Available online at www.sciencedirect.com
ScienceDirect
Image and Vision Computing 26 (2008) 725–730
www.elsevier.com/locate/imavis

Estimating the minimum redundancy in stereo image pair ☆

Wang Yuanqing *

Department of Electronic Science & Engineering, Nanjing University, Nanjing 210093, PR China

Received 14 August 2006; accepted 18 August 2006

Abstract

The raw data in binocular stereo image sequences is twice as that of monocular images, the large amount of information should be reduced. As a result there has been increasing attention given to image compression methods specialized to stereo pairs. Much of this work has concentrated on improving the disparity compensation process and codes the residual image similarly to a monocular image where one view is used to predict another, and the difference is coded. The residual image is usually composed primarily of strong vertical direction edge components surrounded by large areas of near zero intensity. The residual images have different characteristics, but they behave uniquely statistical regularity. This property is demonstrated experimentally in the paper. Two interested statistical variables are described, the one is the total number (N) of the pixels with near zero intensity in the residual image and other is the coordinate displacements $(\Delta x, \Delta y)$ between the left and right image frames for get the residual image. Experimental results indicate that the curve between the parameters N and variables $(\Delta x, \Delta y)$ may be fit by Gaussian function. The maximum of the variable N_m corresponding to the optimal displacements $(\Delta x_{\mathrm{op}}, \Delta y_{\mathrm{op}})$ may be estimated by the Gaussian approximation. An algorithm is further provided to quickly predict the minimal redundancy of the residual image and the corresponding displacement. It is shown how such characteristics may be of great benefit to quickly achieve the higher compression ratio.

Keywords: Stereo image; Stereo image compression; Residual image; Gaussian function; Image characteristics

1. Introduction

Recording stereo video is accomplished using a dual camera system, in which two cameras are aligned along the same horizontal plane and positioned at some distance apart from each other. This video system with stereo pair needs twice the raw data as a monocular imaging system. An uncompressed 640×480 resolution, 30 fps, 24 bpp (8 bit each color for trichromic RGB) stereo video stream would require the network to support a transmission rate of $(640 \times 480$ pixels/frame$) \times (30$ fps$) \times (24$ bpp$) \times (2$ streams$)$ $\approx$442.4 Mbps! Clearly, there are bandwidth limitations introduced by the networks that connect the two sites. In order to create a 3D system with two images that should be transmitted simultaneously, the amount of video data must be reduced by compressing the video stream before transmission and an efficient coding appropriate for stereoscopic images is developed.

For lossy compression, there is also an increased degradation in video quality as the compression rate rises. The goal is to compress the video as much as possible, while still providing an acceptable level of quality. Stereoscopic image sequence compression involves the exploitation of the spatial redundancy between the left and right image frames to achieve the compression ratio higher than that are by the independent compression of the two frames. So far there have been many proposals for image compression methods specialized to stereo pairs. Most of this work has focused on the disparity compensated residual compression where one view is used to predict another, and the difference is coded [1–7]. In the common approach, one channel is compressed monoscopically, using any suitable algorithm. The other channel is used to predict

☆ This work is supported by National Natural Science Foundation of China (60472026).
* Tel./fax: +86 25 83686252.
E-mail address: yqwang@nju.edu.cn

0262-8856/$ - see front matter
doi:10.1016/j.imavis.2006.08.008

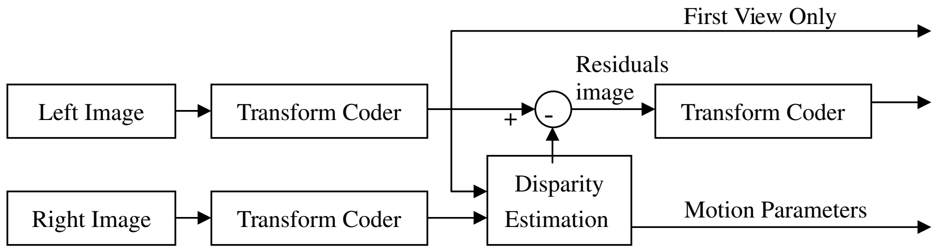

Fig. 1. The disparity compensated coding approach.

the other channel, with the prediction encoded as 'disparity vectors'. Disparity vectors define a warping of one image to approximate the other, and are identical in spirit to the motion compensation vectors of MPEG and other video compression approaches. As shown in Fig. 1. The predicted and actual images are differenced and the prediction error coded and transmitted.

Although disparity compensation and coding have been widely researched for stereo image compression, the characteristics of the residual disparity between left-right images have not been extensively studied. Some researchers are involving in exploit the similar between the left-right stereo image pairs and a variety of disparity estimation methods have been proposed. Moellenho showed in their paper that the disparity-compensated residual image has different correlation characteristics, and is composed primarily of strong vertical-direction edge components surrounded by large areas of near zero intensity [5]. Konrad proposed a disparity estimation method that combines the reliability of feature-based correspondence methods with the resolution of dense approaches. They finished the task by three steps. First, finding feature points in the left and right images using Harris operator. Then, selecting those feature points that allow one-to-one left-right correspondence based on a cross-correlation measure. At the end, using the computed correspondence points to control the computation of dense disparity via regularized block matching that minimizes matching and disparity smoothness errors [6]. Siegel et al., exploit the correlations between 3D-stereoscopic left-right image pairs to achieve high compression factors for image frame storage and image stream transmission [1]. They demonstrated that a reasonable synthesis of one image of a left-right stereo image pair can be estimated from the other uncompressed or conventionally compressed image augmented by a small set of numbers that describe the local cross-correlations in terms of a disparity map.

In order to further achieve the higher compression ratio, we have make researches focused on the disparity characteristics of the residual images for stereo image compression. The residual image is composed primarily of strong vertical-direction edge components surrounded by large areas of near zero intensity, and has Gaussian function characteristics. This paper demonstrates these properties, and outlines how they might be exploited in designing compression algorithms. The rest of the paper is devoted to some details of our experimental results, and a discussion of possibilities for the future.

2. Residual image with spatial displacement

It's obvious that the 3D-stereoscopic image pair is very similar and most of the information contained in either is repeated in the other, which is often described as "highly redundant" or as "highly correlated". By the unique characteristic, there are some existing stereo compression techniques [4].

2.1. Correlation

One of the early ideas was to use a correlation algorithm to generate a three-dimensional map of the environment from the stereo image pairs obtained from the cameras. This map would then be transmitted, along with one of the stereo image pair, and on reception the stereo scene could be reconstructed from the data. The benefit of this type of system is that the viewpoint can be altered, allowing an observer to look at the scene from slightly different angles while retaining true perspective. The primary problem with such a system is that the correlation algorithm is typically computationally expensive, with the faster algorithms often being quite inaccurate. It is simply not possible to do stereo correlation and video compression at a high enough frame rate.

2.2. Residual

A simple stereo compression technique is to compress one of the stereo pair as a normal monoscopic image, and then compress the difference between the two stereo image pairs. Since the left and right images share much of the same information, there will often be large areas where the difference between the two images is small or zero, which means that the difference should compress well.

2.3. Sub-sampling

Using different resolutions for each of the stereo image pairs was attempted and, although the final result could still be stereoscopically fused, the low resolution image was clearly evident. It was decided that sub-sampling one of the stereo image pair does not result in further compres-