**Stereo matching technique using Belief Propagation**

**Chapter-1 Introduction**

Stereo vision is an imaging technique that can provide full field of view 3D measurements in an unstructured and dynamic environment. The basic of stereo vision are similar to 3D perception in a human vision and is based on triangulation of rays from multiple viewpoints. Each pixel in a digital camera collects the light that reaches the camera along a 3D ray. If feature in the world can be identified as a pixel location in an image then this feature lies on the 3D ray associated with that pixel. If multiple cameras are used multiple rays are obtained. The intersection of these rays is the 3D location of the feature.

The key problem to solve stereo vision are to identify which pixel in multiple images match the same world feature .This problem is known as stereo matching or stereo correspondence. Stereo Matching is for a given two or more images of the same scene or object, compute a representation of its shape. Stereo Matching is process of finding disparity or depth information.

The tracing of corresponding phenomena i.e. stereo matching is necessary in many applications like image sequence analysis in entertainment, information transfer and [automated](http://en.wikipedia.org/wiki/Automation) systems. Stereo matching is highly important in fields such as [robotics](http://en.wikipedia.org/wiki/Robotics) to extract information about the relative position of 3D objects in the vicinity of autonomous systems. Other applications 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 not be able to distinguish as a separate object by any other criteria.

Scientific applications for digital stereo vision include the extraction of information from [aerial surveys](http://en.wikipedia.org/wiki/Aerial_survey), for calculation of contour maps or even geometry extraction for 3D building mapping, or calculation of 3D [heliographic](http://en.wikipedia.org/wiki/Heliography) information such as obtained by the ISRO project.

There are many algorithms for stereo matching have been proposed, the computation of disparity still remains challenging in texture less regions, depth discontinuities and occluded areas. The stereo algorithms can be roughly categorized into three classes.

The first class is the area-based algorithms. These algorithms attempt to correlate the gray levels of pixels within a finite neighboring window. A central problem of the area- based algorithms is to find the optimal size of the window. While the window size must be large enough to include enough intensity variation for matching, it must be small enough to avoid the effects of projective distortion.

The second class is the feature- based algorithms. They extract features of interest from the image, such as edges, and match the features. The disadvantage of the feature-based algorithms is that they usually yield sparse disparity maps.

The third class global algorithms such as dynamic programming, graph cuts and belief propagation use global constrains over the entire image. Global algorithms can deal with the texture fewer regions and occluded regions well they performed over the whole images.

In spite of having advantages of global algorithm, there have been few real time stereo matching implementations due to their high complexity. Among global algorithm, dynamic programming can be used to implement real time systems but is not robust since it doesn’t consider vertical consistency

The stereo matching problems can formulate in terms of Markov Random Field as minimum energy function, to find energy minimization function is [NP-](http://en.wikipedia.org/wiki/NP-complete)hard. This means a general solution to this problem will take an unthinkably long time to reach a solution. Belief propagation algorithm is one of the many algorithms which find the approximate solution for minimum energy functions used for stereo matching.

Belief propagation algorithm considers both vertical and horizontal consistency is most robust method in the presence of texture less regions and occlusion .However its high computational complexity and large run time memory make it hard to implement in real time systems

**Chapter-2**

**2.2 research problem**

**Literature Survey:**

**Scope and Limitation**

**Tentative scheme**

[1] Chia-Kai Liang,Chao-Chung Cheng,Yen-Chieh Lai,Liang-Gee Chen and Homer H. Chen in their research paper tittle “ **Hardware -Efficient Belief Propagation** ” is that many powerful optimization algorithm such as graph cuts, Loopy Belief Propagation and Tree reweighted message passing have been developed for finding optimal set of labels to solve early vision problems. Among them Loopy Belief Propagation (LBP) has been widely applied to stereo matching, Image denoising etc. The success of LBP is due to its regularity and simplicity.

It uses a simple message update process to iteratively refine the beliefs of labels for each node. A message sent from one node to another node is updated according to neighboring messages and local energy functions using simple arithmetic operations. BP algorithm requires great amount of memory for storing the messages besides that each message is processed hundreds of times, the saving & loading of messages consumes message considerable bandwidth. Furthermore the messages are sequentially updated each and message is constructed through sequential procedure so it is difficult to utilize hardware parallelism to accelerate BP.

In this paper two techniques are proposed, Tile based BP and fast message construction to address the above said issues. Tile based BP splits the Markov Random Field (MRF) in to many tiles and only stores the messages across the neighboring tiles. The memory and bandwidth required by this technique is only fraction of ordinary B.Th. fast message construction technique is based on the observation that many hypotheses used to construct the messages are repetitive , therefore they only need to be computed once ,this observation allows to reduce the complexity of message construction from quadratic to linear. Moreover it can be easily parallelized.

The execution time of the Tile based BP can be further reduced by terminating the processing of a tile after certain stability or convergence criteria are met. The termination criteria are application dependent and data dependent. Tile base approach can be applied to other message passing scheme also, finding the optimal parameters for applications like stereo matching requires further study

[2] Pedro F.Felzenszwalb and Daniel P .Huttenlocher discussed in the their research paper title “**Efficient Belief Propagation for Early Vision**” is that the MRF models are used for solving early vision problems such as stereo ,optical flow and image restoration.While MRF formulation of these problems gives an energy minimization problem ie.NP-hard, a good approximate algorithms based on graph cuts and BP have been developed for the problem of stereo and image restoration.

Despite substantial advances in both the graph cuts and BP approaches still require several minutes of processing time on fastest desktop computers for solving stereo problems.

In this research paper three new algorithmic techniques that substantially improve the running time of BP for solving early vision problems are discussed. The first technique cost function used in early vision problems enables to compute new message by distance transform technique. The second technique using grid graph, in grid graph the beliefs can be obtained using only half as many message updates. This technique speed up the process of compute the messages. The third technique is multiscale algorithm to perform BP in a coarse to fine manner. In multiscale approach the number of message passing iterations can be small because long ranges of iterations are captured by short paths in coarse scale graphs.

The cost function used were simple while implementing these techniques, it can be applicable to a broad range of cost functions including the use of discontinuity costs that vary based on evidence of a boundary or an occlusion.

[3] Li Zhou, Tao Sun, Yuanzhi Zhan and Jia Wang in their research paper “**Software and Hardware Implementations of Stereo Matching**” discussed about the common ground of stereo vision systems is to model three dimensional (3D) spaces and to render 3D objects, using depth information that is the most important element of stereo vision systems. Stereo matching is one of the most active research topics concerning on the depth information processing capability. It is an important stereo vision technique to extract depth or disparity information from stereo images obtained from slightly different viewpoints, by calculating every pixel’s depth information from stereoscopic images. Stereo matching quality is restricted by real-time processing capability, high computation and algorithm complexity, high processing bandwidth requirement and high algorithm accuracy. Especially for embedded systems, low power consumption, high processing performance, high resolution, and high flexibility are all required, as well as various duration, frequency, viewing distance, screen size, ambient light, *etc.,*

The stereo matching problems is still difficult to automatically predict high quality depth map because of image noise, texture less regions, consistency and occlusions that are inherent in the captured images or video frames. There is optimization potential for both stereo matching algorithm optimization software or hardware implementation in terms of speed, parallelism, data bandwidth, memory storage and, etc.

The Graphic Processing Unit (GPU), Field Programmable Gate Array (FPGA) and Application-Specific Integrated Circuits (ASIC) designs are future research trends in real time embedded stereo vision application systems because of their high parallel processing capabilities and specific powerful calculation supporting components. GPU has more programming flexibility and powerful computation capability, while FPGA and ASIC have high performance, lower power consumption and cost. Stereo matching system could be enhanced fantastically along with software and hardware technology development.

[4]Young-kyu Choi, Williem, and In Kyu Park in their research paper title “**Memory-Efficient Belief Propagation in Stereo Matching on GPU**”proposed and implementa novel memory efficient algorithm of BP in stereo matching on the Graphics Processing Units (GPU)**.** Stereo matching is one of the most fundamental techniques that can be used for 3-dimensional reconstruction. The problem is first formulated as a label assignment problem in a probabilistic graphical model such as Markov random field (MRF). Next, it is solved by well established global energy minimization algorithms such as belief propagation (BP) or graph cuts (GC).However, one of the bottlenecks of such global algorithms is that they require large memory bandwidth and space. In this research there are two aspects one is reducing data size and transfer bandwidth. There are two common approaches to solve this problem, namely reducing number of labels (*i.e.* reduction in disparity range) and reducing number of pixels (*i.e.* reduction in spatial domain). The proposed method follows the second type of approach. These techniques are not exclusive of each other, but can be combined to obtain better result.

In this paper a novel memory efficient algorithm for BP is proposed. The data size and transfer bandwidth is significantly reduced by storing only a part of the whole message. This approach, however, reduces the accuracy of the estimated disparity map. In order to maintain the accuracy, the local messages are reconstructed by taking advantage of the shared memory available in Graphic Processing Units (GPU).

[5] “**Parallelization of Belief Propagation Method on Embedded Multicore Processors for Stereo Vision**” is explained byChi-Hua Lai, Kun-Yuan Hsieh, Shang-Hon Lai, and Jenq Kuen Lee is that Stereo vision has been extensively investigated in the recent years to provide high-quality and high performance applications. One of the most advanced paradigms is to apply the technique in inferring the 3-D position of an object. It is proposed to optimize BP algorithm to provide a real-time performance on GPU. In this paper the parallelization of a belief propagation algorithm on the multicore processors have proposed. The methods used to demonstrate the issues in optimizing the algorithm by exploiting the potential parallelism to expose the architecture benefits. The methodology of analyzing and exploiting parallelism presented in this article is applicable to other stereo vision algorithms.

[6] “**Design of Belief Propagation Based on FPGA for the multistereo CAFADIS CAMERA**” This investigation develops a first FPGA implementation for depth map estimation using the belief propagation algorithm for the CAFADIS plenoptic sensor. The main contribution of this work is the use of FPGA technology for processing the huge amount of data from the plenoptic sensor. FPGA technology features are an important consideration in the CAFADIS camera. The depth reconstruction in real time is ensured due to the extremely high-performance signal processing and conditioning capabilities through parallelism based on FPGA slices and arithmetic circuits and highly flexible interconnection possibilities. Furthermore, the use of a single FPGA can meet the size requirements for a portable video camera. The low cost of FPGA implementation in data processing makes the camera sellable at not too expensive prices in the future.

However, algorithm implementation requires an extremely large internal memory. Such massive amount of storage requirement becomes one of the most crucial limitations for the implementation of Virtex-4, Virtex-5 and Virtex-6 FPGA families and the development platform has to be replaced by a subsequent generation of FPGA. The quantifying results with 16 bit precision have shown performances are really close to the original Matlab programmed algorithm. Results have been compared with other belief propagation algorithms in FPGA and our implementation is comparatively faster.

The design of the belief algorithm was developed using functional VHDL hardware description language and is technology-independent. So, the system can be implemented on any large enough FPGA. Xilinx has just announced the release of 28-nm Virtex-7 FPGAs. These devices provide the highest performance and capacity for FPGAs (up to 65Mb) and they will allow algorithm implementation for larger images.

In the future can possible to implement this architecture in a Virtex-7 and integrate it in a real-time multistereo vision system. The goal is to obtain a fully portable system.

[7] Yu-Cheng Tseng , Nelson Chang and Tian - Sheuan Chang in

“**Low memory cost block-based belief propagation for stereo correspondence**” has presented a new stereo matching algorithm that partitions an image to block and optimizes with belief propagation. Our proposed method could reduce memory storage size by 99% with good performance. The independent property could be helpful for parallel hardware implementation. However, with the block based approach, the discontinuous disparities region occurs in the boundary of edges. Thus, in the future research can enhance the interaction between neighboring blocks such that the independent block could extract useful information from the neighboring finished processing block.

[8] “**Efficient Loopy Belief Propagation using the Four Color Theorem”** The Four-Color Theorem based on the max-product

Belief propagation technique can be used in early computer vision for solving MRF problems where energy is to be minimized.

Methods used in this research yield results that are comparable with other methods, but improve either the speed for large images and/or large label sets (the case of image segmentation, stereo matching and optical own), or both the performance and speed (the case of image denoising).

The Four Color Theorem principle is difficult to apply in cases where

the label set is discrete and no natural order/relation between them can be

inferred. This is the case for stereo matching and optical flow, where the

disparity cost function takes discrete, unrelated values. This causes slower convergence, but is compensated by the low time complexity of the methods,independent of the number of labels. Thus, the proposed methods perform faster than the standard methods considered here, at least for large inputs.

[9] **“Task Parallel Implementation of Belief Propagation in Factor Graphs”**: Parallelizing belief propagation in acyclic factor graphs still remains a challenging problem due to the precedence constraints among the nodes in the graphs. In the meanwhile, task scheduling has been extensively studied and used in parallel

Computing task scheduling is shown to be an efficient tool for a class of linear algebra problems, known as regular applications, on general-purpose multi-core processors. Since belief propagation in acyclic factor graphs is an irregular application, task scheduling is even a more suitable tool for parallelizing it. In this research work, defining a task dependency graph for belief propagation and then using a dynamic task scheduler to

Exploit task parallelism available in the task dependency graph.