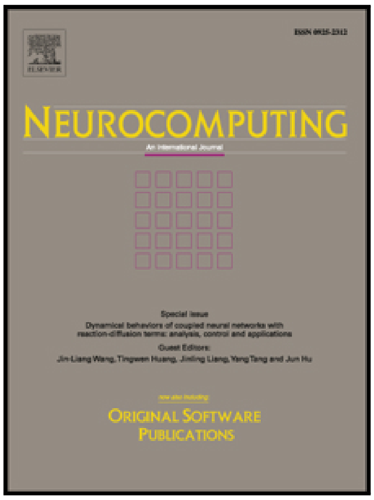
Accepted Manuscript



Affine scale space: an affine invariant image structure to promote the detection of correspondences from stereo images

Biao Zhao

PII: S0925-2312(17)30655-0
DOI: 10.1016/j.neucom.2016.06.094
Reference: NEUCOM 18318

To appear in: *Neurocomputing*

Received date: 23 October 2015
Revised date: 18 February 2016
Accepted date: 12 June 2016

Affine scale space: an affine invariant image structure to promote the detection of correspondences from stereo images

Biao Zhao

Politecnico di Torino - Corso Duca degli Abruzzi, 24, Torino, ITALY

Abstract

Calculating the geometry relationship by the positions of the correspondences from stereo images is a fundamental method to obtain the depth information. Such a method was quite widespread and popular thanks to its efficiency and easily accessed implementation. General speaking, the more density of the correspondences are, the more precisely the depth information can be calculated. Theoretically, a sufficient strengthened correspondences match algorithm can be utilized for depth information calculation under any circumstances. Unfortunately, the updated image feature detections, are all sensitive to the view point changes: following the slight rotating view angles of the stereo images, the number of matched features drastically reduced, resulting in the number of matched features not adequate to cover every details of the stereo images. This disadvantages of feature detections in practice hampers its application for the depth calibration. To tackle the sensitive of view point to stereo images, in this paper, we will propose an affine invariant affine scale space based feature detection, which is more robust to detect the correspondences from stereo images. The purpose of affine scale space is to create a more general approach to the affine invariant image scale representation by modifying the corresponding Gaussian filters in order to cope with the specific change of view point. The affine adaptation of the scale space is to retain a linear relationship with the transiting of the view point. Within the linear relationship, the affine scale space can be established as a more general approach for the detection of correspondences from stereo images. With a better correspondences detection, a more precise depth information can be made.

Keywords:
depth information, stereo images, affine scale space, feature detection

1. Introduction

Obtaining the geometric relationship of the stereo images by analyzing its correspondences [1] is perhaps a most typical way to estimate the depth information. This method is efficient and can be easily accessed by two fixed cameras. It does not rely on any specific devices like Kinect and thus it is more efficient and economical. On the other hand, this traditional method stereo depth estimation is coming to its corner because of the bottleneck of stereo match. A general stereo match is achieved by some regional local and global block matches which are the gradient or intensity values in a size-fixed windows [2]. Because of the ever existing perspective distortions affecting the positions of the neighborhoods, the gradient or intensity values shall be further interfered [3]. In addition, the none correlated pixel values makes the images features almost unpredictable. In most cases, the local extremes will be selected as the representable features, since they are easier to detect and describe. But the modified local regions by the perspective distortion changed this situation. This series view point affections upon the current stereo matches limits its utilization upon the stereo images' depth calibration [4].

However, following the development of visual techniques, we have more options to tackle the affine sensitiveness upon the stereo matches, especially the image feature based stereo matches [5][6]. Image features can often refer to colors, shapes [7][8][9], textures [10], edges [11][12], or any other inherent image features [6]. A valid and robust stereo match should be based on the features stable and robust to the fluctuation of the scene, including the changes of view point and illuminations to some scale [5].

Several state-of-the-art image scale invariant feature detections are based on the derivatives of Gaussian: Laplacian of Gaussian (LoG) and Differential of Gaussian (DoG) [13], which is to find the local extrema among the LoG or DoG filtered images [10]. Usually, the derivatives of Gaussian will result in strong positive responses for dark blobs [14] and strong negative responses for bright blobs [15]. In order to automatically capture the blobs in the image domain, where the scale of the blobs is previous unknown, a multi-scale approach is therefore necessary, which has inspired the creation of scale space.

A scale space can be defined as a collection of several presmoothed images by different sized Gaussian kernels [4]. In general, scale space theory is a framework for multi-scale signal representation to handle with image structures at different scales, by employing a one-parameterized family of smoothed images [16]. This framework provides a scale-invariant representation [17], which is necessary to deal with variations that may occur in image data [18]. Indeed, real world objects, in

Email address: biao_zhao@yahoo.com (Biao Zhao)