چهارمین کنفرانس ملی ایده های نو در مهندسی برق

۲۰و۲۱ آبان ماه ۱۳۹۴ - دانشگاه آزاد اسلامی واحد اصفهان (خوراسگان)

A Review of Object Detection Methods in the Image

*Seyyed Mohammad Reza Hashemi1, Ali Broumandnia 2

1Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran

Email: smr.hashemi@qiau.ac.ir

2Faculty of Computer and Information Technology Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran

Email: Broumandnia@azad.ac.ir

Abstract: *Object detection in images is a process whose goal is to determine the position of the objects in the image along with identifying them. Challenges, facing object detection, such as determination of the lighting, changing the perspective, image resizing, image extension, and rotating the original image, have led to miscellaneous methods. Geometrical methods deal with finding the primary lines, while the visual one aims at the way the pixels are connected. In each method, the discerning features among different image groups should be extracted. This paper reviews several object detection methods, including the intended-area identifying methods along with the area describing ones.*

Keywords: : image detection, geometrical features, feature extraction, object detection

1. Introduction

Image processing science has been formed as the first computers were being produced. There are various conferences, held annually on machine sight, computer sight, pattern recognition, and artificial intelligence, whose main goal is to write algorithms with which one can perform image-processing procedures on pictures, taken by cameras. For segmentation and interpretation purposes, many algorithms have focused on two dimensional T images which may include landscapes, medical images, artificial images, meteorological pictures, text-related images, diagrams and mathematical equation figures, and any image, produced by different types of cameras, telescope, microscope, medical imaging devices, etc. Some of these procedures are related to deleting and promoting the image's quality but its chief part is about object detection. The term object is given to all natural and artificial things, present in the image. Object detection process is about the easiest to the most complicated analysis in which seeing an image leads to concluding it. Object detection could involve image segmentation, identification of artificial and natural objects in the image, changing signs and hand-written words to typed text, determining the patient's status via medical images, determining the status of human body members, identification based on the face, determining an individual's happiness and sadness based on his facial expression, determining the social, economic, and political status of a society, and determining and solving mathematical problems. To detect the shapes, the computer technology is related to computer sight and image processing which deals with identifying conceptual samples of the objects, connected to a certain category such as human detection in public places, face detection, identification of the animals along with their characteristics, cars' detection in highways, and detection of buildings, etc. in digital pictures and films. Some areas of object detection, including face detection as well as identification of the pedestrians, are among the items, widely researched and studied.

Detection Methods Based on Objects' Appearance: The set of pixels, related to an object are firstly located by area descriptor. Afterwards its feature area is extracted via describing methods.

1.1. Descriptors of the Intended Area

Such descriptors distinguish among various areas of the image. Varied methods try to increase the discriminating power so that they can detect certain objects in the images. Some of these detectors are reviewed in what follows.

1.1.1 Corner-Based Harris Descriptors

Corner-Based Harris Descriptor identifies the considered area by means of the matrix, established by second-level moment.

$$\mu = \begin{bmatrix} I_x^2(P) & T_x I_y(P) \\ I_x I_y(P) & I^2(y) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

(1)

I_y and I_x are the first level derivative of the intensity of Image I at position P, in y,x direction respectively. Power scale of corner C is calculated by avoiding the decomposition of the specific amount to the efficiency via second-level moment matrix as below:

$$C = \mathrm{Det}(\mu) - k \times \mathrm{Tr}(\mu)^2 = (AC - B^2) - k \times (A + C)^2$$

(2)

Using this equation, no-maximum amounts are omitted and Harris' corner is identified by positive responses of Function c. Harris point descriptor distributes a large amount of considered points repetitively. The main advantage of this descriptor is its calculation speed while its weakness is that it only determines the location of the intended points. The features of the intended areas such as the scale or orientation are determined for consecutive description. This detector has a constant feature with rotation.

1.1.2 Detectors Based n Hessian Matrix

Hessian Matrix detectors are based on Harris Detector method. The main principle of Hessian is introduced by Hessian Matrix in Equation 3-3. In order to identify the edges and spots via second derivative, it gives a powerful answer:

$$\mu = \begin{bmatrix} I_{xx}(P) I_{xy}(P) \\ I_{xy}(P) I_{yy}(p) \end{bmatrix}$$

(3)

I_{xy} and I_{xx} are the second derivative of the intensity of the image I at position P in directions x and y. And I_{xy} is the combination of derivative x and y in the image orientation. Selection criteria of Hessian points are to use the Hessian matrix, established after omitting the non-maximums. Detectors, based on Hessian Matrix, identify spot-like structures, in correspondence to Laplacian Operator, having a constant feature with rotation.

1.1.3 Hessian and Harris Detectors, Compatible with Scale

The idea of selecting the feature scale due to not having a constant feature is shaped with the scale, itself. The characteristics of scale space have been studied specifically. Based on these researches on scale space, spots with local extreme of Laplacian space S in equation 3-4, are used as a criterion for selecting the scale. In different references this method is often known as Laplacian Hessian or Laplacian Harris detectors. Standard deviation of Gaussian Smoothing is expressed by S for making scale space:

$$S = s^2 \times \left| I_{xx}(p) + I_{yy}(p) \right|$$

(4)

Laplacian Harris and Hessian detectors have shown similar features as comprehensive distribution, both having constant feature with scale.

1.1.4 Gaussian Difference Detector (DoC)

Gaussian Difference Function D is applied on images which have become opaque many times. For local scale S_n and S_{n+1} we have:

$$D(P.S_x) = \left(G(P.S_x) - G(P.S_{xy})\right) \times I(P)$$

(5)

$$G(P.S_x) = G\left((x,y).\, S_x\right) = \frac{1}{2\pi s^2} e^{-(x^2+y^2)/2s^2}$$

(6)

In this equation, the variant G refers to the size of Gaussian scale s; l the intensity of the image at position p; and the asterisk is convulsion operator. Gaussian difference can be measured faster than Laplacian scale, which can determine the local key point accurately. Identification is done with detecting histogram of difference. The detector showed that its behavior is similar to Hessian detector, identifying the similar spot structures. The main variance of DoC detector is that in case of image resizing, it can identify the area.

1.1.5 Edge-Based regions (EBR) and Intensity-Based Regions (IBR):

EBR Detector deals with the edges' behavior around a considered point. In this method the specific photometry quantities of calculation are being used as stopping criterion along the edge. Firstly the position of the considered point (P) and the position of the edge, obtained from stopping criterion (P, P) are defined as a relative frame. The main advantage of this detector is the fastness of its performance, compared to EBSR Detector.

On the other hand, IBR deals with the points around a point in the image with high intensity. Initially, along the radial symmetry of rays, emitted from detected points with high intensity, the specific function of image intensity f = f(x) is generated on several scales, with the stopping criterion, introduced via the local maximum. All stopping points, related to each other in an intended shape, are substituted by an oval. The performance time of this detector is better than EBR.

2. Detectors of the Intended Region

The characteristic of detectors, describing the region, are studied along with local vicinity detectors by available variant features. The concept of constant is that the detectors should be strong against various changes of the image such as causal damages, scale change, change in the lighting intensity, and artificial compressing (like jpeg). It has been seen that detectors' performance excessively depends on the power of region detectors. Wrong identification of the region's position is greatly influential on detectors' performance. Nonetheless, being strong against any position (relatively small) or identification of physical shape's errors are important features of efficient region detectors.

One of these detectors is pixel intensity vectors at the considered region. By means of vectors' interdependence the similarity of the regions is measured. Another important issue is big dimensions of these detectors for adjusting and detecting tasks. There have been many calculative attempts so that by decreasing the dimensions of the detector its discriminative power will be kept.

Detectors can be categorized into three main groups of distribution-based detectors, filter-based ones, and the rest. Distribution-based methods present a certain description of the region's features by the histograms, mostly using geometrical features of the considered points as well as local orientation information. Geometrical features are position and orientation; considered points, the edges and corners; and local orientation information, the gradients. One of the distribution-based detectors is SIFT and the gradient, pca.

2.1 SIFT Detector

It is one of the most popular detectors which has been suggested and has developed an accurate design of detectors and descriptors' combination with high efficiency. The combination of detector and descriptor is called the feature of unchanging compared transformation (SIFT) which includes the detector of an invariant region with scale, called Gaussian difference detector (DoC) as well as an appropriate descriptor, mostly referred to as key SIFT. DoC detector identifies the considered point, repeated a lot. In order to achieve an unvarying description with rotation, firstly the main orientation of the region is obtained by 36 orienting histograms from gradient orientations, which are weighted within Gaussian rotating window. Accordingly, the specific domain of the gradient m and orientation o is calculated for each pixel in the image via pixel difference as the following:

$$m=\sqrt{(I(x+I_y)-I(x-I_y))^2+(I(x,y+I)+I(x,y-1))^2}$$

(7)

$$\varnothing=\tan^{-1}\left(\frac{I(x,y+I)+I(x,y-I)}{(I(x+I,y)-(I(x-I+y)}\right)$$

(8)

The dimensions of repeatability window, estimated by DoG detector, are well determined. It is possible that more than one main orientation exists inside the presented rotating window, in which case several descriptors are established in a location with varying orientations. All weighted gradients are normalized to the main circular orientation by the descriptor. The circular region is divided around the 4*4 patched key point without any overlap with the gradient histogram of the orientations, calculated inside these patches. Smoothing the histogram has been done in order to avoid the sudden changes of the orientation and the number of bins has been decreased to 8 so that the descriptor's size will be limited. As a result an 8*8 = 128 dimension feature vector is established for each key point.

Finally the feature vector is normalized to length unit, and undergoes thresholding to decrease the impacts of linear and non-linear lighting changes. This descriptor's unvarying features with scale are due to unvarying scaled features of DoG detector. Constancy with rotation has been performed by considered main orientation determination. Constant descriptor does not have causal transformation. All the same, SIF calculation might be feasible on other kinds of detectors; therefore, it can use other detectors for scaled invariance as well as casual transformation, becoming complemented during the process.

2.1.1 PCA-SIFT or PCA Gradient

SIFT description method is a key with corrected describing dimensions. Instead of calculating the gradient histogram on DoG points, the primary component analysis (PCA) is being used. A 41*41 patch is extracted in the centralized scale on key points. Afterwards instead of using the histogram, patch description is presented by local gradient orientation with the PCA from the most important specific vectors. It has practically been shown that 20 primary vectors are efficient to appropriately present the patch.

Specific necessary space can be calculated offline. Compared to key SIFT, descriptor's size is decreased 8 times which is generally this method's advantage.

2.1.2 Gradient Local Orientation Histogram (GLOH)

GLOH is a developed form of key SIFT to achieve higher discriminative power. Instead of diving the patches around a key point in a 4*4 arranged network, they are divided in radial and arced networks. Practically 2 radii and 8 arcs divide the patches into 17 positions. A similar idea has been used for the shape content. Gradient orientation of these patches is quantized to 16 bin histograms which in fact results in the 273-dimensional descriptor. These high descriptor dimensions are reduced by PCA so that 128 special vectors, corresponding to 128 of the biggest amounts, are taken into account for the description.

2.1.3 Twisted Images

Twisted images are introduced by 3D textual object detection systems on the shape to simultaneously detect some objects in crowded scenes. Afterwards these descriptors have been developed for 2D images and the usages of the intended finding match. In practice, it uses the domain of twisted image intensity. The next 2D histogram from intensity amounts (I) and their distance from the region center (D) are called the twisted image histogram descriptor. Each row of the 2D descriptor, presenting the histogram of grey area amounts, is in a coiled distance from the center. Eventually the histogram is smoothed and the stage of normalizing is done by invariance or an amount of the produced view. Usually quantifying the intensity histogram is performed in 10 bins and 5 different radial pieces, the result of which is a 50-dimensional descriptor that is invariant and with rotation in the surface.

2.1.4 Shape Content for Area Description

Shape content descriptor uses the distribution of relative points' position. The corresponding orientations in a histogram are selected as the descriptor. Initial points are the internal or external cantor points under the study which can be identified by edge detector and are repetitively sampled in all shape twist. A complete shape presentation can be achieved by gathering all relative positions between two initial sections and paired points' orientations and the dimensions of each descriptor slowly increases as the area size grows. To decrease the inappropriate histogram size, the sample coordinates of the shape are calculated relatively. Each part of the histogram is called a bin.

This descriptor is very sensitive to sample adjacent point position. It has been shown that 5 bins for the radius (logr) and 12 bins for the 0 angle give better results compared to dimensional descriptors. Point distribution is weighted arbitrarily so that the domain histogram can give appropriate results.

3. Conclusion

Object detection has many uses in various areas of computer sight, including retrieving the image and visual supervision. Object detection is that in a certain image, whether there is a certain object or not; and if there is, the position and size of the object is determined. An object detection system can identify real objects in real images by means of object models which have been determined before. So far many characteristics and signs such as color, tissue, and movement have been presented and studied. Perhaps it can be said that shape is one of the most public and dominant of these characteristics. Shape-based approaches extract a demonstration of an object's shape, measuring the shapes' similarity by comparing with this demonstration and in this paper we reviewed and evaluated the detection methods, based on shapes' appearance.

Reference

[1] Schaffalitzky F. and zisserman A., 2002 Multi – view matching for unordered image sets or how do I organize my holiday snaps. Springer Conf. on computer vision LNCS, Vol. 2350, and 414-431.

[2] Carneiro G. and Jepson A., 2002 phase – based local features. Springer Conf. on computer vision LNCY VAOL 2350, 282 – 296.

[3] T.Tuytelaars and L.Gool, Content – based image retrieval based on local affinely invariant regions . springer inter . conf . on visual information and information system . LNCS , Vol 1614.493-500. 1999 .

[4] T.Tuytelaars and L.Gool, Matching separted views based on affine invariant regions . springer inter . journal of computer vision vol 59 . No 1,61-85.2004.

[5] K.Mikolajczk and C.Schmid, A performance evalution of local descript – tors . IEEE Trans . on pattern Analysis and Machine intelligence , vol . 27.1615-1630. 2005 .

[6] R.Davis, Machine vision Theory Algorithms , practicalities . fourth edition Elsevier inc ., 168-185. 2012 .

[7] P.Roth and M.Winter, Survey of appearance – based methods for object recognition . technical report , graz university of technology ICG-TR 01-08. 2008 .

[8] J.Lakshmi and M.Ppunithavalli, A survey performance Evalution of object Detection Techniques in Digital Image processing . International Journal of computer science issues , vol 7,86-94.2010.

[9] Y.ke and R.Sukthankar, PCA-SIFT : A more distinctive representation for local image descriptors . technical report IRP-TR.2003.

[10] R.Duda, P.Hrt and D.Stock, pattern classification . john wiley & sons. 2000 .

[11] A.Johnson and M.Hebert, Using spin-images for efficient multiple model recognition in cluttered 3-d scenes . IEEE Trans on pattern analysis and Machine intelligence vol 21,433-499.1999.

[12] G.Carneiro and A.Jepson, phase – based local features . springer conf . on computer vision LNCY VAOL 2350 , 282 – 296. 2002 .

[13] S.Lazebnik, C.Schmid and J.Ponce, A sparse texture representation model recognition in cluttered 3-d scense . IEEE conf . on computer vision and pattern recognition vol II.319-324. 2003 .

[14] S.BLOGIE, J.MALIK, AND J.PUZHIVHA, SHAPE MATCHING AND OBJECT RECOGNITION USING SHAPE CONTEXT . ieee Trans on pattern analysis and machine intelligence vol 24 , 509-522. 2002 .

[15] T.kadir , D.Boukerroui and M.Brady, Analysis of the scale saliency algorithm . technical report. 2003 .

[16] T.Ojala, M.Pietikainen and D.Haewood, A comparative study of texture measures with classification based on featured distribution . Elsevier pattern recognition vol 29 , no 1,51-59. 1996 .

[17]V. Takala, T.Ahonen and M.Pietikainen, block – based methods for image retrieval using local binary pattermns . springer conf . on image analysis , LNCS Vol 3540,882-891.2005.

[18] T.Ojala , M.Pietikinen and T.Menp, multiresohution gray – scale and notation invariant texture classification with local binarypatterns . IEEE Trans . on pattern analysis and machine intelligence vol . 24 , 971 – 987. 2002 .

[19] J.koenderink and A.Doorn, representation of local geometry in the visual system . springer biological cybernetic vol . 55 , no , 6 , 367 -375. 1987.

[20] M. Mohammadpour, SMR. Hashemi and A. Broumandnia, "Improve Image Re-targeting Algorithm Using Markov Random Field", The second international conference on Pattern Recognition and Image Analysis, 2015

[21] SMR. Hashemi, A.Broumandnia. " A Review of Attention Models in Image Protrusion and Object Detection." The Journal of Mathematics and Computer Science ,Vol 15, Issue 4 2015 , pp ۲۷۳-۲۸۳.

[22] SMR. Hashemi, A. Broumandnia, " A New Method for Image Resizing Algorithm via Object Detection ." International Journal of Mechatronics, Electrical and Computer Technology, Vol 5, Issue 16 2015.

[23] SMR. Hashemi, " Review of algorithms changing image size " Cumhuriyet Science Journal "Vol 36, No 3 Special Issue 2015 (The Second National Conference on Applied Research in Science and Technology)