Signal Processing: *Image Communication* 39 (2015) 159–172

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

journal homepage: www.elsevier.com/locate/image

Sequential block-based disparity map estimation algorithm for stereoscopic image coding ☆

Aysha Kadaikar *, Gabriel Dauphin, Anissa Mokraoui

L2TI, Institut Galilée, Université Paris 13 Sorbonne Paris Cité 99, Avenue Jean-Baptiste Clément 93430 Villetaneuse, France

ARTICLE INFO

Article history:
Received 17 March 2015
Received in revised form
24 July 2015
Accepted 9 September 2015
Available online 21 September 2015

Keywords:
Optimization
Disparity estimation
Rate-distortion
Entropy
Distortion

ABSTRACT

This paper deals with the problem of block-based disparity map estimation for stereoscopic image coding where the estimated map is transmitted to the decoder in order to predict one view from the other. The estimation problem of the disparity map is thus a trade-off between the quality prediction and the binary cost of the disparities to be stored or transmitted. This trade-off is modeled as a joint entropy-distortion metric assuming that the disparity map is encoded with an entropy coder; and one of the two views is fully predicted using this map when applied to the other view. However minimizing this joint metric is a complex combinatorial optimization problem where choices of disparities are all interrelated. A sub-optimal optimization solution is then proposed. It is based on a tree structure which is constructed sequentially whenever a block is matched. The developed algorithm, called Modified *M*-Algorithm (MMA), processes the reference view in a raster scanning order and assumes that the disparities to be selected in the unprocessed area are likely to follow a chosen disparity distribution. This algorithm has the ability at each step of the process not only to retain the *M*-best paths of the tree in terms of entropy-distortion cost but also to explore all possible extensions of each of these *M* paths until reading the last block of the view. Simulations, conducted on stereoscopic images extracted from Middlebury and Deimos datasets, show the advantage of our MMA compared to the conventional Block Matching Algorithm (BMA) with and without regularization both in terms of reducing bitrate and distortion.

1. Introduction

During the last decades, a wide range of multiview applications has emerged still offering an increased immersion to the users such as video games with autostereoscopic displays, 3D-TV or stereo visio-conferencing [1,2]. These applications involve many views of the same scene and thus need fast and efficient coding techniques that take into account the redundancies in the views. In the case of stereoscopic images, a set of two cameras captures the same scene from two slightly different points of views, generally corresponding to the average distance between human eyes. Because of the perspective, objects in the scene are projected at a different position on each of the views. This little shift is called disparity.

To encode stereoscopic images, a simplistic approach consists in coding the two views separately. However this method seems not to be optimal as it does not consider the inter-view redundancies. The typical approach is the disparity compensated coding scheme: (i) the reference view is independently coded; (ii) a stereo matching is performed on the views to estimate a disparity map; (iii) the second view is predicted from the reference view using

☆ Fully documented templates are available in the elsarticle package on CTAN.

* Corresponding author. Tel.: +33 1 49 40 44 55.
E-mail addresses: kadaikar@univ-paris13.fr (A. Kadaikar), gabriel.dauphin@univ-paris13.fr (G. Dauphin), anissa.mokraoui@univ-paris13.fr (A. Mokraoui).

http://dx.doi.org/10.1016/j.image.2015.09.007
0923-5965/

the disparity map; and (iv) a residual error is computed as the difference between the predicted and the original view. Finally, the reference view, the disparity map and the residual error are encoded. A refinement to this scheme, called *closed loop* as opposed to *open loop*, is to estimate the disparity map using the coded-decoded reference view instead of the true reference view.

Estimating the *true* disparity (also referred to as solving the stereo-matching problem) has been at the center of many works. Numerous methods are summarized in [3,4] in the context of the 3-D reconstruction of the scene. Two main approaches can be distinguished: local and global methods. Local methods aim at comparing the similarity between two sets of pixels. These methods can be classified into two broad categories: feature- and area-based approaches [5–7]. They operate very fast, are efficient in highly textured areas but not in untextured or occluded regions. On the other hand, global methods have been introduced to overcome these issues as they are less sensitive to locally ambiguous regions. They estimate the disparity map minimizing only the global energy function on the whole image. This global energy function is generally composed of a term measuring the similarity between two corresponding pixels and a smoothness constraint. The best known global methods are dynamic programming [8,9], graph cuts [10,11], variational methods [12] and belief propagation [13,14].

In the context of stereoscopic image coding, a valuable disparity map is one that has a high ability to predict one view from the other, while minimizing the bitrate needed to be encoded. Such a valuable disparity map may not be similar to the true disparity map. Methods estimating locally a blockwise disparity map are often privileged due not only to their timeliness but also to their small encoding cost compared to a dense disparity map since a single disparity is encoded for each block. The most simple and famous of these local disparity estimation techniques is undoubtedly the Block-Matching Algorithm (BMA). Each disparity is selected by usually minimizing the Sum of Square Differences (SSD) or the Sum of Absolute Differences (SAD) between the intensities of the pixels. Some papers propose to minimize cross-correlation or rank distortion metrics [15,16]. Among these local disparity estimation techniques, most other methods take into account a regularization constraint and encode the resulting disparity map using a predictive technique such as Differential Pulse-Code Modulation (DPCM) [17–20]. This regularization technique helps selecting disparities similar to those of their neighboring blocks. As compared to the BMA, such methods yield more smooth disparity maps at the expense of a little increase in the distortion of the predicted view. Smooth disparity maps tend to be encoded with smaller bitrate, all the more significantly when using DPCM.

This paper extends the dense disparity map strategy developed in reference [21] to block-based disparity map. Our approach is different from the BMA and other previously mentioned estimation techniques in the choice of using a different constraint. Indeed this paper models the trade-off between the quality prediction and the required cost for transmitting the disparities as a joint entropy-distortion metric. The disparity map is then assumed to be encoded with an entropy coder and one of the two views is also assumed to be fully predicted using this estimated disparity map combined to the other view. The minimization of the joint entropy-distortion metric is nevertheless a complex combinatorial optimization problem where the choice of a disparity at any location in the image may favor the choice of that disparity at any other location. Therefore a sub-optimal stereo-matching optimization algorithm is developed. This algorithm processes the reference view in a raster scanning order and assumes that the disparities to be selected in the unprocessed area are likely to follow a chosen disparity distribution. This algorithm, called Modified *M*-Algorithm (MMA), is based on a tree structure having the ability at each step to retain only the *M*-best paths of the tree and to explore all possible extensions of each of these *M* paths.

The remainder of the paper is organized as follows. The stereo-matching optimization problem and its properties are presented in Section 2. A sub-optimal solution to this optimization problem, precisely the proposed MMA, is described in Section 3. Simulation results are provided and discussed in Section 4. Finally, Section 5 concludes our work.

2. Stereo-matching optimization problem

The proposed MMA is derived from a general stereo-matching optimization problem. Thanks to some basic assumptions and using some specific notations, it is possible to express the stereo-matching optimization as finding the disparity map minimizing a global cost function. A simple analysis of this problem gives some insight regarding its complexity.

2.1. Basic assumptions and notations

First we present some working assumptions. The left and right views of the stereoscopic image are assumed to be rectified by having applied a rectification algorithm to the views or by a special care during the images acquisition. The left view is considered as the reference view. Each view is divided into non-overlapped blocks of same size and all pixels within a block have the same disparity. As a result, blocks of the right view matching blocks of the left view are expected to be found in the same scan lines, and disparity is an one-dimension vector (vertical component is equal to zero).

In order to provide better prediction accuracy, disparities with non-integer values are also considered and the sub-pixel intensities are deduced from a weighted combination of neighboring pixel intensities. Lower and upper disparity bounds are also considered as they limit the length of the searching area.

We introduce in the following some notations before presenting the formalization of the stereo-matching optimization. The left and right images (respectively I_L and I_R) are of size $K \times L$ pixels. They are divided into $X \times Y$ non-overlapped blocks each of size $N_X \times N_Y$ pixels. The position of a block is defined by its block-coordinates (i_B, j_B) or its