International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Articles available online http://www.ijoer.in

Vol.3., Issue.3, 2015

REVIEW ARTICLE

ISSN: 2321-7758

IMAGE CHARACTER RECOGNITION AND FEATURE EXTRACTION TECHNIQUES FOR IMAGE PROCESSING: A REVIEW

SHRITOSH KUMAR[1], VISHAL SHRIVASTAV[2]

[1]M.Tech, Student, Department of Computer Science & Engineering, Arya College of Engg. & IT, Jaipur, India

[2]Professor, Department of Computer Science & Engineering, Arya College of Engg. & IT, Jaipur, India

Article Received: 16/04/2015 | Article Revised on:28/04/2015 | Article Accepted on:01/05/2015

ABSTRACT

Feature extraction plays a vital role in the analysis and interpretation of remotely sensed data. The two important components of Feature extraction are Image enhancement and information extraction. Image enhancement techniques help in improving the visibility of any portion or feature of the image. Information extraction techniques help in obtaining the statistical information about any particular feature or portion of the image. This presented work focuses on the various feature extraction techniques and area of optical character recognition is a particularly important in Image processing.

Keywords— Image character recognition, Methods for Feature Extraction, Basic Gabor Filter, IDA, and PCA.

INTRODUCTION

When the input data to an algorithm is too large to be processed and it is suspected to be redundant, then it can be transformed into a reduced set of features vector. This process is called *feature extraction*. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which overfits the training sample and generalizes poorly to new samples. Feature extraction involves reducing the amount of resources required to describe a large set of data.

Feature extraction

The processing helps in maximizing clarity, sharpness and details of features of interest towards information extraction and further analysis. The area of optical character recognition is a particularly important in Image processing:

1. Low-level

a) Edge detection

Edge detection is one of the fundamental steps in image processing, image analysis, image pattern recognition, and computer vision techniques. It is the name for a set of mathematical methods which aim at identifying points in a digital image at which the image brightness changes sharply or, more formally, has discontinuities.

Methods for edge detection: Search-based and zero-crossing

Fig. 1 Canny edge detection applied to a photograph

b) Corner detection

A corner can be defined as points for which there are two dominant and different edge directions in a local neighborhood of the point. It can be also defined as the intersection of two edges.

Methods for corner detection: Moravec corner detection algorithm.

Applications of Corner detection: motion detection, image registration, video tracking, image mosaicing, panorama stitching, 3D modelling, and object recognition.

c) Blob detection

Blob **detection** can be defined as mathematical methods that are aimed at detecting regions in a digital image that differ in properties, such as brightness or color, compared to areas surrounding those regions.

Methods for blob detection: Laplacian of Gaussian, differential methods, local extrema methods

Applications of blob detection: object recognition, object tracking, histogram analysis, segmentation, and texture analysis.

d) Ridge detection

the ridges (or the ridge set) of a smooth function of two variables are a set of curves whose points are, in one or more ways to be made precise below, local maxima of the function in at least one dimension. This notion captures the intuition of geographical ridges. For a function of variables, its ridges are a set of curves whose points are local maxima in dimensions. In this respect, the notion of ridge points extends the concept of a local maximum.

Applications of Ridge detection: intensity landscape, scale invariant skeleton, the Blum's medial axis transform, Applications of Ridge detection, roads in aerial images, blood vessels in retinal images, and three-dimensional magnetic resonance images.

e) Scale-invariant feature transform

Scale-invariant feature transform algorithm to detect and describe local features in images. SIFT key points of objects are first extracted from a set of reference images and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors.

Applications of Scale-invariant feature transform: object recognition, robotic mapping, navigation, Panorama stitching, gesture recognition, video tracking, match moving. human action recognition, and Brain in 3D Magnetic Resonance Images.

Methods for Scale-invariant feature transform detection: Scale-space extrema detection, Keypoint localization, Orientation assignment, Keypoint descriptor, and Theoretical explanation.

2. Curvature

a) Changing intensity

b) Autocorrelation

Autocorrelation is a mathematical tool for finding identifying the missing fundamental frequency in a signal implied by its harmonic frequencies or repeating patterns, such as the presence of a periodic signal obscured by noise.It is often used in signal processing for analyzing series of values or functions, such as time domain signals.

Methods for Autocorrelation detection: Statistics, and Signal processing.

3. Image motion

a) Area based

b) Optical flow

c) *Motion detection*

Motion detection is the process of detecting a change in the surroundings relative to an object or position of an object relative to its surroundings. It can be achieved by either mechanical or electronic methods.

Motion can be detected by: Infrared, Sound, Optics, Vibration, Magnetism, and Radio Frequency Energy

Motion detecting devices include: Sony Computer Entertainment's, HP's Swing, Nintendo's, Microsoft Corporation's, and ASUS Eee Stick.

4. Shape based

a) Thresholding

Threshold technique is used to create binary images from a grayscale image. It is simplest method of image segmentation. Sezgin and Sankur (2004) categorize threshold methods into the following six groups based on the information the algorithm manipulates (Sezgin et al., 2004):

Methods for Threshold detection: Histogram shape-based methods, Multiband thresholding, Object Attribute, Local methods, Clustering-based methods, Spatial methods, and Entropy-based methods

b) Template matching

Template matching is a processing for finding small parts of an image which match a template image in digital image.

Methods for Template matching detection: Feature-based approach, Template-based approach, Motion tracking, and occlusion handling

c) Hough transform

The Hough transform is a sub-problem often arises of detecting simple shapes, such as straight lines, circles or ellipses. It is automated analysis of digital images.

Methods for Hough transform detection: Using the gradient direction to reduce the number of votes, Kernel-based Hough transform, Hough transform of curves, and its generalization for analytical and non-analytical shapes, Circle detection process, Detection of 3D objects, Using weighted features, and Carefully chosen parameter space.

METHODS FOR FEATURE EXTRACTION OR REDUCTION TECHNIQUES

1. Multifactor dimensionality reduction

Multifactor dimensionality reduction (MDR) was designed specifically to identify interactions among discrete variables that influence a binary outcome and is considered a nonparametric alternative to traditional statistical methods such as logistic regression. It is a data mining approach for detecting and characterizing combinations of attributes or independent variables that interact to influence a dependent or class variable. MDR method is a constructive induction algorithm that converts two or more variables or attributes to a single attribute. This process of constructing a new attribute changes the representation space of the data.

1. Nonlinear dimensionality reduction

Nonlinear dimensionality reduction Algorithms that operate on high-dimensional data tend to have a very high time complexity. High-dimensional data, meaning data that requires more than two or three dimensions to represent, can be difficult to interpret. Plot of the two-dimensional points those results from using a NLDR algorithm. One approach to simplification is to assume that the data of interest lie on an embedded non-linear manifold within the higher-dimensional space.

Applications of Nonlinear dimensionality reduction: consider a robot that uses a camera to navigate in a closed static environment. The images obtained by that camera can be considered to be samples on a manifold in high-dimensional space, and the intrinsic variables of that manifold will represent the robot's position and orientation.

2. Isomap

The Isomap algorithm provides a simple method for estimating the intrinsic geometry of a data manifold based on a rough estimate of each data point’s neighbors on the manifold. it is used for computing a quasi-isometric, low-dimensional embedding of a set of high-dimensional data points. It is highly efficient and generally applicable to a broad range of data sources and dimensionalities. It defines the geodesic distance to be the sum of edge weights along the shortest path between two nodes.

3. Latent semantic analysis

A matrix containing word counts per paragraph is constructed from a large piece of text and a mathematical technique called singular value decomposition (SVD) is used to reduce the number of rows while preserving the similarity structure among columns. It is assumes that words that are close in meaning will occur in similar pieces of text. It is a technique in natural language processing, in particular in vectorial semantics, of analyzing relationships between a set of documents and the terms they contain by producing a set of concepts related to the documents and terms.

Applications of Latent semantic analysis are: data clustering, document classification, cross language retrieval, synonymy and polysemy, information retrieval,

Limitations of LSA's:

- LSA cannot capture polysemy.
- The resulting dimensions might be difficult to interpret.
- The probabilistic model of LSA does not match observed data.
- Limitations of bag of words model (BOW).

4. *Partial least squares regression*

Partial least squares regression is used to find the fundamental relations between two matrices (*X* and *Y*), i.e. a latent variable approach to modeling the covariance structures in these two spaces. It was introduced by the Swedish statistician Herman Wold, who then developed it with his son, Svante Wold. It finds a linear regression model by projecting the predicted variables and the observable variables to a new space, it is a statistical method that bears some relation to principal components regression; instead of finding hyperplanes of minimum variance between the response and independent variables.

Applications of Partial least squares regression are: social sciences, econometrics, marketing, strategic management, bioinformatics, sensometrics, neuroscience and anthropology.

5. *Independent component analysis*

Independent component analysis (ICA) is a special case of blind source separation. It is use in signal processing for separating a multivariate signal into additive subcomponents.

The broadest definitions of independence for ICA are:

- Maximization of non-Gaussianity.
- Minimization of mutual information.

The basic establishment of ICA separation of mixed signals gives very good results are based on two assumptions and three effects of mixing source signals.

TABLE I: ICA ASSUMPTIONS AND EFFECTS OF MIXING SOURCE SIGNALS

Assumptions are:	**Effects of mixing source signals are:**
• The values in each source signal have non-gaussian distributions. • The source signals are independent of each other.	• Independence • Normality • Complexity:

6. *Autoencoder*

Autoencoder is based on the concept of sparse coding proposed in a seminal paper by Olshausen et al. in 1996. The aim of an auto-encoder is to learn distributed representation for a set of data, a compressed, typically for the purpose of dimensionality reduction. It is an artificial neural network used for learning efficient coding.

7. *Principal component analysis*

Principal component analysis was invented in 1901 by Karl Pearson, as an analogue of the principal axis theorem in mechanics. PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. PCA can be done by eigenvalue decomposition of a data correlation matrix or singular value decomposition of a data matrix, usually after normalizing the data matrix for each attribute. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance, and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components.

Applications Principal component analysis: Neuroscience

8. *Kernel principal component analysis*

Kernel principal component analysis is use in field of multivariate statistics. It is an extension of principal component analysis (PCA) using techniques of kernel methods. Using a kernel, the originally linear operations of PCA are done in a reproducing kernel Hilbert space with a non-linear mapping.

9. *Multilinear principal component analysis*

MPCA is a multilinear extension of principal component analysis (PCA). The major difference is that PCA needs to reshape a multidimensional object into a vector, while MPCA operates directly on multidimensional objects through mode-wise processing.It is a mathematical procedure that uses multiple orthogonal transformations to convert a set of multidimensional objects into another set of multidimensional objects of lower dimensions. It is a basic algorithm for dimension reduction via multilinear subspace learning.There is one orthogonal transformation for each dimension; hence *multilinear*. This transformation aims to capture as high a variance as possible, accounting for as much of the variability in the data as possible, subject to the constraint of mode-wise orthogonality.

10. Multilinear subspace learning

Multilinear subspace learning is a dimensionality reduction approach for finding a low-dimensional representation with certain preferred characteristics of high-dimensional tensor data through direct mapping, without going through vectorization. MSL aims to learn a specific small part of a large space of multidimensional objects having a particular desired property. MSL methods are higher-order generalizations of linear subspace learning methods such as principal component analysis (PCA), linear discriminant analysis (LDA) and canonical correlation analysis (CCA).

11. Maximum variance unfolding

Maximum variance unfolding creates a mapping from the high dimensional input vectors to some low dimensional Euclidean vector space. It can be viewed as a non-linear generalization of Principal component analysis. It is to exploit the local linearity of manifolds and create a mapping that preserves local neighborhoods at every point of the underlying manifold.

TABLE II: Type of Algorithms for Multilinear subspace learning:

Multilinear extension of PCA	**Multilinear extension of LDA**	**Multilinear extension of CCA**
TTP-based	TTP-based	***TTP-based***
TVP-based	TTP-based	***TVP-based***
	TVP-based	

12. Basic Gabor Filter

A Gabor filter is a complex exponential modulated by a Gaussian function in the spatial domain. A Gabor filter can be represented by the following equation :

$$\Psi (x, y, \lambda, \theta) = \frac{1}{2\Pi SxSy} e^{-1/2\left(\frac{x'^2}{Sx \cdot Sx}+\frac{y'^2}{Sy \cdot Sy}\right)} e^{j2\Pi x'/\lambda}$$

where (x,y) is the pixel position in the spatial domain, λ is the wavelength (a reciprocal of frequency) in pixels, θ is the orientation of a Gabor filter, and S_x, S_y are the standard deviation along the x and y directions respectively. The parameters x' and y' are given as equation 2.3

$x' = x\cos\theta + y\sin\theta$

$y' = -x\cos\theta + y\sin\theta$

The amplitude and phases of Gabor filter bank both provide valuable cues about specific pattern present in images. The amplitude contains directional frequency spectrum information and a phase contains information about the location of edges and image details. Gabor filters with different frequencies and orientations are very effective in capturing local Information present in images. The Gabor features are calculated by convolution of input image with Gabor filter bank.

I(x, y) is a grey-scale ear image of size M* N pixels. The feature extraction procedure can then be defined as a filtering operation of the given ear image I(x, y) with the Gabor filter u, v(x, y) of size u and orientation v .

$G_{u,v}(x, y) = I(x, y) * \Psi(x, y)$

In Gabor feature extraction approach Holistic approach is used in which the features are extracted from the whole image. Gabor filters are applied on images to extract features aligned at particular angle (orientation). The most important parameter of Gabor filter is orientation and frequency. Certain features that share the similar orientation and frequency can be selected and used to differentiate between different ear biometrics depicted in image.

The Gabor feature representation $|o(x,y)|_{m,n}$ of an image I(x,y), for x=1,2,.....N, y=1,2,....M, m=1,2...M , n=1,2,....N, is calculated as the convolution of the input image I(x,y) with Gabor filter bank function $\Psi(x,y, \lambda_m,\theta_n)$.The convolution operation is performed separately for real and imaginary part[37].

$Re(O(x,y))_{m,n} = I(x, y) * Re(\psi(x,y,\lambda_m,\theta_n))$

$Im(O(x,y))_{m,n} = I(x, y) * Im(\psi(x,y,\lambda_m,\theta_n))$

This is followed by the amplitude calculation,

$O(x,y)_{m,n} = ((Re(O(x,y))_{m,n})^2 + (Im(O(x,y))_{m,n})^2)^{1/2}$

Algorithm of Existing Feature Extraction Technique

Input: Gray scale Image Img

Step 1: Loop scale (5 different scales)

Imgresize= image Img is resized with scale.

Step2: Loop theta= [0, 30, 60, 90,120,150,180]

Apply Gabor equation (2.2) with resized image Imgresize and angle theta.

Step3: End Loop

Step4: End Loop

Step 5: For reducing the dimensions use sampling filtering by 25.

Step 6: For reducing the more dimensions use sampling filtering by 25 again.

Step7: Features is given to classifier

Advantage of Gabor Filter Method

(i) Gabor filters have the optimal localization property in spatial and frequency domain.

(ii) Gabor Filters are more tolerant to illumination variation as compare to DCT.

Disadvantage of Gabor Filter Method

- Huge Dimension
- High Redundancy.

Conclusions

In this paper we discuss different type of optical character recognition is a particularly important in Image processing and different type of Feature Extraction or reduction techniques. In this paper give detail of Feature Extraction techniques like PCA, LDA, ICA, NDR, and Gabor etc. . On the basic of different paper review on Feature Extraction we think Gabor Filter is one of the best techniques. It is hoping that this paper bring out understand and inspiration amongst the research group of Feature Extraction.

References

[1.] Comon, Pierre (1994): "Independent Component Analysis: a new concept?", Signal Processing, 36(3):287–314 .

[2.] Comon, P.; Jutten C., (2010): Handbook of Blind Source Separation, Independent Component Analysis and Applications. Academic Press, Oxford UK. ISBN 978-0-12-374726-6.

[3.] Markovsky I. (2012) Low-Rank Approximation: Algorithms, Implementation, Applications, Springer, 2012, ISBN 978-1-4471-2226-5

[4.] Gashler, M. and Martinez, T., Temporal Nonlinear Dimensionality Reduction, In Proceedings of the International Joint Conference on Neural Networks IJCNN'11, pp. 1959–1966, 2011

[5.] The illustration is prepared using free software: E.M. Mirkes, Principal Component Analysis and Self-Organizing Maps: applet. University of Leicester, 2011

[6.] Banafshe Arbab-Zavar, Mark S. Nixon," Robust Log-Gabor Filter for Ear Biometrics", in IEEE, International Conference, IEEE, 2008.

[7.] Steven Cadavid, Mohammad H. Mahoor, Mohamed Abdel-Mottaleb, "Multi-modal Biometric Modeling and Recognition of the Human Face and Ear ", in IEEE International Conference, IEEE, 2013.

[8.] Andrew AS, Nelson HH, Kelsey KT, Moore JH, Meng AC, Casella DP, Tosteson TD, Schned AR, Karagas MR. Concordance of multiple analytical approaches demonstrates a complex relationship between DNA repair gene SNPs, smoking and bladder cancer susceptibility. Carcinogenesis. 2006 May;27(5):1030-7

[9.] Modeling word perception using the Elman network, Liou, C.-Y., Huang, J.-C. and Yang, W.-C., Neurocomputing, Volume 71, 3150–3157 (2008), doi:10.1016/j.neucom.2008.04.030.

[10.] Tenenhaus, M. (2008). "Component-based structural equation modelling"

[11.] Lu, Haiping; Plataniotis, K.N.; Venetsanopoulos, A.N. (2011). "A Survey of Multilinear Subspace Learning for Tensor Data" (PDF). Pattern Recognition 44 (7): 1540–1551. doi:10.1016/j.patcog.2011.01.004.

[12.] Pearson, K. (1901). "On Lines and Planes of Closest Fit to Systems of Points in Space" (PDF). Philosophical Magazine 2 (11): 559–572. doi:10.1080/14786440109462720.

[13.] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology, 24, 417–441, and 498–520.

[14.] Jensen, Jeppe. "Hough Transform for Straight Lines"

[15.] J. M. Park and Y. Lu (2008) "Edge detection in grayscale, color, and range images", in B. W. Wah (editor) Encyclopedia of Computer Science and Engineering, doi 10.1002/9780470050118.ecse603

[16.] D. Lowe (2004). "Distinctive Image Features from Scale-Invariant Keypoints". International Journal of Computer Vision 60 (2): 91. doi:10.1023/B:VISI.0000029664.99615.94.(DOG blob detection with automatic scale selection)

[17.] J. Matas, O. Chum, M. Urban and T. Pajdla (2002). "Robust wide baseline stereo from maximally stable extremum regions" (PDF). British Machine Vision Conference. pp. 384–393.(The MSER blob detector)

[18.] T. M. Phuong, Z. Lin et R. B. Altman. Choosing SNPs using feature selection. Proceedings / IEEE Computational Systems Bioinformatics Conference, CSB. IEEE Computational Systems Bioinformatics Conference, pages 301-309, 2005.

[19.] Fernandes, L.A.F. and Oliveira, M.M., "A general framework for subspace detection in unordered multidimensional data," Pattern Recognition, Elsevier, Volume 45, Issue 9, pp. 3566–3579 (September, 2012).

[20.] Warmuth, M. K.; Kuzmin, D. (2008). "Randomized online PCA algorithms with regret bounds that are logarithmic in the dimension". Journal of Machine Learning Research 9: 2287–2320.

[21.] H. Zha, C. Ding, M. Gu, X. He and H.D. Simon (Dec 2001). "Spectral Relaxation for K-means Clustering" (PDF). Neural Information Processing Systems vol.14 (NIPS 2001) (Vancouver, Canada): 1057–106