the best and the second-best results, respectively.

AU	F1-frame									
	LSVM	APL	DRML	ROI-Nets	DSIN	AU R-CNN	SRERL	MAL	ARL	Our method
1	10.8	11.4	17.3	41.5	42.4	32.1	**45.7**	43.8	43.9	[47.6]
2	10	12	17.7	26.4	39.0	25.9	[47.8]	39.3	**42.1**	39.4
4	21.8	30.1	37.4	66.4	68.4	59.8	59.6	**68.9**	63.6	[70.3]
6	15.7	12.4	29	50.7	28.6	[55.3]	47.1	47.4	41.8	**52.4**
9	11.5	10.1	10.7	8.5	**46.8**	39.8	45.6	[48.6]	40.0	45.4
12	70.4	65.9	37.7	89.3	[90.4]	67.7	73.5	72.7	**76.2**	74.5
25	12	21.4	38.5	88.9	70.8	77.4	84.3	**90.6**	[95.2]	89.0
26	22.1	26.9	20.1	15.6	42.2	52.6	43.6	52.6	[66.8]	**54.8**
Avg	21.8	23.8	26.80	48.5	53.6	51.3	55.9	58.0	**58.7**	[59.2]

4.1. Implementation details

All our experiments were conducted on a computer with a GTX 1080 Ti GPU and 16 GB RAM. We used Chainer1 as our learning framework. In our processing flow, we first used Dlib2 library to get 68 landmarks for each face and then cropped the faces and resized them to 512×512 pixels. Next, we subtract the mean pixel value from all the dataset images. We augmented the dataset in a random order by horizontally flipping the input images. The size of minibatches was set to 8. Since the backbone of our model is the same as that of AU R-CNN [18], we employed the concept of transfer learning, used the pre-trained model on the BP4D [32], and finetuned the last layers. We used Stochastic Gradient Descent (SGD) optimization algorithm and set the learning rate to 10−4. The learning rate was reduced every ten epochs by a factor of 20%.

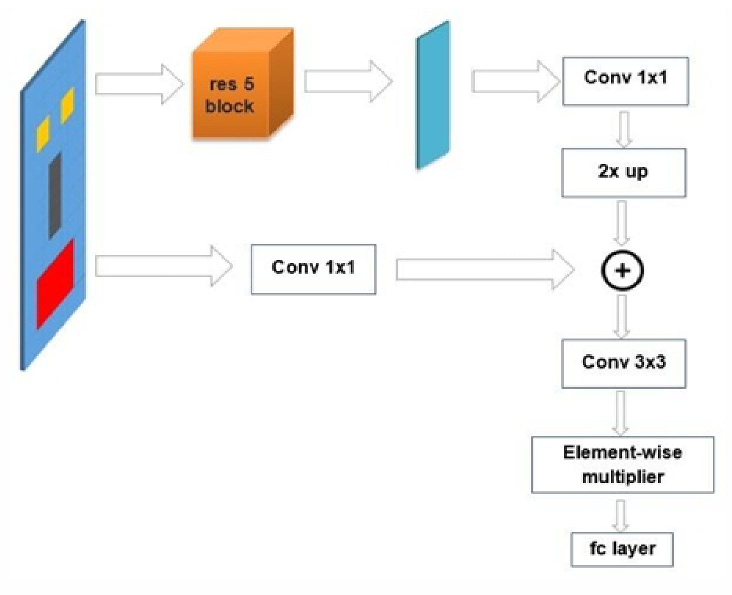


Figure 4. Details of the FPN module. Res5 block is applied to the feature maps in the top flow. These feature maps are then aggregated with the original feature maps.

Standard *L2* norm was used to regularize the network's parameters.

4.1. Evaluation Metrics

We extracted RoIs from the face images and trained our model to treat each of them as a separate bounding box (see Figure 2). In order to evaluate our method, we used the widely used accuracy measure. Because some AUs have low occurrence rates, using the accuracy measure is not enough. Therefore, we also used the F1-frame (F1-score) [33], which is commonly used in the literature and is defined using precision and recall as follows:

$$F1 = 2precision * recall/(precision + recall) \quad (2)$$

All experiments were conducted in a subject exclusive 3-fold cross-validation scheme and accuracy and F1-frame score for all the AUs were calculated. The reported values are the average results (denoted as Avg.) over all experiments.

4.4. Results

In the following, the results of the proposed method are compared with those of similar methods on the DISFA dataset. The results presented in Table 1 and Table 2 are obtained under 3-fold cross-validation setting. The reported values in the tables are F1-frame and accuracy, respectively. Traditional methods like linear support vector machine (LSVM) [34], active patch learning (APL) [35], deep region and multilabel learning (DRML) [36], ROI adaption net (ROI-Nets) [37], DSIN [25], AU R-CNN [18], and the recent successful methods like SRERL [26] and ARL [21] were compared with our method. It should be noted that our comparison includes only those methods that use static twodimensional image representation.

Table 2. Results of different methods in action unit recognition task on DISFA dataset. Reported numbers are accuracies, and bracketed and bold numbers represent the best and the second-best results, respectively.

AU	LSVM	APL	DRML	SRERL	ARL	Our
1	21.6	32.7	53.3	76.2	**92.1**	[94.7]
2	15.8	27.8	53.2	80.9	**92.7**	[93.5]
4	17.2	37.9	60.0	79.1	**88.5**	[88.8]
6	8.7	13.6	54.9	80.4	**91.6**	[91.9]
9	15.0	64.4	51.5	76.5	[95.9]	**95.7**
12	93.8	94.2	54.6	87.9	[93.9]	**92.1**
25	3.4	50.4	45.6	90.9	[97.3]	**93.1**
26	20.1	47.1	45.3	73.4	[94.3]	**91.0**
Avg	27.5	46.0	52.3	80.7	[93.3]	**92.6**

Table 1 shows the performance of different methods in terms of F1-score. Our proposed method outperforms other methods on average of F1-scores. Moreover, our method achieves the best classification results for AUs 1 and 4, and the second-best classification results for AUs 6, 25, and 26. It is observed from the reported results in Table 1 that the other methods also perform differently for different action units. For example, the results of the ARL [21] method are the best for AUs 22 and 26 and the second-best for AUs 2 and 12. The outcome of the same method for AUs 4, 6, and 9 is significantly low compared with that of the other methods. Our proposed method is consistently performing good for all AUs, and although not all of our results are the best, the performance of the method is comparable with that of the others. Generalization capability of our method can be the result of simultaneous utilization of local and global face features through the FPN and SPPNet modules. The proposed method outperforms all other methods for AUs 1, 2, 4, and 6. and is the second best for others in terms of classification accuracy.

As shown in Table 2, there is a significant gap between the two best algorithms (i.e. ARL and our proposed method) and the other methods. This shows the strength of the proposed architecture. On the other hand, our method and the AU RCNN method both use ResNet-101 as the backbone of the models. However, the number of learnable parameters of our method is much less than that of AU R-CNN. This is because we freeze the initial and middle layers and don't train them. Therefore, the proposed method is more efficient than AU R-CNN. Moreover, another advantage of our method is its simplicity compared to the other methods specifically ARL.

4. Conclusion

A common approach for detecting facial expressions is to recognize different facial action units and then use their combination to identify the facial expressions. To this end, some methods use the whole face as a single object to detect and classify action units, while the others detect each action unit separately. Despite the achievements, the latter approaches are prone to misclassification because they miss some useful information. For example, features from the upper face, such as those related to eye and eyebrow gestures, can enhance the detection performance of the AUs in the lower face (e.g., AU 25, AU 26). By using the whole face images, it is possible to capture each AU's global and occurrence-related features. The strengths of each of these two approaches inspired us to use the combination of them to complement each other and eliminate possible flaws. It is known that the constituent area (and scales) and the appearance of each AU may vary. Therefore, we used SPPNet [28] to generalize the proposed model to learn these variations. With an end-to-end trainable framework (SPP-FPNNet), we proposed to combine local features using the SPP module with global features using the FPN module to achieve an efficient approach because of less time has been spent to train the network. This is achieved by freezing the weights of the initial and middle layers of the network, and finetuning the last layers. Our proposed model outperforms the state-of-the-art methods for the well-known challenging DISFA dataset. By training the whole networks parameters, we expect to achieve better results. However, this imposes high computational cost.

As the future line of research, we would like to use temporal features as an integral element in detecting AUs. Moreover, using sequence modeling techniques, more specifically attention mechanism, can be a logical extension of the current work to be able to tackle the temporal dynamics in videos.

References

[1] P. Ekman and W. V. Friesen, Unmasking the face: A guide to recognizing emotions from facial clues. vol. 10, Ishk, 2003.

[2] S. Wang, H. Ding, and G. Peng, "Dual learning for facial action unit detection under non-full annotation." IEEE Transactions on Cybernetics, vol. 52, no. 4, pp. 2225-2237, April 2022.

[3] C. Fabian Benitez-Quiroz, R. Srinivasan, and A. M. Martinez, "Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 5562–5570.

[4] R. Ekman, What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press, USA,1997.

[5] K. Zhao, W.-S. Chu, and A. M. Martinez, "Learning facial action units from web images with scalable weakly supervised clustering," in Proceedings of the IEEE Conference on computer vision and pattern recognition, 2018, pp. 2090–2099.

[6] R. Zhi, M. Liu, and D. Zhang, "Facial representation for automatic facial action unit analysis system," in 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). IEEE, 2019, pp. 1368–1372.

[7] J. J. Bazzo and M. V. Lamar, "Recognizing facial actions using gabor wavelets with neutral face average difference," in Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings. IEEE, 2004, pp. 505–510.

[8] M. Valstar and M. Pantic, "Fully automatic facial action unit detection and temporal analysis," in 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'06). IEEE, 2006, pp. 149–149.

[9] J. Whitehill and C. W. Omlin, "Haar features for facs au recognition," in 7th International Conference on Automatic Face and Gesture Recognition (FGR06). IEEE, 2006, pp. 5–pp.

[10] D. G. Lowe, "Object recognition from local scaleinvariant features," in Proceedings of the seventh IEEE international conference on computer vision, Vol. 2. Ieee, 1999, pp. 1150–1157.

[11] B. Jiang, M. F. Valstar, and M. Pantic, "Action unit detection using sparse appearance descriptors in space-time video volumes," in Face and Gesture 2011. IEEE, 2011, pp. 314–321.

[12] J. J.-J. Lien, T. Kanade, J. F. Cohn, and C.-C. Li, "Detection, tracking, and classification of action units in facial expression," Robotics and Autonomous Systems, vol. 31, no. 3, pp. 131–146, 2000.

[13] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohnkanade dataset (ck+): A complete dataset for action unit and emotionspecified expression," in 2010 ieee computer society conference on computer vision and pattern recognition-workshops. IEEE, 2010, pp. 94101.

[14] M. F. Valstar and M. Pantic, "Fully automatic recognition of the temporal phases of facial actions," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, no. 1, pp. 28–43, 2011.

[15] T. Senechal, V. Rapp, H. Salam, R. Seguier, K. Bailly, and L. Prevost, "Facial action recognition combining heterogeneous features via multikernel learning," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, no. 4, pp. 993–1005, 2012.

[16] A. Rakotomamonjy, F. R. Bach, S. Canu, and Y. Grandvalet, "SimpleMKL," Journal of Machine Learning Research, vol. 9, no. 83, pp. 2491–2521, 2008.

[17] S. Eleftheriadis, O. Rudovic, and M. Pantic, "Joint facial action unit detection and feature fusion: A multiconditional learning approach," IEEE transactions on image processing, vol. 25, no. 12, pp. 5727–5742, 2016.

[18] C. Ma, L. Chen, and J. Yong, "AU R-CNN: Encoding expert prior knowledge into RCNN for action

unit detection," Neurocomputing, vol. 355, pp. 35–47, 2019.
[19] M. Savadi Hosseini and F. Ghaderi, "A hybrid deep learning architecture using 3D CNNs and GRUs for human action recognition," International Journal of Engineering, vol. 33, no. 5, pp. 959–965, 2020.
[20] Z. Shao, Z. Liu, J. Cai, and L. Ma, "Deep adaptive attention for joint facial action unit detection and face alignment," in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 705–720.
[21] Z. Shao, Z. Liu, J. Cai, Y. Wu, and L. Ma, "Facial action unit detection using attention and relation learning," IEEE Transactions on Affective Computing, vol. 13, 2019, pp. 1274–1289.
[22] Y. Chen, G. Song, Z. Shao, J. Cai, T.-J. Cham, and J. Zheng, "Geoconv: Geodesic guided convolution for facial action unit recognition," Pattern Recognition 122 (2022): 108355.
[23] M. Kurmanji and F. Ghaderi, "Hand gesture recognition from RGB-D data using 2D and 3D convolutional neural networks: a comparative study," Journal of AI and Data Mining, vol. 8, no. 2, pp. 177188, 2020.
[24] Z. Liu, J. Dong, C. Zhang, L. Wang, and J. Dang, "Relation modeling with graph convolutional networks for facial action unit detection," in International Conference on Multimedia Modeling. Springer, 2020, pp. 489–501.
[25] C. Corneanu, M. Madadi, and S. Escalera, "Deep structure inference network for facial action unit recognition," in Proceedings of the European Conference on Computer Vision (ECCV), vol. 11216, 2018, pp. 298–313.
[26] G. Li, X. Zhu, Y. Zeng, Q. Wang, and L. Lin, "Semantic relationships guided representation learning for facial action unit recognition," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 8594–8601.
[27] Y. Li, S. Shan, "Meta auxiliary learning for facial action unit detection" IEEE Transactions on Affective Computing, vol. 19, 2021, pp. 14–17.
[28] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," IEEE transactions on pattern analysis and machine intelligence, vol. 37, no. 9, pp. 1904–1916, 2015.
[29] T.-Y. Lin, P. Doll´ar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2117–2125.
[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[31] S. M. Mavadati, M. H. Mahoor, K. Bartlett, P. Trinh, and J. F. Cohn, "Disfa: A spontaneous facial action intensity database," IEEE Transactions on Affective Computing, vol. 4, no. 2, pp. 151–160, 2013.
[32] X. Zhang, L. Yin, J. F. Cohn, S. Canavan, M. Reale, A. Horowitz, P. Liu, and J. M. Girard, "Bp4d spontaneous: a high-resolution spontaneous 3d dynamic facial expression database," Image and Vision Computing, vol. 32, no. 10, pp. 692–706, 2014.

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).
2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);
3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);
4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.
The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

Note