

A New Approach for Stereo Matching Algorithm with Dynamic Programming

Miss. Priyanka M. Lohot
PG Student, Computer Engineering,
Shah & Anchor Kutchhi Engineering College,
Mumbai University.
priya.lohot@gmail.com

Mrs. Bhakti M. Sonawane
Assistant Professor, Computer Engineering Department,
Shah & Anchor Kutchhi Engineering College,
Mumbai University.
bhaktisonawane@rediffmail.com

Abstract— Stereo matching algorithms are one of heavily researched topic in binocular stereo vision. Massive 3D information can be obtained by finding correct correspondence of different points between images captured from different views. Development of stereo matching algorithm is done for obtaining disparity maps i.e. depth information. When disparities computed for scan lines then dense reconstruction becomes time consuming for vision navigation systems. So for pair of stereo images proposed method extracts features points those are at contours in images and then a dynamic program is used to find the corresponding points from each image and calculates disparities. Also to reduce the noise which may lead to incorrect results in stereo correspondence, a new stereo matching algorithm based on the dynamic programming is proposed. Generally dynamic programming finds the global minimum for independent scan lines in polynomial time. While efficient, its performance is far from desired one because vertical consistency between scan lines is not enforced. This method review the use of dynamic programming for stereo correspondence by applying it to a contour instead to individual scan lines. Proposed methodology will obtain the global minimum for contour array in linear time using Longest Common Subsequent (LCS) dynamic programming method with no disparity space image (DSI).

Keywords- Stereo Correspondence, Contour, Dynamic Programming, Disparity space Image, Longest common subsequent.

I. INTRODUCTION

Stereo matching is an important branch of the research area in computer vision. The mobile robot vision navigation system suppose to detect the size and position of obstacle or object from unknown environment in accurate manner and such vision navigation system is carried by mobile robot are based on fast and veracity stereo matching algorithm. In stereo matching algorithm depth information can be obtained by finding correct correspondence of different points between stereo images captured from different views. Stereo correspondence has traditionally been, and continues to be, one of the most heavily investigated topics in stereo vision. The known algorithms for stereo matching can be classified in two basic categories: Area-based algorithms and Feature-based algorithms [2]. Feature-based algorithms (FBA) extract features of interest from the images, such as corners, edge segments or contours, and match them in two or more views. These methods are fast because only a small subset of the image pixels are used, but may fail if the chosen primitives cannot be reliably found in the images; further more they usually only yield Sparse depth map. FBA only calculate points. In Area-based approaches, the system attempts to correlate the grey levels of image patches in the views being considered, assuming that they present some similarity. The resulting depth map can then be interpolated. The underlying assumption appears to be a valid one for relatively textured areas; however, it may prove wrong at occlusion boundaries and within featureless regions.

Alternatively, the map can be computed by directly fitting a smooth surface that accounts for the disparities between the two images. This is a more principled approach since the problem can be phrased as one of optimization; however, the smoothness assumptions that are required may not always be satisfied.

So the disparity map exported by area-based algorithms can be dense enough. As a consequence, there are a large number of spots in the disparity map and the algorithm is time-consuming. Feature-based algorithms only calculate the points which acquired by certain rules. So it produces a sparse disparity map. The vision navigation of mobile robot just needs to detect the outline of the obstacle without reconstructing the surface of the obstacle precise. So, as above considerations, we first extract the contour of the scene, then calculate the disparity of the contour. When solving the disparity using traditional matching algorithm (such as SAD, SSD, etc.), the optimal disparity for each pixel is chosen individually. A noisy disparity map which will generate the wrong construction may be created, which will result in wrong detecting of the obstacle. Aiming to solve this problem, Ohta proposed a stereo correspondence method based on the dynamic programming [5]. After initial matching costs were computed by traditional matching algorithm, the dynamic programming algorithm is applied to global optimization and makes the quality of the disparity map obviously improved. In order to eliminate stripes in disparity map, ground control points (GCPs) are proposed by Stephen [8]. But all of these algorithms are applied only to corresponding scan lines of stereo pairs and not suitable for the contour. So a stereo matching algorithm of contour based on dynamic programming is proposed in this paper. Stereo vision is one of the most actively researched area in computer vision. A large number of stereo correspondence algorithms have been developed. In order to gauge progress in the area, Scharstein et al. wrote a paper which provides an update on the state of the art in dense two frame stereo correspondence algorithms under known camera geometry [2]. The algorithms produce dense disparity maps which are useful in a number of applications including view synthesis, robot navigation, image-based rendering and tele-presence. For researchers in stereo vision one of the most useful outputs of the paper is a quantitative

test bed for stereo correspondence algorithms available, from vision.middlebury.edu/stereo. This paper reviews stereo correspondence algorithms and also proposes a new approach for dynamic programming on contour.

II. EXSITING SYSTEM

This section reviews already discussed stereo correspondence algorithms mostly which have received attention in the Middlebury evaluation [1].

Since the excellent taxonomy presented by Schartein and Szeliski many new methods have been proposed [2]. Matching methods can be grouped into those producing sparse output and those giving a dense result. Feature based methods stem from human vision studies and are based on matching segments or edges between two images, thus resulting in a sparse output. This disadvantage, dreadful for many purposes, is counterbalanced by the accuracy and speed obtained. In order to categorize and evaluate the stereo correspondence algorithms that produced dense output, a context has been proposed [2]. According to this, dense matching algorithms are classified in local and global once. Local methods (area-based) trade accuracy for speed. The disparity computation at a given point depends only on intensity values within a finite window. Global methods (energy-based) are time consuming but very accurate. Their goal is to minimize a global cost function that combines data and smoothness terms.

A. Disparity Space Image

Stephen represents a data structure called as the Disparity-space image, or DSI. The data structure explores the occlusion and stereo problem and it facilitated in development of a dynamic programming algorithm that uses occlusion constraints. The DSI is an explicit representation of matching space [8].

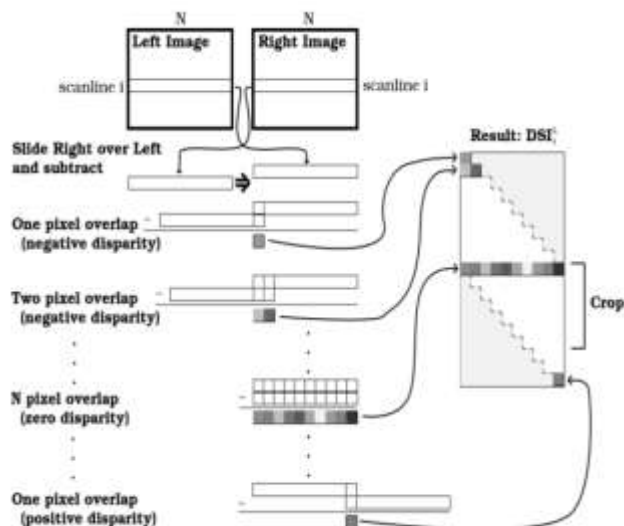


Figure 1. This figure describes how disparity space image is generated [8]. DSI representation for i th scanline in following way: Select the i th scanline of the left and right images, sL sR respectively, and slide them across one another one pixel at a time. At each step, the scanlines are subtracted and the result is entered as the next line in the DSI. The DSI representation stores the result subtracting every pixel in sL and sR and maintains the spatial relationship between the match points. As such, it may

consider as matching space, with x along the horizontal, and disparity along the vertical.

The corresponding epipolar scan line from the left and right images are used. The scan line from the left image is held still as the scan line from the right image is shifted across. After each pixel shift, the scan line is subtracted. The result from overlapping pixels is placed in the resulting DSI.

III. PROPOSED METHOD

Objective of proposed work is to compute disparity maps for stereo image pair simultaneously and efficiently. The motivation behind proposed algorithm is to use the powerful and efficient optimization tool provided by dynamic programming, and instead applying it on individual scan line use it on contours which will more suited for vision navigation system. Disparity computed at one contour pixel depends on the disparity at all others contour pixels. Thus the proposed dynamic programming is truly semi global optimization algorithm and is not 1D optimization method because it operates across both vertical and horizontal dimensions.

Dynamic programming is a classical global optimization algorithm on stereo correspondence. The traditional dynamic programming is applied to corresponding scanlines to find lowest cost rout in the disparity space image (DSI) to optimize result. Further, a contour based dynamic programming with disparity space image (DSI) used to obtain optimal disparity [7]. This paper introduces an improved stereo matching algorithm of contour based with longest common subsequent (LCS) dynamic programming. This will reduce computational complexity comparing to DSI and more accurate result comparing to DSI on contours.

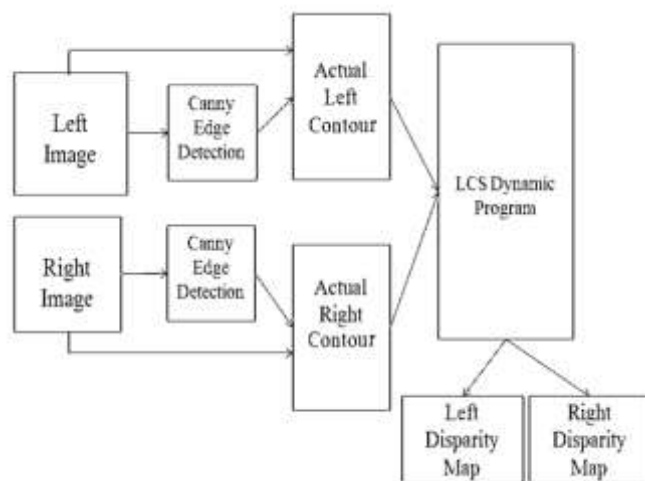


Figure 2 The general block diagram for execution of algorithm.

The general block diagram for proposed system is shown in Figure 2 which has three main sections as canny edge detection, actual contour extraction and dynamic programming with Longest Common Subsequent (LCS) method. This method proposed new stereo matching technique based on dynamic programming algorithm. Instead applying dynamic programming on individual scan line this method uses dynamic programming directly on contour array to find global optimization and obtains optimal disparity. Traditional Scan line based method requires polynomial time for execution as well as disparity computed using search space technique on

contour i.e. DSI on contour also takes polynomial time for execution; this new method uses dynamic programming directly on contour array which will execute in liner time. Sum of points needed to be computed will be considerably reduced when algorithm will compute stereo correspondence with high speed and more accurate than dense reconstruction.

A. Canny edge Detection

The Canny edge detector is an edge detection operator that uses a multi-stage algorithm to detect a wide range of edges in images. Edges or contours in images are areas with strong intensity contrasts. As the contour is the place where the great changes have taken place on the image plane, we can remove the area of near gray level effectively. Edge detecting an image significantly reduces the amount of data and filters out useless information, while preserving the important structural properties in an image.

There are many traditional algorithms such as Sobel operator, Prewitt operator, Laplace operator and so on. This entire operator can perform at high speed. But detected edges are always incomplete and uncontinuous and unsuited to be used directly. By contrast, The canny operator is an optimization algorithm for high signal to noise and precision of detection, extensive used [12].

Apply canny edge detection algorithm for contour pixel extraction of left and right image. The algorithm runs in 5 steps, apply each step on separate image of stereo pair ie. Left and Right image.

- Smoothing:
Blurring of the image to remove noise.
- Finding gradients:
The edges should be marked where the gradients of the image has large magnitudes.
- Non-maximum suppression:
Only local maxima should be marked as edges.
- Double thresholding:
Potential edges are determined by thresholding.
- Edge tracking by hysteresis:
Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

In proposed algorithm canny edge detection has been taken directly. Which when applied on left and right images, generated output gives detail edges with intensity 1 for all contours and every non contour pixel appear with intensity 0. This resultant canny left and right images pair used further for getting stereo correspondence on contour.

B. Actual Contour Intensity

Canny edge detection obtains same intensity for all contour pixels. In solving problem for matching algorithm it is necessary to acquire actual contour intensity. Here mapping of canny edge with gray level image come to picture for stereo pair is shown in following Figure 3.

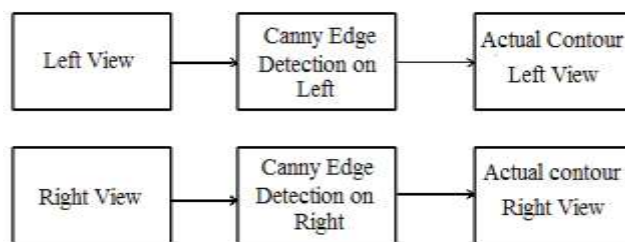


Figure 3 Computing actual contour pixels from gray scale image.

Contour extraction mainly took place to obtain exact shape and size information. In order to find correspondence in reference image the actual intensity is needed. Therefore actual contour intensities extraction took place.

C. Dynamic programming Principal

Stereo matching through dynamic programming (DP) can be considered a semi-global method [8]. This algorithm gets the global minimum for independent scanlines in polynomial time. Problems with dynamic programming stereo include the selection of the right cost for occluded pixels and the difficulty of enforcing inter-scanline consistency, although several methods propose ways of addressing the latter. Another problem is that the dynamic programming approach requires enforcing the monotonicity or ordering constraint. This constraint requires that the relative ordering of pixels on a scan line remain the same between the two views, which may not be the case in scenes containing narrow foreground objects.

The stereo problem in this context becomes that of finding the minimum cost path, or path of least resistance, through a cost matrix such as the one shown in Figure 4. This matrix is constructed by calculating the differences of all the pixels in the reference scan line and all the pixels of the corresponding scan line over a range of disparity levels.

The name dynamic programming comes from a mathematical process through which a problem is solved by breaking it into smaller pieces. The smaller problems are solved first and the collection of all the smaller problems is equivalent to the problem as a whole. The process is used to find the minimum cost path through a cost matrix derived from the image scanlines. First the possible paths through the cost matrix are defined using some constraints.

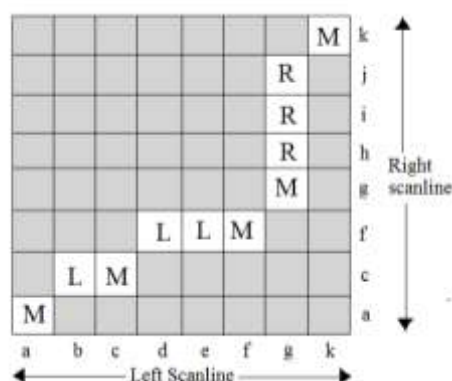


Figure 4 Stereo matching using dynamic programming [8].

The main constraint used in DP is the ordering constraint. The cost of a point is defined in a similar manner to the cost in Equation 1.

$$E(f) = E_{\text{data}}(f) + E_{\text{smooth}}(f) \dots\dots\dots(1)$$

$$E_{\text{data}}(f) = \sum_p D_p(f_p),$$

$$E_{\text{smooth}}(f) = \sum_p \sum_{q \in N(p)} V_{pq}(f_p, f_q)$$

Where $N(p)$ is the neighborhood pixels of p and $D_p(f_p)$ is the data cost, which is a measure of the dissimilarity between pixel p in the left image and pixel p & f_p in the right image. $V_{pq}(f_p, f_q)$ is the smoothness cost that favors the piecewise smooth objects. Taking these costs together, they arrive at the optimal disparities by minimizing the total energy in formula (1). The minimum cost to reach a particular point is then calculated and used to calculate the minimum cost of reaching the next point of any path that moves through the first point. One advantage of global and semi-global methods is that, to a certain degree, they are insensitive to weakly-textured areas. In DP the path will not stray far from the disparities at the edges of the texture less area, because that would increase the cost of the path.

D. Compute Disparity by Dynamic programming

Dynamic programming is recursion without repetition. Dynamic programming example uses a multidimensional array to store the result of recursive sub problems. To compute disparity by dynamic programming between two pixel values need to develop a recursive definition. Suppose we have two pixel arrays which represent all pixels on contour positions and to compute matching pixels values dynamic programming approach will apply. Instead computing Disparity Space Image for left and right image we will apply LCS dynamic programming on contour arrays of left and right image and get matching pixel value for respective contour array. This is done by computing intensity difference of left and right contour array. Here, the polynomial time require for computation of DSI based dynamic programming is more complex than linear time required LCS based dynamic programming. The new approach of LCS based will ensures that computed value will be optimal and more accurate than DSI.

In global optimization, the constraint on the disparity map d are formulated into an equation 1 $E(f)$ which is den minimize over all image pixels. A typical objective function has the following form:

$$E(d) = E_{\text{data}}(d) + E_{\text{smooth}}(d) \dots\dots\dots(2)$$

The data term, $E_{\text{data}}(d)$, measures how well the disparity function d agrees with the input image pair. Using the disparity space formulation,

$$E_{\text{data}} = \sum C(I_i, d(i, j)) \dots\dots\dots(3)$$

where, C is a (initial or aggregated) matching cost. The smoothness term $E_{\text{smooth}}(d)$ encodes the smoothness assumptions made by the algorithm. Modeling is started on objective function of the type in equation (2). The data term in equation (3) can be written as,

$$E_{\text{data}}(d) = m(d_{(i,j)}) \dots\dots\dots(4)$$

let i, j be the coordinates of the pixel in the left image and $d_{(i,j)}$ be the value of disparity map d at pixel. Let $m(d_{(i,j)})$ be the matching penalty for assigning disparity $d_{(i,j)}$ to pixel, in our frame work $m(d_{(i,j)})$ can be sum of absolute differences (SAD) to compare image contours,

$$m(d_{(i,j)}) = \text{SAD}(i, j, d) = |I_l(i+u, j+v) - I_r(i+u, j+v-d)| \dots\dots\dots(5)$$

This means that for every contour pixel in the left image search is performed along the same contour in right image that best matches it, where $d \in h$. Note, if an image has j columns and number of possible disparity values is h , then straightforward dynamic programming take $O(jh^2)$ time, but running time can be reduced to $O(jh)$. Let, $s(d_n)$ be the smoothness penalty for assigning disparity d_n and λ is disparity penalty.

Thus the energy function required to be optimized is given by equation (2) as follows,

$$E(d_{(i,j)}) = m(d_{(i,j)}) + \lambda s(d_n) \dots\dots\dots(6)$$

Then the minimum value of the energy in equation (6) can be written as,

$$E(d_{(i,j)}) = \min(m(d_{(i,j)}) + \lambda s(d_n)) \dots\dots\dots(7)$$

The optimal disparity assignment for $d(i, j)$ in terms of LCS dynamic programming can be written as follows, Let $x = \{x_1, x_2, x_3, \dots, x_m\}$ and $y = \{y_1, y_2, y_3, \dots, y_n\}$ be left and right contour sequences and let $z = \{z_1, z_2, z_3, \dots, z_k\}$ be any LCS of x and y .

- If $x_m = y_n$, then $z_k = x_m = y_n$ implies that z_{k-1} is an LCS of x_{m-1} and y_{n-1} .
- If $x_m \neq y_n$, then $z_k \neq x_m$ implies that z is an LCS of x_{m-1} and y .
- If $x_m \neq y_n$, then $z_k \neq y_n$ implies that z is an LCS of x and y_{n-1} .

The output of z is used for disparity map computation further. The proposed algorithm is simple to describe & implement but also much more efficient. The obtained result of LCS on contour is better than DSI results, which is shown in results and discussions of this paper.

IV. RESULT AND DISCUSSIONS

The algorithms have been tested on the benchmark Middlebury Database; there are many stereo pairs which are used in our experiments and results. Ground truth disparity images present in Middlebury database are used for comparing results. The error is computed as the percentage of pixels is far from the true disparity by more than one. These statistics are collected for non occluded contour pixels. Every stereo pair is given as input to algorithm which will first extract the contour of image, this is because algorithm mainly introduced for pick

and place operations where size, shape and distance of object is needed. Further stereo matching algorithm is applied on actual contour extracted image.

Result is computed for different image pair from Middlebury database to find depth information in which object near is appear more brighter whereas object which is far is appear darker.

The Figure below is the computed results for Midd 2 stereo images pair from Middlebury database for DSI and LCS mechanism on contour pixels. Detail evaluation is done in this section.

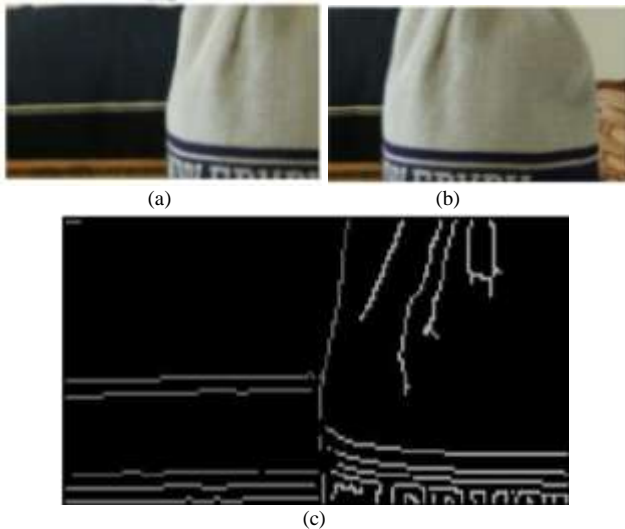


Figure 5 Midd 2 stereo pair and disparity map.

Figure 5(a) is left view image and Figure 5(b) is right view image. Both images are given as input.. In left view object which is near is more shifted towards right and in right view that object shifted to left. On which canny edge detection and actual contour extraction algorithm is applied to generate disparity map. Figure 5(c) is ground truth disparity map obtained by dense disparity map given in Middlebury database for evaluation of accuracy. Cap is near so it more brighter comparing to background in disparity map image Figure 5(c).



Figure 6 Midd 2 disparity map by DSI.

Figure 6 gives disparity map computed using disparity space image (DSI) mechanism, in which we can see differentiating between both images is difficult as we cannot easily identify in the image that which object is near or far, improved mechanism with longest common subsequent (LCS) we can see the object differs than one another.

Figure 7 gives disparity map by LCS where object behind cap is far then computed disparity for that object is darker as

comparing to object which is near to camera is appear more brighter, refer highlighted area by red oval shape. Here, accuracy is achieved by LCS method is more comparing to DSI method of disparity computation.



Figure 7 Midd 2 disparity map by LCS.

The quality will be achieved from computing accuracy of obtained results; this can be done by computing bad pixels percentage and accuracy percentage. For each image pair, we report percentage for all contour pixels. Only non occluded pixels are considered in method using LCS. For bad pixels percentage minimum is considered as best value. If minimum bad pixels percentage is best, alternately maximum percentage of accuracy become best value for given stereo pair.

TABLE I. EVALUATION FOR 2 PAIRS DIFFERENT IMAGES.

Stereo pair.	RMS value		Bad Pixels Percentage %		Accuracy Percentage	
	DSI	LCS	DSI	LCS	DSI	LCS
Wood 1	7.65	3.22	6.56	6.85	93.44	93.15
Midd 1	1.04	1.06	9.87	1.008	90.13	98.99
Midd 2	1.06	1.07	9.92	1.002	90.08	98.99
Baby 1	5.46	5.37	4.81	4.51	95.19	95.49
Baby 2	4.72	4.61	3.89	3.83	96.11	96.17
Baby 3	8.47	8.57	7.46	7.80	92.54	92.2
Bowling 1	9.41	9.32	8.05	8.04	91.95	91.96
Flower	1.03	9.75	8.84	8.14	91.16	91.86
Wood 2	8.17	7.87	6.91	6.73	93.09	93.27

The above table shows different values computed for 2 pair images from Middlebury database. We can see from above table bad pixels percentage is decreases for LCS comparing to DSI. RMS values for both DSI and LCS methods are also calculated. Table above shows detail evaluation for all different images from Middlebury database. The accuracy percentage achieved by LCS method is better comparing to DSI mechanism on contour.

V. CONCLUSION

Proposed method is a new stereo matching algorithm based on dynamic programming approach on contours. Instead applying dynamic programming on individual scan lines this method applies it directly on contour array to find global optimization and to get optimal disparity. Traditional scan lines based method require polynomial time for its execution as well as disparity computed using search space based technique on contour such as DSI on contour also requires polynomial time for its execution; this proposed new algorithm uses LCS dynamic program on contour array which will executes each subsequence in linear time for length of other contour sequence. Sum of points that needed to be computed will be considerably reduced so this algorithm computes stereo correspondence with high speed and more accuracy than other.

Furthermore, this proposed method acts as a guideline for dense stereo matching. Use of LCS dynamic programming for stereo matching on contour is done as reference work for dense matching.

ACKNOWLEDGMENT

I thank to our college 'Shah & Anchor kutchhi Engineering college', for providing its resources and other facilities; helping us to come up with the topic in excellent manner. I sincerely acknowledge my internal guide Mrs. Bhakti Sonawane for her invaluable guidance in giving form and substances to this paper. Her guidance and invaluable suggestions has been major motivating factor. It is her enthusiasm and sincere efforts that have given a sense of direction and purpose to this.

I also thank the principal sir Prof. Vinit Kotak and the Head of Department, Prof. Uday Bhawe sir, for their able guidance and inputs in implementing the presented topic. And lastly our non-teaching staff for making our premises one where working can be a pleasure.

REFERENCES

- [1] <http://vision.middlebury.edu/stereo/2010>
- [2] Scharstein D, and Szeliski R, "A taxonomy and evaluation of dense two- frame stereo correspondence algorithms", International Journal of Computer Vision, 2002, 47(1-3):7-42.
- [3] Fua, P., "A Parallel Stereo Algorithm that Produces Dense Depth Maps and Preserves Image Features", Machine Vision and Applications, 1993, 6(1), 35-49.
- [4] Hannah, M.J, "A System for Digital Stereo Image Matching", Phot. Eng. and Rem. Sens, 2001, 55(12), 1765-1770.
- [5] Ohta, Y., and Kanade, T., "Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming", IEEE transactions on Pattern Analysis and Machine Intelligence, 1985, Vol. 7 No. 2, pp. 139-54.
- [6] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. In International Conference on Computer Vision, pages 377-384, 1999.
- [7] L.Junqin, D. Yan., X. Hu., Q. Xing., B. Yang., "Dynamic Programming algorithm for Stereo Correspondence of Cntour", IEEE 5th International conference on Image and Signal Processing, 2012.
- [8] Stephen S Intille, and Aaron F Bobick, "Disparity-space images and large occlusion stereo", 3rd European Conf on Computer Vision (ECCV'94), 1994, pp.179-86.
- [9] Elisabetta Binaghi, ignazio Gallo, Giuseppe Marino, Mario Raspanti, "Neural Adaptive stereo matching", Pattern recognition Letters 25, pp. 1743-1758, 2004.
- [10] L. Kotoulas, A. Gasteratos, G. Ch. Sirakoulis, C. Georgaulas, and I. Andreadis, "Enhancement of Fast Acquired Disparity Maps using a 1-D Cellular Automation Filter", Proceedings of the fifth LASTED International Conference on Visualization, Imaging and Image Processing, Benidorm, Spain, September 7-9, 2005.
- [11] Fua, P., "A Parallel Stereo Algorithm that Produces Dense Depth Maps and Preserves Image Features", Machine Vision and Applications, 1993, 6(1), 35-49.
- [12] Canny J, "A Computational Approach to Edge Detection", IEEE Trans on PAMI, 1986, pp 679-698.