

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

## Feature Extraction Methods: A Review

P.Prathusha<sup>1</sup>, S.Jyothi<sup>2</sup>

Asst. Professor, Department of Computer Science Engineering, RGM CET, Nandyal, AP, India<sup>1</sup>

Professor, Department of Computer Science, SPMVV, Tirupathi, India<sup>2</sup>

**ABSTRACT:** Extracting features from input data is vital for successful classification task and machine learning tasks. Feature extraction, obviously, is a transformation of large input data into a low dimensional feature vector, which is an input to classification or a machine learning algorithm. The task of feature extraction has major challenges discussed in this paper. The quest for universal learning is one of the most significant, fascinating, and revolutionary intellectual developments of all time. In present day scenarios the problems are very complex, have high dimensional data and are heterogeneous. The challenge is to learn and extract knowledge from that data to make correct decisions. The objective of this paper is to give an overview of methods used in feature extraction for various applications, give an overview of modern challenges in feature extraction.

**KEYWORDS:** colour features, texture features, shape features, fuzzy features, Large scale machine learning, filter methods, wrapper methods, embedded methods, feature selection, feature subset selection, dimensionality reduction, fuzzy feature extraction

### I. INTRODUCTION

The huge amount of data generated has grown multiple times with millions or trillions of data sets generated every day with internet and social media. Feature extraction is a critical task which involves huge amount of data as input and transforming it into optimal feature set. Feature extraction starts from an initial set of measured data and builds derived values (features) intended to be informative and non-redundant, facilitating the subsequent learning and generalization steps, and in some cases leading to better human interpretations.[118].

#### 1.1 Terminology and principal representations

Let  $X$  denote the input feature space and  $Y$  denotes the set of output labels. The sample is represented as

$$S = \{(x_1, y_1), \dots, (x_m, y_m)\}$$

Let  $D$  is another feature space and Let Function  $F$  maps  $X$  to  $D$ . Let  $h \in D$  is applied to every  $x_i \in X$  the feature extracted sample is  $h(S) = \{(f(x_1), y_1), \dots, f(x_m, y_m)\}$

#### Feature Engineering

Feature Engineering comprises of human intervention and human logic in extracting features for any application.

#### Feature re-weighting

A method of assigning best weight for each feature is known as Feature Re-weighting. If weights are  $\{0,1\}$  it is variable selection..

#### Feature extraction in software

Common numerical programming environments such as Matlab, SciLab, NumPy, and R Language provide some of the simpler feature extraction techniques. More specific algorithms are often available as publicly available scripts or third-party add-ons.[118]

#### 1.2 Regional features versus global features

Most of the existing algorithms have two approaches, either regional features [67],[68],[69],[70] or global features [63],[65],[66] exclusively. For global approach [63],[65],[66], a global feature vector is extracted from an image, such

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

as color histogram, color correlogram, edge direction histogram, and so on. The global features are advantageous in classifying simple scene categories such as "sunset," "mountain," "building," etc. For regional features [67],[68],[69],[70], an image is segmented into several regions and represented as a set of visual feature vectors, each of which represents one homogenous region.[71] gives improved feature extraction in image annotation by integrating global, regional and contextual features .[72] also addresses the combination of local and global features for object recognition

## II. BASIC FEATURE EXTRACTION TECHNIQUES

Feature Extraction basically uses texture, color and shape feature extraction methods .

### 2.1 Texture Feature Extraction

Human visual systems uses texture for Image recognition and interpretation. Textures are complex visual patterns composed of entities, or sub patterns that have characteristic brightness, color, slope, size, etc. Texture can be regarded as a similarity grouping in an image [126] .

### 2.2 Spatial and Spectral Features :

Depending on the domain from which the texture feature is extracted, textures are classified into **spatial texture feature extraction** methods and **spectral texture feature extraction methods** [14]. Spatial texture features are extracted by computing the pixel statistics or finding the local pixel structures in original image domain, whereas spectral transforms an image into frequency domain and then calculates feature from the transformed image

**Table.1 Spatial versus Spectral features**

Texture Method	Advantages	Disadvantages
Spatial Texture feature extraction	Interpretation is easy. No loss of information	Noise Sensitivity
Spectral Texture feature Extraction	Needs less computation	Needs square image regions of sufficient size Tranformation is a must

### 2.3 Texture Analysis

Texture analysis is characterization of regions of an image by their texture content. Texture analysis attempts to quantify intuitive qualities described by terms such as rough, smooth, silky, or bumpy as a function of the spatial variation in pixel intensities.[121] For example, smooth texture areas, the range of values in the neighbourhood around a pixel will be a small value; in areas around rough texture, the range will be larger. Standard deviation of pixels in a neighbourhood gives the degree of variability of pixel values in that region. Texture analysis is used in remote sensing, automated inspection, medical image processing etc . Texture analysis is used to find the texture boundaries, called **texture segmentation**

Approaches to texture analysis are usually categorized as

- Structural,
- Statistical,
- Model-based and
- Transform

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

**Structural** approaches [128] [127] represent texture by well-defined primitives (*micro texture*) and a hierarchy of spatial arrangements (*macro texture*) of those primitives.[123]

**Statistical** approaches [141] do not attempt to understand explicitly the hierarchical structure of the texture. Instead, they represent the texture indirectly by the non-deterministic properties that govern the distributions and relationships between the grey levels of an image [123]

**Model based** texture analysis [129][130][131] [132][133][134] using fractal and stochastic models, attempt to interpret an image texture by use of, respectively, generative image model and stochastic model.

**Transform methods** of texture analysis, such as Fourier [138] Gabor [139][140] and wavelet transforms [137] [136][135] represent an image in a space whose co-ordinate system has an interpretation that is closely related to the characteristics of a texture (such as frequency or size).

## 2.4 LDP, DLEP, MDLEP, CMDLEP

### LDP - Local directional Pattern

Local Directional Pattern (LDP) [152] uses the edge response of neighbouring pixel to encode the texture of an Image. A binary value 1 or 0 is assigned depending on the presence of the edge

$$LDP_n = \sum_{i=1}^8 f_i(m_i - m_k) * 2^i, \quad f_i(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1)$$

### DLEP - Directional Local Extrema Patterns

In this method [153] two neighbouring pixel intensities in a direction are compared with the pixel centre position to assign either 1 or 0. It describes the spatial structure of the local texture based on the local extrema of centre pixel  $x_c$ .

$$I'(x_i) = I(x_c) - I(x_i); \quad i = 1, 2, \dots, 8 \quad (2)$$

### MDLEP - Magnitude Directional Local Extrema Pattern

In [154] they proposed MDLEP to increase the performance by considering the magnitudes of local patterns

### CMLDEP- Combined Magnitude Directional Local Extrema Patterns

This method is proposed in [155] which outperforms MDLEP both in terms of precision and recall. The combination of Gabor and MDLEP explores more information present in the image compared to LBP and other related local pattern operators.

### Local Binary Pattern and its variations

LBP was introduced by author Ojala [145]. LBP considers the value of central pixel as threshold and the difference between center pixel and its surrounding neighbours is taken into consideration to assign 0 or 1.

$$LBP_{x,y} = \sum_{p=0}^{p-1} y(x_p - x_c) 2^p, \quad y(c) = \begin{cases} 1 & c \geq 0 \\ 0 & c < 0 \end{cases} \quad (3)$$

Where  $x_c$  represents the gray value of central pixel and  $x_p$  represents the grey value

Analysis of LBP operator and its extension versions was suggested by K. Subbareddy [146]. This paper explains about the improvement of LBP operator for following enhancements

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

- 1) Selection of its neighbourhood
- 2) Enhancement of its robustness
- 3) Improvement of its discriminative capability
- 4) Extension to three dimensional data
- 5) Combination with different other approaches

**Improved LBP (ILBP)** improves the discriminative ability of LBP ;Improved LBP (ILBP) which compares all the pixels (including the central pixel) with the mean intensity of all the pixels in the patch, The authors Tan and Triggs [151] extended the original LBP to a version with 3-value codes, called **Local Ternary Patterns (LTP)**.LTP improves the robustness of LBP on noisy images, **Gabor wavelet based LBP(GWB-based LBP)** and **Center-Symmetric LBP CS-LBP** combine other methods with LBP to bring in new advantages. **3DLBP** extends the LBP operator to describe 3D volume data. Wenchao Zhang proposes a novel face representation[147], **Local Gabor Binary Pattern Histogram Sequence(LGBPHS)**, which is insensitive to illumination variations due to lighting, expression, and aging. A **Center-Symmetric LBP (CS-LBP)** [148], was proposed by only comparing pairs of neighbouring pixels which are in the same diameter of the circle. The authors zaho and Pietikanein first, modelled [149] the textures with **volume local binary patterns (VLBP)**, which are an extension of the local binary patterns (LBP) operator widely used in ordinary texture analysis [150],

### III. PERCEPTUAL TEXTURAL FEATURES

Perceptual texture features was suggested by authors Abbadeni, Noureddine [124][125].

**GTDM: Grey-Tone Difference Matrix**

GTDM is also related to human perception of textures. Grey-Tone Difference Matrix (GTDM) was suggested by Amadasum . It defines texture measures correlated with human perception of textures.

The perceptual textural features are Coarseness, Directionality, busyness and contrast described briefly below.

#### 3.1 Coarseness

Coarseness measures the size of the primitives that constitute the texture. A coarse texture is composed of large primitives and is characterized by a high degree of local uniformity of grey-levels. A fine texture is constituted by small primitives and is characterized by a high degree of local variations of grey-levels.

$$Cs = \frac{1}{\frac{1}{2} \times (\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} Max_x(i,j)/n + \sum_{j=0}^{m-1} \sum_{i=0}^{n-1} \frac{Max_y(i,j)}{m}} \quad (4)$$

#### 3.2 Directionality

Directionality is a global property in an image. It measures the degree of visible dominant orientation in an image.

#### 3.3 Contrast

Contrast measures the degree of clarity with which one can distinguish between different primitives in a texture. A well contrasted image is an image in which primitives are clearly visible and separable.

#### 3.4 Busyness

Busyness refers to the intensity changes from a pixel to its neighbourhood: a busy texture is a texture in which the intensity changes are quick and rush; a non busy texture is a texture in which the intensity changes are slow and gradual.

#### 3.5 GLCM texture features

A GLCM is a histogram of co-occurring grey scale values at a given offset over an image.[122] uses glcm method to extract texture features from crop images

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

## 3.6 Transformations:

Some transformations are given below

### 3.6.1 Fourier Transform

Fourier transform performance is poor in practice, due to its lack of spatial localization.

### 3.6.2 Hough transformation

The Idea of the hough transformation is it converts image space into a parameter space. This conversion is a mapping that analyses the image looking for a special sort of parameterized features usually lines or circles. It's beauty lies in the many different shapes that can be detected together with their parameters.[11]

### 3.6.3 Wavelet transforms advantages

Wavelets best suited for texture analysis in a specific application. Wavelets are used by [135][136][137] . Arivazhagan and Ganesan proposed an approach to classify the texture that uses wavelet transform[143].

The advantages offered by wavelets are

1. to represent textures at the most suitable scale,
2. wide range of choices exists for the wavelet function

Disadvantages:

The problem with wavelet transform is that it is not translation-invariant.

### 3.6.4 Gabor filters

Gabor filters [142][144] provide means for better spatial localization; they are limited in practice because there is usually no single filter resolution at which one can localise a spatial structure in natural textures. Gabor wavelets are the mathematical model of visual cortical cells of mammalian brain and using this, an image can be decomposed into multiple scales and multiple orientations. [120]

## IV. SHAPE FEATURE EXTRACTION TECHNIQUES

There are 3 main classification methods . [14]

**4.1 Contour-based methods and region-based methods** [15]. It is proposed by MPEG-7. It is based on the use of shape boundary points rather than shape interior points.

structural approaches and global approaches are further classifications under each category.

Sub-class is based on whether the shape is represented as a whole or represented by segments/sections (primitives).



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

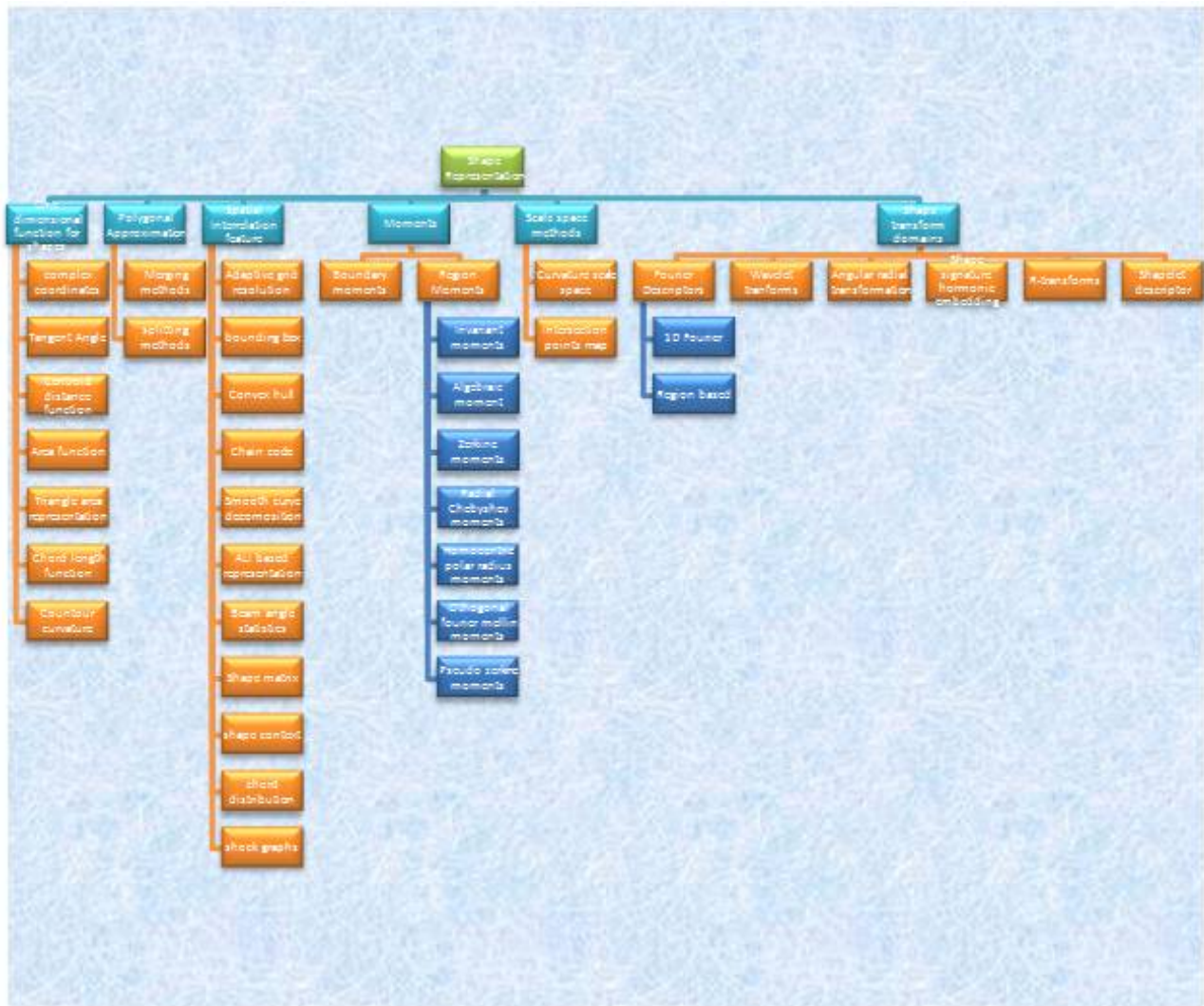


Fig.1 An overview of shape description techniques

**4.2 Space domain and transform domain** Methods in space domain match shapes on point (or point feature) basis, while feature domain techniques match shapes on feature (vector) basis[16].

**4.3 Information preserving (IP) and non-information preserving (NIP).** IP methods allow an accurate reconstruction of a shape from its descriptor, while NIP methods are only capable of partial ambiguous reconstruction. For object recognition purpose, IP is not a requirement.

#### 4.4 Zerkine Moments

[12] [13] Simple moments are not orthogonal, thus it is difficult to reconstruct the image from them. To overcome this drawback and to get the distribution of gray levels in the color image combination of Zernike and Hu moments is used. Zernike moments are used for less information redundancy. It allow a better description of objects than simple moments. These moments are joined together with Geodesic descriptors to get boundary of the object. .

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

## 4.5 Geodesic Descriptors

Geodesic descriptors are invariant under geodesic isometrics, and quasi-invariant to shape articulations and bendings. This is relevant to perform robust retrieval on articulated shapes, such as animal or human with varying poses. The Geodesic descriptors are calculated after detection of the center of the Object [12][13].

## V. COLOR FEATURES

Color is the vital discriminative feature for images. The color spaces are RGB, LUV, HSV and HMMD [112]. Color feature can be extracted from images or regions using color space. Many important color features have been proposed in the literatures, including color histogram [113][83], color moments(CM) [114], color coherence vector (CCV) [115] and color correlogram [89][116].[77] –[98] [117] are based on color features.

The common moments are mean, standard deviation and skewness, can be defined as follows:

$$\mu_i = \frac{1}{N} \sum_{j=1}^N f_{ij} \quad (5)$$

$$\sigma_i = \left\{ \frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2 \right\}^{1/2} \quad (6)$$

$$\gamma_i = \left\{ \frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3 \right\}^{1/3} \quad (7)$$

Where  $f_{ij}$  is the color value of the  $i$ th color component of the  $j$ th image pixel and  $N$  is the total number of pixels in the image.

$\mu_i, \sigma_i, \gamma_i$  ( $i=1,2,3$ ) denote the mean, standard deviation and skewness of each channel of an image respectively

## VI. FEATURE EXTRACTION VERSUS FEATURE SELECTION

Feature extraction methods apply a transformation on the original feature vector to reduce its dimension from  $n$  to  $m$  where  $m < n$  where as Feature selection methods select a small subset of original features. Feature *Extraction* methods are transformative apply a transformation to data to project it into a new feature space with lower dimension. Examples are PCA, and SVD where as Feature Selection methods choose features from the original set based on some criteria to filter out unimportant or redundant features. Feature selection: reduce dimensionality by selecting subset of original variables. Example: forward or backward feature selection. Feature selection on micro array data is experimented by [179].

### 6.1 Feature Selection Algorithms

**Feature selection process :** It consists of four basic steps (shown in Fig. 1), namely, subset generation, subset evaluation, stopping criterion, and result validation [ 32].

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

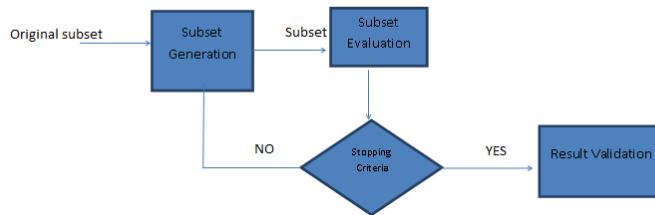


Figure.2. Four steps of feature selection

When the input data to an algorithm is too large to be processed and it is suspected to be redundant (e.g. the same measurement in both feet and meters, or the repetitiveness of images presented as pixels), then it can be transformed into a reduced set of features (also named a feature vector). This process is called feature selection. If the input data of an algorithm is large for processing and it is redundant then it is transformed into a reduced. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.[118].

### 6.1.1 Filter methods

A generalized algorithm for filter methods [32] is modelled in figure 2. For a data set X, the algorithm starts the search from a given subset  $Subset_N$  (an empty set, a full set, or any randomly selected subset) and searches through the feature space by a particular search strategy. Each generated subset S is evaluated by an independent measure M and compared with the previous best one. If it is found to be better, it is regarded as the current best subset. The search iterates until a predefined stopping  $Criterion_{Stop_{cr}}$  is reached. The algorithm outputs the last current best subset  $Y_{opt}$  as the final result.[35][36][37].

<b>FILTER ALGORITHM</b>	
<b>Input:-</b>	
X (F <sub>0</sub> ,F <sub>1</sub> ,.....,F <sub>n-1</sub> )	//a training data with n features
Subset <sub>N</sub>	//a subset from where to start the search
Stop <sub>cr</sub>	//a stopping Criterion
<b>Output:-</b>	
Y <sub>opt</sub>	// an Optimal subset
<b>Algorithm:-</b>	
begin	
Initialize	
Y <sub>opt</sub> = Subset <sub>N</sub> ;	
New <sub>opt</sub> = Evaluate (Subset <sub>N</sub> , X ,M); //Evaluate subset by a Independent Measure M	
do begin	
S = Selectsubset(X) ; //generate a subset for evaluation	
New = Evaluate(S, X, M); //Evaluate the current subset S by M	
If(New is better than New <sub>opt</sub> )	
New <sub>opt</sub> = New;	
Y <sub>opt</sub> = S;	
End until(Stop <sub>cr</sub> is met);	
Return Y <sub>opt</sub> ;	
End	

Fig.3. Generalized Filter method



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

## 6.1.2 Wrapper Methods

A generalized wrapper algorithm (shown in figure 4) is very similar to the generalized filter algorithm except that it utilizes a predefined mining algorithm A instead of an independent measure M for subset evaluation.

<b>WRAPPER ALGORITHM</b>	
<b>Input:-</b>	X (F <sub>0</sub> ,F <sub>1</sub> ,.....,F <sub>n-1</sub> ) //a training data with n features
	Subset N //a subset from where to start the search
Stop <sub>cr</sub>	//a stopping Criterion
<b>Output:-</b>	Y <sub>opt</sub> // an Optimal subset
<b>Algorithm:-</b>	begin
	Initialize
	Y <sub>opt</sub> =Subset N;
	New <sub>opt</sub> =Evaluate (Subset N, X ,MinAlg);//Evaluate subset by a mining //Algorithm MinAlg
	do begin
	S=Selectsubset(X) ; //generate a subset for evaluation
	New=Evaluate(S, X, MinAlg); //Evaluate the current subset S by A
	If(New is better than New <sub>opt</sub> )
	End until(Stop <sub>cr</sub> is met);
	New <sub>opt</sub> =New;
	Y <sub>opt</sub> =S;
	Return Y <sub>opt</sub> ;
	end

Figure.4 Wrapper Methods

## 6.1.3 Hybrid or Embedded algorithms

Many researchers have proposed hybrid or embedded algorithms which combine the features of both wrappers and filter methods[33][26][27][28][29][30][31].[32] gave a generalized algorithm followed in embedded or hybrid algorithm. The hybrid algorithm starts a search with given subset either empty or with few features and iterates at each cardinality. In each round for a best subset with cardinality c, it searches through all possible subsets of cardinality c + 1 by adding one feature from the remaining features. Each newly generated subset S with cardinality c + 1 is evaluated by an independent measure M and compared with the previous best one. If S is better, it becomes the current best subset Y'opt at level c + 1. At the end of each iteration, a mining algorithm A is applied on Y'opt at level c + 1 and the quality of the mined result is compared with that from the best subset at level c. If Y'opt is better, the algorithm continues to find the best subset at the next level; otherwise, it stops and outputs the current best subset as the final best subset. The quality of results from a mining algorithm provides a natural stopping criterion in the hybrid model.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

```

HYBRID ALGORITHM
Input:-
X (F0,F1,.....,Fn-1) //a training data with n features
Subset N //a subset from where to start the search
Stopcr //a stopping Criterion
Output:-
Yopt // an Optimal subset
Algorithm:-
Begin
    Initialize
    Yopt=Subset N;
    C0 = Cardinality(Subset N) ; //Calculate the cardinality of subset N
    θbest = Evaluate (Subset n, X, M);
    σbest = Evaluate(Subset N, X , MinAlg); //Evaluate Subset N by a mining algorithm MinAlg
    for c=c0+1 to N begin
        for i=0 to N-c begin
            S=YoptU { Fj}; //generate a subset with cardinality c for evaluation
            σ = evaluate(S, X, M); //Evaluate the current subset S by M
            If(σ is better than σbest)
                σbest=σ;
                YoptI = S;
            End
            σ = Evaluate (YoptI, X , MinAlg); // Evaluate YoptI by MinAlg
            If(σ is better than σbest)
                Yopt = YoptI;
                σbest=σ;
            else
                break and return Yopt;
            end;
        return Yopt; end;

```

Figure 5 . A generalized embedded algorithm

Criteria	Filter methods	Wrapper Methods	Hybrid methods
<b>Classification Algorithm</b>	Independent of classification algorithm	Dependent on Classification algorithm	Dependent on classification algorithm
<b>Computational Cost</b>	Less for huge data sets	More for huge data sets	In comparison with wrapper methods it is less
<b>speed</b>	Faster than wrappers	Slower than filter methods	Faster than wrapper methods
<b>Generality</b>	Result has more generality than wrappers method	As classifier is involved there is lack of generality	As classifier is involved there is lack of generality
<b>Superiority</b>	They are less optimal	In supervised learning problems they are superior	Performance is reduced if more irrelevant features are considered in target set
<b>Selection of subset</b>	Selects large subset of data	Accuracy and Recognition rate is higher than filter methods	These methods are least prone to over fitting

Figure .6.Comparisons of filter, Wrapper and Hybrid Methods

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

## Feature Extraction Techniques for Speech Recognition

Speech is the way of communication between human beings. It is defined as a process automatically recognizes the spoken words of person based on given speech signal information. It is also known as Automatic Speech Recognition or computer speech recognition and speech to text conversion. To recognize the speech feature extraction and word recognition these two steps are followed. After feature extraction, feature matching is performed for word recognition. Ghule, Kishori R., and R. R. Deshmukh[75] reviewed about some techniques like Mel frequency Cepstral Coefficient MFCC, Linear prediction coding LPC, Linear prediction cepstral coefficient (LPCC), Discrete Wavelet Transform (DWT). MFCC is the most popular feature extraction technique.

### Correlation based feature selection techniques

A feature should be highly correlated to the class and not much correlated to any other feature of the class.[34]

Entropy is a measure of uncertainty of a random variable. It can be defined by the following

$$H(X) = - \sum P(x_i) \log_2 (P(x_i)) \quad (8)$$

And the entropy of X after observing values of another variable Y is defined as

$$H(X/Y) = - \sum P(y_j) \sum P(x_i/y_j) \log_2 (P(x_i / y_j)) \quad (9)$$

Here,  $P(x_i)$  is the prior probabilities for all values of X, and  $P(x_i/y_j)$  is the posterior probabilities of X when values of Y are given Entropy is a measure of correlation between two variables. [18][19][20].

## VII. DIMENSIONALITY REDUCTION

This techniques extract features by projecting input data into a lower dimensional subspace. Principal Component Analysis[169], Random projection and Linear Discriminant Analysis are some methods.

Reducing the dimensionality of a dataset is an important and often challenging task.[172][173] This can be done by either reducing the number of features, a task called feature selection, or by reducing the number of patterns, called data reduction.

This methods were applied to classification methods [156][157][158]. The reduction of the patterns in a dataset, a lot of different techniques have been developed over the past few years. These include Statistical methods such as Chi-square [159], Similarity [160][171] and Aggregation [161]), Geometrical methods [162], Clustering (both classic [163] and fuzzy [164]), Autoregressive (AR) techniques [165], PCA [160][164], Singular Value Decomposition (which is one of the most popular methods) [164][166], Wavelets (DWT) [165], Independent Component Analysis [160], Fourier Decomposition (DFT [165] and a variation of FA [167]), SOM [168] and an alternative data reduction approach (ALEV) [168].

### Diagonal features for character recognition

Diagonal features are used for extracting the features of a character. Diagonal feature computation consists of two steps: character image is divided into different zones Diagonal Features are computed for each zone of the character. Features are also computed from each image by considering zones in horizontal and vertical directions.[74]

### Orthogonal Random Features and structured orthogonal Random features (SORF)

Felix, X. Yu presents Random Fourier Features: in Gaussian kernel approximation, replacing the random Gaussian matrix by a properly scaled random orthogonal matrix significantly decreases kernel approximation error. These techniques called as Orthogonal Random Features (ORF)[17], and provide theoretical and empirical justification for this behaviour. Motivated by this discovery, researchers further proposed Structured Orthogonal Random Features (SORF), which uses a class of structured discrete orthogonal matrices to speed up the computation. The method

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

reduces the time cost from  $O(d^2)$  to  $O(d \log d)$ , where  $d$  is the data dimensionality, with almost no compromise in kernel approximation quality compared to ORF.

## VIII. FAST CLUSTERING BASED FEATURE SELECTION ALGORITHM FOR HIGH DIMENSIONAL DATA

The main concept of the FAST clustering feature selection algorithm is to cluster subsets with most similar characteristics while removing irrelevant subsets from that cluster. This FAST algorithm has both efficiency and accuracy for finding required set. The FAST algorithm has two steps the first step is it uses minimum spanning tree to divide data into different clusters and the in second step, it removes irrelevant sets and gives accurate and efficient result. All clusters formed using FAST algorithms are independent of each other. It is discussed further in [39][40][41][42][43][44][45][46][47][48].

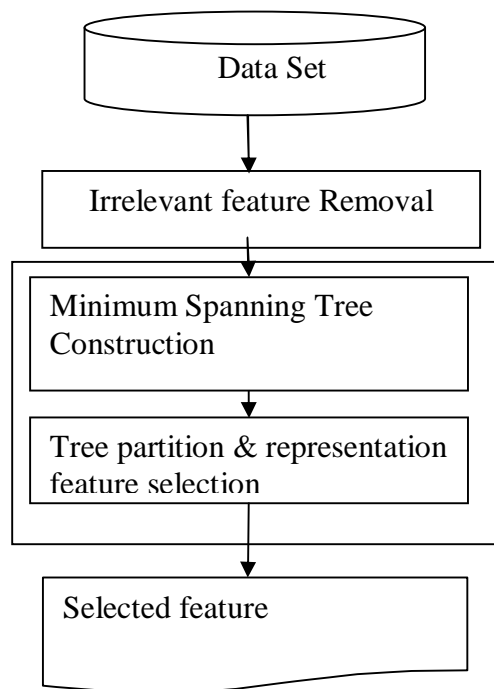


Figure. 7. Basic Framework of FAST algorithm

## IX. CHALLENGES IN FEATURE EXTRACTION METHODS

Few specified below are the upcoming challenges in feature extraction methods.

### 9.1 Large Scale Machine Learning

In most of the domains today, data is arriving faster than the algorithm is able to learn from it. To avoid wasting this data, we must switch from the traditional "one-shot" machine learning approach to systems that are able to mine continuous, high-volume, open-ended data streams as they arrive. The authors Domingos, Pedro, and Geoff Hulten [3][4][5] have identified developed an approach to building stream mining algorithms that satisfies all of them. The approach is based on explicitly minimizing the number of examples used in each learning step, while guaranteeing that user-defined targets for predictive performance are met. So far, Hulten, Geoff, and Pedro Domingos [6][7][8] have applied this approach to four major (and widely differing) types of learner: decision tree induction, Bayesian network

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

learning, k-means clustering, and the EM algorithm for mixtures of Gaussians. The researchers ultimate goal is to develop a set of primitives such that any learning algorithm built using them scales automatically to arbitrarily large data streams.[9] has given good insights of future problems in machine learning.[3] This paper introduced Hoeffding trees, a method for learning online from the high-volume data streams that are increasingly common. Hoeffding trees allow learning in very small constant time and have strong guarantees of high asymptotic similarity to the corresponding batch trees. VFDT is a high-performance data mining system based on Hoeffding trees.

Domingos, Pedro, and Geoff Hulten [4] [5] derives an upper bound for learner's loss as a function of number of examples used in each step of the algorithm. Then minimizes each step while guaranteeing the no change in model with infinite data. It is empirically applied to K-means algorithm. Bottou, Léon, Frank E. Curtis, and Jorge Nocedal provides a review and commentary on the past, present, and future of numerical optimization algorithms in the context of machine learning applications[10]

## 9.2 Fuzzy Based feature extraction

The authors Pal, Nikhil R., Vijay K. Eluri, and Gautam K. Manda [57] and Nagalakshmi, G., and S. Jyothi [99] discusses about fuzzy based feature extraction method. Sammon's non linear projection method lacks predictability and is ineffective for large data sets This paper [58] Chaudhari, Narendra S., and Avishek Ghosh introduces predictability implements an extension of Sammon's algorithm using fuzzy logic approach.

Clifton, David A., and Bruce W. Pilsworth gives the review of existing methods as histogram method and parametric Gaussian method.[59]. Histograms produce a density estimate of a series of data . From the density estimate, a probability distribution can be derived by multiplying each bin height by its bin width. Parametric Gaussian method assumes the data to be Gaussian. But the data may be skewed or multimodal which gives inaccurate results.

**kernel method [59]:** The kernel method is an improvement over the existing two methods for creating a density estimate. It produces highly smooth density estimates for highly skewed and multi-modal distributions, and thus exhibits the quality of robustness that is required.

**Linguistic label method[59]:** The concept behind this method is a novel one, and contrasts with the strictly statistical concepts used by the Kernel Method. It is based on the idea of computing with words by intuitively categorizing data using linguistic labels. Three Linguistic label methods are "FD1" "FD2" and "FD3" .The difference in these methods lies in the length of the interval in the series of data to find fuzzy features.

A novel feature extraction criterion, fuzzy maximum margin criterion (FMMC), is proposed by means of the maximum margin criterion (MMC) and fuzzy set theory by authors Cui, Yan, and Liya Fan[61]. The between-class and within-class fuzzy scatter matrices are redefined by incorporating the membership degrees of samples. Also the use of generalized singular value decomposition (GSVD) to the criterion makes the algorithm more effective. Yang, Wankou, et al [62] proposes a new method of feature extraction and recognition, namely, the fuzzy inverse Fisher discriminate analysis (FIFDA) based on the inverse Fisher discriminate criterion and fuzzy set theory.

## 9.3 Privacy Aware Machine Learning

Privacy has become a major concern for all the machine learning tasks. Machine Learning Algorithms operate on sensitive and private data. Obviously data protection, safety , Information security, privacy and fair use of data plays utmost important role in health data science, social network data etc. The goal of analysing large amounts of data for information extraction collides with the privacy of individuals. Hence, in order to protect sensitive information, the effects of the right to be forgotten on machine learning algorithms need to be studied more extensively. The "right to be forgotten" has been recently discussed with a particular focus on removing personal data and sensitive (personal) information from automated analysis if requested by individual. [176] while many of the technological issues have been apparently solved, the legal aspects of the collection and processing of vast amounts of data using machine learning algorithms has been neglected [174].

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

Some of the following methods are used with sensitive data

**Anonymization.** The data set is transformed into a derived set that blurs the sensitive attributes without actually removing them altogether.[177]

**Pseudonymization.** Related to anonymization, Pseudonymization works by removing sensitive attributes and replacing them with a placeholder, while keeping the internal logical structure of the data set, i.e., records with the same sensitive attributes get assigned the same pseudonym.[177]

**Functional Encryption.** Functional encryption allows the calculation of certain mathematical operations on the encrypted values [177]

This document web page[178] give a study of statistical risk minimization problems under a version of privacy in which the data is kept confidential even from the learner

## 9.4 Role of saliency Maps in feature extraction

Saliency maps are reviewed by authors Verma, Jyoti, and Vineet Richhariya [56]. There is a gap in the view-based object recognition literature between the image-based systems and the feature-based systems. While the image-based systems work with complex objects, e.g. faces, they are highly sensitive to occlusion, scale, and deformation. The feature-based systems, rely on highly sensitive feature extraction processes that perform poorly in the presence of surface markings or texture. Saliency map graph offers a robust, transformation invariant, multi scale representation of an image that not only captures the salient image structure, but provides the locality of representation required to support occluded object recognition[53]

Detection of salient image regions is useful for applications like image segmentation, adaptive compression, and region-based image retrieval.[50] this paper gives a novel method to determine salient regions in images using low-level features of luminance and color.[51]. Sample sub windows at random positions with random sizes to extract features method is used by Moosmann, Franck, Diane Larlus, and Frederic Jurie.[52]

Existing CNN-based methods operate at the patch level instead of the pixel level. Resulting saliency maps are typically blurry, especially near the boundary of salient objects.[53] This paper introduces deep network consists of two complementary components, a pixel level fully convolutional stream and a segment-level spatial pooling stream.

## 9.5 Scalability

Usually a machine learning algorithm or feature extraction algorithm is tested on small data and then extended to large data sets. When the algorithm is extended to large data sets lot of problems come into play such as curse of dimensionality, computational complexity and processing time and learning time increases exponentially. To overcome this problem it has been obviously understood that algorithms should be competent for large data sets also.

[99] – [111][182] discusses about dimensionality reduction in clustering and feature selection using feature weights.

Is the algorithm chosen scalable? On What parameter data is driving?

## X. RELEVANT AND OPTIMAL FEATURES

Based on a review of previous definitions of feature relevance, [180] [181] classified features into three disjoint categories, namely, strongly relevant, weakly relevant, and irrelevant features.

Let  $F$  be a full set of features,  $F_i$  a feature, and  $S_i = F - \{F_i\}$ .

These categories of relevance can be formalized as follows.

**Definition 1 (Strong relevance)** A feature  $F_i$  is strongly relevant iff  $P(C | F_i, S_i) \neq P(C | S_i)$ .

**Definition 2 (Weak relevance)** A feature  $F_i$  is weakly relevant iff  $P(C | F_i, S_i) = P(C | S_i)$ , and  $\exists S'_i \subset S_i$ , such that  $P(C | F_i, S'_i) \neq P(C | S'_i)$ .

**Corollary 1 (Irrelevance)** A feature  $F_i$  is irrelevant iff  $\forall S'_i \subseteq S_i, P(C | F_i, S'_i) = P(C | S'_i)$ .

The relevant and optimal features are necessary for good result of classification.



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

## Suitable Feature extraction algorithm

Feature extraction is highly subjective in nature, it depends on what type of problem you are solving. There is no generic feature extraction algorithm which works in all cases. Like classifiers, it is not possible to say which is the best algorithm for feature selection or extraction. It depends on your application. The questions you can answer and choose the best algorithm are

- What kind of problem are you solving? (classification, regression, clustering, etc.)
- Do you have a huge data?
- Does your data have very high dimensionality?
- Is the data labeled?
- What do you want to perform feature extraction or feature selection?
- Which method do you want to use a supervised or unsupervised method?

## XI. CONCLUSION

This paper gives a brief review of the existing methods present in feature extraction in different applications. The method that you choose for feature extraction for your application depends on many factors. Also this paper discusses about the latest challenges in feature extraction.

## REFERENCES

- [1] Guyon, Isabelle, and André Elisseeff. "An introduction to variable and feature selection." *Journal of machine learning research* 3.Mar (2003): 1157-1182.
- [2] Yager, Ronald R., and Liping Liu, eds. *Classic works of the Dempster-Shafer theory of belief functions*. Vol. 219. Springer, 2008.
- [3] Domingos, Pedro, and Geoff Hulten. "Mining high-speed data streams." *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2000.
- [4] Domingos, Pedro, and Geoff Hulten. "A general method for scaling up machine learning algorithms and its application to clustering." *ICML*. Vol. 1. 2001.
- [5] Domingos, Pedro, and Geoff Hulten. "Learning from infinite data in finite time." *Advances in neural information processing systems*. 2002.
- [6] Hulten, Geoff, and Pedro Domingos. "Mining complex models from arbitrarily large databases in constant time." *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2002.
- [7] Hulten, Geoff, Pedro Domingos, and Yeuhi Abe. "Mining massive relational databases." *Proceedings of the IJCAI-2003 workshop on learning statistical models from relational data*. 2003.
- [8] Hulten, Geoff, Laurie Spencer, and Pedro Domingos. "Mining time-changing data streams." *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2001.
- [9] Domingos, Pedro. "Structured machine learning: Ten problems for the next ten years." *Proc. of the Annual Intl. Conf. on Inductive Logic Programming*. 2007.
- [10] Bottou, Léon, Frank E. Curtis, and Jorge Nocedal. "Optimization methods for large-scale machine learning." *arXiv preprint arXiv:1606.04838* (2016).
- [11] Aichert, André. "Feature extraction techniques." *Camp Medical Seminar*. Vol. 13. 2008.
- [12] Bane, Shilpa, and D. R. Pawar. "Survey on Feature Extraction methods in Object Recognition." *International Journal of Computer Science and Information Technologies*, vol 5 (2014): 3224-3226.
- [13] Bencharef, O., et al. "Color objects recognition system based on artificial neural network with Zernike, Hu & Geodesic descriptors." *Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2012 6th International Conference on*. IEEE, 2012.
- [14] Yang, Mingqiang, Kidiyo Kpalma, and Joseph Ronsin. "A survey of shape feature extraction techniques." (2008): 43-90.
- [15] Zhang, Dengsheng, and Guojun Lu. "Review of shape representation and description techniques." *Pattern recognition* 37.1 (2004): 1-19.
- [16] Horn, Berthold. *Robot vision*. MIT press, 1986.
- [17] Felix, X. Yu, et al. "Orthogonal random features." *Advances in Neural Information Processing Systems*. 2016
- [18] Doshi, Mital. "Correlation Based Feature Selection (Cfs) Technique To Predict Student Performance." *International Journal of Computer Networks & Communications* 6.3 (2014): 197.
- [19] Hall, Mark Andrew. "Correlation-based feature selection for machine learning." (1999).
- [20] Yu, Lei, and Huan Liu. "Feature selection for high-dimensional data: A fast correlation-based filter solution." *Proceedings of the 20th international conference on machine learning (ICML-03)*. 2003.
- [21] L. Ladla and T. Deepa " Feature Selection Methods And Algorithms" *International Journal on Computer Science and Engineering (IJCSE)* vol.3(5) pp. 1787-1797 2011
- [22] Rohini, P. "Integration of MDLEP and Gabor Function as a Feature Vector for Image Retrieval System." *IUP Journal of Computer Sciences* 10.1/2 (2016): 62.

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijrset.com](http://www.ijrset.com)

Vol. 6, Issue 12, December 2017

- [23] Rao, L. Koteswara, D. Venakata Rao, and P. Rohini. "Integration of Color and MDLEP as a Feature Vector in Image Indexing and Retrieval System." *Proceedings of the International Congress on Information and Communication Technology*. Springer Singapore, 2016.
- [24] Rohini P, KoteswaraRao L, VenkataRao D, "Combination of CDLEP and Gabor Transform for CBIR" Conference proceeding s by Springer AISC at IC3T – 2015 during July 2015
- [25] Geng, Xiubo, et al. "Feature selection for ranking." *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2007.
- [26] Ghareb, Abdullah Saeed, Azuraliza Abu Bakar, and Abdul Razak Hamdan. "Hybrid feature selection based on enhanced genetic algorithm for text categorization." *Expert Systems with Applications* 49 (2016): 31-47.
- [27] Kabir, Md Monirul, Md Shahjahan, and Kazuyuki Murase. "A new local search based hybrid genetic algorithm for feature selection." *Neurocomputing* 74.17 (2011): 2914-2928.
- [28] Apolloni, Javier, Guillermo Leguizamón, and Enrique Alba. "Two hybrid wrapper-filter feature selection algorithms applied to high-dimensional microarray experiments." *Applied Soft Computing* 38 (2016): 922-932.
- [29] Talavera, Luis. "An evaluation of filter and wrapper methods for feature selection in categorical clustering." *Advances in Intelligent Data Analysis VI* (2005): 742-742.
- [30] Das, Sanmay. "Filters, wrappers and a boosting-based hybrid for feature selection." *ICML*. Vol. 1. 2001.
- [31] Xing, Eric P., Michael I. Jordan, and Richard M. Karp. "Feature selection for high-dimensional genomic microarray data." *ICML*. Vol. 1. 2001.
- [32] Liu, Huan, and Lei Yu. "Toward integrating feature selection algorithms for classification and clustering." *IEEE Transactions on knowledge and data engineering* 17.4 (2005): 491-502.
- [33] Sivagaminathan, Rahul Karthik, and Sreeram Ramakrishnan. "A hybrid approach for feature subset selection using neural networks and ant colony optimization." *Expert systems with applications* 33.1 (2007): 49-60.
- [34] Yu, Lei, and Huan Liu. "Feature selection for high-dimensional data: A fast correlation-based filter solution." *Proceedings of the 20th international conference on machine learning (ICML-03)*. 2003.
- [35] Somol, Petr, et al. "Adaptive floating search methods in feature selection." *Pattern recognition letters* 20.11 (1999): 1157-1163.
- [36] Ghareb, Abdullah Saeed, Azuraliza Abu Bakar, and Abdul Razak Hamdan. "Hybrid feature selection based on enhanced genetic algorithm for text categorization." *Expert Systems with Applications* 49 (2016): 31-47.
- [37] Dash, Manoranjan, et al. "Feature selection for clustering-a filter solution." *Data Mining, 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on*. IEEE, 2002.
- [38] Sutar, Swapnil A., and Devendra P. Gadekar. "Survey on Fast Clustering Based Feature Selection Algorithm for High Dimensional Data." *International journal of science and research* 3.12 (2014): 691-694.
- [39] Song, Qinbao, Jingjie Ni, and Guangtao Wang. "A fast clustering-based feature subset selection algorithm for high-dimensional data." *IEEE transactions on knowledge and data engineering* 25.1 (2013): 1-14.
- [40] Yu, Lei, and Huan Liu. "Feature selection for high-dimensional data: A fast correlation-based filter solution." *Proceedings of the 20th international conference on machine learning (ICML-03)*. 2003.
- [41] Dash, Manoranjan, Huan Liu, and Hiroshi Motoda. "Consistency based feature selection." *Pacific-Asia conference on knowledge discovery and data mining*. Springer Berlin Heidelberg, 2000.
- [42] Chopade, Nitin B., and Beena S. Khade. "A New Clustering Based Algorithm for Feature Subset Selection."
- [43] Feng, Dingcheng, Feng Chen, and Wenli Xu. "Efficient leave-one-out strategy for supervised feature selection." *Tsinghua Science and Technology* 18.6 (2013): 629-635.
- [44] Deng, Houtao, and George Runger. "Feature selection via regularized trees." *Neural Networks (IJCNN), The 2012 International Joint Conference on*. IEEE, 2012.
- [45] Sun, Yijun, Sinisa Todorovic, and Steve Goodison. "Local-learning-based feature selection for high-dimensional data analysis." *IEEE transactions on pattern analysis and machine intelligence* 32.9 (2010): 1610-1626.
- [46] Saha, Sriparna, et al. "Feature selection and semi-supervised clustering using multiobjective optimization." *SpringerPlus* 3.1 (2014): 465.
- [47] Munieswari, R., and S. Saranya. "A Survey on Feature Selection Using FAST Approach to Reduce High Dimensional Data." *International Journal of Engineering Trends and Technology (IJETT)* 8.5 (2014).
- [48] **Jesna Jose,"Fast for Feature Subset Selection Over Dataset" International Journal of Science and Research (IJSR), Volume 3 Issue 6, June 2014.**
- [49] Srivastava, Shweta, Nikita Joshi, and Madhvi Gaur. "A review paper on feature selection methodologies and their applications." *International Journal of Computer Science and Network Security (IJCSNS)* 14.5 (2014): 78.
- [50] Achanta, Radhakrishna, et al. "Salient region detection and segmentation." *Computer Vision Systems* (2008): 66-75
- [51] Gao, Dashan, and Nuno Vasconcelos. "Discriminant saliency for visual recognition from cluttered scenes." *Advances in neural information processing systems*. 2005.
- [52] Moosmann, Franck, Diane Larlus, and Frederic Jurie. "Learning saliency maps for object categorization." *International Workshop on The Representation and Use of Prior Knowledge in Vision (in ECCV'06)*. 2006.
- [53] Shokoufandeh, Ali, Ivan Marsic, and Sven J. Dickinson. "View-based object recognition using saliency maps." *Image and Vision Computing* 17.5 (1999): 445-460.
- [54] Li, Guanbin, and Yizhou Yu. "Deep contrast learning for salient object detection." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
- [55] He, Zhiyong, Xin Chen, and Lining Sun. "Saliency mapping enhanced by structure tensor." *Computational intelligence and neuroscience* 2015 (2015): 3.
- [56] Verma, Jyoti, and Vineet Richhariya. "A Survey Paper on an Efficient Salient Feature Extraction by Using Saliency Map Detection with Modified K-Means Clustering Technique." *International Journal of Communication and Computer Technologies* 1.5: 2278-9723.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

- [57] Pal, Nikhil R., Vijay K. Eluri, and Gautam K. Mandal. "Fuzzy logic approaches to structure preserving dimensionality reduction." *IEEE Transactions on Fuzzy Systems* 10.3 (2002): 277-286.
- [58] Chaudhari, Narendra S., and Avishek Ghosh. "Feature Extraction Using Fuzzy Rule Based System." *IJCSA* 5.3 (2008): 1-8.
- [59] Clifton, David A., and Bruce W. Pilsworth. "Critical Comparison of Methods of Fuzzy Feature Extraction."
- [60] Grabisch, Michel. "Fuzzy integral for classification and feature extraction." *Fuzzy measures and integrals: Theory and Applications* (2000): 415-434.
- [61] Cui, Yan, and Liya Fan. "Feature extraction using fuzzy maximum margin criterion." *Neurocomputing* 86 (2012): 52-58.
- [62] Yang, Wankou, et al. "Feature extraction using fuzzy inverse FDA." *Neurocomputing* 72.13 (2009): 3384-3390.
- [63] Chang, Edward, et al. "CBSA: content-based soft annotation for multimodal image retrieval using Bayes point machines." *IEEE Transactions on Circuits and Systems for Video Technology* 13.1 (2003): 26-38.
- [64] Li, Jia, and James Ze Wang. "Automatic linguistic indexing of pictures by a statistical modeling approach." *IEEE Transactions on pattern analysis and machine intelligence* 25.9 (2003): 1075-1088.
- [65] Mei, Tao, et al. "Coherent image annotation by learning semantic distance." *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008.*
- [66] Barnard, Kobus, et al. "Matching words and pictures." *Journal of machine learning research* 3.Feb (2003): 1107-1135.
- [67] Jeon, Jiwoon, Victor Lavrenko, and Raghavan Manmatha. "Automatic image annotation and retrieval using cross-media relevance models." *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval.* ACM, 2003.
- [68] Jeon, Jiwoon, Victor Lavrenko, and Raghavan Manmatha. "Automatic image annotation and retrieval using cross-media relevance models." *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval.* ACM, 2003.
- [69] Zha, Zheng-Jun, et al. "Joint multi-label multi-instance learning for image classification." *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. IEEE, 2008.*
- [70] Blei, David M., and Michael I. Jordan. "Modeling annotated data." *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval.* ACM, 2003.
- [71] Wang, Yong, et al. "Combining global, regional and contextual features for automatic image annotation." *Pattern Recognition* 42.2 (2009): 259-266.
- [72] Lisin, Dimitri A., et al. "Combining local and global image features for object class recognition." *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on. IEEE, 2005.*
- [73] Lowe, David G. "Distinctive image features from scale-invariant keypoints." *International journal of computer vision* 60.2 (2004): 91-110.
- [74] Jindal, Anita, Renu Dhir, and Rajneesh Rani. "Diagonal Features and SVM Classifier for Handwritten Gurumukhi Character Recognition." *International Journal of Advanced Research in Computer Science and Software Engineering* 2.5 (2012).
- [75] Ghule, Kishori R., and R. R. Deshmukh. "Feature Extraction Techniques for Speech Recognition: A Review." *International Journal of Scientific & Engineering Research* 6.5 (2015): 2229-5518.
- [76] Panchanathan, Sethuraman, et al. "The role of color in content-based image retrieval." *Image Processing, 2000. Proceedings. 2000 International Conference on. Vol. 1. IEEE, 2000.*
- [77] Chen, Wei-Ta, Wei-Chuan Liu, and Ming-Syan Chen. "Adaptive color feature extraction based on image color distributions." *IEEE Transactions on image processing* 19.8 (2010): 2005-2016.
- [78] Pei, Soo-Chang, and Ching-Min Cheng. "Color image processing by using binary quaternion-moment-preserving thresholding technique." *IEEE Transactions on Image Processing* 8.5 (1999): 614-628.
- [79] Patil, Sandip S., and Atul V. Dusane. "Use of color feature extraction technique based on color distribution and relevance feedback for content based image retrieval." *International Journal of Computer Applications* 52.17 (2012).
- [80] Maheshwary, P. R. I. T. I., and N. A. M. I. T. A. Srivastava. "Retrieval of remote sensing images using color, texture and spectral features." *International Journal of Engineering Science and Technology* 2.9 (2010): 4306-4311.
- [81] Pujari, Jagadeesh, and P. Hiremath. "Content based image retrieval based on color texture and shape features using image and its complement." *International Journal of Computer Science and Security* 1.4 (2007): 25-35.
- [82] Saad, Michele. "Low-level color and texture feature extraction for content-based image retrieval." *Final Project Report, EE K381* (2008): 20-28.
- [83] Zhao, Qiang, et al. "Stone images retrieval based on color histogram." *Image Analysis and Signal Processing, 2009. IASP 2009. International Conference on. IEEE, 2009.*
- [84] Hafner, James, et al. "Efficient color histogram indexing for quadratic form distance functions." *IEEE transactions on pattern analysis and machine intelligence* 17.7 (1995): 729-736.
- [85] Swain, Michael J., and Dana H. Ballard. "Color indexing." *International journal of computer vision* 7.1 (1991): 11-32.
- [86] Ganar, Apurva N., C. S. Gode, and Sachin M. Jambhulkar. "Enhancement of image retrieval by using colour, texture and shape features." *Electronic Systems, Signal Processing and Computing Technologies (ICESC), 2014 International Conference on. IEEE, 2014.*
- [87] Han, Ju, and Kai-Kuang Ma. "Fuzzy color histogram and its use in color image retrieval." *IEEE Transactions on image Processing* 11.8 (2002): 944-952.
- [88] Rao, Aibing, Rohini K. Srihari, and Zhongfei Zhang. "Spatial color histograms for content-based image retrieval." *Tools with Artificial Intelligence, 1999. Proceedings. 11th IEEE International Conference on. IEEE, 1999.*
- [89] Rasheed, Waqas, et al. "Image retrieval using maximum frequency of local histogram based color correlogram." *Multimedia and Ubiquitous Engineering, 2008. MUE 2008. International Conference on. IEEE, 2008.*
- [90] Johnson, Garrett M., et al. "Derivation of a color space for image color difference measurement." *Color Research & Application* 35.6 (2010): 387-400.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijrset.com](http://www.ijrset.com)

Vol. 6, Issue 12, December 2017

- [91] Lin, Chuen-Horng, Rong-Tai Chen, and Yung-Kuan Chan. "A smart content-based image retrieval system based on color and texture feature." *Image and Vision Computing* 27.6 (2009): 658-665.
- [92] Huang, Jing, et al. "Image indexing using color correlograms." *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. IEEE, 1997.
- [93] Paschos, George, Ivan Radev, and Nagarajan Prabakar. "Image content-based retrieval using chromaticity moments." *IEEE Transactions on Knowledge and Data Engineering* 15.5 (2003): 1069-1072.
- [94] Shao, Hong, et al. "Image retrieval based on MPEG-7 dominant color descriptor." *Young Computer Scientists, 2008. ICYCS 2008. The 9th International Conference for*. IEEE, 2008.
- [95] Talib, Ahmed, et al. "A weighted dominant color descriptor for content-based image retrieval." *Journal of Visual Communication and Image Representation* 24.3 (2013): 345-360.
- [96] Yang, Nai-Chung, et al. "A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval." *Journal of Visual Communication and Image Representation* 19.2 (2008): 92-105.
- [97] Wang, Xiang-Yang, Hong-Ying Yang, and Dong-Ming Li. "A new content-based image retrieval technique using color and texture information." *Computers & Electrical Engineering* 39.3 (2013): 746-761.
- [98] Yang, Nai-Chung, et al. "A fast MPEG-7 dominant color extraction with new similarity measure for image retrieval." *Journal of Visual Communication and Image Representation* 19.2 (2008): 92-105.
- [99] Nagalakshmi, G., and S. Jyothi. "Prawn Recognition on Fuzzy Based Feature Vectors with Image Segmentation." *International Journal of Advanced Research in Computer Science and Software Engineering* 6.5 (2016).
- [100] Berchtold, Stefan, Christian Böhm, and Hans-Peter Kriegel. "The pyramid-technique: towards breaking the curse of dimensionality." *ACM SIGMOD Record*. Vol. 27. No. 2. ACM, 1998.
- [101] Berchtold, Stefan, et al. "A cost model for nearest neighbor search in high-dimensional data space." *Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*. ACM, 1997.
- [102] Berchtold, S., Keim, D.A., Kriegel, H.-P.: The X-tree: An index structure for high dimensional data. Proc. 22nd VLDB Conf. (1996) 28–39
- [103] Ravi Kanth, K. V., Divyakant Agrawal, and Ambuj Singh. "Dimensionality reduction for similarity searching in dynamic databases." *ACM SIGMOD Record*. Vol. 27. No. 2. ACM, 1998.
- [104] Katayama, Norio, and Shin'ichi Satoh. "The SR-tree: An index structure for high-dimensional nearest neighbor queries." *ACM Sigmod Record*. Vol. 26. No. 2. ACM, 1997.
- [105] Kruskal, Joseph B. "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis." *Psychometrika* 29.1 (1964): 1-27.
- [106] Kruskal, Joseph B. "Multidimensional scaling and other methods for discovering structure." *Statistical methods for digital computers* 3 (1977): 296-339.
- [107] Hristescu, Gabriela, and Martin Farach-Colton. "Cofe: A scalable method for feature extraction from complex objects." *Lecture notes in computer science* (2000): 358-371.
- [108] Bryant, Devon Leno. *Scalable Audio Feature Extraction*. Diss. University of Colorado Colorado Springs, 2014.
- [109] McKay, Cory, Ichiro Fujinaga, and Philippe Depalle. "jAudio: A feature extraction library." *Proceedings of the International Conference on Music Information Retrieval*. 2005.
- [110] Lecerf, Loic, and Boris Chidlovskii. "Scalable feature extraction from noisy documents." *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*. IEEE, 2009.
- [111] Masud, Mohammad M., Latifur Khan, and Bhavani Thuraisingham. "A scalable multi-level feature extraction technique to detect malicious executables." *Information Systems Frontiers* 10.1 (2008): 33-45.
- [112] Stanchev, Peter, David Green Jr, and Boyan Dimitrov. "High level color similarity retrieval." (2003).
- [113] Jain, Anil K., and Aditya Vailaya. "Image retrieval using color and shape." *Pattern recognition* 29.8 (1996): 1233-1244.
- [114] Flickner, Myron, et al. "Query by image and video content: The QBIC system." *computer* 28.9 (1995): 23-32.
- [115] Pass, Greg, and Ramin Zabih. "Histogram refinement for content-based image retrieval." *Applications of Computer Vision, 1996. WACV'96., Proceedings 3rd IEEE Workshop on*. IEEE, 1996.
- [116] J. Huang, S. Kuamr, M. Mitra, et al., "Image indexing using colorcorrelogram", In Proc. CVPR, (1997), pp. 762-765
- [117] Zhang, Dengsheng, Md Monirul Islam, and Guojun Lu. "A review on automatic image annotation techniques." *Pattern Recognition* 45.1 (2012): 346-362.
- [118] ] [https://en.wikipedia.org/wiki/Feature\\_extraction](https://en.wikipedia.org/wiki/Feature_extraction)
- [119] Nixon, Mark S., and Alberto S. Aguado. *Feature extraction & image processing for computer vision*. Academic Press, 2012.
- [120] Arivazhagan, Sundaram, L. Ganesan, and S. Padam Priyal. "Texture classification using Gabor wavelets based rotation invariant features." *Pattern recognition letters* 27.16 (2006): 1976-1982.
- [121] ] <https://in.mathworks.com/help/images/texture-analysis.html>
- [122] Shidnal, Sushila. "A texture feature extraction of crop field images using GLCM approach." *IJSEAT* 2.12 (2015): 1006-1011.
- [123] Materka, Andrzej, and Michal Strzelecki. "Texture analysis methods—a review." *Technical university of lodz, institute of electronics, COST B11 report, Brussels* (1998): 9-11.
- [124] Prathusha.Perugu et.al A Innovative method for Image retrieval using perceptual textural features International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 5, September- October 2012, pp.246
- [125] Abbadeni, Noureddine. "Computational perceptual features for texture representation and retrieval." *IEEE transactions on image processing* 20.1 (2011): 236-246.
- [126] Rosenfeld, Azriel, and Avinash C. Kak. "Digital picture processing, vol. 2." (1982): 55-190.
- [127] Levine, Martin D. *Vision in man and machine*. McGraw-Hill College, 1985.
- [128] Haralick, Robert M. "Statistical and structural approaches to texture." *Proceedings of the IEEE* 67.5 (1979): 786-804.



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

- [129] Cross, George R., and Anil K. Jain. "Markov random field texture models." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1 (1983): 25-39.
- [130] Pentland, Alex P. "Fractal-based description of natural scenes." *IEEE transactions on pattern analysis and machine intelligence* 6 (1984): 661-674.
- [131] Chellappa, Rama, and Shankar Chatterjee. "Classification of textures using Gaussian Markov random fields." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 33.4 (1985): 959-963.
- [132] Derin, Haluk, and Howard Elliott. "Modeling and segmentation of noisy and textured images using Gibbs random fields." *IEEE Transactions on pattern analysis and machine intelligence* 1 (1987): 39-55.
- [133] Manjunath, B. S., and Rama Chellappa. "Unsupervised texture segmentation using Markov random field models." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13.5 (1991): 478-482.
- [134] Strzelecki, M., and A. Materka. "Markov random fields as models of textured biomedical images." *Proc. 20th National Conf. Circuit Theory and Electronic Networks KTOiUE'97*. 1997.
- [135] Lu, Chun S., Pau C. Chung, and Chih F. Chen. "Unsupervised texture segmentation via wavelet transform." *Pattern Recognition* 30.5 (1997): 729-742.
- [136] Laine, Andrew, and Jian Fan. "Texture classification by wavelet packet signatures." *IEEE Transactions on pattern analysis and machine intelligence* 15.11 (1993): 1186-1191.
- [137] Mallat, Stephane G. "Multifrequency channel decompositions of images and wavelet models." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37.12 (1989): 2091-2110.
- [138] Rosenfeld, Azriel, and Joan S. Weszka. "Picture recognition." *Digital pattern recognition*. Springer, Berlin, Heidelberg, 1980. 135-166.
- [139] Daugman, John G. "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters." *JOSA A* 2.7 (1985): 1160-1169.
- [140] Bovik, Alan C., Marianna Clark, and Wilson S. Geisler. "Multichannel texture analysis using localized spatial filters." *IEEE transactions on pattern analysis and machine intelligence* 12.1 (1990): 55-73.
- [141] Anuradha, K. "Statistical Feature extraction to classify oral cancers." *Journal of Global Research in Computer Science* 4.2 (2013): 8-12.
- [142] Manjunath, Bangalore S., and Wei-Ying Ma. "Texture features for browsing and retrieval of image data." *IEEE Transactions on pattern analysis and machine intelligence* 18.8 (1996): 837-842.
- [143] Arivazhagan, S., and L. Ganesan. "Texture classification using wavelet transform." *Pattern recognition letters* 24.9 (2003): 1513-1521.
- [144] Arivazhagan, Sundaram, L. Ganesan, and S. Padam Priyal. "Texture classification using Gabor wavelets based rotation invariant features." *Pattern recognition letters* 27.16 (2006): 1976-1982.
- [145] Ojala, Timo, Matti Pietikäinen, and David Harwood. "A comparative study of texture measures with classification based on featured distributions." *Pattern recognition* 29.1 (1996): 51-59.
- [146] k.subbareddy and Dr.A.P.Siva Kumar "Analysis of LBP and its variations –A Survey" in International Conference for IETE journal 2016
- [147] Zhang, Wenchao, et al. "Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition." *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*. Vol. 1. IEEE, 2005.
- [148] Heikkilä, Marko, Matti Pietikäinen, and Cordelia Schmid. "Description of interest regions with center-symmetric local binary patterns." *ICVGIP*. Vol. 6. 2006.
- [149] Zhao, Guoying, et al. "Rotation-invariant image and video description with local binary pattern features." *IEEE Transactions on Image Processing* 21.4 (2012): 1465-1477.
- [150] Daugman, John G. "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression." *IEEE Transactions on Acoustics, Speech, and Signal Processing* 36.7 (1988): 1169-1179.
- [151] Tan, Xiaoyang, and Bill Triggs. "Enhanced local texture feature sets for face recognition under difficult lighting conditions." *IEEE transactions on image processing* 19.6 (2010): 1635-1650.
- [152] Jabid, Taskeed, Md Hasanul Kabir, and Oksam Chae. "Local directional pattern (LDP) for face recognition." *Consumer Electronics (ICCE), 2010 Digest of Technical Papers International Conference on*. IEEE, 2010.
- [153] Murala, Subrahmanyam, R. P. Maheshwari, and R. Balasubramanian. "Directional local extrema patterns: a new descriptor for content based image retrieval." *International journal of multimedia information retrieval* 1.3 (2012): 191-203.
- [154] Reddy, P. Vijaya Bhaskar, and A. Rama Mohan Reddy. "Content based image indexing and retrieval using directional local extrema and magnitude patterns." *AEU-International Journal of Electronics and Communications* 68.7 (2014): 637-643.
- [155] Rohini, P. "Integration of MDLEP and Gabor Function as a Feature Vector for Image Retrieval System." *IUP Journal of Computer Sciences* 10.1/2 (2016): 62.
- [156] Frohlich, Holger, Olivier Chapelle, and Bernhard Scholkopf. "Feature selection for support vector machines by means of genetic algorithm." *Tools with artificial intelligence, 2003. proceedings. 15th ieee international conference on*. IEEE, 2003.
- [157] Chao, Shi, and Chen Lihui. "Feature dimension reduction for microarray data analysis using locally linear embedding." *Proceedings Of The 3rd Asia-Pacific Bioinformatics Conference*. Vol. 1. 2005.
- [158] Cao, L., et al. "Saliency analysis of support vector machines for gene selection in tissue classification." *Neural Computing & Applications* 11.3 (2003): 244-249.
- [159] Li, Xiao-Bai, and Varghese S. Jacob. "Adaptive data reduction for large-scale transaction data." *European Journal of Operational Research* 188.3 (2008): 910-924.
- [160] Thireou, Trias, et al. "Evaluation of data reduction methods for dynamic PET series based on Monte Carlo techniques and the NCAT phantom." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 569.2 (2006): 389-393.
- [161] Skyt, Janne, Christian S. Jensen, and Torben Bach Pedersen. "Specification-based data reduction in dimensional data warehouses." *Information Systems* 33.1 (2008): 36-63.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 6, Issue 12, December 2017

- [162] Song, Hao, and Hsi-Yung Feng. "A global clustering approach to point cloud simplification with a specified data reduction ratio." *Computer-Aided Design* 40.3 (2008): 281-292.
- [163] Okun, Oleg, and Helen Priisalu. "Unsupervised data reduction." *Signal Processing* 87.9 (2007): 2260-2267.
- [164] Russell, Stephen, and Victoria Yoon. "Applications of wavelet data reduction in a recommender system." *Expert Systems with Applications* 34.4 (2008): 2316-2325.
- [165] Polat, Kemal, and Salih Güneş. "A novel data reduction method: distance based data reduction and its application to classification of epileptiform EEG signals." *Applied Mathematics and Computation* 200.1 (2008): 10-27.
- [166] Vanlanduit, S., et al. "Reduction of large frequency response function data sets using a robust singular value decomposition." *Computers & structures* 84.12 (2006): 808-822.
- [167] Vanherzeele, J., et al. "Data reduction using a generalized regressive discrete Fourier series." *Journal of sound and vibration* 298.1 (2006): 1-11.
- [168] Toplak, Werner. "A comparison of SOM and ALEV for data reduction purposes in transport telematics." *Proc. of 26th IASTED International Conference on Artificial Intelligence and Applications, CD Proceedings ISBN. 2008.*
- [169] Yang, Jian, et al. "Two-dimensional PCA: a new approach to appearance-based face representation and recognition." *IEEE transactions on pattern analysis and machine intelligence* 26.1 (2004): 131-137.
- [170] Otsu, Nobuyuki. "Mathematical studies on feature extraction in pattern recognition." *Electrotechnical Laboratory, Researches* 818 (1981): 220.
- [171] Itoh, Hayato, et al. "Dimension reduction methods for image pattern recognition." *International Workshop on Similarity-Based Pattern Recognition*. Springer, Berlin, Heidelberg, 2013.
- [172] Lee, Kichun, Alexander Gray, and Heeyoung Kim. "Dependence maps, a dimensionality reduction with dependence distance for high-dimensional data." *Data Mining and Knowledge Discovery* 26.3 (2013): 512-532.
- [173] Wang, Jingyan, et al. "Iterated large-margin discriminant analysis for feature dimensionality reduction in medical image retrieval." *International Journal of Biomedical Engineering and Technology* 7.2 (2011): 116-134.
- [174] Kieseberg, Peter, et al. "Protecting anonymity in data-driven biomedical science." *Interactive Knowledge Discovery and Data Mining in Biomedical Informatics*. Springer Berlin Heidelberg, 2014. 301-316.
- [175] De Hert, Paul, and Vagelis Papakonstantinou. "The proposed data protection Regulation replacing Directive 95/46/EC: A sound system for the protection of individuals." *Computer Law & Security Review* 28.2 (2012): 130-142.
- [176] Malle, Bernd, et al. "The right to be forgotten: towards machine learning on perturbed knowledge bases." *International Conference on Availability, Reliability, and Security*. Springer International Publishing, 2016.
- [177] <https://ercim-news.ercim.eu/en107/special/privacy-aware-machine-learning-and-the-right-to-be-forgotten>
- [178] [https://stanford.edu/~jduchi/projects/DuchiJoWa12\\_nips.pdf](https://stanford.edu/~jduchi/projects/DuchiJoWa12_nips.pdf)
- [179] Hira, Zena M., and Duncan F. Gillies. "A review of feature selection and feature extraction methods applied on microarray data." *Advances in bioinformatics* 2015 (2015).
- [180] Blum, Avrim L., and Pat Langley. "Selection of relevant features and examples in machine learning." *Artificial intelligence* 97.1 (1997): 245-271.
- [181] John, George H., Ron Kohavi, and Karl Pflieger. "Irrelevant features and the subset selection problem." *Machine learning: proceedings of the eleventh international conference*. 1994.
- [182] Swapna, Ch Swetha, V. Vijaya Kumar, and J. V. R. Murthy. "A Novel Approach for Feature Selection." *Information Systems Design and Intelligent Applications*. Springer, New Delhi, 2015. 877-885.