Scientific
Research
Publishing

3D Depth Measurement for Holoscopic 3D Imaging System

Eman Alazawi, Mohammad Rafiq Swash, Maysam Abbod

Department of Electronic and Computer Engineering, College of Engineering, Design and Physical Sciences
Brunel University, West London, UK
Email: Rafiq.swash@brunel.ac.uk

Abstract

Holoscopic 3D imaging is a true 3D imaging system mimics fly's eye technique to acquire a true 3D optical model of a real scene. To reconstruct the 3D image computationally, an efficient implementation of an Auto-Feature-Edge (AFE) descriptor algorithm is required that provides an individual feature detector for integration of 3D information to locate objects in the scene. The AFE descriptor plays a key role in simplifying the detection of both edge-based and region-based objects. The detector is based on a Multi-Quantize Adaptive Local Histogram Analysis (MQALHA) algorithm. This is distinctive for each Feature-Edge (FE) block *i.e.* the large contrast changes (gradients) in FE are easier to localise. The novelty of this work lies in generating a free-noise 3D-Map (3DM) according to a correlation analysis of region contours. This automatically combines the exploitation of the available depth estimation technique with edge-based feature shape recognition technique. The application area consists of two varied domains, which prove the efficiency and robustness of the approach: a) extracting a set of setting feature-edges, for both tracking and mapping process for 3D depthmap estimation, and b) separation and recognition of focus objects in the scene. Experimental results show that the proposed 3DM technique is performed efficiently compared to the state-of-the-art algorithms.

Keywords

Holoscopic 3D Image, Edge Detection, Auto-Thresholding, Depthmap, Integral Image, Local Histogram Analysis, Object Recognition and Depth Measurement

1. Introduction

The development of 3D technologies for the acquisition and visualisation of still and moving 3D images remains a great challenge. Holoscopic 3D Imaging also known as Integral Imaging is a type of autostereoscopic 3D

technology that offers completely natural 3D colour effects almost like the real world [1]. It has the ability to produce a true volumetric representation of a real scene in the form of a planar intensity distribution using a microlens array [2]. This technique can be seen as an evolutionary step in providing promising applications in the areas of 3DTV, 3D cinema, medicine, robotic vision, biometrics, military, digital design, and digital games [3] where 3D depth plays [4] [5] a key role. Recently, there have been some developments in holoscopic 3D imaging and display system, in particular 3D reformatting [6], aperture stitching [7], 3D pixel mapping [8] [9], 3D display [10], 3D search and retrieval [11] and computer graphics [12] including medical image visualisation [13] and pre-processing [14].

Today's 3D display technology supplied to the viewers is based on stereovision techniques, which require the viewers to wear a spatial pair of glasses to perceive the left and right eye image via spatial or temporal multiplexing [1]-[3].

Many research groups have proposed multi-view autostereoscopic 3D displays where the viewers will not need to wear glasses; however, it requires the viewers to stay within a defined distance range from the display and it also fails to simulate and produce natural viewing experience. As a result, such techniques are liable to cause eye strain, fatigue, and headaches during continuous viewing while the viewer's focus on the image [3] [15].

Currently, advances in micro-lens manufacturing technologies as well as increases in processing power and storage capabilities mean that holoscopic 3D imaging has become viable for future 3D imaging and display technology. Thus, it has attracted a great attention of scientists and researchers from very various disciplines.

The fundamental principal of holoscopic 3D imaging is to capture the spatio-angular distribution of the light rays from the object using a lens array [16] [17]. Generally, the lens arrays are categorised into two main types based on application requirements. The unidirectional micro-lens array is a simplified version that offers a single directional 3D depth and motion parallax; whereas, the omnidirectional micro-lens array offers complete 3D depth and motion parallax [18]. Hundreds of micro-lenses are packed together to make up a single micro-lens array. The spatio-angular distribution of light rays can thus be sampled as a form of micro-images also known as an Elemental Image (EI) array. A planar detector surface records the holoscopic image as a 2D distribution of intensities. A 2D array of viewpoints integrates the holoscopic 3D image due to the structure of the encoding lenticular or micro-lens array. Therefore, each lenslet views the scene at a slightly different angle than its neighbor, so the capture of the scene is from many viewpoints. It contains of all the information necessary for 3D depth information to be spatially represented, and then displays via a similar decoding array [19]. The principle concept for capturing and displaying the image is illustrated in **Figure 1**.

This paper presents a 3D depth estimation algorithm with feature-based technique to generate robust geometric 3D structures from a holoscopic 3D image. It facilitates correspondence matching for a small set of regularly spaced pixels in the EI array. The AFE descriptor is implemented using MQLHA in the extracted stage. The extracted features are the combination of the edge and point features that have the ability to handle relatively featureless surface contours.

2. Related Work

The quality of the segmentation for object recognition results assumes vital significance for evaluating which targets the objectives (*i.e.* identify precise object contours and coarser) of object presence detection algorithm have met. This paper addresses the 3D depth estimation challenge in holoscopic 3D imaging system, which improves the recognition by accurately identifying the object's contours.

The 3D imaging technologies overcomes the 2D imaging limitations in terms of object size processing. They offer more low-level features and cues compared to the 2D imaging technology and are thus attracting researchers' interest in the 3D area.

Manolache *et al.* [19] were the first to report work on depth extraction from holoscopic 3D imaging systems using point spread functions to represent the correlated holoscopic 3D imaging system. An inverse problem tackles the depth estimation; however, the image inverse problem proves to be ill-posed. Cirstea *et al.* [20] used two regularization methods capable of presenting realistic reconstructions to cure the ill-posedness.

A practical feature-based matching approach for obtaining unidirectional 3D depth measurement from the holoscopic 3D image through disparity analysis and viewpoint images extraction was explored and presented by Wu, *et al.* [21]. To form the Viewpoint Image (VPI) sample pixels from different Elemental Images (EI) are used
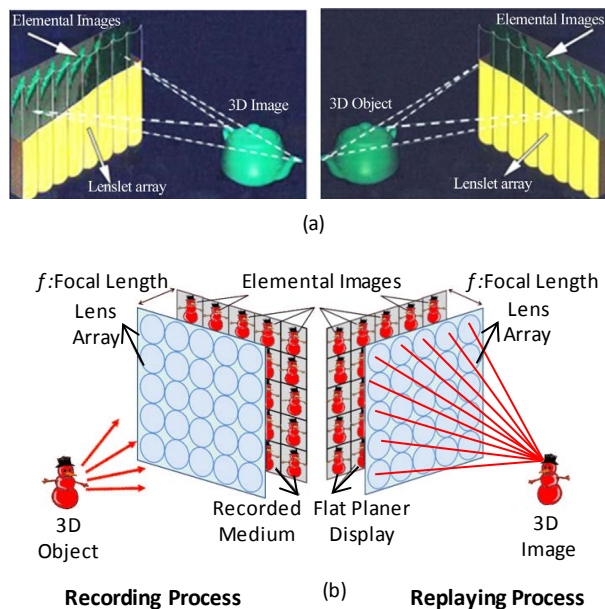
Figure 1. (a) Unidirectional holoscopic 3D imaging system using a cylindrical micro-lens type: (left) Recording process and (right) Replay process. (b) Omnidirectional holoscopic 3D imaging system using a spherical micro-lenses array. (Left) Recording and (Right) Playback process).

rather than a macro block of pixels corresponding to a micro-lens unit. Each VPI presented a two-dimensional parallel recording of the 3D scene from a particular direction. The calculation of the object depth was from the disparity estimation between these VPIs' displacements using a depth equation. The resulting 3D images contain enormous amounts of non-useful information and homogeneous regions.

Zarpalas, *et al.* [22] proposed two depth extraction methods using the extraction and matching of a set of feature points and then trying to fit a surface to the reconstructed features that aims to optimally connect them. In the first method, a H3DI grid attempts to identify the surface points as a subset of the intersections between the pixels' projection rays to the appropriate surfaces. The disadvantages of this method are that it produces non-smooth and non-uniform sampling of the 3D space solutions. In the second method the authors derived their frameworks from [21] and employed graph cuts to match the surface on the extracted features by estimating the disparities between the pixels of the given pair of VPIs. It is necessary to solve multiple optimizations (*i.e.* one for every pair) and merge all depth-maps to produce a final depthmap, which disregards the piecewise nature of the depth. Later, Zarpalas *et al.* [23] proposed a depth estimation algorithm formulated as an energy optimization global matching problem, where the solution was to use the anchoring graph cuts technique. The algorithm successfully overcomes the unidirectional holoscopic 3D imaging system problems and produces a smooth scene structure. In this technique, to separate the objects from the background, a foreground mask must be employed. The authors were unsuccessful when trying to generate foreground masks directly from the holoscopic 3D imaging system, as it is hard to calculate depth on large non-informative and homogeneous regions [23]. Moreover, due to the energy minimization algorithm (*i.e.* global methods), this technique would be computationally expensive.

Recently, the authors' in [24] have extended the work presented by [21] to obtain a more precise object depth in terms of sustaining an acceptable object contour and noise reduction. The authors introduced an automatic masking procedure with depthmap calculation, which results in the removal of any errors from the background. This process is fast and effective in identifying miscalculated areas in the background, which can be accurately removed. However, this was unsuccessful in some object depth regions that have closed colour similarity to other regions and textureless regions with features similar to the background. More recently, Alazawi *et al.* [25], overcame the deficiency of spatial information associated with local descriptors. The proposed technique enhanced the accuracy of the depth-through-disparity algorithm by taking into account that the depth is piecewise continuous in the space.

The authors created automatic descriptors, which concentrated on extracting reliable sets of features to exploit

the rich information contained in the central VPI. Furthermore, the approach was built on a new corresponding and matching technique based on an adaptive aggregation window size. A novel approach to setting feature-point blocks and matching of the corresponding features was achieved automatically. The advantages of this framework was that it solves the disadvantages identified in [21]-[24], namely the problems in the depthmap estimation from unidirectional holoscopic 3D images. The approach overcomes errors that commonly appear in the depth-through-disparity algorithm, *i.e.*, uncertainty and region homogeneity at the image location and dissimilar displacements within the matching block around the object borders. Moreover, the coherent object segmentation/separation technique was adopted to enhance the 3D object scene where perceiving the object depth calculation is more difficult and less accurate than in enormous non-instructive and uniform regions. Recently, Alazawi *et al.* [26] presented a novel approach to estimate and extract an accurate full parallax 3D depth measurement from omnidirectional holoscopic 3D imaging systems. The work was done for real and computer-generated holoscopic 3D images, which confirmed the efficiency, robustness, and speediness [25] of the algorithm without the need to use further enhancement processes.

Recently, an innovative approach has been proposed for 3D depth measurement [27] with the optimum self-similarity between the extracted VPIs was exploited to predict a best match score for a set of feature-edge blocks for the matching and corresponding process. A novel feature-based edge detection algorithm was used to extract reliable and efficient feature-edge automatically. The extracted feature edges are utilized to produce 3D depth-based-edge map. The produced 3D map is simultaneously capable for localization and mapping depth information of the object/objects in the scene to estimate an accurate depth map. The approach was successfully implemented on unidirectional holoscopic 3D imaging system, which is a horizontal directional imaging system. This was an initial stage to evaluate the successfulness of the proposed approach.

This framework is an extension of the author's previous work [27]. Inclusiveness and generalization of the algorithm is the key to the success of proposed approach, *i.e.* able to be implemented on different systems. In this work the focus is on the generalization capability of the approach to be implemented on an omnidirectional holoscopic 3D imaging system to display full-colour images with continuous parallax. This is more sophisticated and useful than unidirectional holoscopic 3D imaging system. The novel framework offers a significant enhancement for the 3D depthmap estimation compared with the above state-of-the-art techniques. The efficiency of the approach depends upon a prudent choice of features. Then the feature-edges are optimized and are utilized to assist and estimate 3D depthmap from holoscopic 3D imaging system.

The proposed AFE detection algorithm can extract object boundaries and discard non-boundary edges caused by textures on object surfaces that affect the accuracy of 3D measurement. **Figure 2** illustrates the block diagram of the proposed approach for generating 3D depth map from the existing data automatically without human interaction, particularly since human interaction could result in human errors and also be time-consuming.

3. Proposed 3D Depthmap Estimation Method

In this approach, the model-building task of the 3D depthmap estimation depends on either the explicit recovery or indirect inference of the 3D features embodied in the depth data. The 3D objects can be adequately modelled in terms of feature extraction from the plane patches and spatial connection among the plane patches to identify the structure of the 3D world and provide the essential constraints of matching. The efficacy of the approach crucially depends upon the robust estimation of these features and their connection. The proposed approach is an unsupervised edge detection algorithm that automatically determines thresholds in image edge extraction. The non-boundary feature-edges are not desired in the depthmap estimation stage. The initial extracted feature-edges are optimized by other phases of the proposed AFE detector to extract the reliable feature-edge that is used for the depthmap estimation stage.

The three-stages of modeling the proposed approach for generating an accurate 3D depth cues map that is used to estimate and extract 3D object are shown in **Figure 2**. The proposed 3D depth estimation approach automatically builds up a new version of 3D depth-based-edge technique for holoscopic 3D imaging system. Edges typically occur on the boundary between two different regions in the image, and therefore, are viewed as a contour across which the image intensity value undergoes a sharp variation that can be exploited to aid segmentation and 3D object recognition [28]. Due to this fact, the task of finding and localizing these changes in intensity is required for edge detection and subsequent extraction of 3D surface boundary points/feature-edges. The extracted features are the combination of the edge and point features that have the ability to handle relatively featureless surface contours.
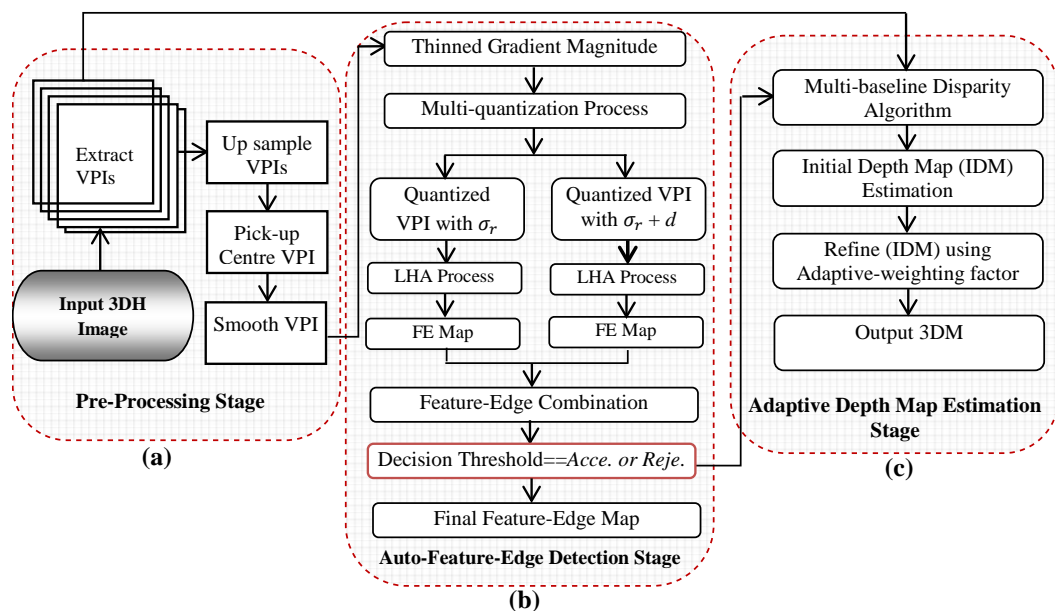
Figure 2. Illustration of the stages in the proposed approach (a) pre-processing process, (b) auto-feature detection process, and (c) adaptive depth map estimation process.

3.1. Pre-Processing Stage

The key idea behind this stage is the resampling of the collected data into the form of a VPI for the subsequent process. The holoscopic 3D imaging records a set of 2D images from a 3D scene on a photographic film placed at the focal plane of the micro-lens sheet (**Figure 1(a)**). The individual 2D images are referred to as EIs. Each EI contains information from a slightly different view of the object. A vital step in modeling for the 3D depth map for 3D object localization system is to extract 3D cues from images of different VPIs.

To this end, the "Viewpoint Image Extraction" is an important and initial stage, which transforms the EIs into VPIs by exploiting the strong interval correlation between pixels displaced by one micro-lens. To obtain single orthographic (VPI) view from Unidirectional or Omni-directional H3DI data, will require sampling of all pixels in the same location under different micro-lens, resulting in a single view direction of the recorded scene. This image is called the viewpoint image (VPI) containing information of the scene from one particular view direction. The author's previous work [25] [26] has clearly explained the transformation and selection process of the VPIs from both unidirectional and omnidirectional holoscopic 3D images. The construction of VPIs is illustrated in **Figure 3**. The extracted low resolution VPIs are up-sampled by N times using quadratic interpolation in horizontal and vertical directions from omnidirectional holoscopic 3D images whereas only horizontal direction from unidirectional holoscopic 3D images. N is ascertained by knowing the farthest object from the camera.

3.2. Auto-Feature-Edge Detection Stage

A Modeling Auto-Feature-Edge (AFE) detector algorithm for holoscopic 3D imaging system is necessary to produce an accurate true volume. The main aim of using an AFE is to have a single detector that is able to integrate 3D cues and locate objects in a scene. The feature detection stage detects both edge-based and region-based features using the automatic threshold selection method for image analysis to provide an efficacious feature. For each FE block this will be an individual process of looking for large contrast changes. The FE points are pixels at or around which the image values are subject to sharp variations.

There are three steps to setting features through edge detection. First, suppress noise, then detect and enhance edges, and finally localise and extract accurately features points. These extracted features are used for the integration of 3D cues to locate objects. Thresholding also plays a key role for feature detection due to its simplicity, high speed of operation, and ease of implementation [29]. This process uses an automatic bi-level (one threshold value) thresholding algorithm to set and segment features on the center (reference) VPI into regions corresponding to the object(s) and background *i.e.* object in the scene. Robust feature edge information is not always
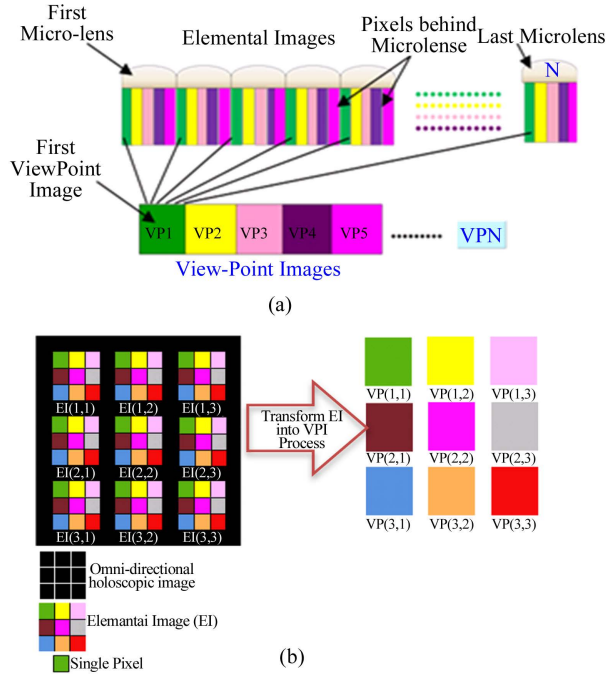
Figure 3. The principle of reconstruction of VPIs: (a) shows the extraction process for unidirectional holoscopic 3D images and (b) shows VPI reconstruction fro omnidirectional holoscopic 3D images (For simplicity, assume there are 3 × 3 pixels under each micro-lens, and extract one pixel from the same position under different micro-lenses are placed in orderly fashion, forming on VPIs.

achievable when using a single technique across a wide range of image content. To extract distinctive Feature-Edge (FE) blocks, *i.e.*, large contrast changes (gradients), the integration of more than one technique is a suitable approach to enhance the performance of edge-based features shape recognition. **Figure 2(b)** shows the principle of the AFE threshold algorithm adopted in the Multi-Quantize Adaptive Local Histogram Analysis (MQALHA) algorithm [30]. The processing of the local histograms of small non-overlapping blocks of the thinned gradient magnitude acts as an additional noise rejection step and automatically determines the local threshold for each block [29]. The processing can be divided into the following two major phases as shown in **Figure 4**.

1) Feature Edge Enhancement Phase

First or second order derivative masks in the spatial domain are used in the edge enhancement phase to smooth the image and calculate the potential edge features. The following explains the main steps of the phase:

a) Smooth the VPI to reduce the noise by convolving the VPI with a discrete Gaussian window = 3 from the 2D Gaussian function: (1)

$$G(x, y) = \frac{1}{2\pi\sigma} \exp\left[-\left(x^2 + y^2\right)\big/2\sigma^2\right] \tag{1}$$

where, the parameter σ indicates the width of the Gaussian distribution that defines the effective spread of the function.

b) The boundary between two regions is detected using the gradient magnitude increase as the contrast between the FE on a smoothed VPI (x , σ) and the background by computing the magnitude $m(x)$ and orientation $\theta(x)$ of the greatest change in intensity in the small neighborhood of each pixel [30] as pixel differences:

$$\overrightarrow{m(x)} = \sqrt{\left[\text{VPI}(x+1, y) - \text{VPI}(x-1, y)\right]^2 + \left[\text{VPI}(x+1, y) - \text{VPI}(x-1, y)\right]^2} \tag{2}$$

$$\theta(x) = \tan^{-1}\left[\frac{\text{VPI}(x, y+1) - \text{VPI}(x, y-1)}{\text{VPI}(x+1, y) - \text{VPI}(x-1, y)}\right] \tag{3}$$
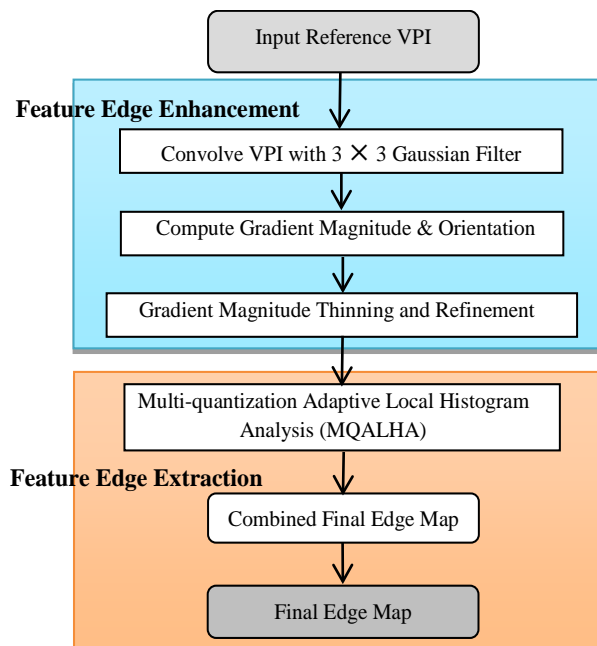
Figure 4. Enhancement and extraction of the feature edge steps.

Therefore, for the smoothed VPI the gradient can be written as:

$$\mathrm{VPI}_\sigma = \nabla\big[G_\sigma(x) * \mathrm{VPI}(x)\big] = [\nabla G_\sigma](x) * \mathrm{VPI}(x) \tag{4}$$

c) Thin the contour to a single pixel width, thus, the non-maxima suppression process is performed to maintain only Fes, where their gradient magnitude is the local maximum. Each pixel in the gradient magnitude VPI_σ is set to zero, unless it is a local maxima, otherwise it is kept unchanged. To eliminate the weak intensities of FE pixels, we replace each pixel's intensity in the thinned gradient magnitude VPI by the average of its surrounding pixels in a 3×3 window.

2) Feature Edge Extraction Phase

The reliable feature edges are selected in this phase by processing the computed gradient of VPI. The aim of this step is to be able to increase the continuity of the FEs for the multi-resolution aspect to enhance the computational speed and to enhance the technique's efficiency so it can be used with a range of noise levels. The following set out in more detail this edge extraction process.

3.2.1. Multi-Quantization Adaptive Local Histogram Analysis (MQALHA) Algorithm

The goal of this step is to increase the FEs' continuity, improve the computational speed of the multi-resolution aspect, and optimize the method for use with various noise levels. The MQALHA algorithm was used to design and implement a new first derivative operator based on the edge detection method to provide the best trade-off between detection and edge localization and resolution [31]. The MQALHA algorithm achieves the multi-resolution task through local histogram analysis by producing low and high edge maps from two differently quantized thinned gradient magnitude VPIs. The MQALHA flow diagram for detecting and extracting feature-edges is shown in **Figure 2(b)**. The attractive attribute of this technique is the estimation of noise from the histogram of the thinned gradient magnitude image.

The following two steps demonstrate the MQALHA algorithm:

First Step: Image Quantization

Using the scheme presented in [20], the multi-quantization process scheme is used to automatically quantize the refined thinned gradient magnitude VPI into two 12-level intensities. In the gradient magnitude histogram the small peak fluctuations were smoothed using a 1-D Gaussian filter with a standard deviation of 0.5. The first local maximum in the smoothed histogram is located and the grey level set to σ_r that corresponds to the position of the first maximum. The first quantized VPI is generated using this estimated value to quantize the thinned

gradient magnitude VPI. The σ_r value is moved a distance (d) right so that the second VPI is produced from a higher starting point than that in the 12-level quantizer. The value of estimated σ_r is the grey level that is combined with a Full-Width-at-Quarter-Maximum (FWQM) height of the distribution, which is used to automatically predict the distance (d) and model the noise inhabitants in the gradient magnitude image. In this work due to the VPIs having a lower resolution (from orthographic projection image) than the perspective projection images, the higher value of ($\sigma_r + d$) was experimentally selected to correspond to the position of a Full-Width-at-Quarter-Maximum (FWQM) of peaks in the noise distribution as shown in **Figure 5(a)**. Due to the noise level in the image, the second quantized starting point will increase or decrease along with σ_r. The error noise can be reduced by the adaptive quantization process at this stage, which is an advantage as this means that larger local histogram smoothing operators are not needed. This results in the overall edge extraction algorithm having a greater computational speed.

To robustly represent the local histogram without losing information the quantization step is required, and it is needed to enable the processing to be conducted on a small block.

Second Step: ***Optimal Threshold***

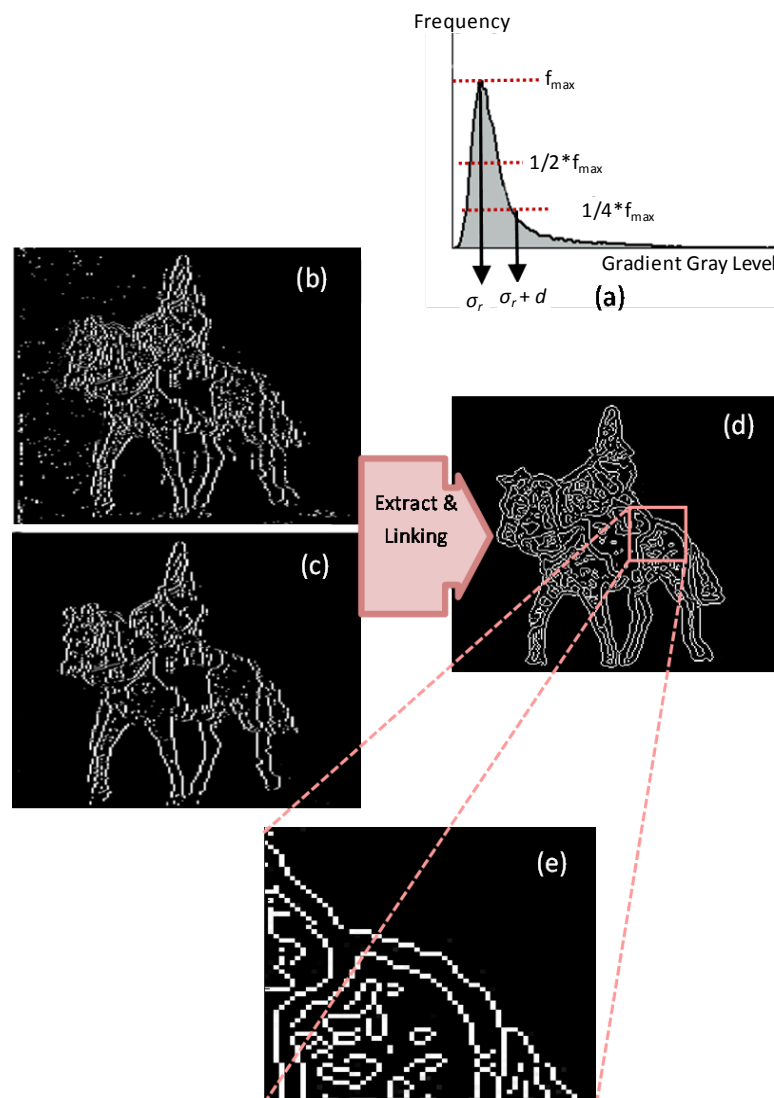**Figure 5.** AFE detection stage, (a) first starting point (σ_r) (low threshold) setting strong edges and second starting point (σ_r + d) (high threshold) setting weak edges, (b) low thresholding result on "Horseman" real-world VPI (noisy map), (c) high thresholding re-sultant on "Horseman" (without noise) (d) applying extract and link process to set final EF map, and (e) magnified section of a final FE map.

To extract the two edge maps at different resolutions, local histogram smoothing and thresholding are undertaken on the quantized VPIs, which are divided into 4×4 non-overlapping blocks [30]. A 1-D Gaussian kernel of width $W = 3$ is used to smooth the blocks that are computed from the local histograms of the blocks. Then depending on the smoothed histogram's shape each block is classified as either a background block or EF block. Background blocks are those where the histogram is unimodal, and in these all the pixels are set to a value of zero. If the block is not in the background, take the smoothed histogram and differentiate it before taking the first valley's position as the threshold value for that block. For all pixels where the threshold value is exceeded by the quantized grey levels they are considered to be edge pixels and set to a value of one [31]. In the lower threshold FE map the thresholded block has the VPI of σ_r and is saved, while in the higher threshold FE map the thresholded block has the VPI of $(\sigma_r + d)$ and is saved for further processing. The auto threshold results are shown in **Figure 5(b)**, **Figure 5(c)**.

3.2.2. Concatenating and Linking Edge Process/(Combination Process)

The edge combination process is applied to combine the two EF maps shown in **Figure 5(b)**, **Figure 5(c)** to form a single feature integration map "final edge map" in **Figure 5(d)**, **Figure 5(e)** by extracting and linking the discontinuous segment FEs obtained at a higher threshold with those from lower threshold FE maps. The process is started by scanning the higher threshold FE map to find the end of an edge segment (*i.e.* endpoint). The FE endpoints of the lower threshold FE map are examined for possible connected successors using a local neighborhood window size of 3×3 in the direction of the FE endpoint (*i.e.* in the Non-maxima suppression stage direction). The successor edge point is determined by examining the values of the three probable neighbors of the lower threshold edge map identified from the direction of the endpoint. **Figure 6** shows the process where the successor edge point is generated by considering the values of three probable neighbors. Then, if one of the successor pixels of the lower threshold edge map contains an edge, the algorithm will extract and link it to the endpoint of the high threshold EF map. Then the newly identified edge pixel will become the new endpoint.

3.3. Hybrid Depthmap Estimation Stage

The objective of this stage is the implementation of the resultant FE map with a practical approach for obtaining depth through extraction VPI and disparity analysis. The approach has been presented in [25] to generate a 3D map shown in **Figure 2(c)**. Feature detection and matching are essential and crucial components used to align images, perform video stabilization, object recognition, and computer vision applications [29] [30]. In this approach the generation of 3D M consists of five sequential phases.

The following steps detail the five phases for the generation:

1) *Feature Descriptors Phase*: For any object in an image, attracting points on the object can be extracted to provide a feature description of the object. These descriptions are procured by the use of training images. A feature edge is considered as a "good" interesting point if its local maxima are greater than the present threshold. Low contrast areas and feature edge points are discarded. Interesting points are extracted from the second stage using the training center VPI to provide a feature description of the object and are stored in a database. Feature descriptors have been designed so that Euclidean Distances (EDs) in feature space are directly used to grade possible matches. The simplest matching technique in corresponding and matching feature blocks is to set a threshold and eliminates all matches not within this threshold value. However, setting the matching threshold faces some difficulties as it dependent partially on the variation of imagery being processed. Accordingly, a novel simple auto threshold technique has been used to detect feature points in the matching process to avoid the need for manual parameter specification and hence time saving. Ideally, its value should be adjusted depending
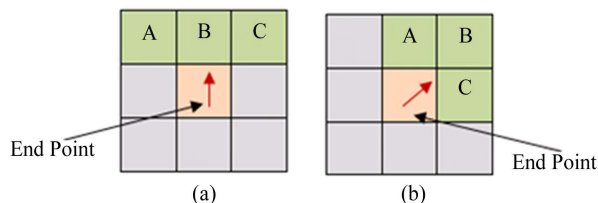

Figure 6. Probable successor location of the end point. (a) The three possible successors pointed by 90° central pixel and (b) 45°.

on the different regions of the feature space. It should be simple to implement and time efficient. Low thresholds result in incorrect matches being considered as correct whereas high thresholds result in correct matches being missed.

2) *Feature Tracking and Matching Phase*: For the extracted 2D VPIs, it is often necessary to track the displacement of individual features from VPI to VPI. The selected desirable features to track are often obtained in the same way as proper features to match. In the matching process the small displacements from pairs of VPIs (*i.e.* reference VPI and target VPI) have been calculated using the sum of square difference (SSD) cost function. To speed the process (thus saving time) in the feature tracking and matching phase a learning algorithm has been used to build special purpose recognizers to rapidly search for features in entire VPIs [25]. Moreover, the implementation of training sets of feature points on the reference VPI has led to an increase in the stability of the 3D solutions. The disparity displacement (*d*) is extracted, which is the physical distance between corresponding pixels in pairs of VPIs. To determine the initial disparity d_c^* of the central pixel "*c*" of the window block, we calculate the score disparity (cost match function) for all neighboring pixels of "*c*" in horizontal direction for unidirectional holoscopic 3D images and both horizontal and vertical directions for omnidirectional holoscopic 3D images using Equation (5). The cost match function for the initial disparity map in horizontal direction for 3DM estimation $\forall$ d in *x* direction is mathematically described as:

$$C(c,d) = SSD(d) = \sum_{x,y \in w} \left[\hat{I}_1(c) - \hat{I}_2(c+d) \right] \tag{5}$$

where, $C(c, d)$ is the cost function, d is a horizontal displacement vector between pixel *p* of the window (block) w in the I_1 center VPI and $(c + d)$ in the I_2 target VPI. The Multi-Baseline Algorithm (MBA) matching score function is used, where there are a number of stereo pairs (P) [21] [24] [25] to estimate the initial disparity map. The initial matching decision is carried out using the built up score function from all the VPIs pairs with different baselines:

$$C(cb,d)_{\text{initial}} = SSSD(d) = \sum_{i=2}^{K} \sum_{c \in w} \left[\hat{I}_1(cb) - \hat{I}_i(cb+d) \right]^2 \tag{6}$$

where, *SSSD* is the sum of *SSD* which minimizes the depth map (D) and P is the pairs of the VPIs. To determine the initial disparity of the central pixel *c* (*xl, yl*), the score for all neighboring pixels (*nb*) is calculated using different values of $d (d = 0, 1, \cdots, d\max)$; then the value of d_c^* is chosen that gives the pixel (*c*) the minimal cost among all computed disparity cost functions of *cb*. The d_c^* has the minimum given by:

$$d_c^* = \arg\left\{ \min_{d \in R} \left\{ \text{score}(cb, \boldsymbol{d}) \right\} \right\} \tag{7}$$

where, *cb* represents the window around pixel "*c*" and *R* is the search area. Disparity is estimated in the "winner-takes-all" method by only selecting the disparity label with the lowest cost. The disparity is verified whether it is dominant in the block cantered on pixel *c* (*cb*). If it is, then this disparity is considered as the final disparity and no further refinement is required; if it is not, further refinement is required.

3) *Feature Disparity Refinement Phase*: The resulting disparity map described above is not optimal because it still contains some noise and errors. Therefore, an extra step is necessary to detect and remove erroneous estimates on occlusions. The aim of this phase is to reduce artefacts and to resample the disparity maps to correct inaccurate full resolution disparity values and handle occlusion areas. For this, the following adaptive weighting factor approach by [25] has been used for refinement of the disparity map to prevent multiple neighboring edges from being selected as interesting for the same feature and to obtain large support areas in untextured zones, while adapting shapes near the object boundaries. Moreover, unreliable pixels are properly extrapolated from surrounding high confidence pixels. The adaptive weighting factor is a combination of three confidence terms:

$conf(wc) = \exp\left(\left\| -I_{cb} - \overline{I_{nb}} \right\| \right) / \sigma^2$ that represents the high variance of the center block *cb* and neighboring blocks *nb*, where, $I_{cb}, \overline{I_{nb}}$ represent the center pixel value *cb* and average value of the pixels within the block *nb*, respectively. Using small *w* allows for a reduction in the variance within each block. The distance term.

$dist(wc) = \sqrt{(i_{cb} - i_{nb})^2 + (j_{cb} - j_{nb})^2}$ is calculated as the spatial Euclidean distance between the coordinates.

where the high variance is closer to the centre block "*cb*" and "*wc*" is the weighting factor for the center point

"*c*". The colour difference term is calculated as the Euclidean distance between neighbor blocks in the CIELab colour space of pixel *cb* (C_{cb} = [L_{cb}, a_{cb}, b_{cb}]) and pixel *nb* (C_{nb} = [L_{nb}, a_{nb}, b_{nb}]), which is expressed as: $c_{diff(wc)} = \sqrt{\left(L_{cb} - L_{nb}\right)^2 + \left(a_{cb} - a_{nb}\right)^2 + \left(b_{cb} - b_{nb}\right)^2}$. The aggregated cost is computed as a weighted factor of cb for the different neighboring blocks and can be written as:

$$W\left(nb, cb\right) = conf\left(wc\right) \cdot dist\left(wc\right) \cdot c\_diff\left(wc\right) \tag{8}$$

The voting scheme used was from the support region; therefore, each pixel p collects votes from reliable neighbors as: $\text{Vote}_c\left(d\right) = \left\{n \middle| n \in U\left(c\right) \wedge d_n = d\right\}$, where, "*n*" is a consistent pixel from a reliable neighbor in the support aggregation window of the center pixel "*c*" and the disparity *dn* contributes one vote, which is accumulated in the set Votep (*d*). The final disparity of the *cb* is decided by the maximum majority weighted vote number $d_c^* = \arg\ msx\left|Vote_c\left(d\right)\right|$ According to Equation (6) and Equation (8) the filtered cost match function $C\left(bc, d\right)_{\text{Out}}$ at a *cb* block resolves into two terms: C_{initial} is the visual cost match function of the centre block "*cb*" and C_{smooth} indicates the smooth cost match function of the nearest neighbor block "*nb*" and is defined as:

$$C\left(cb, d\right)_{\text{Out}} = C\left(cb, d\right)_{\text{initial}} + C_{\text{smooth}} \tag{9}$$

By use of smoothing term Equation (6) and (8), Equation (9) can be rewritten as:

$$C\left(cb, d\right)_{\text{Out}} = C\left(cb, d\right)_{\text{initial}} + \sum_{nb \in n(bp)} W\left(nb, cb\right) \times \min\left\{C\left(nb, d\right)\right\} \tag{10}$$

where, $\sum_{nb \in n(cb)} W\left(nb, cb\right) \times \min\left\{C\left(nb, d\right)\right\}$ represents the smooth cost match function of the nearest neighbor block *nb*. With this improvement, the disparity map is more accurate as the calculation of the new disparity only involves the FE blocks belonging to the same region (object).

4) 3D Depthmap Calculation Phase: The depth equation $D = (d \cdot \psi \cdot F)/\Delta$ is derived through geometrical analysis of the optical recording process giving the mathematical relationship between the object depth and the corresponding VPI pair displacement "*d*", where, "*D*" is the corresponding depth to be calculated, and Ψ and *F* are the pitch size and the focal length of the recording micro-lens, respectively. *d* is the disparity of the object point within two extracted VPIs and Δ is the sampling distance between the two VPIs. The free-noise 3DM is calculated from the VPIs by establishing the corresponding disparity from Equation (10).

5) Feature 3D Map Smoothing: Finally a further disparity enhancement process is used to maintain a good trade-off between accuracy and processing time. A median filter is applied to further smooth the resulting 3D depthmap. The median filter is a robust method, often used to remove impetuous noise from an image [32]. A median window size filter of 5 × 5 is calculated by first systematizing all the pixel values from the neighboring pixels into a numerical order and then replacing the considered pixel with the middle ordered pixel and then the pixels, which are characterized as mismatches are rejected.

4. Experimental Results Sections

This section presents the results of the above described approach on two data sets with comparisons to the state-of-the-art 3D depthmap estimation from the H3DI system to quantitatively evaluate the performance of the proposed algorithm. The efficiency measurement and its core differences of this technique have been proved by implementing this algorithm on synthetic database, which provides the knowledge of the ground truth. The AFE detector algorithm is tested with both on real-world and computer generated holoscopic 3D content using unidirectional and omnidirectional holoscopic 3D images.

A feature extraction module with a robust edge detection grouping technique often involves the identification of all feature edges in the disparity map analysis. The FEs map is more distinctive and more robust to camera VPI changes. These edges provide robust local feature cues that can effectively help to provide geometric constraints and establish robust feature correspondences between VPIs. **Figure 5** shows the processing results of the AFE descriptor on the real-world image "Horseman". The result highlights the main contribution achieved in this approach. **Figure 8** and **Figure 9** illustrate then resulting 3D depth map of the proposed algorithm, where

the different objects are correctly perceived and distinguished from other objects in the scenes. In this section a quantitative comparison has been used that provided a fair way to evaluate the performance of the obtained results from the disparity maps.

To evaluate the efficiency of the proposed descriptor, *i.e.* AFE descriptors, Section 4.1 provides description and properties of synthetic database [33] and real world holoscopic 3D images. Section 4.2 discusses the evaluation performed on feature edge detection accuracy. To evaluate the proposed approach, comprehensive comparisons to state-of-the-art techniques [21] [25] are presented in Section 4.3.

4.1. Benchmark Holoscopic 3D Content Descriptions

Two types of data have been used into evaluate the performance of the proposed technique to evaluate and compare the performance with counterpart techniques.

1) Synthetic Holoscopic 3D Images

In this assessment 32 synthetic H3DIs were used to evaluate the performance of the AFE detection algorithm with known dense ground-truth disparity maps. A sample of 28 dataset images are available to the public in [33] and are varied from simple to complex object scenes and have unidirectional disparity. The 3D object scenes have divergent texture qualities and were transformed to fit the Field of View (FOV) of the computer-generated micro-lenses array. The micro-lens array consisted of 99 cylindrical lenses with 7×700 pixels resolution, lens pitch size = 0.5 mm, focal length $f = 4$ mm, lens height pitch = 50 mm with resolution 700 pixels, and the resolution of each 3 DHI size = 693×700 pixels [23] [33]. Due to space limitations, we cannot display the results of all 28 image sequences, therefore some images that have a variety of texture and intensity complexities have been presented in this paper. A further test was carried out on another four computer generated H3DIs which were obtained from the different project laboratories. The computer generated holoscopic 3D image of balls has been displayed, which contains several objects at different depths and a blank background. So they provide higher reliability of effectiveness for the evaluation. The pitch size (ψ), focal length (f) and the radius of curvature (r) of the micro-lens array used in this recording are 0.6 mm, 1.237 mm and 0.88 mm, respectively. A further test is carried out on omnidirectional holoscopic 3D images produced at Brunel University that contains several objects at different depths and with different background colours and which were obtained from a H3D virtual camera containing a micro-lenses array with a 150×150 micro-lens. The resolution of the image is 4500×4500 pixels, with 150×150 EIs with a resolution of 30×30 (900) pixels each, corresponding to a 1:1 ratio and therefore each VPI = 150×150 pixels. The micro-lens pitch size was (P) = 90 μm, focal length (f) = 0.001 μm and pixel pitch (ρ) = 3.1 μm.

2) Real World Holoscopic 3D Images

The proposal algorithm has been tested on real data of unidirectional holoscopic 3D images, which were commonly used with currently state-of –the art techniques [21]-[25]. The "Horseman" and "Tank" image size is 1280×1264 pixels and have 160 cylindrical lenses, therefore offering 160 elemental images of size 8×1264 pixels and 8 VPIs of size 160×1264 pixels. The first one has non-informative and homogeneous regions, while the second one has some textureless regions (*i.e.* lacks features). The "Palm" image is considered as a flat image, with size 5628×3744 pixels, captured with 84 cylindrical lenses, the width of each lens 67 pixels, providing 84 elemental images each of size 67×3744 pixels and 67 VPIs each of size 84×3744 pixels. Therefore, the depth is very narrow in this image. Further experiments are carried out in 3D VIVANT laboratory at Brunel University by the author in order to evaluate the general ability of the proposed approach. The test is carried out on thirteen "Box-Tags" that were captured at various distances (200 mm) from capture-camera, starting from 600 mm and ending at 3000 mm. The image consists of multiple depth planes as the distance between each Tag is invariant. The box was placed at a 45-degree slant from the Tags and its thickness was measured in the experiments with the camera starting at a distance of 600 mm from the "Box".

4.2. Feature Edge Evaluation

To evaluate the performance of the proposed feature edge detection technique, two commonly used feature detection descriptors were used to select local features in VPIs. The evaluation was carried out on real and synthetic unidirectional and omnidirectional holoscopic 3D images with different photometric transformation and scene types. Speeded up Robust Features (SURF) descriptor proposed by Bay *et al.* in 2006 [34] and Scale Invariant Feature Transform (SIFT) descriptor proposed by Lowe in 2004 [35] were used. SURF is a fast descriptor

and has good performance similar to SIFT; moreover it performs well with illumination changes [36]. **Table 1** shows the qualitative evaluation of these widely-known descriptors (*i.e.* SIFT and SURF) and the proposed technique. Furthermore, the proposed AFE detection algorithm is also compared with author's previous work in [25] [26] and with a manual threshold algorithm in [21]. The average number of detected feature points was used in this evaluation includes two types of images (*i.e.* 48 synthetic and real holoscopic 3D images). SIFT located 695 feature points, SURF 720, the author's previous work [25] 2910, manual threshold 2875 and the proposed AFE was proficient enough to create 4736 feature points. The results proved that the proposed AFE detection algorithm produced more reliable features. This is due to the extraction and linking edge scheme, which is an efficient way to close the gaps between closely spaced unconnected edges by linking one pixel gap in any one of the four directions. This operation helps increase the accuracy of the proposed feature detection algorithm, and significantly reduced the errors in edge location. Experimental results demonstrate that the proposed algorithm is robust and more accurate, where all real-edge pixels that represent the object boundaries are preserved, through the optimization of missing-edges process. This significantly increases the accuracy of the estimated depthmap on textureless regions. The overall edge fusion result is shown in **Figure 5(d)**, **Figure 5(e)**, in which nearly all object boundaries are extracted and non-boundary edges are eliminated. The Mean Absolute Percentage Error (MAPE) was used to evaluate the efficiency of the proposed algorithm and has been defined as [36]:

$$\text{MAPE} = \frac{1}{N}\sum\nolimits_{(x,y)}^{N}\frac{\left|A_{\text{depth}}(x,y)-E_{\text{depth}}(x,y)\right|}{A_{\text{depth}}}\times 100\% \tag{11}$$

where, $A_{\text{depth}}(x, y)$ and $E_{\text{depth}}(x, y)$ are the actual depth and estimated depth values at (x, y) pixel position, respectively, and N is the amount of compared pixels. It should be noted, that actual and estimated depths were calculated on the foreground only. It is worth pointing out that, the MAPE evaluation has been performed on synthetic H3DI database only, due to their ground truth availability.

Based on the experimental results, the SIFT descriptor MAPR is 4.48%, while SURF 4.59% and AFE detector 3.55% for the same quantity of the detected FEs (700 - 750). The evaluation results on **Table 1** (the average detected features on synthetic and real H3DIs. Note that, MAPR comparison has been performed on only synthetic datasets) clearly indicate that the proposed descriptor out-perform existing descriptors to generate more reliable feature-edges than them. As mentioned above, this is due to the extract and link endpoints on edge combination process in Section 3.2, which led to an increase in the number of feature-edges. Furthermore, it increased the reliability and robustness of these feature-edges points. Although, experimentally notable that there is a slight increase of MAPR of the proposed AFE 4.72% on total detected FEs. However, the MAPE percentage (for the same amount of detected FEs) indicated that SIFT is superior in the matching process than other techniques, but limits the performance of feature matching techniques in terms of speed and scalability.

The proposed algorithm had a superior performance for identifying more reliable feature-edges and is more robust to noise which is very important features to make the estimated depth map more reliable. Therefore, it overcame most of the difficult situations experienced in disparity analysis techniques, *i.e.* untrackable (object border, object occlusion and reappearance) regions by identifying strong feature-edge before the matching process. This is mainly acting when two different displacements exist within the matching block in matching analysis between pairs of VPIs. In other words, the performance of the AFE detector technique is of great importance in achieving correct depth estimation on object borders. In terms of speed (run time), the proposed AFE technique was superior than SIFT and slightly slower than SURF, therefore, the current SIFT cannot be used for real-time image or video retrieval and matching [37], although, SURF is fast and has as good performance as SIFT, but it is not stable to rotation and illumination changes [36].

4.3. Depth Accuracy Comparison

This section provides a few chosen sets of results from varied captured images with multiple objects. The goal was to select a variety of practically relevant situations where long-range and different plane depths are important. The evaluation of depth map results is based on error measures using the existing ground truth depth map of the synthetic database in [33]. The remaining data sets were synthetically generated at Brunel University. These results compare the performance of the proposed technique to the state-of-the-art methods for evaluating the effectiveness of each technique. Although, the actual depth map of real world holoscopic 3D images are not available, the depthmap generated from the proposed approach immediately compared to the real-word H3DIs

Table 1. Comparison the detected feature points.

Algorithm	Number of Detected Features	MAPE %	Matching Time (s)
SIFT	690	4.48	7.167
SURF	720	4.59	3.756
Authors Previous Work [14] [15]	2910	5.06	9.665
Manual Threshold [10]	2875	9.78	20.592
Proposed AFED (Auto Feature Edge Detection)	4736 700 - 750	4.72 3.55	10.895 4.086

which are used in [21] [22] [24] [25] methods. Moreover, the proposed technique used the parameters learnt in the above methods (**Table 2**) to give greater reliability to the proposed method.

The results in **Figure 7** show a comparison of estimated 3DMs of objects in the scene from synthetic H3DIs of the proposed algorithm to methods from the literature. The evaluation has been done using Equation (11) to compare with the results available in the methods from the literature. The smaller average depth error value is the better. **Table 3** shows the error comparison results. The MAPE is 8.44% for the proposed algorithm, 9.27% for [25], 11.60% of [22], 27.75% for [24] and 38.28% for [21]. The results indicated that the proposed technique was superior to the state-of-the-art algorithms and achieved a comparable performance in terms of depth estimation on the foreground noise reduction and shape recognition. Moreover, the proposed approach obviously outperforms the comparative methods in terms of shape recognition of the objects' contours and object recognition by separating the objects from the scene. This has been done successfully through the obtained robust feature boundaries as shown in **Figure 5(d)**, where all other methods failed to generate any form of depth map without creating foreground mask of the object contour. This can significantly lead to increased algorithm's computational time. **Figure 8** and **Figure 9** shown the results of the proposed algorithm on unidirectional holoscopic 3D images where the estimated depth object's contours are always obtained with high accuracy compared to [21] [25] algorithms. The proposed method is more distinctive and more robust to camera VPI changes (*i.e.* VPI reliance on FE detection changes as the VPI changes, and typically reflects the geometry of the scene). This provides effectively local feature information that is used to provide geometric constraints and create robust feature correspondences between VPIs in both terms: 3D depth estimation and object shape recognition.

4.4. Experimental Results

The performance of the proposed approach has been evaluated on a test-set comprising of a wide variety of realword and synthetic dataset images to quantitatively compare the effects of various cues. In this section the reported results are based on the following contributions of the proposed AFE algorithm:

A feature extraction module with a robust edge detection grouping technique that often involves the identification of all feature edges in the disparity map analysis. The FEs map is more distinctive and more robust to camera VPI changes, and these edges provide robust local feature information that can effectively help to provide geometric constraints and establish robust feature correspondences between VPIs. **Figure 8** illustrates the achieved 3DM results of the proposal algorithm for some real and synthetic images, where the different objects are correctly perceived and distinguished from other objects in the scenes, while the 3D map estimation from real world and synthetics omnidirectional holoscopic 3D images are illustrated in **Figure 10**. The red colour indicates the estimated objects' depth captured near to the camera, while the dark blue colour estimated the far objects' depths. The results shown, the 3D objects are in correct relative positions. A comparison of the accuracy of 3D depth map estimation of the proposed method and state of the art method [21] [24] [25] are shown in **Figure 7**. The high precision details in the objects' contours on real-word and synthetic H3DIs from the proposed algorithm showed that the proposed method outperformed other state-of-the-art algorithms. However, there was detection of depth discontinuities due to potential ambiguities in matching of FEs resulting in object surfaces at different depths shown in **Figure 8(c)**, **Figure 9(c)** and **Figure 10(c)**. The reason for these ambiguities relates to the following process: the thinning edge term (non-maximum suppression) that inhibits multiple responses by suppressing any neighbor pixel value (sets it so that it is equal to 0) to a single FE. Even so neighboring connectivity may provide important feature information, which is more than the explicit 3D measurements in depth estimation

Table 2. Main parameters and statistical data (S.D.).

Parameters	S. D
Pairs no. of VPI	6
Window size	17×17
No. of neighbour blocks-NBN	8
Max disparity-R	20
Probability voting adaptive window shape size	3
2D Gaussian filter size	3×3
Colour threshold pixel neighbour	2
Gaussian standard deviation starting σ_r	0.8
Similarity measurement method	SSD

Table 3. Comparison withstate-of-the-art methods.

Methods	MAPE %
[10] Local method (Manual threshold)	38.281
[11] Global method	11.609
[13] Local method (Manual threshold)	27.25
[14] Auto local method	9.278
Proposed approach/auto local method	8.44
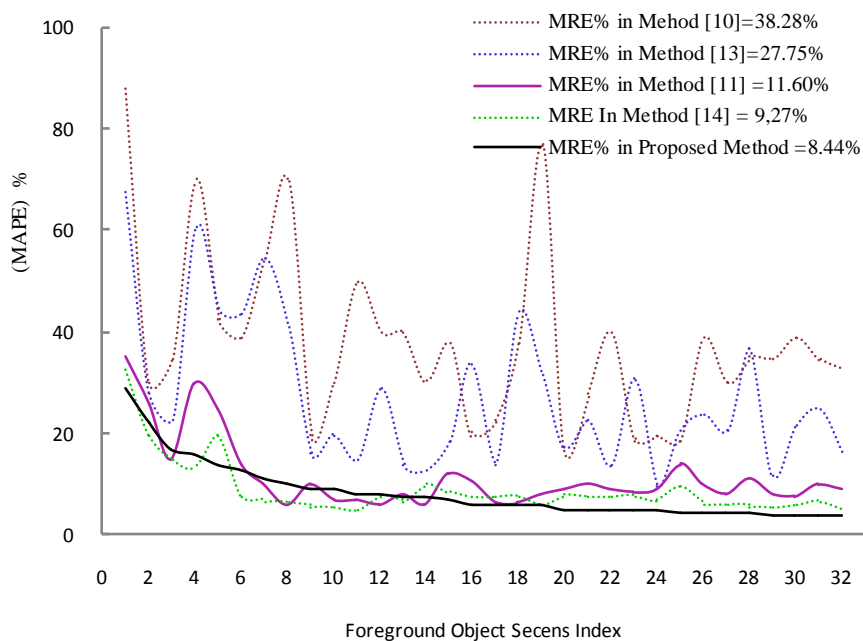
Figure 7. Comparison the proposed approach accuracy to literature methods on synthetic's database using the objective evaluation Mean Absolute Percentage Error (MAPE) for each object's in the scene.

from the extracted VPIs technique. Furthermore, although the viewpoint is constrained, the image resolution is low so not many details are visible; and indeed, the VPIs' quality varies (brightness, contrast, and sharpness) due to factors including illumination and focusing. Therefore, feature detection and extraction is constrained by
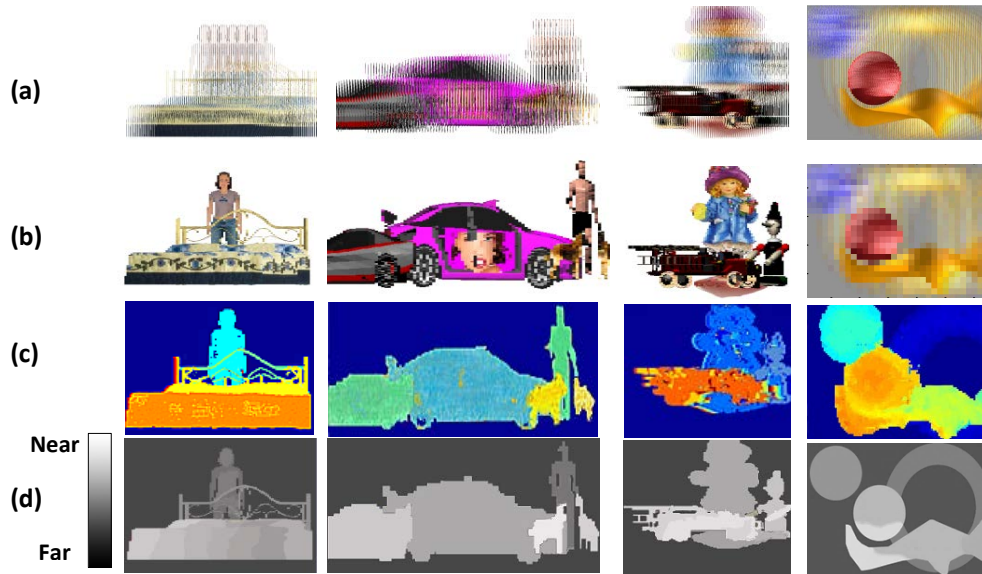
Figure 8. Illustration of the proposed method performanceon various unidirectional holoscopic 3D images with the depths of the scene layers start gradually from bright to dark on grayscale (*i.e.* bright indicate the object near camera captured and the darker indicates the background of the scene).
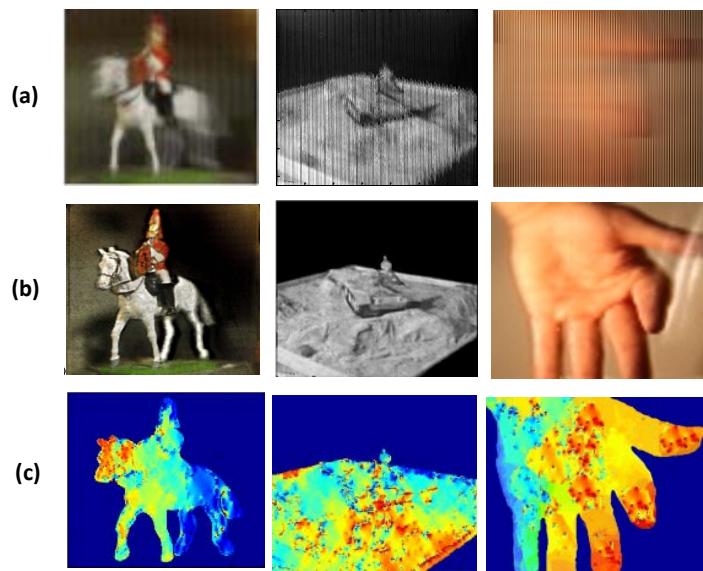

Figure 9. Resulting images of proposed method on various real world unidirectional holoscopic 3D images (a) Horsman, (b) Tank and (c) Palm.

the VPI and the resolution.

However 3D object segmentation is out of the scope of this proposed work. The proposed AFE algorithm can be used to present a simple image segmentation technique for object recognition from holoscopic 3D imaging system. This can create a separate object segment to form separate objects in the scene from the background for non-overlapping objects. In this technique, the results of the FE detection are the guide to identify regions and then convert every connected FE into a set of connected segments (contour). Therefore, the detected edge pixels are used to automatically extract the seeds required for growing a region through pixel aggregation of the interior pixel area of each object. The advantages of this method are that the segmentation is obtained without a prior knowledge of the number, position, size, or shapes of the objects and there is no requirement for a homogeneity criterion threshold in the region growing process. Furthermore, this mask can be used as a filtering
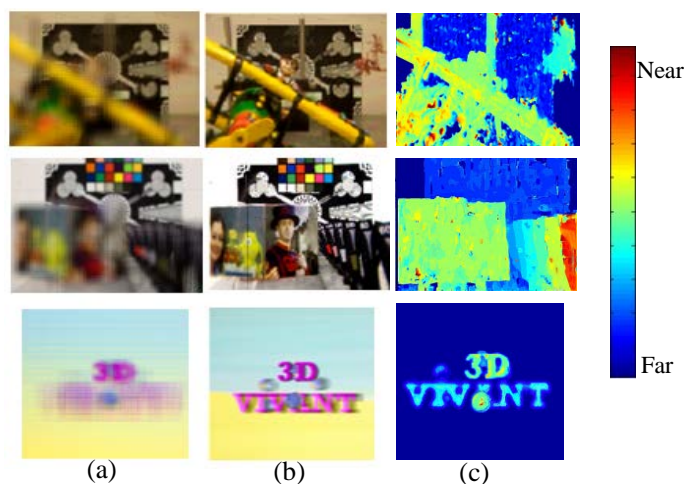
Figure 10. Resulting images of proposed method on omnidirectional holoscopic 3D images.

process to improve the accuracy of the extracted depth map from different methods by removing background noise and separating foreground. For more details see the author's previous work [27].

5. Conclusion

The proposed method achieved a better performance of 3D depth measurement estimation directly from disparity maps of the sampled VPIs. Consequently, setting and extracting efficient and informative features played a significant role in the performance of a precise depthmap estimation that had well-defined boundaries. This paper described a computational method to edge detection, feature-edge extraction, and object identification techniques for a holoscopic 3D imaging system where the edges were the main features. It functioned well with both real and synthetic unidirectional and omnidirectional holoscopic 3D images since we formulated the feature edge detection as a 3D object recognition problem. We devised an optimization scheme that resolved the 3D depth-contour problem. The solution to this was by combining a semi-global disparity map estimation technique with an FE detection technique to automatically generated the 3D depthmap, where the depth cues were derived from an interactive labelling of feature-edges using a novel auto feature-edge detection algorithm. Objects at different depths had been successfully detected and it had been clearly shown that the algorithm successfully overcame the depth contours problem for holoscopic 3D images and produced an accurate object scene. In addition, it improved the performance of the depth estimation related to generalisation, speed and quality. The use of the feature-edge to estimate an accurate 3D object contour map and the recognition and separation of objects in the optimization process along with the smoothness constraints were the basis of its success. Experiments identifying the precise object contours and which objects were present in a scene showed that the proposed method outperformed other state-of-the-art algorithms as illustrated in **Figure 9**. However, the depth estimation still needed further improvement on its efficiency when producing super resolution depth maps. This work could be improved and extended in terms of collection of training feature-edge information under different VPI resolutions, which could be utilized for obtaining even better performance.

References

[1] Hong, J., Kim, Y., Choi, H.-J., Hahn, J., Park, J.-H., Kim, H., Min, S.-W., Chen, N. and Lee, B. (2011) Three-Dimensional Display Technologies of Recent Interest: Principles, Status, and Issues. *Applied Optics*, **50**, H87-H115. http://dx.doi.org/10.1364/AO.50.000H87

[2] Zhu, Y. and Zhen, T. (2009) 3D Multi-View Autostereoscopic Display and Its Key Technologies. *Proc. APCIP Conference on Image Processing*, **2**, 31-35.

[3] Onural, L. (2007) Television in 3-D: What Are the Prospects? *Proceedings of the IEEE*, **95**, 1143-1145. http://dx.doi.org/10.1109/JPROC.2007.896490

[4] Alazawi, E., Aggoun, A., Abbod, M., Fatah, O.A. and Swash, M.R. (2013) Adaptive Depth Map Estimation from 3D Integral Image. 2013 *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting* (*BMSB*),

London, 5-7 June 2013, 1-6. http://dx.doi.org/10.1109/BMSB.2013.6621736

[5] Alazawi, E., Abbod, M., Aggoun, A., Swash, M.R., Fatah, O.A. and Fernandez, J. (2014) Super Depth-Map Rendering by Converting Holoscopic Viewpoint to Perspective Projection. 2014 *3DTV-Conference*: *The True Vision—Capture*, *Transmission and Display of* 3*D Video* (3*DTV-CON*), Budapest, 2-4 July 2014, 1-4. http://dx.doi.org/10.1109/3DTV.2014.6874714

[6] Fatah, O.A., *et al.* (2013) Generating Stereoscopic 3D from Holoscopic 3D. *3DTV-Conference*: *The True Vision— Capture*, *Transmission and Display of* 3*D Video* (3*DTV-CON*), Aberdeen, 7-8 October 2013, 1-3.

[7] Swash, M.R., Fernández, J.C., Aggoun, A., Abdulfatah, O. and Tsekleves, E. (2014) Reference Based Holoscopic 3D Camera Aperture Stitching for Widening the Overall Viewing Angle. 2014 *3DTV-Conference*: *The True Vision— Capture, Transmission and Display of* 3*D Video* (3*DTV-CON*), Budapest, 2-4 July 2014, 1-3. http://dx.doi.org/10.1109/3DTV.2014.6874765

[8] Swash, M.R., Aggoun, A., Abdulfatah, O., Fernández, J.C., Alazawi, E. and Tsekleves, E. (2013) Distributed Pixel Mapping for Refining Dark Area in Parallax Barriers Based Holoscopic 3D Display. 2013 *International Conference on* 3*D Imaging*, 3-5 December 2013, 1-4. http://dx.doi.org/10.1109/IC3D.2013.6732101

[9] Swash, M.R., Abdulfatah, O., Alazawi, E., Kalganova, T. and Cosmas, J. (2014) Adopting Multiview Pixel Mapping for Enhancing Quality of Holoscopic 3D Scene in Parallax Barriers Based Holoscopic 3D Displays. 2014 *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, Beijing, 25-27 June 2014, 1-4.

[10] Swash, M.R., Aggoun, A., Abdulfatah, O., Li, B., Fernández, J.C., Alazawi, E. and Tsekleves, E. (2013) Moiré-Free Full Parallax Holoscopic 3D Display Based on Cross-Lenticular. 3*DTV-CON*: *Vision beyond Depth AECC*, Aberdeen, 7-9 October 2013.

[11] Swash, M.R., Aggoun, A., Abdulfatah, O., Li, B., Fernández, J.C. and Tsekleves, E. (2013) Dynamic Hyperlinker for 3D Content Search and Retrieval. 2013 *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting* (*BMSB*), London, 5-7 June 2013, 1-4. http://dx.doi.org/10.1109/BMSB.2013.6621760

[12] Swash, M.R., Aggoun, A., Abdulfatah, O., Li, B., Fernández, J.C. and Tsekleves, E. (2013) Omnidirectional Holoscopic 3D Content Generation Using Dual Orthographic Projection. 2013 *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting* (*BMSB*), London, 5-7 June 2013, 1-4. http://dx.doi.org/10.1109/BMSB.2013.6621674

[13] Makanjuola, J., Aggoun, A., Swash, M., Grange, P., Challacombe, B. and Dasgupta, P. (2012) 3D-Holoscopic Imaging: A Novel Way to Enhance Imaging in Minimally Invasive Therapy in Urological Oncology. *Journal of Endourology*, **26**, P1-A572.

[14] Swash, M.R., Aggoun, A., Abdulfatah, O., Li, B., Jacome, J.C., Alazawi, E. and Tsekleves, E. (2013) Pre-Processing of Holoscopic 3D Image for Autostereoscopic 3D Display. 5*th International Conference on* 3*D Imaging* (*IC*3*D*), 3-5 December 2013, 1-5. http://dx.doi.org/10.1109/ic3d.2013.6732100

[15] Lawton, G. (2011) 3D Displays without Glasses: Coming to a Screen near You Computer. *Computer*, **44**, 17-19.

[16] Park, J.-H., Hong, K. and Lee, B. (2009) Recent Progresses in Three-Dimensional Information Processing Based on Integral Imaging. *Applied Optics*, **48**, H77-H94. http://dx.doi.org/10.1364/AO.48.000H77

[17] Park, J.-H., Baasantseren, G., Kim, N., Park, G., Kang, J.-M. and Lee, B. (2008) View Image Generation in Perspective and Orthographic Projection Geometry Based on Integral Imaging. *Optics Express*, **16**, 8800-8813. http://dx.doi.org/10.1364/OE.16.008800

[18] Aggoun, A., Tsekleves, E., Swash, M.R., Zarpalas, D., Dimou, A., Daras, P., Nunes, P. and Soares, L.D. (2013) Immersive 3D Holoscopic Video System. *IEEE MultiMedia*, **20**, 28-37.

[19] Manolache, S., Kung, S.Y., McCormick, M. and Aggoun, A. (2003) 3D-Object Space Reconstruction from Planar Recorded Data of 3D-Integral Images. *Journal of VLSI Signal Processing Systems*, **5**, 5-18.

[20] Cirstea, S., Aggoun, A. and McCormick, M. (2008) Depth Extraction from 3D-Integral Images Approached as an Inverse Problem. *IEEE International Symposium on Industrial Electronics*, Cambridge, 30 June-2 July 2008, 798-802.

[21] Wu, C., McCormick, M., Aggoun, A. and Kung, S.Y. (2008) Depth Mapping of Integral Images through Viewpoint Image Extraction with a Hybrid Disparity Analysis Algorithm. *Journal of Display Technology*, **4**, 101-108. http://dx.doi.org/10.1109/JDT.2007.904360

[22] Zarpalas, D., Biperis, I., Fotiadou, E., Lyka, E., Daras, P. and Strintzis, M.G. (2011) Depth Estimation in Integral Images by Anchoring Optimization Technique. *IEEE International Conference on Multimedia & Expo* (*ICME*), Barcelona, 11-15 July 2011, 1-6. http://dx.doi.org/10.1109/icme.2011.6011887

[23] Zarpalas, D., *et al.* (2012) Anchoring Graph Cuts towards Accurate Depth Estimation in Integral Images. *Journal of Display Technology*, **8**, 405-417. http://dx.doi.org/10.1109/JDT.2012.2190493

[24] Fatah, O., Aggoun, A., Nawaz, M., Cosmas, J., Tsekleves, E., Swash, M. and Alazawi, E. (2012) Depth Mapping of

Integral Images Using Hybrid Disparity Analysis Algorithm. *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, Seoul, 27-29 June 2012, 1-4. http://dx.doi.org/10.1109/BMSB.2012.6264257

[25] Alazawi, E., Aggoun, A., Fatah, O., Abbod, M. and Swash, R.M. (2013) Adaptive Depth Map Estimation from 3D Integral Image. *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, London, 5-7 June 2013, 1-6. http://dx.doi.org/10.1109/bmsb.2013.6621736

[26] Alazawi, E., Aggoun, A., Abbod, M., Swash, R.M. and Fatah, O. (2013) Scene Depth Extraction from Holoscopic Imaging Technology. *The True Vision—Capture*, *Transmission and Display of* 3*D Video* (3*DTV-CON*), Aberdeen, 7-8 October 2013, 1-4.

[27] Alazawi, E., Cosmas, J., Swash, M.R., Abbod, M. and Abdulfatah, O. (2014) 3D-Interactive-Depth Generation and Object Segmentation from Holoscopic Image. *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, Beijing, 25-27 June 2014, 1-5. http://dx.doi.org/10.1109/bmsb.2014.6873494

[28] Marr, D. and Hildreth, E. (1980) Theory of Edge Detection. *Proceedings of the Royal Society of London. Series B*: *Biological Sciences*, **207**, 187-217. http://dx.doi.org/10.1098/rspb.1980.0020

[29] Khallil, M. and Aggoun, A. (2006) Edge Detection Using Adaptive Local Histogram Analysis. *IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, **2**, 757-760. http://dx.doi.org/10.1109/icassp.2006.1660453

[30] Khalil, A., Aggoun, A. and Elmabrouk, A. (2003) Edge Detector Using Local Histogram Analysis. *Proceedings of SPIE*, **5150**, 2141-2151. http://dx.doi.org/10.1117/12.503023

[31] Aggoun, A. and Khallil, M. (2007) Multi-Resolution Local Histogram Analysis for Edge Detection. *IEEE International Conference on Image Processing*, **3**, 45-48. http://dx.doi.org/10.1109/icip.2007.4379242

[32] Vega-Rodríguez, M.A., Sánchez-Pérez, J.M. and Gomez-Pulido, J.A. (2002) An FPGA-Based Implementation For Median Filter Meeting the Real-Time Requirements of Automated Visual Inspection Systems. *Proceedings of the* 10*th Mediterranean Conference on Control and Automation*, Lisbon, 9-12 July 2012.

[33] Holoscopic 3D Dataset, Accessed Date: 27th of May 2016, Login: Guest. ftp://ftp.iti.gr/pub/holoscopy/integral_database/database/data/

[34] Bay, H., Tuytelaars, T. and Van Gool, L. (2006) SURF: Speeded Up Robust Features. In: Leonardis, A., Bischof, H. and Pinz, A., Eds., *Computer Vision—ECCV* 2006, Springer, Berlin, 404-417. http://dx.doi.org/10.1007/11744023_32

[35] Lowe, D. (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, **60**, 91-110.

[36] Panchal, P.M., Panchal, S.R. and Shah, S.K. (2013) A Comparison of SIFT and SURF. *International Journal of Innovative Research in Computer and Communication Engineering*, **1**, 323-327.

[37] Liu, C., Yuen, J. and Torralba, A. (2011) SIFT Flow: Dense Correspondence across Different Scenes and Its Applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (*TPAMI*), **33**, 978-994.