Abstract

The process of analyzing images of real world scenes taken from two different viewpoints in order to estimate depth is called stereovision. The main problem in stereovision analysis is to find corresponding points in images taken from different perspectives. Several algorithms are available in the current literature for finding the corresponding pixels in stereo images. These algorithms are classified into area-based, feature-based and pixel-based techniques. In this work, merits and demerits of each of these techniques are discussed. A brief review of some of the algorithms in each technique is presented.

Area-based (window-based) techniques are preferred for correspondence as they provide matching at every pixel using intensities of pixels in the vicinity of corresponding points. These techniques do not need pre-processing to extract features and post-processing such as interpolation to obtain the disparity map. Stereo matching algorithms in area-based techniques are developed with an assumption that all the pixels in the corresponding windows have equal disparity. However, this assumption is violated due to occlusion, depth discontinuities and projective distortion. In addition, the pixel intensities in the corresponding windows of stereo images may differ due to different camera parameters and different contrast. In this work, we developed algorithms to solve the correspondence problem in the presence of these phenomena.

Two stereo matching algorithms using fuzzy-set theoretic approach are developed. In these algorithms, the contrast of stereo images is assumed as equal. The first algorithm establishes the correspondence, even if the pixel intensities differ in the corresponding windows due to different camera parameters. The second algorithm establishes the correspondence in the presence of the occluded pixels in the corresponding windows. Similarity measure, viz., weighted normalized cross correlation (WNCC) is used in the

second algorithm. This algorithm is successful, even if the pixel intensities in the corresponding windows differ due to different camera parameters. Properties of the fuzzy-set theory based similarity measures used in these algorithms are discussed.

Success of stereo matching algorithms in area-based techniques depends on similarity measure used for matching. Linear correlation measures such as normalized cross correlation (NCC), sum of squared difference (SSD) and sum of absolute difference (SAD) depend on the absolute value of pixel intensities. In this work, rank correlation measures based on non-parametric statistics are considered for stereo matching. These measures use relative magnitudes of pixel intensities. Hence, the contrast of stereo images need not be equal. We compared the performance of Ordinal measures (κ), Spearman rank correlation (r_s) and Kendall rank correlation (τ) in terms of minimum false matches in the presence of salt and pepper noise, sensitiveness to insignificant data and discriminatory power (ability to reject two non-corresponding windows). In our experiments, superiority of Kendall rank correlation is established.

Area-based techniques are computationally expensive. In addition, these techniques fail to establish the correspondence in the poorly textured region due to insufficient intensity variation in the corresponding windows. Feature-based techniques fail to establish the correspondence due to sparse/no features in poorly textured region. In this work, a hybrid approach is considered for stereo matching having the advantages of both feature-based and area-based techniques. In this approach, the pixel coordinates are used for matching instead of pixel intensities. Rotationally invariant Zernike moments are used to match the edge pixels in stereo images. These matched features are used as control points to calculate the affine model of stereo images. Dense matching can be obtained using this affine model of stereo pair by considering the pixel coordinates. The use of this approach is limited to stereo images satisfying linear affine geometry.