The Weighted Feature Selection Method

*Ikhlas Watan Ghindawi **Mustafa S. Kadhm ***Duaa Enteesha Mhawi
Ikhlas_watan@yahoo.com muit.salam@sadiq.edu.iq dododuaaenteesha@yahoo.com

* Computer Science Department - Al Mustansiriya University
** Computer Engineering - Imam Ja'afar Al-Sadiq University
***TechniquesMiddle technical university technical institute for administration

Abstract

Feature selection problem is a critical issue in multiple important fields like : Pattern Recognition, Classification ,and Images Matching. This paper develop a method to select an optimum number of important features from feature set. This paper present the Weighted Feature Selection Method that gives each extracted feature a weight value depending on it's location and it's neighbors.

KeyWords
Image Processing , Feature Selection, Feature Generation , Feature Extraction

الخلاصة

مشكلة تحديد العناصر هي مشكلة حرجة في مجالات مهمة متعددة مثل: التعرف على الأنماط، والتصنيف، ومطابقة الصور. يطور هذا البحث طريقة لاختيار العدد الأمثل من الصفات الهامة من مجموعة كبيرة من الصفات . يقدم هذا البحث طريقة لتحديد الصفة المهمة في الصور من خلال اعطاء كل صفة قيمة معينة تمثل وزنها اعتمادا على موقعها والجيران.

1- Introduction

Feature selection is an imperative information handling venture in design acknowledgment and order issues, and it has been effectively connected to many fields [1]. The procedure of feature selection is as follows: in accordance with pre-designed selection criteria, the most important features of the given input data are selected by the optimal operations under the prefixed criterion and the remaining features are removed from the input to reduce the data amount. In addition the feature selection, as an important pre-processing operation, plays an important role in the data pre-processing: Feature selection can eliminate the redundancy, interference, noise and less important data in input, feature selection can improve the accuracy ,after the feature selection operation, a large number of non-relevant data which contain many interferential components are removed , and feature selection can improve the operational efficiency[2,3].

Because of the importance of feature selection , it has been a hot point and difficult issues in this research area. How to determine an effective feature selection criterion, to extract the important features and to verify the validity, reliability, applicability, robustness and computational time[4,5] .This research integrated a weighted method for feature selection, which, in principle, work with any image processing system using feature selection approach.

2. The Related Works

Feature selection has been broadly used to diminish calculation time and enhance precision. One popular technique is RELIEF [6] that assigns weights to a particular feature based on the differences between the values of nearest neighbor pairs. Cao et al. [7] further developed this strategy by learning feature weights in piece space, pruning ceaselessly pointless features. Hermes and Buhmann [8]begun by building a SVM classifier utilizing every accessible element, and recursively evacuating those that have minimal effect on the choice capacity if expelled. So also, Avidan [9] utilized a voracious consecutive forward determination strategy to discover a subset of highlights and bolster vectors that rough the SVM arrangement got utilizing every single accessible component.Multi-class SVM was used in [10] to select the most informative features for face recognition. The proposed SVM-DFS canspeed up classification without degrading the matching accuracy. Mahamud and Hebert [11] pproposed discriminative object parts determination and utilized contingent dangers as the separation measure in closest neighbor look. Dorko and Schimid [12] presented a technique for choosing most discriminative object part classifiers in view of probability proportion and common data. None of these methodologies concentrates on pivot invariance or uses the extra data presented by particularly planned and marked preparing sees.

3.The Proposed Methodology

feature determination has been widely examined in machine learning and statistics in the course of the most recent couple of years. The issue is characterized as finding a subset of features that is adequate to encode (e.g. unsupervised) or foresee (e.g. supervised) target names.This method, work on choosing the subset of image features that is considered as important features . This section describes a proposed method in following steps:

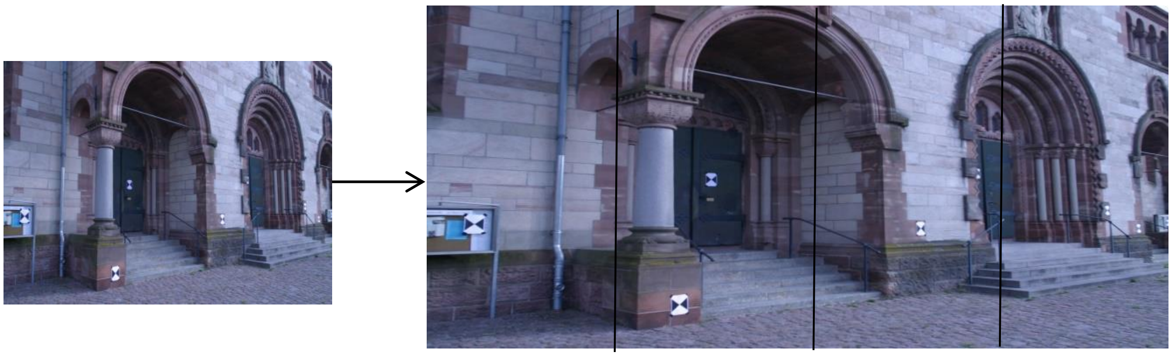
o

1- Divide each image to multiple equal regions or Sub images.
2- For each Sub Image :
   a. Noise removal using Averaging filter.
   b. Find the features using SIFT Algorithm.
   c. Select randomly 8 features.
   d. Giving weight value for each feature by :
      i. Build cluster for each selected feature by k_MeansAlgorithm
      ii. The feature that has the biggest cluster , take the highest weight value
3- From each Sub Image will take two features, which has the highest weight value.

The following sub sections review all steps of weighted feature selection method in details.

3.1 Divide the Image

The primary point of this paper is to create calculation that ideally select a subset has k features from N image features . This subset is distributed to whole image without exception, so the image is divided into eight sub images and work later on each sub image separately to get the highest possible accuracy in detect and select the important features.512×512 images have been divided to eight sub images (A_i) each one have 64×64 pixels. Show Figure 1.



(a) Image (b) $A_1..A_8$ Sub Images

Figure 1: Divide an image to eight sub images

3.2 Remove The Noise

Noise removing is the way toward decreasing commotion from an digital image. The primary motivation behind noise removing is to improve the quality of an image. Generally images are debased with the noise when they are transmitted or amid image obtaining process. In the spatial area, sifting relies upon position and its neighbors. Average filter the most straightforward and the least demanding technique to actualize the smoothing of images i.e. limiting the degree of power variety among neighboring pixels. It is additionally used to limit noise in images as often as possible. The second step is separating and remove the noise from Sub Image utilizing Average Filter .In this step the filtering concept is to replace the processing pixel element value in an image matrix with the average value of its neighbors including itself. Repeatedly it removes those pixel values which are not representing their surroundings. Average filter is somewhat equivalent to the convolution filter. This filter also relies on kernel to represent the shape and size of the vicinity to be sampled while mean is calculated. Average Filter reads the input image and sets the image header information for the output image. It applies the smoothing algorithm and sets the output image as input for the next iteration if it's not the last iteration and writes the output image after completion of all the iterations. The mean algorithm works by adding all the surrounding or neighbor pixel values and takes the mean or average of those values. The resulting value is placed in the central pixel.

3.3 Detect the Features

In this step, the SIFT detector has been used to extract the features from each sub image. The SIFT finder utilizes as keypoints image structures which look like "blobs". By looking for blobs at numerous scales and positions, the SIFT detector is invariant (or, more accurately, covariant) to translation, rotations, and re scaling of the image. The SIFT (Scale Invariant Feature Transform) is an approach for extracting distinctive invariant features from images. It is
widely used in feature extraction. SIFT descriptor detects the extreme points through the whole scale space. Four major stages in SIFT are (1)scale space local extrema detection. (2)accurate keypoints localization.(3)orientation assignment.(4)keypoints descriptor. First step, the features locations are determined as the local extreme a of Difference of Gaussians (DOG pyramid).

Let I (x, y) be an image and S (x, y,σ) the scale space of I,

which is defined as

$$S(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

Where * is the convolution operation in x and y and G (x, y, σ) is a variable-scale Gaussian defined as

$$G(x; y; \sigma) = 1/2\pi\sigma * e^{-(x2+y2)/2\sigma 2} \quad (2)$$

The scale-space extreme a detection starts with the detection of local maxima and minima of D(x,y,σ) where

$$D(x,y,\sigma) = (G(x, y, k) - G(x, y,\sigma)) * I(x, y)$$

$$= S(x, y, k) - S(x, y, \sigma) \quad (3)$$

The detection is performed by searching over all scales and image locations in order to identify potential interest points that are invariant to scale and orientation. Second step is to accurately localize the keypoints. This is performed by rejecting those keypoints, which have low contrast or are poorly localized along an edge. Unstable extreme as are detected by considering a threshold over the extremum of the Taylor expansion of D(x,y, σ) .

Once the SIFT- feature location is determined , next step is to assign orientation to each feature based on local image gradient. The gradient's absolute value and direction are given by

$$m(x,y) = \sqrt{(S(x+1,y) - S(x-1,y))^2 + (S(x,y+1) - S(x,y-1))^2} \quad \textbf{(3)}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{S(x,y+1) - S(x,y-1)}{S(x+1,y) - S(x-1,y)}\right) \quad \textbf{(4)}$$

The final stage of the SIFT approach is to build descriptor for each keypoints. The gradient magnitudes and orientations within each feature are computed by and weighted by appropriate Gaussian window, and the coordinate of each pixel and its gradient orientation are rotated relative to the keypoints orientation.

3.4 Weighted Extracted Features

In this step, the weight values of each extracted feature from the Sub Image are given. By selecting randomly eight features from feature set and implement the K- Means Algorithm to build a cluster for each of these selected features. The size of cluster determine the value of feature's weight . The weight value equal to number of features in cluster, where each Sub Image produced eight clusters and will select the two features , which has the highest weight values. Finally , this method will produce 16 features from whole image.K-means clustering algorithm is a simple unsupervised learning algorithm. This clustering algorithm attempts to partition *n* given data objects into a set of diverse clusters whereas every data object belongs to a specific cluster that has the nearest mean. K-means algorithm includes two separate stages. In the assignment stage, the algorithm chooses *k* centers (here we choice k=8). Then, it attempts to assign every data object or observation to the closest center. The Euclidean distance is commonly considered to compute the distance between the cluster centroids or centers and each observation. When all of the observations are assigned to the given *k* clusters, the primary stage is accomplished and the early clustering is done. In the next stage, the algorithm recalculates the means of the prior created clusters to be the centers of the new clusters. This procedure continues iteratively until the convergence of the criterion function. The criterion function is presented as follow:

$$E = \sum_{i=1}^{n} (xi - \widetilde{xi})^2 \qquad (4)$$

where E denotes the sum of the squared error (SSE) for all data objects

4. Experimental Results

For implement the method the JPG images of historic building with 512×512 pixels have been used . Figure 2 show some examples of these images.

Figure 2: Some example images

To study the accuracy of the proposed strategy , it implemented 20 times on each image in the data base , the result proved that, 93% of extracted features from each image are similar for the 20 times . Also the execution time is constant and does not change at all 20 execution times.

5. Conclusions

This paper has been presented a developed method for identifying the important features of any image . For this purpose, SIFT image features have been extracted after divide the image into eight Sub Images. Furthermore, the K- Means Algorithm has been used to cluster predefined features and give it a weight value . This method deal with the weighted features as a basic and fundamental contents of image. The results show that, this method can detect the important features with high accuracy without ignore any feature and gives the guarantees to ignore any noise or irrelevant features. This method detect the features which are suitable for image matching and classification.

6. References

[1] X. Chen, T. Fang, H. Huo, et al. Graph-based Feature Selection forObject-Oriented Classification in VHR Airborne Imagery [J]. IEEE Transactions onGeoscience and Remote Sensing, 2011, 49 (1): 353-365.

[2] Y.J. Li, D.F. Hsu and S.M. Chung. Combining Multiple Feature SelectionMethods for Text Categorization by Using Rank-Score Characteristics [C].21st InternationalConference on Tools with Artificial Intelligence (ICTAI). 2009: 508-517.

[3] Q.W. Wang, B.Y. Li and J.L. Hu. Feature Selection for Human ResourceSelection based on Affinity Propagation and SVM Sensitivity Analysis [C]. World Congresson Nature & Biologically Inspired Computing (NaBIC), 2009: 31-36.

[4] K.C. Khor, C.Y. Ting and S.P. Amnuaisuk.Forming an Optimal Feature Set forClassifying Network Intrusions Involving Multiple Feature Selection Methods [C]. 2010, International Conference on Information Retrieval & Knowledge Management (CAMP),2010: 179-183.

[5] H. Nguyen, K. Franke and S. Petrovic.Improving Effectiveness of IntrusionDetection by Correlation Feature Selection [C]. International Conference on Availability,Reliability, and Security (ARES), 2010: 17-24.

[6] K. Kira and L. A. Rendell. The feature selection problem:Traditional methods and a new algorithm. In AAAI, pages129–134, 1992.

[7] B. Cao, D. Shen, J.-T.Sun, Q. Yang, and Z. Chen. Featureselection in a kernel space. In International Conference onMachine Learning, 2007.

[8] L. Hermes and J. Buhmann. Feature selection for supportvector machines. In ICPR, 2000.

[9] S. Avidan. Joint feature-basis subset selection.In CVPR,2004.

[10] G. Dorko and C. Schmid.Selection of scale-invariantparts for object class recognition. In IEEE InternationalConference on Computer Vision, pages 634–639, 2003.

[11]S. Mahamud and M. Hebert. The optimal distance measurefor object detection. In IEEE Computer SocietyConference on Computer Vision and Pattern Recognition,pages 245–255, June 2003.

[12] Z. Fan and B. Lu.Fast recognition of multi-view faceswith feature selection. In IEEE International Conferenceon Computer Vision, pages 76–81, Oct. 2005.