Signal Processing: *Image Communication* 23 (2008) 379–390

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Symmetric distributed coding of stereo omnidirectional images[☆]

Vijayaraghavan Thirumalai *, Ivana Tosic, Pascal Frossard

Ecole Polytechnique Fédérale de Lausanne (EPFL), Signal Processing Laboratory (LTS4), Lausanne 1015, Switzerland

ARTICLE INFO

Article history:
Received 29 March 2008
Accepted 4 April 2008

Keywords:
Distributed source coding
Omnidirectional images
LDPC codes
Camera networks
Stereo imaging

ABSTRACT

This paper presents a distributed coding scheme for the representation of 3D scenes captured by a pair of omnidirectional cameras with equivalent computational resources and transmission capabilities. The images are captured at different viewpoints and are encoded independently. A joint decoder exploits the correlation between images for improved decoding quality. The distributed coding is built on the multi-resolution representation of spherical images, whose information is split into two partitions. The encoder then transmits one partition after entropy coding, as well as the syndrome bits resulting from the channel encoding of the other partition. The joint decoder exploits the intra-view correlation by predicting one partition from the other partition. At the same time, it exploits the inter-view correlation using block-based disparity estimation between images from different cameras. Experiments demonstrate that the distributed coding solution performs better than a scheme where images are handled independently. Furthermore, the coding rate stays balanced between the different cameras, which permits to avoid hierarchical relations between vision sensors in camera networks.

1. Introduction

Camera networks find widespread usage in several applications that rely on the effective representation of scenes or the analysis of 3D information. These networks normally consist of several cameras distributed in the scene, and pose several problems like the coding of multi-view images, the reconstruction of the 3D structure from multiple views, or the multi-view object recognition, for example. This paper focuses on the compression of multi-view images and particularly stereo omnidirectional images. The images captured from different viewpoints are usually correlated, which permits to reduce the coding rate by exploiting efficiently the redundancy between the different views. Instead of joint encoding that unfortunately requires communication between cameras, we rely on the Slepian–Wolf theorem [30] and design a distributed coding scheme where images are encoded independently, but decoded jointly in order to exploit the correlation between the images, as illustrated in Fig. 1.

Most of the research carried out on distributed coding for multi-view images or videos propose solutions based on coding with side information. In this case, one of the cameras is chosen as the primary source and its output is encoded independently. The other cameras represent secondary sources whose rate can be drastically reduced if the joint decoder uses the primary source as side information. Such a coding scheme obviously does not balance the transmission rate between the encoders. However, it is often interesting in practice to rather avoid hierarchical relations between sensors and to distribute the coding and transmission cost equally among the sensors. In this paper, we therefore concentrate on symmetric coding scheme, where all cameras are equally important in the representation of the 3D scenes.

We consider a scenario, where two catadioptric cameras are distributed in the 3D scene, as shown in Fig. 1. Each catadioptric camera samples the plenoptic

---

[☆] This work has been partly supported by the Swiss National Science Foundation, under grant 20001-107970/1.

* Corresponding author.
E-mail addresses: vijayaraghavan.thirumalai@epfl.ch (V. Thirumalai), ivana.tosic@epfl.ch (I. Tosic), pascal.frossard@epfl.ch (P. Frossard).

0923-5965/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.image.2008.04.005

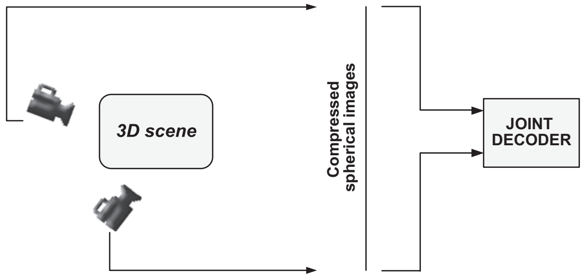

Fig. 1. Distributed coding of the 3D scenes. The correlated images are compressed independently and are decoded jointly.

function that represents the entire visual information seen by the observer [3]. The catadioptric camera is realized by a convex, reflective parabolic mirror placed above and parallel to the camera approximating an orthographically projecting lens. In such a case, the ray of light incident with the focus of the parabola is reflected to a ray of light parallel to the parabola's axis. This construction is equivalent to a purely rotating perspective camera [13]. We propose to work directly in the spherical domain by appropriately mapping the captured omnidirectional images on the sphere through inverse stereographic projection. This mapping permits us to obtain the light field in its natural radial form and to avoid the potential discrepancies that may arise due to Euclidean assumptions in perspective imaging. As the spherical image captures full 360° view, it is particularly suitable for representing the 3D scene. The wide field of view permits to reconstruct the associated 3D structure by processing a significantly smaller number of multi-view images, compared to the number of views in a perspective camera network performing the same task. The reduction of the total number of views is furthermore advantageous for determining the camera arrangement that achieves the complete representation of the 3D scene. Certainly, the quality of the reconstructed 3D structure depends on the resolution of the catadioptric sensor. In order to maintain a good reconstruction quality, the omnidirectional camera requires higher resolution image sensors [5]. Fortunately, in modern camera technologies the resolution of the image sensors has been drastically increased during the last years, offering a variety of high resolution devices at affordable costs.

In this paper, we propose a transform domain symmetric distributed coding scheme for representing a 3D scene captured by the stereo omnidirectional cameras. The correlated omnidirectional images initially undergo a multi-resolution decomposition based on the spherical Laplacian pyramid (SLP), which brings the advantage of shift invariance. The resulting sets of coefficients are quantized and then split into two correlated partitions. The quantized coefficients of the first partition are entropy coded, and sent to the decoder. The second partition is encoded using the nested scalar quantization (NSQ) [41], which is a binning scheme that encompasses a scalar quantizer and a coset encoder. It outputs the coset bin indexes and permits to reduce the coding rate compared to encoding the quantized coefficients directly. The coset bin indexes are further encoded using a Slepian–Wolf encoder based on multi-level LDPC codes [20,21], in order to achieve further compression. The resulting syndrome bits are finally transmitted to the joint decoder.

The joint decoder estimates the quantized coefficients of the second partition from the quantized coefficients of the first partition, by exploiting intra-view correlation. Furthermore, the joint decoder takes benefit of the correlation between views by performing block-based disparity estimation (DE) on the sphere [35], which matches similar blocks of solid angles from two omnidirectional images, directly in the spherical domain. Therefore, the proposed scheme efficiently combines the intra and inter Wyner–Ziv image coding, which allows for a balanced coding rate between cameras. Such a strategy proves to be beneficial with respect to independent processing of omnidirectional images and shows only a small performance loss compared to joint encoding of the different views. Moreover, we exploit the inter-view correlation by block-based DE, which estimates the displacement between the corresponding objects without using epipolar geometry constraint. Hence, the block-based DE technique used in our scheme does not require any camera parameters which are usually required in the techniques based on epipolar geometry, to perform the correspondence matching (e.g., [4]). This is certainly beneficial in camera networks where the camera parameters are not given or when camera network calibration is not achievable in practice. The proposed scheme therefore provides a low-complexity coding solution for the representation of 3D scenes, which does not require complex setup nor hierarchical organization between vision sensors.

The rest of the paper is organized as follows. Section 2 overviews the related work in distributed coding with a special focus on camera networks. Section 3 presents the distributed coding algorithm adapted to omnidirectional images. Section 4 presents in more details the Wyner–Ziv coding strategy, while Section 5 describes the joint decoding scheme. Section 6 finally presents the experimental results that demonstrate the benefits of the proposed solution. Section 7 concludes this paper.

2. Related work

The first information-theoretical results on distributed source coding (DSC) appeared already in the late seventies. In particular, it has been shown that independent coding of correlated sources can achieve the same rate-distortion bound as joint encoding if a joint decoder can efficiently exploit the correlation between the sources [30]. Rate-distortion bounds have been established later for the particular case of coding with side information [38]. However, most results presented in [30,38] have remained non-constructive for about three decades. Practical DSC schemes have been designed only recently, by establishing a relation between the Slepian–Wolf theorem and channel coding [25]. Subsequently, several practical DSC systems have been presented using different channel codes, e.g., Turbo codes [10,1] or LDPC codes [19].