Commentary: < Trainable COSFIRE Filters for Keypoint Detection and Pattern Recognition >

<Guodong Chen>
<z5224156>

1. Introduction

With the increasing application of computer vision, more attention has been paid on keypoint detection and pattern recognition. Meanwhile, there are still some shortages on these two parts. This paper is to introduce a trainable filter named combination of shifted filter responses (COSFIRE). The method is not only simple and straightforward to implement in theory, but also effective and robust in practice.

Methods for keypoint detection perform pool in selectivity when it comes to the shape properties and lack of robustness to contrast variations. For example, when a dissimilarity measure for keypoint descriptors is based on Euclidean distance, we may find highly similar for two patterns. However, lower similarity scores can be produced for the shape properties and contrast variations. In that case, COSFIRE are introduced to solve above problems.

Many users may benefit from COSFIRE. Firstly, from the perspective of computer engineer, the method can be easily implemented as the filter is simple. Secondly, for doctors, this technique can help them perform fine tests. Finally, COSFIRE can be used in self-driving cars to detect and recognize traffic signs, which mean fewer traffic accidents may happen and more free time for all drivers.

In addition, COSFIRE also make some difference to current practice in the field. The application of this method can supply an efficient filter to extract feature for handwritten digits. Apart from that, this technique has advantage in detecting and recognizing traffic signs. At the same time, trained COSFIRE filter is also introduced to medical field like, detecting vascular bifurcations in retinal fundus images.

2. Methods

There are 6 main parts are included in Method part: Detection of Orientations by 2D Gabor Filters, Configuration of a COSFIRE Filter, Blurring and Shifting Gabor Filter Responses, Response of a COSFIRE Filter, Achieving Invariance and Detection of More Complex Patterns.

For the first one orientations detection by 2D Gabor Filters, the proposed COSFIRE filter is built by applying the response of Gabor filters as input. As for the second one configuration of a COSFIRE Filter, it indicates process of obtaining four parameter values for COSFIRE Filter. The next is blurring and shifting Gabor filter responses, it blurs the responses from Gabor filter first, then responses are shifted together in the filter center. For the fourth one response of a COSFIRE filter, this is defined as weighted geometric mean of blurring and shifting Gabor filter responses. Then, the fifth one achieving invariance is to keep invariance in four aspects: refection, scale, contrast inversion and rotation. Finally, Detection of More Complex Patterns, its configuration result comes from local prototype pattern.

Several computer vision methods are applied in the section: Gabor filter, Thresholding, SIFT, blurring and shifting. Gabor filter is used to detect lines which represent the dominant orientations in the neighborhood of the specified area of interest. Thresholding is used in three areas. Firstly, t1 decides which level of a Gabor filter should use to show the presence of a line or an edge. Secondly, t2 only used in configuration phrase. Lastly, t3 control the responses of COSFIRE filter which can decrease the filter to a certain fraction. For SIFT, is used in scale-invariant approaches, which keeps the scale is invariant. For blurring and shifting, it centralizes the Gabor filter responses.

For these methods, Gabor filter used in paper is effective as it analyses essential orientation information of the image. However, Gabor filter is not irreplaceable as other orientation-selective filter can also be used, like Sobel filter. Similarly, it can be found that thresholding is also a suitable method for image processing. As for SIFT, although it performs well in practice, SURF could be better than SIFT. Because SURF can be more efficient than SIFT. For blurring and shifting part, some specific method can be used in the two part for achieving a better performance.

3. Results

The experiment result for all three applications is above 95%. For the detection of vascular bifurcations in retinal fundus, the recall is about 98% and the precision is approximately 96%, the performance of COSFIRE filters is better than the others reported method. For the second application handwritten digits recognition, it achieves about 99% correct classification, which is close to the best-performed method in application. In addition, the last application traffic signs detection and recognition, it gets 100% for both recall and precision, which performs best considering accuracy and computational complexity.

Several evaluation strategies are applied to three applications. They are precision, recall, harmonic mean and recognition rate. Precision is the number of relevant and classified correctly objects divided by the total number of relevant objects. Recall is the number of relevant and classified correctly objects divided by total number of relevant objects. Harmonic mean is equal to 2*(Recall * Precision) / (Recall + Precision). Recognition rate is the number of correctly recognized objects divided by total objects.

There are some findings for all three applications. For the first one, vascular bifurcations detection in retinal fundus, when using sufficient filers, vascular bifurcation can be detected. Moreover, if additional features can be detected by filters, that may result in higher precision. In addition, the result, without any improvements, is better than others which is reported previously. As for the second application recognition of handwritten digits, even if without any optimization and preprocessing/or postprocessing operation is applied for COSFIRE filters, it still have an effective performance. That means this model is robust. For the last one traffic signs detection and recognition, when absolute threshold is 0.04 and validity ratio is 0.5, all 48 traffic scenes are detected, which is a perfect result.

From the above findings, although without any optimization of the filter, it also achieve a effective and robust result for every three applications. it can be more convincing for potential users to adopt proposed COSFIRE filters if it can use some metadata for testing. Because in three applications, the dataset seems not big enough.

4. Conclusions

There are many strengths of proposed COSFIRE filters mentioned in this paper. Firstly, this method is easy and simple to implement. It does not require many complex formation, calculation and transformation. Secondly, robustness and effectiveness are important metrics for this method. Robustness means the method is suitable for many different situations without much adjustment. The method also has a good perform on evaluation. On the other hand, the paper also can be improved in some aspect. For one thing, it may mention some more blur or shift method can be applied. For another thing, the size of dataset for three applications seem to be not big enough, which mean the result may be overfitting. Thus, a larger size dataset should be used for this method.

In addition, some remaining issues that still need to be addressed before the problem stated in the introduction can be considered solved. One of the issues is that get permission to use private image before we analyst it.

There are some issues can be used for future research. Initially, there are many other computer vision tasks which can also apply the proposed trainable COSFIRE filters for detection and recognition such as tracking, pedestrians, geometric stereo calibration, architectural symbols and image retrieval. Secondly, more features can be included in the properties of a COSFIRE filter, like color feature and texture feature for a local prototype pattern. Finally, one potential direction can be extended from 2D COSFIRE filter to 3D COSFIRE filer, which can be applied in more situation and can be more robust.

References