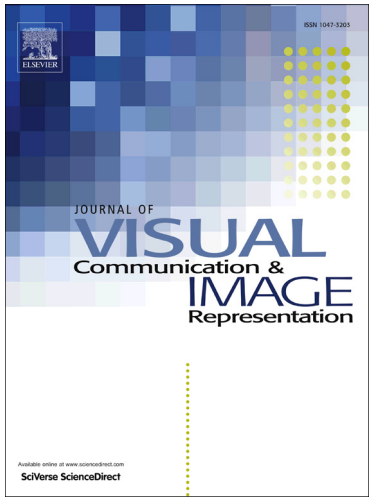

Convolutional Neural Network based Deep Conditional Random Fields for Stereo Matching

Zhi Wang, Shiqiang Zhu, Yuehua Li, Zhengzhe Cui

PII: S1047-3203(16)30181-X
DOI: http://dx.doi.org/10.1016/j.jvcir.2016.08.022
Reference: YJVCI 1851

To appear in: *J. Vis. Commun. Image R.*

Received Date: 28 June 2016
Revised Date: 9 August 2016
Accepted Date: 26 August 2016

Please cite this article as: Z. Wang, S. Zhu, Y. Li, Z. Cui, Convolutional Neural Network based Deep Conditional Random Fields for Stereo Matching, *J. Vis. Commun. Image R.* (2016), doi: http://dx.doi.org/10.1016/j.jvcir.2016.08.022

Convolutional Neural Network based Deep Conditional Random Fields for Stereo Matching

Zhi Wang*, Shiqiang Zhu, Yuehua Li, Zhengzhe Cui

State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, No.38, Zheda Road, Hangzhou, China, 310027

*11325067@zju.edu.cn

Abstract: Stereo matching has been studied for many years and is still a challenge problem. The Markov Random Fields (MRF) model and the Conditional Random Fields (CRF) model based methods have achieved good performance recently. Based on these pioneer works, a deep conditional random fields based stereo matching algorithm is proposed in this paper, which draws a connection between the Convolutional Neural Network (CNN) and CRF. The object knowledge is used as a soft constraint, which can effectively improve the depth estimation accuracy. Moreover, we proposed a CNN potential function that learns the potentials of CRF in a CNN framework. The inference of the CRF model is formulated as a Recurrent Neural Network (RNN). A variety of experiments have been conducted on KITTI and Middlebury benchmark. The results show that the proposed algorithm can produce state-of-the-art results and outperform other MRF-based or CRF-based methods.

Keywords: Stereo matching, conditional random fields, Convolutional Neural Network.

1. Introduction

Stereo matching, as a long-standing problem in the research of computer vision [1], has been used in a variety of fields, such as 3d reconstruction, navigation for autonomous driving system, panoramic stereo imaging [2]and DoF rendering[3]. There are generally two kinds of techniques for stereo matching: local methods (also known as cost filtering methods) and global methods (also known as energy minimization methods). The cost filtering methods are widely used in the early times because of its simple implementation. Many local stereo matching algorithms have been presented, such as bilateral filter method (BF) [4], guided filter method (GF) [5], non-local cost aggregation method [6]. The core idea of this kind of methods is to determine the disparity of the pixel based on the neighbouring pixels around it [7]. The local methods usually make the assumption that all pixels in the local support window have