Towards Perception Based Image Retrieval

B. Mohammed Imran and M.M. Sufyan Beg

Department of Computer Engineering
Jamia Millia Islamia (A Central University), New Delhi 110025, India
imran.fuz@gmail.com, mbeg@jmi.ac.in

Abstract. To deal with the rising need of perception based image retrieval, we present a novel approach of fuzzy image retrieval. Indeed, with an abrupt increase in crimes, the whole world is looking forward for an intelligent image retrieval system, which retrieves facial features of criminals as input queries. The present work is focused towards retrieving fuzzy images, with features in perceptions as inputs. We begin with retrieving fuzzy geometric shapes by describing its features in natural language propositions as query. Zadeh proposed computing with words (CW) which deals with perceptions, wherein the natural language is a major source of perceptions. That is, the perception based information is the inputs for image retrieval. We devise our image retrieval system with query processing, search module and Lucene. Moreover, the ranking of retrieving fuzzy objects is based on the highest value of query relevance vector. We found that the fuzzy geometric shapes are retrieved correctly based on the perception based query.

Keywords: Extended fuzzy logic, f-image retrieval, f-principle, perception based image retrieval, f-geometry.

1 Introduction

With a sharp increase in crime throughout the world, terrorism in particular is creating nuisance to the whole mankind by some identified miscreants. Such miscreants escape even from the surveillance camera located at the high security zones. However, there are many chances, in which the on-lookers provide some identification about the miscreants. Certainly, such information in natural language mostly in perception contains fuzziness in it. Undoubtedly, the perception based information on features play an important role in retrieving similar facial features from a criminal database. In fact, the criminal database consists of several types of eyes, nose, forehead, ears, lips, cheek, chin, eyebrows, hair style etc are stored along with the descriptions. Nonetheless, this might help a lot in making automated sketches of miscreants in a quicker way. Further, it is possible to retrieve the faces of criminals from the crime database in near future. It is with this idea that this approach of f-image retrieval is being proposed here. f-image retrieval is an image retrieval method in which the objects for image retrieval are perception based query in natural language. Since, the basis of image retrieval is from text retrieval, we implement some of the techniques of text retrieval in our work.

N. Meghanathan et al. (Eds.): CCSIT 2012, Part III, LNICST 86, pp. 280–289, 2012.

This paper is organized as follows, in section 1.1, we discuss about the related work. In section 2, we look into the concept of f-geometry, so that it becomes more noticeable about our work. In section 3, we present a detailed discussion on f-image retrieval system along with the process involved. Section 4, we present some experimental results of retrieved images in snapshots. Finally, we conclude in section 5, along with future directions.

1.1 Related Work

Several researchers and institutes have contributed their research on image retrieval [11]. With Google on top of searching and information retrieval system, but, lacks capabilities of reasoning, deduction and concept of relevance [3, 4]. In [14], work on image retrieval using power law transformation is performed. However, their work is an application of content based image retrieval. In [15], presents an approach of recognizing faces using parallel neural network, in addition to fuzzy clustering and parallel neural networks. In [16], presents their work on information extraction for face recognition using fuzzy logic. In [17], discusses their efforts in a new approach for object tracking system on images in motion [18]. To the best of our knowledge, we have not found any work like f-image retrieval using natural language yet.

2 *f*-Geometry

In the world of Euclidean geometry (Weg), the instruments for drawing are precisiated pen, rulers and compass. This can be called as the crisp geometry C. In the world of fuzzy geometry (Wfg), the instruments used for drawing are only the unprecisiated spray pen. This can be called as fuzzy geometry f-C or fuzzy version of crisp geometry. Ultimately, the geometric objects drawn in Wfg are irregular in appearance. The formulation of f-definitions for fuzzy geometric shapes is in the seminal paper of f-geometry [1]. Additionally, the extent of validity to a crisp geometric shape called as f-validity is estimated for a set of basic geometric objects in another paper [2]. However, a brief discussion on f-geometry will make a better understanding of our work. The detailed information can be found in [1, 2]. Zadeh's f-geometry belongs to unprecisiated fuzzy logic [6, 13], which is different in both spirit and substance from the earlier fuzzy geometries belonging to precisiated fuzzy logic [1, 2, 5, 10].

3 *f* - Image Retrieval System

An f-image retrieval system retrieves the f-geometric objects from the image database on querying the features of f-geometric objects. However, the images in the database are stored if it has membership value at least 0.1 and image description. The architecture of f-image retrieval shows the modules and the process that takes place in Figure1.

We have shown the architecture of an f-Image retrieval system with the following modules such as: Query processing, Search module, Creation of relevance matrix and Image Database.

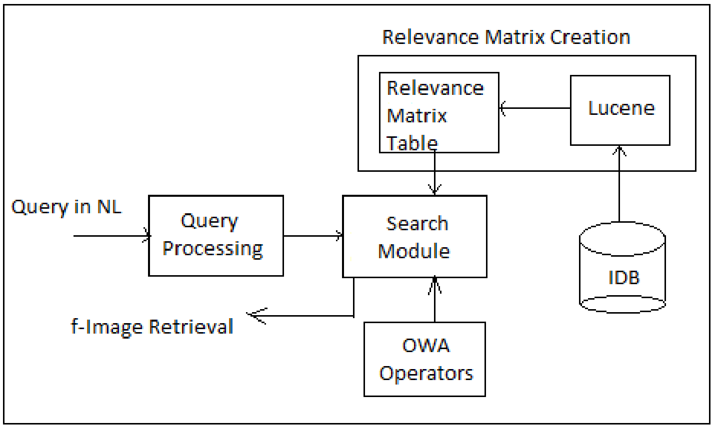

Fig. 1. The architecture of an *f*-image retrieval

3.1 Query Processing

Query Processing (QP) is the first module of our proposed f- image retrieval system. First of all, we begin with QP, which processes the input query by (i) stop word removal and (ii) stemming. The stop word removal eliminates quite often repeated words which may have a very little meaning to play a vital role in the query. Usually, they are articles and prepositions, *like a, an, the, for* etc. But, it has no effect on the semantics of the word. Secondly, stemming is the other process carried out in QP. Ultimately, stemmer increases the performance of searching. Stemmer is required, because, the queries arriving to an image retrieval system is usually not in an artificial language, but, mostly in natural language propositions. So, the stemmer converts a word to its canonical form [19]. Although, it has no major effect on the semantics of the word, like cats, catty, cat like words usually refers to a pet animal cat. At this juncture, we look forward to retrieve fuzzy geometric objects from an image database.

3.2 Search Module

Search module is the combination of several individual units, which includes relevance matrix and Ordered Weighted Averaging operators. The relevance matrix is produced using Lucene, which will be discussed in section 3.3; we discuss the operation of relevance matrix and OWA operators in this section.

3.2.1 Relevance Matrix

The relevance matrix R, created by Lucene, is a two dimensional matrix containing the tf-idf (Term Frequency- Inverse Document Frequency) values for a term in the column and image id in the row. These terms are from the image descriptors belonging to an image. Let R_{ik} be a variable that takes the values between [0, 1] indicates the relevance of an image *i* for the associated keyword *k* and vice versa. The search module we use interacts with an image database that contains f-geometric shapes along with the descriptions of their shapes. The relevance matrix contains the composition values of tf-idf in it. This concept of relevance matrix is very well envisaged by an Example [8].

3.2.2 Term Frequency- Inverse Document Frequency

Term Frequency- Inverse Document Frequency (TF-IDF) is a statistical evaluation to compute the importance of a term in an image description, in a collection of images in a database. Generally, this is used in text mining and information retrieval. Since, we are concerned with image retrieval, so, we have stored a collection of fuzzy valid geometric shapes in the form of images along with their descriptions. The weight of each term in an image increases relatively to the number of occurrences in an image. TF-IDF is higher for the rarely used terms and lower for the more occurrences of the terms in the image description. The various TF-IDF techniques are often used in searching as an important tool for indexing as well as to compute the relevance of a term in a given query.

Term frequency (TF) of a term is computed as the frequency of a term in a given image description. The frequency of a word in a given document is taken as a good measure of importance of that word in the given image description.

$$f_{TF} = \frac{f_w}{f_{w\ \max}}$$

Inverse Document Frequency (IDF) is a weighting scheme that makes rare words more important than common words. In other words, it points out the discriminatory power of a given term and is based on the intuition that a term's rareness across the collection is a measure of importance.

$$f_{IDF} = \log \frac{\rho}{\rho_w}$$

where f_w and $f_{w\ max}$ refers to the frequency of keyword *w* word in the description of image and maximum frequency of any word in the database respectively. ρ_w refers to the number keywords occurs in an image description and ρ is the total number of keywords in image description in the group of images.

3.2.3 Query Formulation Using OWA Operators

The query formulation is needed to resolve the complications that occur in the classification of images. Because, there might be cases where the user uses a set of keywords in query for image retrieval, but, unaware about the preciseness of using AND operator or OR operator to combine these keywords. Use of AND operator may require all the keywords to be present in the image description, therefore, returning a few images. Whereas, usage of OR operator may require any of the keywords present in the image description, returning too many images. In such situations, it is desirable to resolve these complications using OWA operator, described in the next section.

3.2.3.1 Ordered Weighted Averaging (OWA) operators. The OWA operator is the central concept of information aggregation. It was originally introduced by Yager [9]. It facilitates the means of aggregation in solving problems associated with multi criteria decision making. The discussion on calculation of weight is clearly mentioned in [12].

3.2.3.2 OWA operator based queries. We implement linguistic operator so that the user has an option of choosing "most", "at least half", and "as many as possible". In addition, we have provided two more options "AND" and "OR" in addition to the three linguistic operator, along with the user query. For example, if the user selects "most" then, most of the keywords associated with the image description in the database will be retrieved. Let us see common query Q be given as:

$$Q=\text{"linguistic operator"}\ (t_{j1}, t_{j2}, t_{j3,\ldots\ldots,}t_{jn})$$

where the linguistic operator "most", "at least half" and "as many as possible" take the parameters *a* and *b* as (0.3, 0.8), (0, 0.5) and (0.5, 1) respectively.

The n_{th} element of the Query Relevance (QR) vector will be found as:

$$QR_k = \sum_{j=1}^{n} w_i . x_i$$

where x_i is the ith largest in the set of elements of ($R_{K,j1}$, $R_{K,j2}$,$R_{K,j3....}R_{K,jn}$) and w_i are the weights of the OWA operation. However, the ranking of images is based on the highest value of query relevance vector. So, the next rank is for the second highest value QR vector.

3.3 Creation of Relevance Matrix

The relevance matrix is created by Lucene, a text search engine library discussed in this sub-section.

3.3.1 Lucene

Lucene provides search capabilities to various text based information retrieval systems. It indexes and makes searchable any kind of data which is present in a textual format. Moreover, Lucene does not worry about the source, the format and the language, when the data is in text. Nevertheless, Lucene facilitates indexing and searching of data stored in text files, word document files, PDF files, other textual format, web pages, and remote web servers etc. This facilitates full-text search potential, which is normally not available in many databases. The same is performed in our work, wherein the indexing and the searching of image descriptions in the database is carried out. However, the creation of indices and searching is explained in detail in [7].

The process of relevance matrix creation is performed with the help of Lucene. It fetches the image description terms from the image database and indexes them. The Relevance matrix contains the TF-IDF values for each term in each image descriptor as discussed in previous section. The detail discussion of creation of relevance matrix is given below in steps.

3.3.2 Creation of Lucene Index

Step 1: A Lucene index is created which contains details about the terms present in the image descriptors. Later, this index is used to construct the actual Relevant Matrix.

Step 2: Since the no terms can be very large hence a single relevance matrix table is divided into separate relevance matrix viz, t_a, t_b, t_c, t_d,....t_z. Here each table contains tf-idf values of terms beginning with the alphabet that names the table.

3.4 Image Database

We have made a collection of fuzzy geometric objects stored in the database. Each image has undergone fuzzy validity check, wherein each f-geometric object is estimated using f-algorithm. Additionally, some f-geometric objects with multiple attributes like triangle, rectangle and square require f-theorem too. Therefore, we have incorporated two criteria for storing f-images. Firstly, an f-geometric object should have membership value at least 0.1. Secondly, the possible ways in which an f-geometric object can be described in natural language are stored for every image in the image database. Certainly, the terms used in describing the images are perception based, but, not measurement based.

4 Experimental Results

The significance of our work can well be visualized by the inputs and their corresponding outputs. However, we have presented the outputs in terms of snapshots. Moreover, we show some of the process that has taken place in individual modules as snapshots. For example, the relevance matrix table containing the tf-idf values, the image database with fuzzy geometric shapes and their possible descriptions and the query relevance vector.

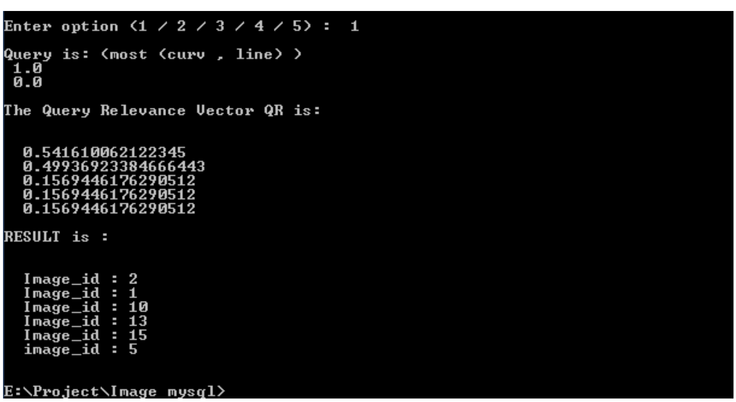

Fig. 2. Snapshot for the linguistic query 'the curved line'.

Figure 2 shows the query relevance vector of the terms used in the input query. For example, we have used a very simple query ‘the curved line’ manifests the application of stemming and stop word removal in the second line. It means that only the term ‘curv’ and ‘line’ are taken as search contents. Additionally, the OWA operator ‘most’ is applied here. It refers to retrieve the corresponding image, whose description contains the most number of query terms in image description. Moreover, the query relevance vector is the sum of the products of tf-idf and the weights. By this value, the image associated with image id: 2 is retrieved, which is shown in the output snapshots. Nevertheless, the query relevance vector generates its value in a ranking order. The highest value among the ranked order is the optimal output among the relevant images that is found to be as 0.5416. Nevertheless, the ranking of several images could be possible whenever there is a very large corpus of images in our database.

mysql> select image_id,name,Image_desc from image;

image_id	name	Image_desc
1	Line	exact line, straight line, line, perfect line, line like object, straight object
2	Curved Line	curving line, curved object, concave line, convex line, sinewave like object
3	Zig Zag line	small corner object
4	Circle	circle, circular,round,moon like structure,ball
5	Inexact circle	slightly egg shaped circle, oval shaped circle, elliptical, ellipse
6	Inexact circle	circle with slight outer projection.
7	Triangle	Tricorner shaped, structure with three corners, triangle, eqilateral triangle, good triangle, absolute triangle
8	Improper Triangle	like a triangle, almost a triangle
9	Convex Triangle	Convex triangle ,tri-sided object
10	Concave Triangle	Concave triangle, triangles with bulgy lines,tri-sided
11	Square	Four edged equi-sided four corner, exact square, four sided, right angle
12	Rectangle	Four edged inequal sided proper corner, rectangle, four sided, right angle
13	Inexact square	improper four corner, object with four curvy lines,
14	Right angle triangle	three sided with right angle, right angle triangle
15	Inexact rectangle	improper four corner, object with four curvy lines,

15 rows in set (0.00 sec)

Fig. 3. Snapshot of the image and their descriptors in the database

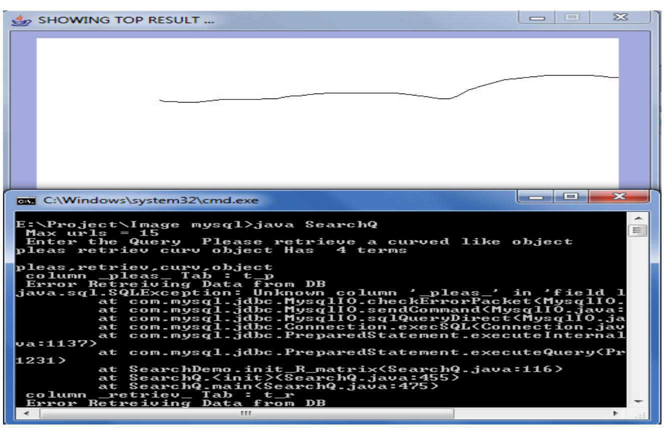

Fig. 4. Snapshot of the output for the query ‘please retrieve a curved like object’

In Figure 3, the image description terms associated with image_id is shown. However, for the sake of identity, we have mentioned the image_id as approximate name of the image in the database as well. For example, we have described the features of an inexact rectangle, which can be formed using four fuzzy lines and four fuzzy right angles. So, we have mentioned four improper corner and four curvy lines. But, there are no such standard terms available for the fuzzy geometric shapes.

Figure 4 shows the output of our f-image retrieval system in an applet. But, we provide the input query in terms of perceptions in the command prompt. In Figure 4, we have entered the input query in natural language containing perceptions, but, not the exact term. For example, we enter the inputs as 'please retrieve a curved like line' which provides us a curvy like line as output. Although, the process of individual modules explained in the architecture is performed for all the types of queries.

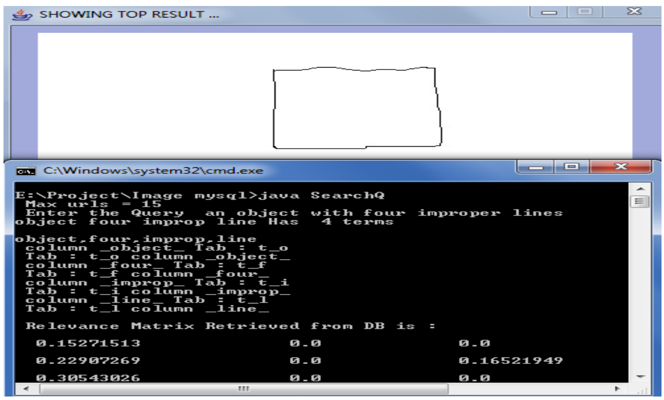

Fig. 5. Snapshot of the output for the query 'an object with four improper lines'

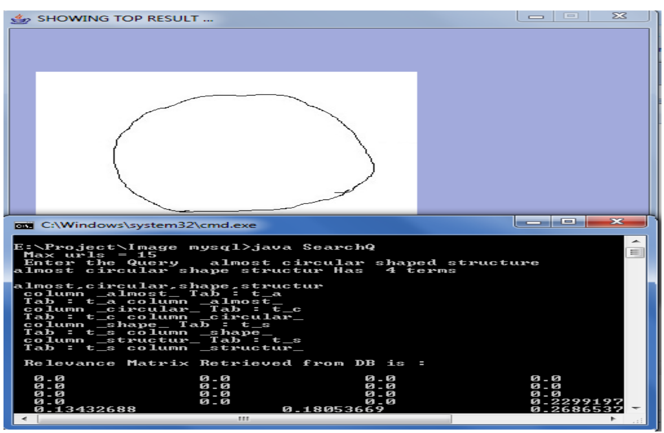

Fig. 6. Snapshot of the output for the query 'almost circular shaped structure'

In Figure 5, the output from f-image retrieval system for the natural language query ‘an object with four improper lines’ retrieves an improper square like object. For such a query, the stop word removal removes ’an’ and ‘with’, making only four terms as significant terms for searching.

In Figure 6, the output of a perception based query is shown. The query has a vague term ‘almost circular’ is searched by our search engine, later on application of OWA operators on the terms in the documents fetches an appropriate image. The image retrieved is found to be an improper circle. Similarly, the same process is followed for all the types of queries in this approach. In Figure 7, we have posed a query asking for ‘an object which is likely to be a triangle’ was our query. This query has much of fuzziness in it, thereby, considering the fuzziness in the term as well as the triangle like structure is retrieved.

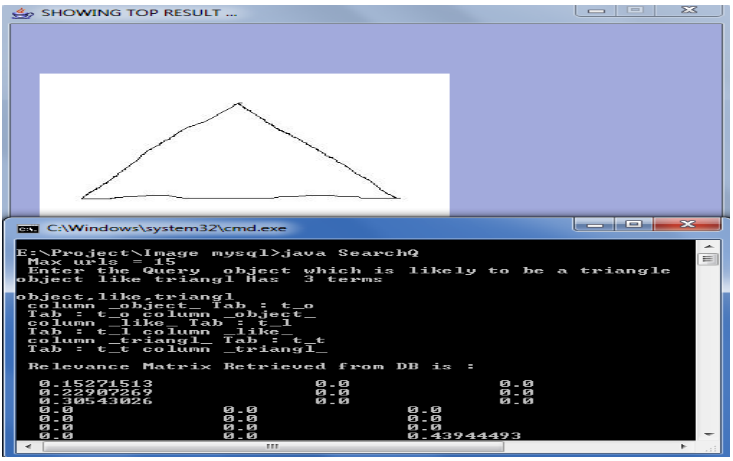

Fig. 7. Snapshot of the output for the query ‘object which is likely to be a triangle’

Nevertheless, if the same queries are posed on to image retrieval system like Google, the images retrieved are found to be inappropriate. Moreover, the extension of our work in closed and open domain system will enhance image retrieval for imprecise, vague and fuzzy queries.

5 Conclusion

In this paper, we have presented a novel approach of f-image retrieval using linguistic queries. Since, bivalent logic cannot support the natural language queries in image retrieval systems. So, our work is a model, which reveals the concept of computing with words. Conversely, the approach we have implemented can be extended in retrieving inexact shapes with queries in natural language. Especially, the facial features of different criminals can be retrieved. However, a large database consists of several facial features need to be described in possible ways in natural language. Further, semantic enhancement can be added in query processing and in the search module. In general, our work has wide applications in intelligent image retrieval.

References

1. Imran, B.M., Beg, M.M.S.: Elements of Sketching with Words. In: Proc. of IEEE International Conference on Granular Computing (GrC 2010), San Jose, California, USA, August 14-16, pp. 241–246 (2010)
2. Imran, B.M., Beg, M.M.S.: Towards Computational forensics with f-geometry. In: World Conference on Soft Computing 2011. San Francisco State University, California (2011)
3. Imran, B.M., Beg, M.M.S.: Image Retrieval by Mechanization and f-Principle. In: Proc. Second International Conference on Data Management (ICDM 2009), IMT Ghaziabad, India, February 10-11, pp. 157–165 (2009)
4. Imran, B.M., Beg, M.M.S.: Automation of PNL-based f – Image Retrieval System. In: Proc. of International Conference on Computational Vision and Robotics (ICCVR 2010), Bhubaneswar, India, pp. 291–298 (2010)
5. Imran, B.M., Beg, M.M.S.: Elements of Sketching with Words (An Extended paper). International Journal of Granular Computing, Rough Sets and Intelligent Systems (IJGCRSIS) 2(2), 166–178 (2011)
6. Zadeh, L.A.: Toward Extended Fuzzy Logic - A First Step, Fuzzy Sets and Systems. Journal of Information Sciences, 3175–3181 (2009)
7. Gospodnetic, O., Hatcher, E.: Lucene in Action. Manning (2004)
8. Beg, M.M.S.: User Feedback Based Enhancement in Web Search Quality. International Journal of Information Sciences 170(2-4), 153–172 (2005)
9. Yager, R.R.: On OWA aggregation operators in multi-criteria decision making. IEEE Transactions on Systems, Man, and Cybernetics 18, 183–190 (1988)
10. Imran, B.M., Beg, M.M.S.: Fuzzy Identification of Geometric Shapes. In: Meghanathan, N., et al. (eds.) CCSIT 2012, Part III. LNICST, vol. 86, pp. 269–279. Springer, Heidelberg (2012)
11. Salton, G., Buckley, C.: Term weighting approaches in automatic text retrieval. Information Processing and Management 24(5), 513–523 (1988)
12. Beg, M.M.S., Ahmad, N.: Subjective Enhancement and Measurement of Web Search Quality. In: Nikravesh, M., Azvine, B., Yager, R., Zadeh, L.A. (eds.) Enhancing the Power of the Internet, pp. 95–130. Springer, Heidelberg (2004)
13. Zadeh, L.A.: From Fuzzy Logic to Extended Fuzzy Logic—The Concept of f-Validity and the Impossibility Principle, Plenary session. In: FUZZ-IEEE 2007, Imperial College, London, UK (July 24, 2007)
14. Han, Q., Song, C., Zhang, M., Niu, X.: Proc. International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IIHMSP 2008, Harbin, China, pp. 87–90 (August 2008)
15. Lu, J.: A Method of Face Recognition Based on Fuzzy c-Means Clustering and Associated Sub-NNs. IEEE Trans. on Neural Networks 18(1), 150–160 (2007)
16. Vishwakarma, P., Pandey, S.: Fuzzy based Pixel wise Information Extraction for Face Recognition. IACSIT International Journal of Engineering & Technology 2(1), 117–123 (2010)
17. Garcia, J., Monila, J.M., Besada, J.A., Portillo, J.I., Casar, J.R.: Robust Object Tracking with Fuzzy Shape Estimation. In: Proc. of the Fifth International Conference on Information Fusion 2002, Washington, USA, pp. 67–71 (2002)
18. Qi, W., Li, X., Tang, L., Xu, Z.: Shape Recognition by Fuzzy Distance Measure. In: Proc. International Conference Intelligent Computing and Intelligent Systems (ICIS 2010), Xiamen, China, pp. 737–741 (2010)
19. Lovins, J.B.: Development of a stemming algorithm. Mechanical Translation and Computational Linguistics 11, 22–31 (1968)