THINK INDIA JOURNAL

ISSN:0971-1260
Vol-22- Issue-14-December-2019

Efficient Object Detection And Match Using Feature Classification

[1st] **Prof.Raghavendra Rao**
Dept. Computer Science and Engineering
Sri Sairam College of Engineering
raghavendrarao.cse@sairamce.edu.in

[2nd] **Mahaveer Sahu**
Dept. Computer Science and Engineering
Sri Sairam College of Engineering
sahumahaveer1999@gmail.com

[3rd] **Shashank Singh**
Dept. Computer Science and Engineering
Sri Sairam College of Engineering
shashank.singh.singh96@gmail.com

[4th] Siddhant Khemka
Dept. Computer Science and Engineering
Sri Sairam College of Engineering
siddhantkhemka2311@gmail.com

[5th] Sonal R
Dept. Computer Science and Engineering
Sri Sairam College of Engineering
sonalr0311@gmail.com

***Abstract*—This paper presents a new approach for efficient object detection and matching in images and videos. We propose a stage based on a classification scheme that classifies the extracted features in new images into object features and non-object features. This binary classification scheme has turned out to be an efficient tool that can be used for object detection and matching.**

I. INTRODUCTION

The capability of detecting and registering objects in a video can be done by fixed camera or by a moving camera. The camera can be a hand-held or a robotic camera, or an onboard camera. The challenging problems to be solved are object detection, 3D projection, feature Extraction and matching and image registration. The common strategy used to solve the problem of object detection is to use the reference model. In this model objects can be represented as a templet or a set of relevant features. At run time, input images are matched with the object template or features in order to register the object with the current image. Therefore, if the object is planar then the registration aims to compute the homographic transform between a reference frame and the current frame. If the object is 3D then the registration aims to compute its 3D pose or projection matrix with respect to the camera. In all cases, a set of feature matches should be computed before carrying the registration process.

II. THEORITICAL BACKGROUND

In this paper, we introduce a classification stage that classifies the image features into object features and nonobject features. We use a binary Support Vector Machine (SVM) classifier. The training of the SVM classifier is carried out offline, while the classification of the extracted feature points in each video frame runs online and in real time. The proposed classification procedure is useful not only for increasing the speed and reliability of feature matching but also for reducing the computational load associated with the robust estimation of the registration parameters. The registration parameter is based on the robust technique RANSAC whose computational load will be too high if the percentage of false matches is high

Problem Definition

Our problem can be stated as follows. Given an object represented by one or more set of reference features, we would like to detect the occurrence of that object in a new image and if the detection is positive this object should be registered with the current new image via feature matching. Obviously, the associated feature matching is more challenging than stereo matching where the number of features in both images can be roughly the same and where many simplifying geometric constraints can be used. The object is represented by about 18 feature points whereas the input image contains hundreds of extracted features. Thus feature matching between these two sets of features will be very difficult.

III. LITERATURE SURVEY

A. Fast key point Recognition in Ten Lines of Code

While feature point recognition is a key component of modern approaches to object detection, existing approaches require computationally expensive patch pre-processing to handle perspective distortion. In this paper, we show that formulating the problem in a Naive Bayesian classification framework makes such pre-processing unnecessary and produces an algorithm that is simple, efficient, and robust. Furthermore, it scales well to handle large number of classes. To recognize the patches surrounding key points, our classifier uses hundreds of simple binary features and models class posterior probabilities.We make the problem computationally tractable by assuming independence between arbitrary sets of features. Even though this is not strictly true, we demonstrate that our classifier nevertheless performs remarkably well on image datasets containing very significant perspective changes.

B. Robust Real-Time Object Detection

visual object detection framework that is capable of processing images extremely rapidly while achieving high detection rates. There are three key contributions.The first is the introduction of a new image representation called the "Integral Image" which allows the features used by our detector to be computed very quickly. The second is a learning algorithm, based on AdaBoost, which selects a small number of critical visual features and yields extremely efficient classifiers.The third contribution is a method for combining classifiers in a "cascade" which allows background regions of the image to be quickly discarded while spending more computation on promising object-like regions. A set of experiments in the domain of face detection are presented. The system yields face detection performance comparable to the best previous systems. Implemented on a

ISSN:0971-1260
Vol-22- Issue-14-December-2019

conventional desktop, face detection proceeds at 15 frames per second.

C. *SoftPOSIT: Simultaneous Pose and Correspondence Determination*

The problem of pose estimation arises in many areas of computer vision, including object recognition, object tracking, site inspection and updating, and autonomous navigation using scene models. We present a new algorithm, called Soft POSIT, for determining the pose of a 3D object from a single 2D image in the case that correspondences between model points and image points are unknown. The algorithm combines Gold's iterative Soft Assign algorithm for computing correspondences and DeMenthon's iterative POSIT algorithm for computing object pose under a full-perspective camera model. Our algorithm, unlike most previous algorithms for this problem, does not have to hypothesize small sets of matches and then verify the remaining image points. Instead, all possible matches are treated identically throughout the search for an optimal pose. The performance of the algorithm is extensively evaluated in Monte Carlo simulations on synthetic data under a variety of levels of clutter, occlusion, and image noise. These tests show that the algorithm performs well in a variety of difficult scenarios, and empirical evidence suggests that the algorithm has a run-time complexity that is better than previous methods by a factor equal to the number of image points. The algorithm is being applied to the practical problem of autonomous vehicle navigation in a city through registration of a 3D architectural models of buildings to images obtained from an on-board camera.

D. *A Fast Local Descriptor for Dense Matching*

In this paper, we introduce a local image descriptor that is inspired by earlier detectors such as SIFT and GLOH but can be computed much more efficiently for dense wide-baseline matching purposes. We will show that it retains their robustness to perspective distortion and light changes, can be made to handle occlusions correctly, and runs fast on large images. Our descriptor yields better wide-baseline performance than the commonly used correlation windows, which are hard to tune. Too small, they do not bring enough information. Too large, they become vulnerable to perspective variations and occlusion. Therefore, recent methods tend to favour small correlation windows, or even individual pixel differencing and rely on global optimization techniques such as graph-cuts to enforce spatial consistency. They are restricted to very textured or high-resolution images, of which they typically need more than three. Our descriptor overcomes these limitations and is robust to rotation, perspective, scale, illumination changes, blur and sampling errors.

E. *Image registration method: A Survey*

This paper aims to present a review of recent as well as classic image registration methods. Image registration is the process of overlaying images (two or more) of the same scene taken at different times, from different viewpoints, and/or by different sensors. The registration geometrically align two images (the reference and sensed images).The reviewed approaches are classified according to their nature (area-based and feature-based) and according to four basic steps of image registration procedure: feature detection, feature matching, mapping function design, and image transformation and resampling. Main contributions, advantages, and drawbacks of the methods are mentioned in the paper. Problematic issues of image registration and outlook for the future research are discussed too. The major goal of the paper is to provide a comprehensive reference source for the researchers involved in image registration, regardless of particular application areas.

IV. EXISTING SYSTEM

The capability of detecting and registering objects in a video sequence captured by either a fixed camera or a moving camera is the corner stone in many computer vision applications. The camera can be a hand-held camera, a robotics camera, or an onboard camera. To this end, many challenging problems should be solved, namely object detection, 3D object pose, feature extraction and matching, and image registration. The problem of object detection has been studied by many researchers. Supervised techniques based on learned appearances has been used to detect objects whose class can be described statistically such as faces, facade windows, and vehicle rears. These techniques include Adaptive Boosting and Active Appearance Models..

V. PROPOSED SYSTEM

We are proposing here by detecting and matching an arbitrary object with an image. By using that the use of feature classification makes the whole process of matching and registration faster and more robust

Advantages

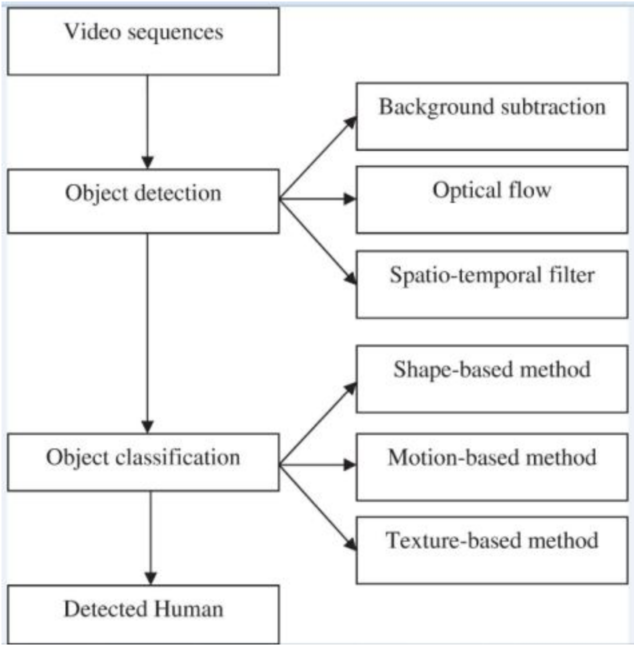
1) Features form a small subset of the input image features. This filtering procedure is very useful not only for increasing the reliability of feature matching but also for reducing the computational load associated with subsequent processing

2) A good separation is achieved by the hyper plane that has the largest distance to the neighboring data points of both classes.

VI. SYSTEM FRAMEWORK

The entire process in this project can be represented in the following diagram.

ISSN:0971-1260
Vol-22- Issue-14-December-2019



VII. MODULES

A. Feature Extraction

Feature points are extracted in the object reference template and in the input image. For this purpose, we use the SIFT (Scale Invariant Feature Transform) points. SIFT points are invariant to image scaling and rotation, and partially invariant to change in illumination and 3D rotation. Each SIFT point has a descriptor computed from an histogram of local oriented gradients around the key point and stored in a 128 dimensional vector. Feature classification: Before putting image SIFT.

B. Feature Classification

Before putting image SIFT points into correspondence with the object SIFT points, it will be advantageous to filter out the image SIFT points such that the majority of the retained image points belong to the object in question. Recall that the object features form a small subset of the input image features. This filtering procedure is very useful not only for increasing the reliability of feature matching but also for reducing the computational load associated with subsequent processing. To this end, we use a binary classification scheme in which every extracted image feature is classified into one of these two classes: object class and non-object class. In our case, this classifier is built using Support Vector Machines with two classes of features. The two-class SVM is built offline on two sets of training features. The positive training features (object features) are extracted in some given reference images of the object (real images and/or warped images). The negative training examples are composed of all feature points that are not belonging to the object (e.g., feature points belonging to the background). Support Vector Machines (SVMs) are a set of related supervised learning methods used for classification and regression. Input data are viewed as two sets of points in an n-dimensional space, an SVM will construct a separating hyperplane in that space, one which maximizes the margin between the two data sets. To calculate the margin, two parallel hyperplanes are constructed, one on each side of the separating hyperplane, which are "pushed up against" the two data sets. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the neighboring datapoints of both classes, since in general the larger the margin the lower the generalization error of the classifier. SVMs were extended to classify data sets that are not linearly separable through the use of non-linear kernels. In our work, we use nonlinear SVMs with polynomial kernels of degree two.

C. Feature Matching

At this stage, we have two sets of features. The first set is given by the reference features. The second set is given by the input image features that are classified as object features. In order to put the features in these two sets into correspondence we use the first Nearest Neighbour classifier. To make the process fast, we use the KD-tree technique. Moreover, a threshold is used in order to discard some matches. To register the object (homography or 3D pose computation) the RANSAC technique is used.

CONCLUSION

This paper presented an approach for detecting and matching an arbitrary object with an image. We have shown that the use of feature classification makes the whole process of matching and registration faster and more robust.

REFERENCES

[1] M. Ozuysal, P. Fua, and V. Lepetit, "Fast keypoint recogni- tion in ten lines of code," in Computer Vision and Pattern Recognition, 2007

[2] P. Viola and M. Jones, "Robust real-time object detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137–154, 2004

[3] P. David, D. DeMenthon, R. Duraiswami, and H. Samet,"Softposit: Simultaneous pose and correspondence determination," in European Conference on Computer Vision, 2002.

[4] E. Tola, P. Fua, and V. Lepetit, "A fast local descriptor for dense matching," in Computer Vision and Pattern Recognition, 2008.

[5] B. Zitova and J. Flusser, "Image registration methods: a survey," Image and Vision Computing, vol. 21, pp. 977–1000, 2003.

[6] T. Cootes, G. Edwards, and C. Taylor, "Active appearance models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 681–684, 2001.

[7] D. Lowe, "Distinctive image features from scale invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–100, 2004.

[8] D. Meyer, F. Leisch, and K. Hornik, "The support vector machine under test," Neurocomputing, vol. 55, pp. 169–186, 2003.