

Visual stereo matching combined with intuitive transition of pixel values

Kwangmu Shin¹ · Daekeun Kim¹ · Kidong Chung¹

Received: 10 February 2014 / Revised: 3 June 2015 / Accepted: 18 September 2015 /

Published online: 28 September 2015

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The objective of stereo matching is to find the corresponding pixels from similar two or more images. However, it is difficult problem to get precise and consistent disparity under a variety of real world situations. In other words, the color values of stereo images are easily influenced by radiometric factors such as illumination direction, illumination color, and camera exposure. Therefore, conventional stereo matching methods can have low performances under radiometric conditions. In this paper, we propose a novel stereo matching approach that is robust in controlling various radiometric variations such as local and global radiometric variations. We designed a hybrid stereo matching approach using transition of pixel values and data fitting. Transition of pixel values is utilized for the coarse stereo matching stage, and polynomial curve fitting is used for the fine stereo matching stage. Experimental results show that the proposed method has better performances compared to the stereo matching algorithms of comparison group under severely different radiometric conditions between stereo images. Consequentially, we demonstrate that the proposed method is less sensitive to various radiometric variations, and shows an outstanding performance in computational complexity.

Keywords Multimedia · Visual processing · Algorithms · Stereo matching · Pixel values

1 Introduction

Computer stereo vision technology, namely stereo matching, has been an important topic and has been constantly developed in the field of computer vision for more than three decades [17]. It is still one of the most active research areas. In practical applications, stereo matching plays an important role in many fields such as multimedia, robotics, autonomous vehicles, virtual reality, and security [21, 23]. In particular, stereo matching is highly significant in the field of

✉ Kwangmu Shin
sin@pusan.ac.kr

¹ Department of Computer Engineering, Pusan National University, Mountain 30, Jangjeondong, Geumjeonggu, Busan 46241, Republic of Korea

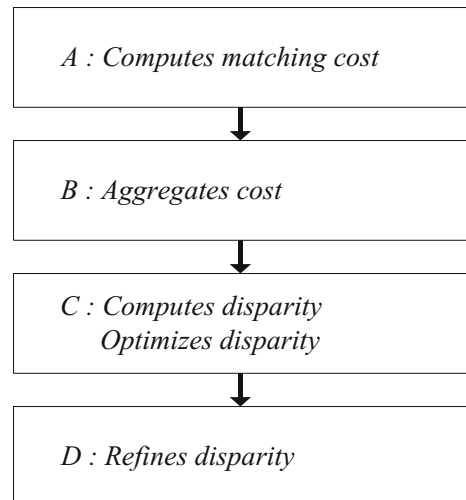
robotics, as it is vital to the extraction of information about the relative position of 3D objects in the vicinity of autonomous systems. Other application for robotics include object recognition, where depth information allows for the system to separate occluding image components, such as one chair in front of another, which the robot may otherwise not be able to distinguish as a separate object by any other criteria. Thus, the stereo vision may be fused with other multimedia-related technology [2, 16, 24].

In stereo vision, two cameras that are displaced horizontally from one another are used to obtain two differing views of a scene. This is a manner similar to human binocular vision. In other words, stereo images are two images of the same scene taken from different viewpoints. By comparing these two images, the relative depth can be obtained. The depth is in type of disparities which are inversely proportional to the differences in distance to the objects. In detail, the disparities can be acquired through locating for each pixel of an image, the corresponding pixel of the other image. A map of all pixel displacements in an image is a disparity map. This process is called stereo matching.

Figure 1 shows general procedures in the stereo matching algorithm [25]. In *step A*, a cost of every individual pixel is assigned to all possible disparities. In *step B*, an assumption is made that neighboring pixels share the same disparity. An aggregation of initial pixel-wise matching costs is carried out over a support region around each pixel. In *step C*, an optimal disparity value is selected for each pixel. Local methods usually employ a winner-takes-all strategy that the disparity with the lowest aggregated cost is chosen. Global methods optimize an energy function defined over all image pixels by concurrently imposing a smoothness constraint. In *step D*, there are goals at correcting imprecise disparity values and handling occlusion areas. Generally used approaches include scan-line optimization, median filtering, subpixel estimation, region voting, peak removal, and etc.

Substantial issues of stereo matching are to provide high accuracy and fast execution in a variety of environments. In fact, many studies have been conducted to date. Scharstein et al. [25] and Szeliski [27] provide broad reviews of the state-of-the-art methods in stereo matching algorithms. A variety of algorithms for stereo matching can be generally classified as either local or global methods. The global methods compute all disparities of images concurrently by optimizing the global energy function that includes a data term and a smoothness term. The

Fig. 1 Overview of general stereo matching algorithm



methods typically skip the cost aggregation step and supports piecewise smooth disparity selection. The methods can generally obtain precise disparity maps; however, the methods are mainly complex and computationally expensive. Thus, the methods are still usually much slower than local methods, since global optimization is a NP-hard problem while local matching runs in polynomial.

Recent studies relating to the global method are as follows. Yang [31] proposed a non-local solution to avoid being adversely affected by the local nature of traditional window-based cost aggregation algorithms. In other words, the matching cost values are aggregated adaptively based on pixel similarity on a tree structure derived from the stereo image pair to preserve depth edges. The nodes of this tree are all the image pixels, and the edges are all the edges between the nearest neighboring pixels. The similarity between any two pixels is decided by their shortest distance on the tree. Global energy minimization used for disparity optimization commonly requires large computational effort with high memory capacity. Veksler [30] reduced the search space using the local stereo matching method. A graph cut technique was used to minimize the energy function. As a result, this method could effectively reduce the memory capacity, but the computational complexity was not reduced. In addition, Kolmogorov et al. [15] utilized a graph cut for global energy optimization. Belief propagation was used for global energy optimization in [22, 26, 32].

On the other hand, the local methods calculate the disparity of the pixel, based on the support window cost aggregation [11, 12]. Consequentially, the local methods have simple design and are more efficient in aspects of computational complexity than the global methods.

In the following, we can see recent research work of the local method. Min et al. [20] reduced the search range using a subset of informative disparity hypotheses. However, this method cannot get precise results at depth discontinuities as the aggregation windows located on depth edges represent pixels from different depths. Veksler [29] proposed a cost aggregation method with size-adaptive windows in order to solve this problem. In another approach, Kang et al. [14] utilized a multiple-window strategy. This method decides the optimal aggregation window from a set of pre-defined windows of the same size that are located at different positions. Hosni et al. [8] used locally adaptive support weights to compute the probability that the center pixel and a neighbor pixel might belong to the same region. Zhang et al. [35] separately carried out horizontal and vertical passes for cost aggregation using orthogonal integral images. Hosni et al. [9] took an approach to estimate a support region of pixel via color segmentation. This method calculates the geodesic distance from all pixels to the center pixel of the window in a square support window. Pixels of low geodesic distance are given high support weights. As a result, it has a significant effect in the stereo matching. Zhang et al. [34] proposed a robust voting scheme to refine initial estimates based on a piecewise smoothness prior, improving the quality in occluded regions and low-textured regions effectively. The refinement is guided by the segmentation result of input images. Unreliable initial estimates are detected and rejected using an efficient left-right consistency check.

A typical stereo matching estimates disparities between corresponding pixels in stereo images. In the process, there is an important assumption that the corresponding pixels should have similar color values [25]. In other words, it is assumed that the object surface is a Lambertian surface [1, 3, 36]. In the Lambertian, the color of each 3D point acquired from different cameras will be constant. The reason is that Lambertian surface reflects the incident light in all directions with the same strength making the camera's viewpoint invariant. However, in most real-world situations, the objects are not Lambertian and reflect light with

view dependency. In detail, if two images are captured under changing radiometric effects such as illumination and camera exposure, they may have different color values. Therefore, a typical stereo matching isn't commonly able to provide a precise disparity map [4].

Recent studies relating to the effects of radiometric variations are as follows. Miled et al. [19] developed a spatially varying multiplicative model to account for brightness changes induced between left and right views. The depth estimation problem is then formulated as a constrained optimization problem in which an appropriate convex objective function is minimized under various convex constraints modelling prior knowledge and observed information. Weijer et al. [28] carries out the grey-edge algorithm that employs the average color of edge differences for the color normalization. Jung et al. [13] performs the adaptive color transformation that finds pseudo-corresponding pixels based on the rank matching. Then, this method transforms the color of each pixel to be consistent with that of the corresponding pixel. Hirschmuller et al. [7] employed color invariant matching costs. The normalized cross correlation compensates for the gain and the bias in color values between stereo images.

In this paper, we present a novel stereo matching approach that is robust in controlling various radiometric variations such as local and global radiometric variations. The local and global radiometric variations indicate the effect of illumination and camera exposure. Furthermore, we considered the computational complexity of the proposed method. Therefore, the proposed method is performed based on local stereo matching. The key contribution of this work is that the proposed method presents a new approach in stereo matching. That is, we designed a hybrid stereo matching approach using transition of pixel values and data fitting. Transition of pixel values is utilized in the coarse stereo matching stage, and data fitting is used in the fine stereo matching stage.

The remainder of this paper is structured as follows. Section 2 describes stereo matching algorithms of comparison group. Section 3 presents a hybrid stereo matching using transition of pixel values and data fitting. Section 4 presents experimental results. Finally, Section 5 concludes this paper.

2 Stereo matching algorithms of comparison group

In this section, we describe the adaptive support-weight and the adaptive normalized cross-correlation method in more detail. The two methods are used as a comparison group in our experiments. As mentioned in Section 1, there are many kinds of stereo matching algorithms. However, we chose the two methods as a comparison group for the following reasons. In our experiments, we categorized the comparison group into two types depending on characteristics of stereo matching algorithms. The characteristics of stereo matching algorithms are based on consideration of radiometric effects such as illumination and camera exposure. First, the adaptive support-weight method was selected as a comparison group that does not take into account radiometric effects. The reason is that this method demonstrates a good performance under common environments without radiometric effects, and has been reasonably verified experimentally [33]. Next, the adaptive normalized cross-correlation method was chosen as a comparison group that takes into account radiometric effects. This method demonstrates an outstanding performance from the stereo images taken under different radiometric effects. That is, it is significantly robust and accurate with radiometric effects [5, 6].

2.1 Adaptive support-weight approach

The adaptive support-weight approach (ASW) [33] exploits the support-weights of the pixels in a given support window using color similarity and geometric proximity. The ASW method is composed of three parts with adaptive support-weight computation, dissimilarity computation based on the support-weights, and disparity selection.

$$w(p, q) = \exp\left(-\left(\frac{\Delta c_{pq}}{\gamma_c} + \frac{\Delta g_{pq}}{\gamma_p}\right)\right) \quad (1)$$

Equation 1 describes the support-weight of the ASW method. This process is the most important part in the ASW method and entirely based on the contextual information within a given support window. Δc_{pq} and Δg_{pq} represent the color difference and the spatial distance between pixel p and q , respectively. γ_c is related with the color similarity. Finally, γ_p is related with window size.

Second, the dissimilarity between pixels is measured by aggregating raw matching costs with the support weights in both support windows. This process takes into account the support-weights in both reference and target support windows. After the dissimilarity computation, the disparity of each pixel is simply selected by the winner-takes-all method without any global reasoning. The winner-takes-all method simply picks the lowest matching cost.

To summarize, the ASW method assigns an adaptive weight to each pixel in the support window, according to how a support pixel is likely to lie on the same disparity with the center pixel. The more likely a support pixel is to lie on the same disparity with the center pixel, the higher the weight is changed. Basically, the assignment of an adaptive weight for each support pixel amounts to changing the support window in terms of size, shape, and center offset. Therefore, the weight computation is significant, since it directly decides the support window. The adaptive support weight of a pixel is defined as inversely proportional to the color dissimilarity and, the spatial distance between this current pixel and the center pixel. Consequently, the ASW method shows an outstanding performance under common environments without radiometric effects. However, the ASW method also has a problem that does not solve an uncertainty effectively which is caused by nearby pixels at different disparities but with similar colors. As previously mentioned, the weight function utilizes a color similarity term and a spatial proximity term which includes important assumptions with implication. If the color of support pixels have high similarity with the color of center pixel, the support pixels are likely to have the same disparity with the center pixel. Also, if the color of support pixels are spatially close to the center pixel, the support pixels are likely to have the same disparity with the center pixel. But the ASW method overlooks that the assumptions can be ignored in various test images. As a result, this point serves to increase the ambiguity of disparity within a support window.

2.2 Adaptive normalized cross-correlation approach

The normalized cross-correlation method (NCC) [18] is a well-known similarity measure between two pixels with neighbors. Applying this NCC method directly to stereo matching of general image pairs would result in two important problems. There could be a complicated nonlinear relationship between two corresponding pixels between stereo images, thereby causing the NCC method to stop working. Therefore, applying the NCC method to raw stereo images in a simple way does not work well because the diverse radiometric variations are not

taken into consideration. The second problem is that the supporting windows in the left and right images do not appear accurately because of the view changes. Consequentially, the NCC method usually produces a fattening effect near the object boundaries similar to conventional window-correlation-based matching measures.

Thus, the adaptive normalized cross-correlation approach (ANCC) [5, 6] tried to find a solution to these problems. The nonlinear relationship that exists between corresponding pixel color values because of various unknown radiometric variations is transformed into a linear one by employing log-chromaticity color space. Next, in order to reduce the fattening effect and increase the accuracy between matching windows, it defines a modified NCC method measure that utilizes the adaptive weighting scheme.

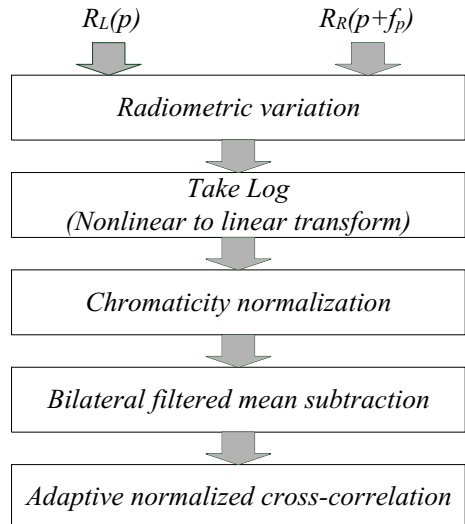
Figure 2 depicts the overview of the ANCC method. The principle of this method is that the color formation model is modeled and incorporated into a new stereo correlation measure. This method considered the color formation process in an explicit manner instead of using the raw color value for handling the diverse radiometric variations that occur between stereo images. This method provides a new data cost that is insensitive to the diverse radiometric variations. Also, it reduces the problems faced with window-based stereo methods.

And this method subtracted the bilateral filtered value instead of the simple window mean value for coherent normalization around window pixels. Taking into consideration of these bilateral filtered weights again, this method defined a new correlation measure as Eq. 2.

$$ANCC_{\log Chrom_R}(f_p) = \frac{\sum_{i=1}^M w_L(t_i)w_R(t_i)[R_L'''(t_i)] \times [R_R'''(t_i)]}{\sqrt{\sum_{i=1}^M |w_L(t_i)R_L'''(t_i)|^2} \times \sqrt{\sum_{i=1}^M |w_R(t_i)R_R'''(t_i)|^2}} \quad (2)$$

Equation 2 describes adaptive normalized cross-correlation for the *logChrom_R* channel. Adaptive normalized cross-correlation for the *logChrom_G* and *logChrom_B* channels can be computed in a similar manner. *M* represents *m* x *m* window. *t* represents each pixel. *w_L* represents

Fig. 2 Overview of the adaptive normalized cross-correlation approach (RL(p) is the value of the RED channel in the left image at pixel p and RR(p + fp) is the value of the RED channel in the right image at pixel p + fp)



the corresponding weight vector of each pixel in the left image. w_R is the weight vector in the right image. R_L^m represents the value after performing linear transformation, chromaticity normalization, and bilateral filtered mean subtraction in the left image. R_R^m is the value in the right image.

Consequently, the ANCC method doesn't vary with illumination change and camera gamma correction. And the fattening effect can be reduced because this method incorporates the spatial weight information adaptively. Although the ANCC method is robust and accurate from the stereo images taken under different radiometric effects, the performance and complexity of this method is highly dependent on window size. There is a problem that this method cannot control a local variation of brightness. Lastly, the ANCC method has another important problem. As described earlier, this method uses the color formation model explicitly in order to control various radiometric effects. The operation of the color formation model can be referred to as pre-processing. As a result, this pre-processing scheme may have an adverse effect on the overall complexity of the stereo matching.

3 Hybrid stereo matching using transition of pixel values and data fitting

3.1 Coarse stereo matching stage - computation of matching cost utilizing pixel values transition

In most images, pixels adjacent to certain pixels have similar color information. This characteristic is retained even if there are radiometric effects such as illumination variations and camera exposure variations. Thus, we can confirm that the difference between value of the specific pixel and the neighboring pixels is similar. A condition is that it should be the corresponding region in stereo image pair. Coarse stereo matching method has been triggered from a simple and intuitive idea, as seen above.

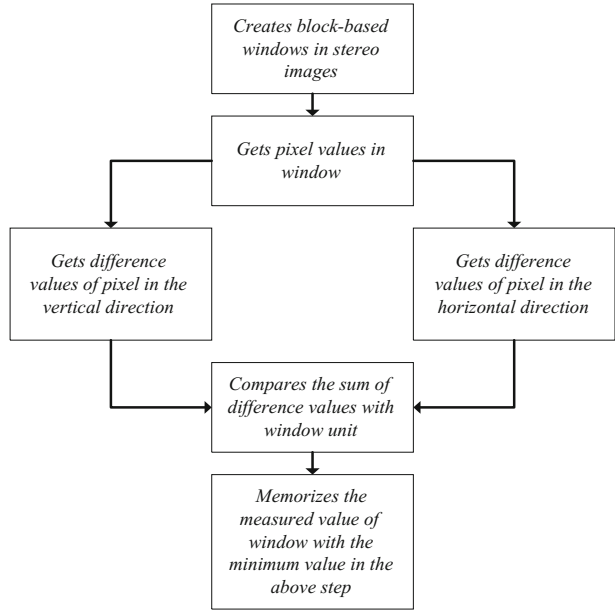
Figure 3 shows the flowchart of the coarse stereo matching method utilizing pixel values transition. This method set block-based windows in stereo images and obtains pixel values in both windows. And then, this method compares the window of the right image with the window of the left image while the window of the right image is moving. We obtain difference values of pixel in the vertical direction and in the horizontal direction, respectively. Next, this method compares the sum of difference values with window units. Lastly, this method memorizes the measured value of the window with the minimum value in the above step.

Figure 4 depicts algorithm of coarse stereo matching method using pixel values transition. In Fig. 4, the difference value between $A-B$ and $A'-B'$ is added to the difference value between $C-D$ and $C'-D'$. That is, these calculation processes are performed in the horizontal direction. Also, the difference value between $A-C$ and $A'-C'$ is added to the difference value between $B-D$ and $B'-D'$. These calculation processes are carried out in the vertical direction.

$$\sum_{(i,j) \in W} \left(\left| I_1(i,j) - I_1(i,j+1) \right| - \left| I_2(i,j) - I_2(i,j+1) \right| \right) + \sum_{(i,j) \in W} \left(\left| I_1(i,j) - I_1(i+1,j) \right| - \left| I_2(i,j) - I_2(i+1,j) \right| \right) \quad (3)$$

Equation 3 illustrates in more details the algorithm of the coarse matching method of Fig. 4. Equation 3 is used in the calculation of each block unit. Information of each block can be obtained from the stereo images. Also, the information includes a component of the vertical

Fig. 3 Flowchart of coarse stereo matching stage

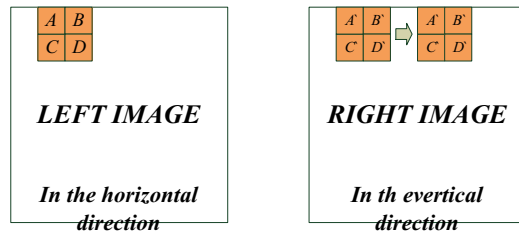


and horizontal directions. In this equation, i and j indicate the coordinate of a pixel, W indicates the target window and I_1 and I_2 indicate the left image and the right image, respectively.

3.2 Fine stereo matching stage - computation of matching cost utilizing data fitting

Figure 5a and b show the Baby1 stereo images which have undergone different exposure conditions. The left image was acquired with a level of illumination of 2 and exposure of 2. And the right image was acquired with a level of illumination of 2 and exposure of 1. Figure 5c is the disparity map of coarse stereo matching method. In the same manner, Fig. 6 shows the result of coarse stereo matching method carried out for the Bowling2.

In Figs. 5c and 6c, we can see that coarse stereo matching method extracts invalid disparity in the boundary region. Part of the boundary region is marked with a rectangular shape. In other words, the invalid disparity exceeds the maximum disparity of the image or has value of 0. The reason is that coarse stereo matching method uses only the information of



$$\min[(|A - B| - |A' - B'|) + (|C - D| - |C' - D'|) + (|A - C| - |A' - C'|) + (|B - D| - |B' - D'|)]$$

Fig. 4 Algorithm of the coarse stereo matching stage

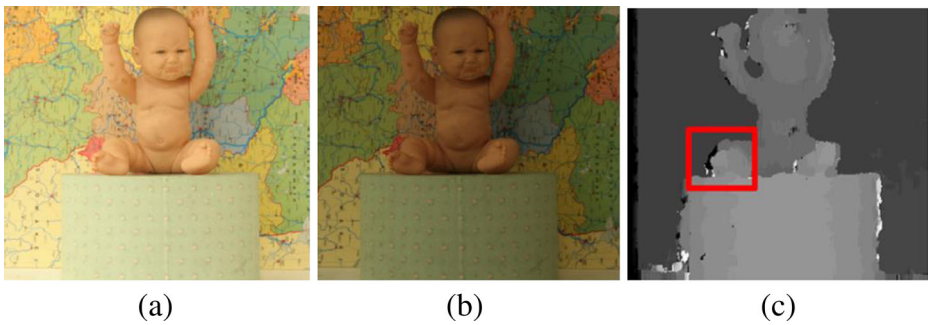


Fig. 5 Result of the coarse stereo matching stage on Baby1 image pair with varying exposure. **a** The left image with illumination (2)-exposure (2). **b** The right image with illumination (2)-exposure (1). **c** The coarse stereo matching method

the neighboring pixels. As a result, this method has the limitation that does not take into account the overall changes.

We have combined the fine stereo matching method with the coarse stereo matching method in order to compensate the limitation. In the fine stereo matching method, we utilize polynomial curve fitting of one kind of data fitting. Before polynomial curve fitting is applied, we have tried various algorithms to find the appropriate method. The methods include Sobel, Prewitt, Roberts, and the Canny method that can approximate the gradient magnitude of the image. However, we could not obtain improved performance as compared to the coarse stereo matching method. Consequentially, polynomial curve fitting was selected from experimental verification.

The objective of curve fitting is to find the parameters of a mathematical model that describes a set of data in a way that minimizes the difference between the model and the data. The most common approach is the polynomial least squares method, a well-known mathematical procedure for finding the coefficients of polynomial equations that are a best fit to a set of X, Y data. A polynomial equation expresses the dependent variable Y as a polynomial in the independent variable X . Those coefficients can be used to predict values of Y for each X . The best fit simply means that the differences between the actual measured Y values and the Y values predicted by that equation are minimized. We also use the polynomial least squares method with the best fit model.

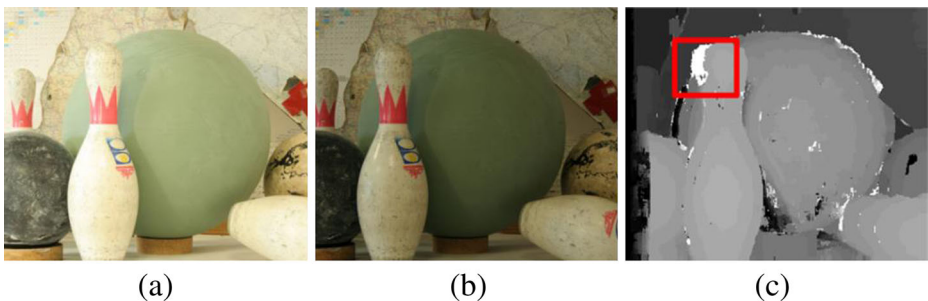


Fig. 6 Result of the coarse stereo matching stage on Bowling2 image pair with varying exposure. **a** The left image with illumination (2)-exposure (2). **b** The right image with illumination (2)-exposure (1). **c** The coarse stereo matching method

Figure 7 illustrates a flowchart of the fine stereo matching method utilizing data fitting. This method utilizes the average of the pixel values within the target window in the vertical direction. If the window size is 7×7 , we obtain 7 kinds of the average value in the vertical direction. The reason for selecting the vertical direction is as follows. Typical images that contain objects have mainly components in the vertical direction, because objects are generally standing vertically. In Section 4.1, we can see it through test bed images. Therefore, we can confirm that typical images containing objects mainly include changes in the vertical direction. The average values are used as the data of y coordinate, and the data of x coordinate is a natural number that is monotonically increasing. The polynomial curve fitting is performed based on this data. As a result, we can find the nearest polynomial to the data.

In detail, if the polynomial curve fitting is performed in the target window of the left image, a particular polynomial expression is created. We store each of the coefficients according to the order. And then, the polynomial curve fitting is performed as well as the left image in the right image. Window position of the right image is based on the window position of the left image, and the fitting is carried out continuously within a certain search range. We compare polynomial expressions obtained in the window of both images, and calculate the difference in the coefficients having a highest order. Lastly, we decide the corresponding regions that have the smallest difference in the coefficients of both windows. Figure 8 depicts the algorithm of fine stereo matching method using data fitting.

Figure 9 shows the overview of the hybrid stereo matching using transition of pixel values and data fitting. That is, this figure depicts the final algorithm of the proposed method.

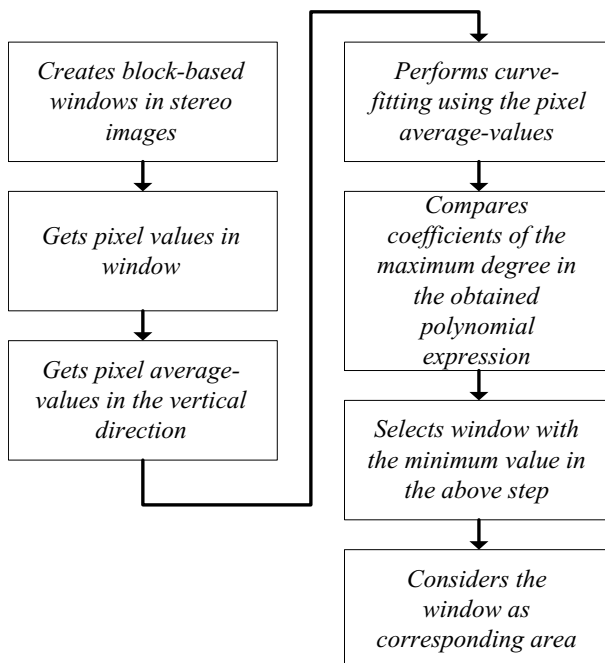
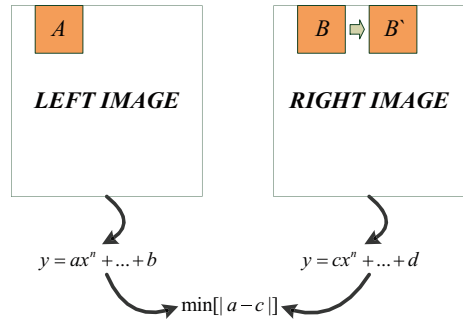


Fig. 7 Flowchart of fine stereo matching stage

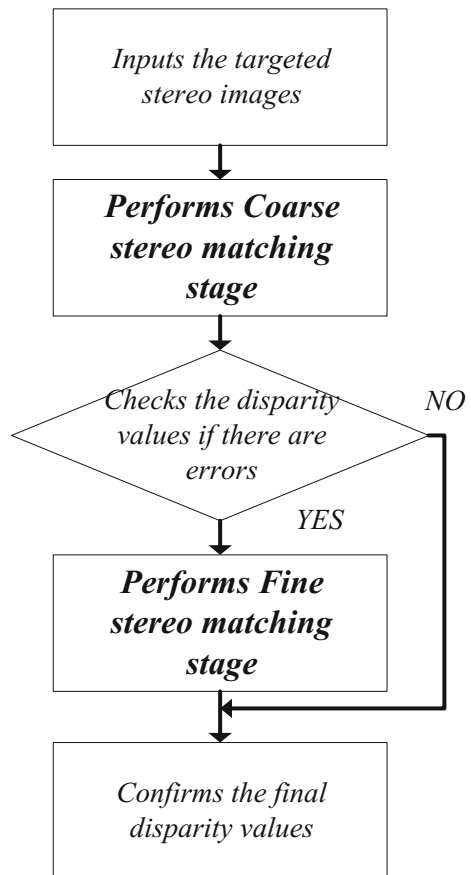
Fig. 8 Algorithm of the fine stereo matching stage



4 Experimental results

In this section, we will evaluate the performance of the stereo matching algorithms on the Middlebury stereo datasets [10]. For the comparison of the proposed method with others, we used a variety of images such as the test bed images (Baby1, Baby2, Bowling1, Bowling2, Flowerpots). There are three different illuminations (indexed as 1, 2, 3) and three different exposures (indexed as 0, 1, 2) in each data set. Each dataset provide images in three

Fig. 9 Overview of the proposed method



resolutions: full-size (width: 1240..1396, height: 1110), half-size (width: 620..698, height: 555) and one-third-size (width: 413..465, height: 370). We used one-third-sized images. And the datasets contain multiple views (7 views in total). We selected two of these views (view 0 as the left view and view 2 as the right view). In our experiments, support window of size 7×7 is selected for the study.

We compared results of the proposed method with those of the conventional stereo matching methods: the adaptive support-weight method (ASW) [33], the adaptive normalized cross-correlation method (ANCC) [5, 6] and the ground truth (measurement of image). The comparison group was classified into two kinds depending on features of the stereo matching algorithms. The features of the stereo matching algorithms are based on consideration of radiometric effects such as illumination and camera exposure. The ASW method was selected as a comparison group that does not take into consideration radiometric effects. On the other hand, the ANCC method was chosen as a comparison group that takes into consideration radiometric effects. The ANCC method is an essential comparison group in terms of considering the radiometric effects. Thus, this method can be compared with the proposed method directly. In our experiments, this method was widely used as a comparison group more often than the ASW method. Additionally, there are more grounds for the reasons why the two methods were selected as the comparison group. The grounds were described in Section 2. We performed experiments under various conditions in order to improve the reliability of the test.

The experiment consists of camera exposure variations (both subjective evaluations and objective evaluations), illumination variations (both subjective evaluations and objective evaluations), and execution time.

4.1 Camera exposure variations - subjective evaluations

In this section, we evaluate the effects of the camera exposure variations with subjective evaluations. The subjective evaluations indicate that the disparity maps of the test stereo matching algorithms are compared with the ground truth disparity map. The disparity maps are gray scale images whose intensities represent the depth information. The darker the pixel is, the further the object is from the viewer. In order to test the effects of camera exposure variations, we fixed the level of illumination to 2, and changed only the level of exposure from 2 to 1. In other words, this experiment simulates a global variation of brightness.

Figure 10a and b depict the Baby1 stereo images which have undergone different exposure conditions. The left image was acquired with a level of illumination of 2 and exposure of 2. And the right image was acquired with a level of illumination of 2 and exposure of 1. Figure 10c is the ground truth disparity map. Figure 10d is the disparity map of the ASW method for the stereo image pair in Fig. 10a and b. Figure 10e is the disparity map of the ANCC method. Figure 10f is the disparity map of the proposed method. Likewise, Fig. 10e and f also use the stereo image pair in Fig. 10a and b. In the same style, Figs. 11, 12, 13 and 14 depict the results of the test stereo matching algorithms carried out for the Baby2, Bowling1, Bowling2 and Flowerpots, respectively.

As in the case of Fig. 10, in this global variation of brightness, the ASW method yields the worst performance. That is, the ASW method provides a low quality disparity map since this method takes serious impact with different exposure conditions. The extreme exposure variations make images either very dark or bright. These effects make indistinct image features such as edges. And the proposed method yields better performance than the ANCC method in terms of quality of the disparity map. The ANCC method is fairly stable and precise under

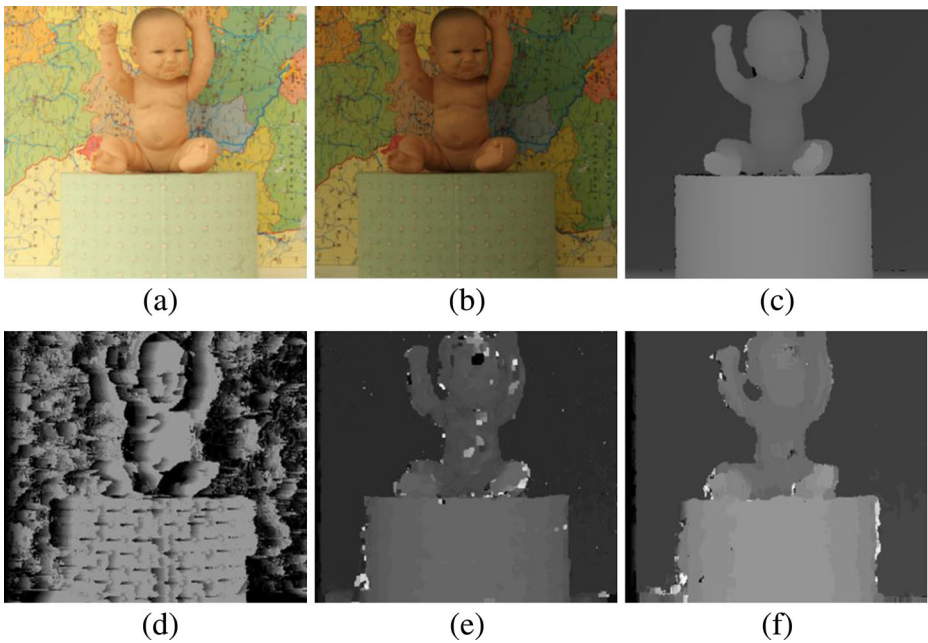


Fig. 10 Results of the test stereo matching algorithms on Baby1 image pair with varying exposure. **a** The left image with illumination (2)-exposure (2). **b** The right image with illumination (2)-exposure (1). **c** Ground truth. **d** ASW method. **e** ANCC method. **f** The proposed method

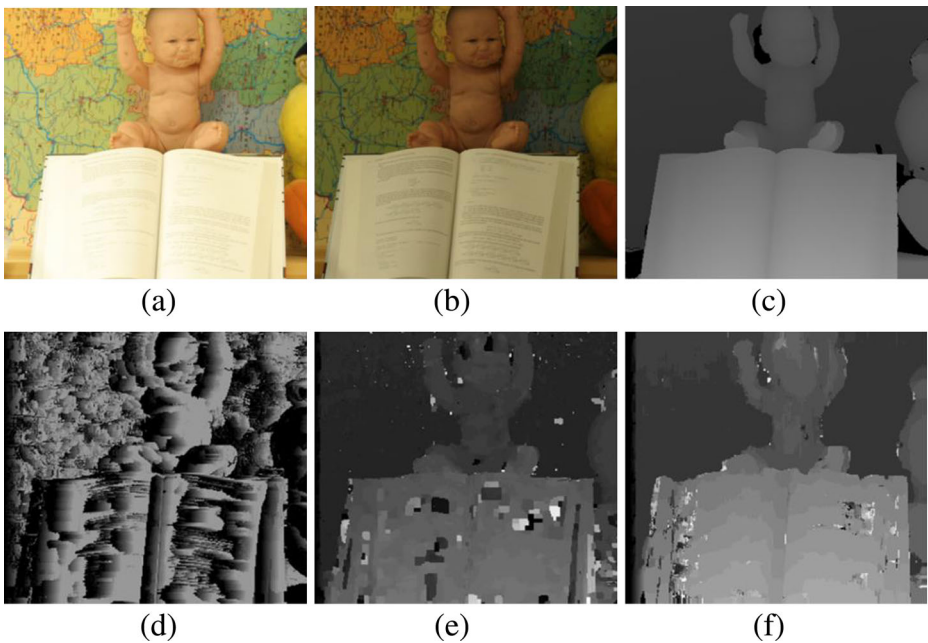


Fig. 11 Results of test stereo matching algorithms on Baby2 image pair with varying exposure. **a** The left image with illumination (2)-exposure (2). **b** The right image with illumination (2)-exposure (1). **c** Ground truth. **d** ASW method. **e** ANCC method. **f** The proposed method

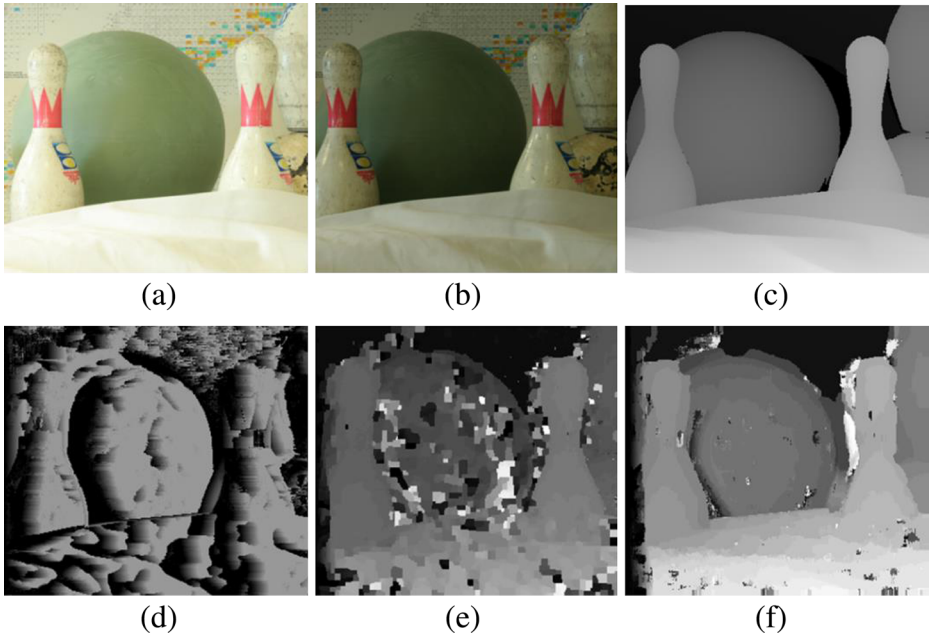


Fig. 12 Results of test stereo matching algorithms on Bowling1 image pair with varying exposure. **a** The left image with illumination (2)-exposure (2). **b** The right image with illumination (2)-exposure (1). **c** Ground truth. **d** ASW method. **e** ANCC method. **f** The proposed method

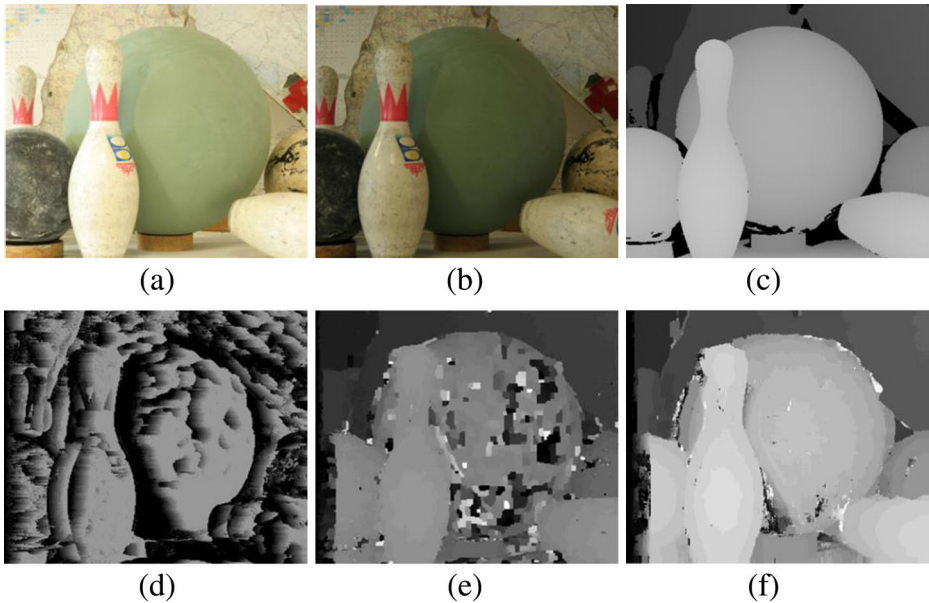


Fig. 13 Results of test stereo matching algorithms on Bowling2 image pair with varying exposure. **a** The left image with illumination (2)-exposure (2). **b** The right image with illumination (2)-exposure (1). **c** Ground truth. **d** ASW method. **e** ANCC method. **f** The proposed method

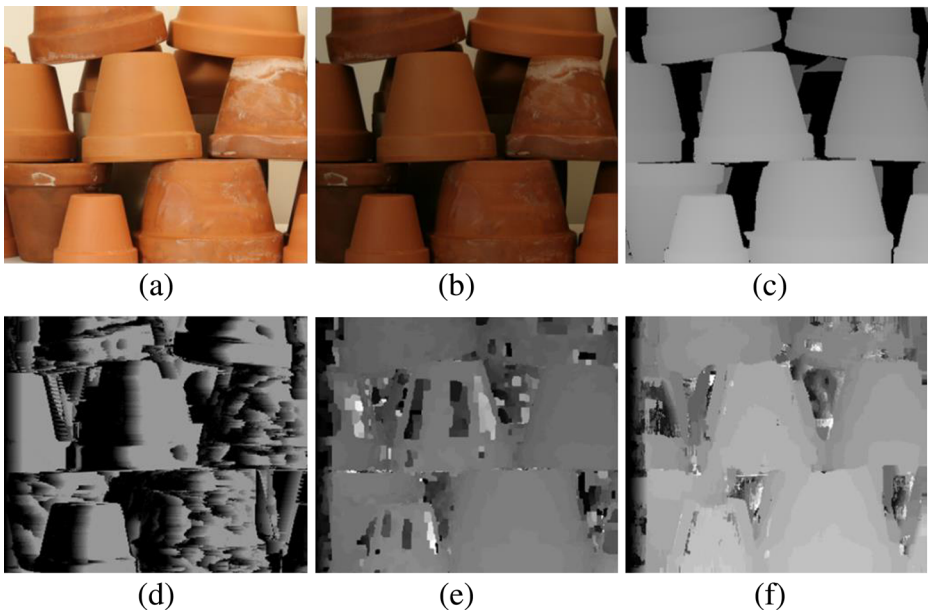


Fig. 14 Results of test stereo matching algorithms on Flowerpots image pair with varying exposure. **a** The left image with illumination (2)-exposure (2). **b** The right image with illumination (2)-exposure (1). **c** Ground truth. **d** ASW method. **e** ANCC method. **f** The proposed method

different radiometric effects. However, the ANCC method with log-chromaticity color is somewhat unstable for the near-saturated color region, which has about (255,255,255) or (0,0,0) RGB color values [5, 6]. For this reason, the ANCC method partially has the performance degradation. In Figs. 11, 12, 13 and 14, we can see the results in consistency. That is, Figs. 11, 12, 13 and 14 show results similar to Fig. 10. As a result, the proposed method shows outstanding performance compared to other methods in different exposure conditions. Furthermore, the proposed method represents fairly accurate disparity maps as compared to the ground truth.

4.2 Camera exposure variations - objective evaluations

In Section 4.2, we evaluate the effects of camera exposure variations with objective evaluations. The conditions of illumination and exposure are the same as Section 4.1. Also, we use the same test bed images as Section 4.1. The objective evaluations utilize a peak signal-to-noise ratio (PSNR) for comparison between the disparity maps of the test stereo matching algorithm and the ground truth disparity map. PSNR is an expression for the ratio between the maximum possible value of a signal and the power of distorting noise that affects the quality of its representation. Commonly, PSNR can be used to establish quantitative measures to compare the effects of image enhancement algorithms on image quality. In results of PSNR, the higher the value of PSNR the smaller the difference between the disparity map of the test stereo matching algorithm and the ground truth disparity map. We can determine whether the test algorithm has a more accurate disparity map. The ASW method has been excluded from the experiment of this section. The reason is as follows. We have sufficiently verified in the experiments of Section 4.1 that this method produces the worst performance.

Figure 15 depicts the comparison of PSNR values between the ANCC method and the proposed method under different exposure conditions. The average PSNR value of the ANCC

method is 15.77 and the proposed method is 18.05. That is, the proposed method is about 13 % higher than the ANCC method. Consequentially, we can see that the proposed method has higher values than the ANCC method in all subjects. The proposed method shows a better performance compared to the ANCC method.

4.3 Illumination variations - subjective evaluations

In this section, we evaluate the effects of illumination variations with subjective evaluations. In order to test the effects of illumination variations, we fixed the level of exposure to 1, and changed only the level of illumination from 3 to 1. In other words, this experiment simulates a local variation of brightness.

The ASW method has been excluded from the experiment to evaluate the effect of the illumination variations. The reason is as follows. The ASW method is targeted at general environments without considering radiometric conditions. We have previously verified in the experiments of Section 4.1 that this method yields the worst performance. Moreover, illumination variations can result in various local radiometric effects, which are one of the most difficult factors among the radiometric variations.

Figure 16a and b show the Baby1 stereo images which have undergone different illumination conditions. The left image was acquired with a level of illumination of 3 and exposure of 1. And the right image was acquired with a level of illumination of 1 and exposure of 1. Figure 16c is the ground truth disparity map. Figure 16d is the disparity map of the ANCC method for the stereo image pair in Fig. 16a and b. Figure 16e is the disparity map of the proposed method. Likewise, Fig. 16e uses the stereo image pair in Fig. 16a and b. In the same style, Figs. 17, 18, 19 and 20 show the results of the test stereo matching algorithms carried out for the Baby2, Bowling1, Bowling2 and Flowerpots, respectively.

As in the case of Fig. 16, in this local variation of brightness, the proposed method yields better performance than the ANCC method in terms of quality of the disparity map. The ANCC method has certain limitations that the method cannot handle the multiple illumination conditions as well as non-Lambertian reflectance objects [5, 6]. Therefore, the ANCC method partially shows the performance degradation in illumination variations. As described in the earlier text, illumination changes can cause a variety of local radiometric changes. It is a very

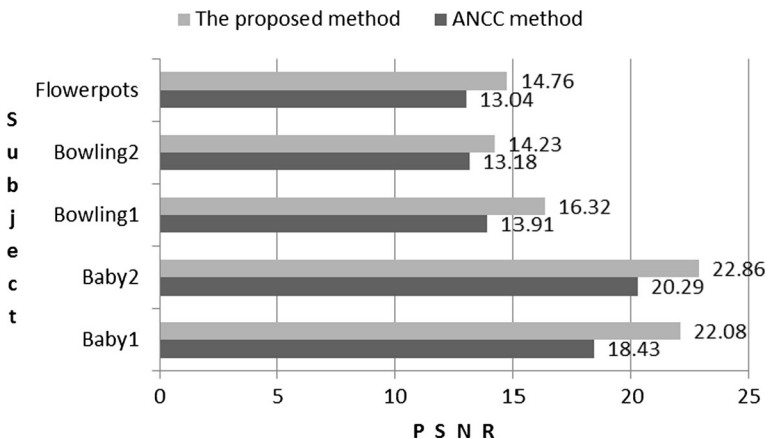


Fig. 15 Comparison of PSNR values with different exposure conditions

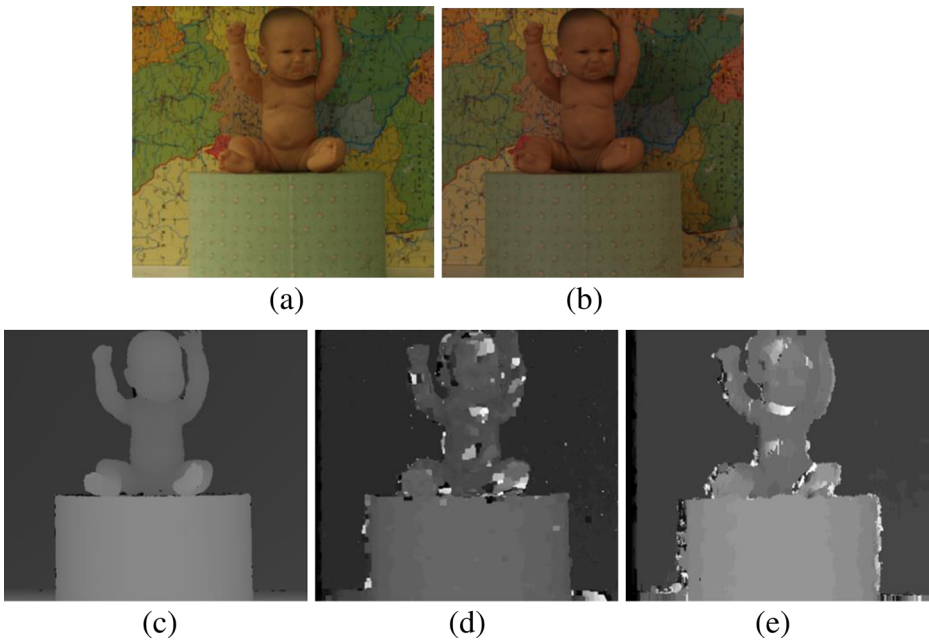


Fig. 16 Results of test stereo matching algorithms on Baby1 image pair with varying illumination. **a** The left image with illumination (3)-exposure (1). **b** The right image with illumination (1)-exposure (1). **c** Ground truth. **d** ANCC method. **e** The proposed method

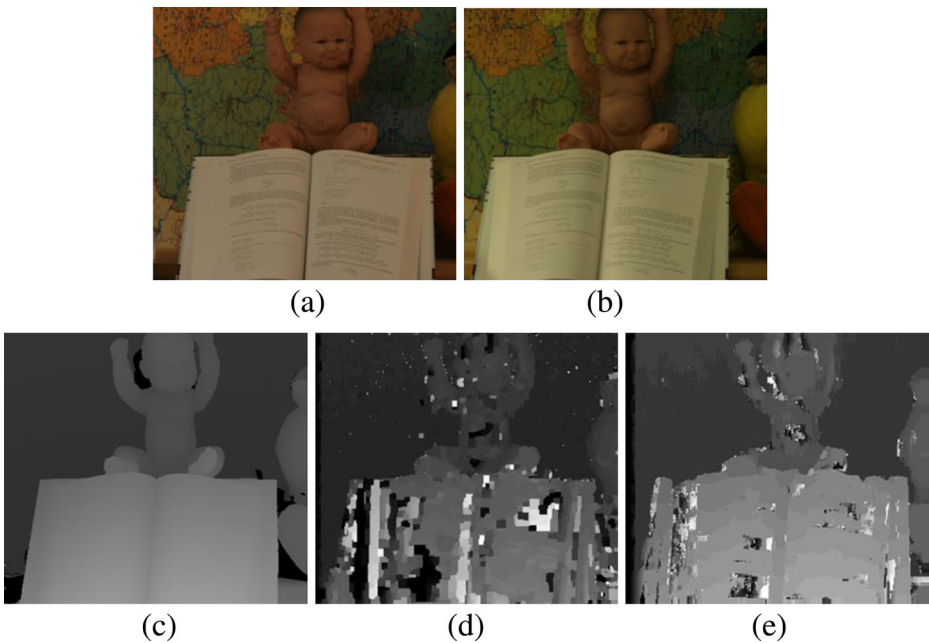


Fig. 17 Results of test stereo matching algorithms on Baby2 image pair with varying illumination. **a** The left image with illumination (3)-exposure (1). **b** The right image with illumination (1)-exposure (1). **c** Ground truth. **d** ANCC method. **e** The proposed method

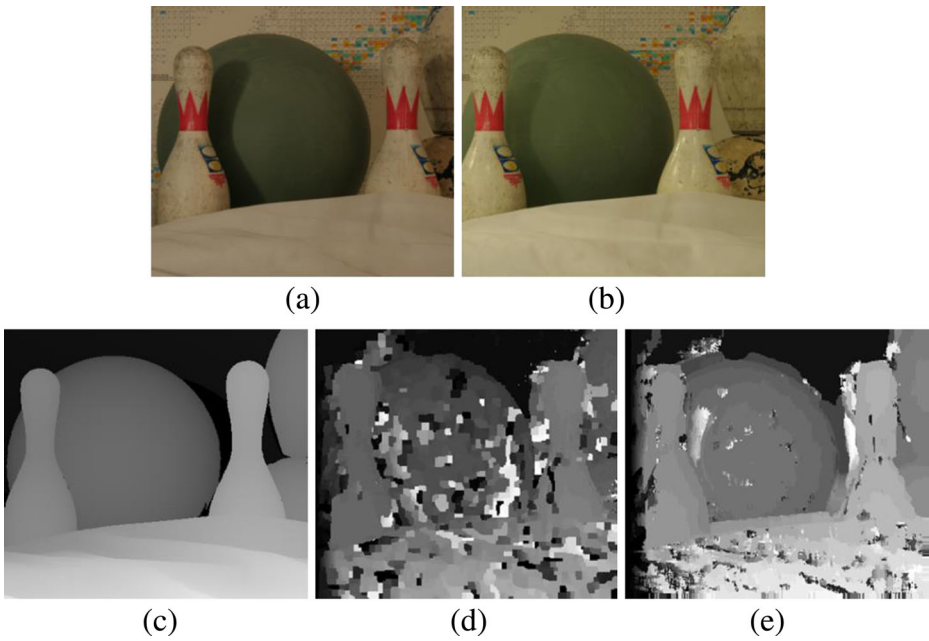


Fig. 18 Results of test stereo matching algorithms on Bowling1 image pair with varying illumination. **a** The left image with illumination (3)-exposure (1). **b** The right image with illumination (1)-exposure (1). **c** Ground truth. **d** ANCC method. **e** The proposed method

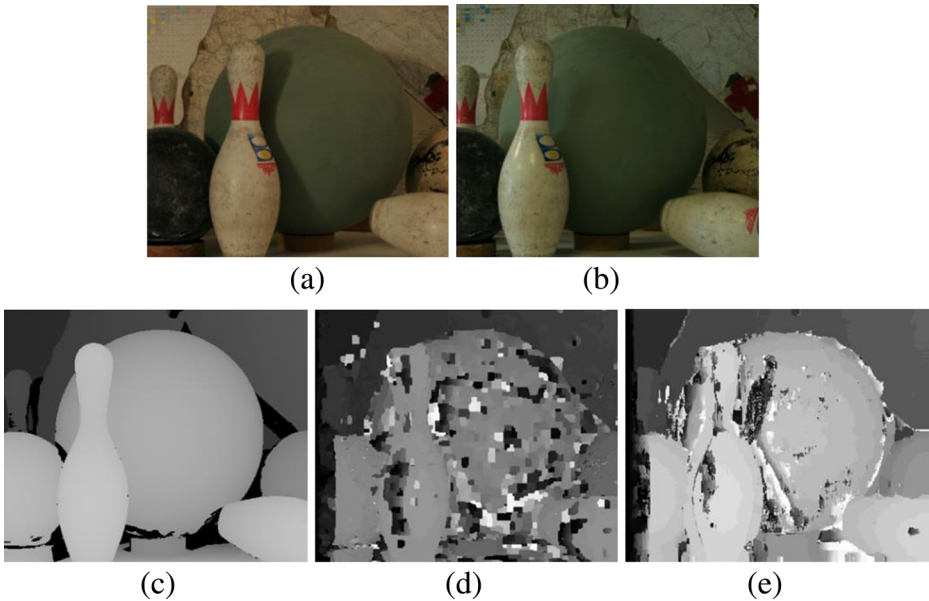


Fig. 19 Results of test stereo matching algorithms on Bowling2 image pair with varying illumination. **a** The left image with illumination (3)-exposure (1). **b** The right image with illumination (1)-exposure (1). **c** Ground truth. **d** ANCC method. **e** The proposed method

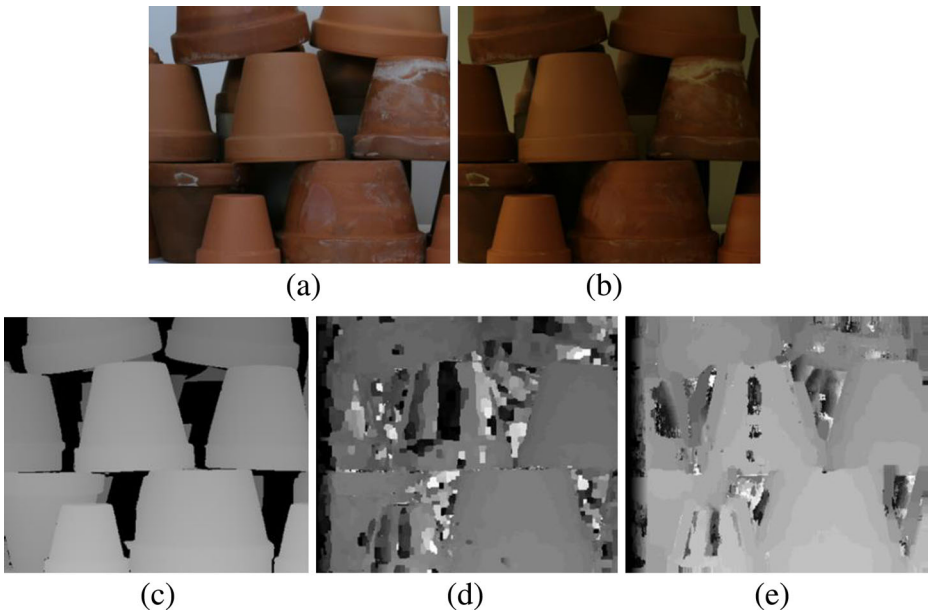


Fig. 20 Results of test stereo matching algorithms on Flowerpots image pair with varying illumination. **a** The left image with illumination (3)-exposure (1). **b** The right image with illumination (1)-exposure (1). **c** Ground truth. **d** ANCC method. **e** The proposed method

difficult factor among the radiometric variations for the stereo matching problem. Thus, both methods show relatively low performance under different exposure conditions. In Figs. 17, 18, 19 and 20, we can see the results in consistency. That is, Figs. 17, 18, 19 and 20 show results similar to Fig. 16. As a result, the proposed method outperforms the ANCC method in different illumination conditions. Furthermore, the proposed method shows significantly precise disparity maps as compared to the ground truth.

4.4 Illumination variations - objective evaluations

We evaluate the effects of camera illumination variations with objective evaluations. The conditions of illumination and exposure are the same as Section 4.3. Also, we use the same test bed images as Section 4.3. The objective evaluations utilize PSNR for comparison between the disparity map of the test stereo matching algorithm and the ground truth disparity map.

Figure 21 depicts the comparison of the PSNR values between the ANCC method and the proposed method under different illumination conditions. The average PSNR value of the ANCC method is 15.3 and the proposed method is 17.28. That is, the proposed method is about 12 % higher than the ANCC method. Consequentially, we can see that the proposed method has higher values than the ANCC method in all subjects. The proposed method shows a better performance compared to the ANCC method.

4.5 Execution time

Section 4.5 evaluates the execution time with different exposure conditions. The conditions of illumination and exposure are the same as Section 4.1. Also, we use the same test bed images

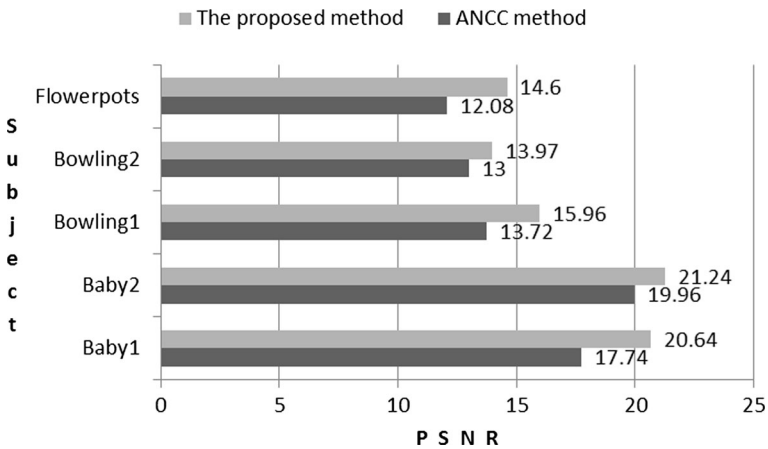


Fig. 21 Comparison of PSNR values with different illumination conditions

as Section 4.1. The experimental system environments are as follows. The CPU is an Intel core2-duo 2.4GHz, capacity of RAM is 3GB, operating system is Windows 7 and C++ is used as the programming language.

Figure 22 compares the execution time between the ANCC method and the proposed method under different exposure conditions. The average execution time of the ANCC method is 126.4 (sec) and the proposed method is 67.6 (sec). That is, the ANCC method is about 1.9 times slower than the proposed method. As a result, we can see that the proposed method is faster than the ANCC method in all subjects. The proposed method indicates a better performance compared to the ANCC method in execution time. The difference of execution time can be explained as follows. The ANCC method uses the computationally heavy graph-cut optimization. On the contrary, the proposed method employs the simple winner-takes-all approach. Moreover, the ANCC method utilizes the color formation model as pre-processing. These reasons may have a negative impact on the overall complexity of the stereo matching.

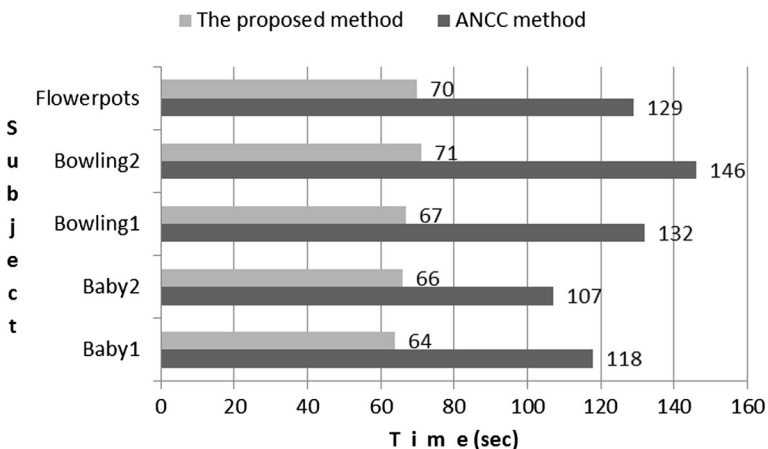


Fig. 22 Comparison of execution time with different exposure conditions

5 Conclusions

In this paper, we proposed a novel stereo matching approach that is robust in controlling various kinds of radiometric conditions such as local and global radiometric variations. We presented a hybrid stereo matching approach that uses the coarse and the fine stereo matching method. Transition of pixel values is utilized for the coarse stereo matching method, and polynomial curve fitting is used for the fine stereo matching method. Experimental results show that the proposed method has a better performance compared to the stereo matching algorithms of comparison group under severely different radiometric conditions between stereo images. As a result, we verified that the proposed method is less sensitive to various radiometric variations. Furthermore, this method shows an outstanding performance in execution time. In the future work, the proposed method can be improved utilizing adaptive window methods in accuracy of disparity map. Application of adaptive search range can contribute to a further computational complexity reduction. Moreover, we think that the proposed method can be more strictly validated through the extension of test bed images.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Brown MZ, Burschka D, Hager GD (2003) Advances in computational stereo. *IEEE Trans Pattern Anal Mach Intell* 25(8):993–1008
2. Chaudhary A, Vatwani K, Agrawal T, Raheja JL (2012) A vision-based method to find fingertips in a closed hand. *J Inf Process Syst* 8(3):399–408
3. Faugeras OD (1993) *Three-dimensional computer vision: a geometric viewpoint*. MIT Press
4. Gijsenij A, Gevers T, van de Weijer J (2011) Computational color constancy: survey and experiments. *IEEE Trans Image Process* 20(9):2475–2489
5. Heo YS, Lee KM, Lee SU (2008) Illumination and camera invariant stereo matching. *IEEE Conf Comput Vis Pattern Recognit (CVPR)*: 1–8
6. Heo YS, Lee KM, Lee SU (2011) Robust stereo matching using adaptive normalized cross-correlation. *IEEE Trans Pattern Anal Mach Intell* 33(4):807–822
7. Hirschmuller H, Scharstein D (2007) Evaluation of cost functions for stereo matching. *IEEE Conf Comput Vis Pattern Recognit (CVPR)*: 1–8
8. Hosni A, Bleyer M, Gelautz M, Rhemann C (2009) Local stereo matching using geodesic support weights. *IEEE International Conference on Image Processing (ICIP)*, pp 2093–2096
9. Hosni A, Bleyer M, Gelautz M, Rhemann C (2010) Geodesic adaptive support weight approach for local stereo matching. *Proc of the Computer Vision Winter Workshop*, pp 60–65
10. <http://vision.middlebury.edu/stereo/>
11. Hu X, Zhang C, Wang W, Gao X (2010) Disparity adjustment for local stereo matching. *IEEE Conf Comput Inf Technol (CIT)*: 1388–1392
12. Iqbal M, Morel O, Meriaudeau F (2010) Choosing local matching score method for stereo matching based on polarization imaging. *IEEE Conf Comput Autom Eng (ICCAE)* 2:334–338
13. Jung IL, Chung TY, Sim JY, Kim CS (2013) Consistent stereo matching under varying radiometric conditions. *IEEE Trans Multimedia* 15(1):56–69
14. Kang SB, Szeliski R, Chai J (2001) Handling occlusions in dense multi-view stereo. *IEEE Conf Comput Vis Pattern Recognit (CVPR)*: 103–110
15. Kolmogorov V, Zabih R (2001) Computing visual correspondence with occlusions using graph cuts. *IEEE Conf Comput Vis Pattern Recognit (CVPR)*: 508–515
16. Malhotra R, Jain A (2012) Fault prediction using statistical and machine learning methods for improving software quality. *J Inf Process Syst* 8(2):241–262

17. Marr D, Poggio T (1979) A computational theory of human stereo vision. *Proc R Soc Lond* 204(1156): 301–328
18. Mattoccia S, Tombari F, Di Stefano L (2008) Fast full-search equivalent template matching by enhanced bounded correlation. *IEEE Trans Image Process* 17(4):528–538
19. Miled W, Pesquet JC, Parent M (2009) A convex optimization approach for depth estimation under illumination variation. *IEEE Trans Image Process* 18(4):1574–1608
20. Min D, Lu J, Do MN (2011) A revisit to cost aggregation in stereo matching: how far can we reduce its computational redundancy?. *IEEE International Conference on Computer Vision (ICCV)*, pp 1567–1574
21. Morikawa C, Aizawa K (2012) Iconic visual queries for face image retrieval. *J Converg* 3(3):39–46
22. Na IT, Choi JH, Jeong H (2009) Robust fast belief propagation for real-time stereo matching. *IEEE Conf Adv Commun Technol (ICACT)* 2:1175–1179
23. Ng JKY (2012) Ubiquitous healthcare: healthcare systems and applications enabled by mobile and wireless technologies. *J Converg* 3(2):15–20
24. Ohkawara T, Aikebaier A, Enokido T, Takizawa M (2012) Quorums-based replication of multimedia objects in distributed systems. *Human-Centric Comput Inf Sci* 2(11):1–16
25. Scharstein D, Szeliski R (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int J Comput Vis* 47(1):7–42
26. Sun J, Zheng NN, Shum HY (2003) Stereo matching using belief propagation. *IEEE Trans Pattern Anal Mach Intell* 25(7):787–800
27. Szeliski R (2011) *Computer vision: applications and algorithms*. Springer, Germany
28. van de Weijer J, Gevers T, Gijssenij A (2007) Edge based color constancy. *IEEE Trans Image Process* 16(9): 2207–2214
29. Veksler O (2003) Fast variable window for stereo correspondence using integral images. *IEEE Conf Comput Vis Pattern Recognit (CVPR)*: 556–561
30. Veksler O (2006) Reducing search space for stereo correspondence with graph cuts. *British Conference on Machine Vision*, p 709–719
31. Yang Q (2012) A non-local cost aggregation method for stereo matching. *IEEE Conf Comput Vis Pattern Recognit (CVPR)*: 1402–1409
32. Yang Q, Wang L, Ahuja N (2010) A constant-space belief propagation algorithm for stereo matching. *IEEE Conf Comput Vis Pattern Recognit (CVPR)*: 1458–1465
33. Yoon KJ, Kweon IS (2006) Adaptive support-weight approach for correspondence search. *IEEE Trans Pattern Anal Mach Intell* 28(4):650–656
34. Zhang K, Lu J, Lafruit G, Lauwereins R, van Gool L (2009) Accurate and efficient stereo matching with robust piecewise voting. *IEEE International Conference on Multimedia and Expo (ICME)*, pp 93–96
35. Zhang K, Lu J, Lafruit G (2009) Cross-based local stereo matching using orthogonal integral images. *IEEE Trans Circuits Syst Video Technol* 19(7):1073–1079
36. Zhou X, Boulanger P (2012) Radiometric invariant stereo matching based on relative gradients. *IEEE International Conference on Image Processing (ICIP)*, pp 2989–2992



Kwangmu Shin received his diploma degree (M.S.) in Computer Engineering from Pusan National University in 2007. He works on his Ph.D. degree from Pusan National University. His scientific interests include Image Processing and Vision, Video Coding and Transmitting, Embedded Linux.



Daekeun Kim received his diploma degree (M.S.) in Computer Engineering from Pusan National University in 2012. His scientific interest include Image Processing and Vision.



Kidong Chung received his M.S. and Ph.D. degrees in Computer Science from Seoul National University in 1975 and 1986, respectively. He currently works as professor from Pusan National University. His scientific interests include Image Processing and Vision, Video Coding, File System, Multimedia System and Communications.