MOTION DETECTION BY REGION MATCHING

APPROVED:

Dean of the Graduate School

MOTION DETECTION BY REGION MATCHING

by

PRADEEPKUMAR V. ANNAIYAPPA, B.E.E.E.

THESIS
Presented to the Faculty of the Graduate School of
The University of Texas at El Paso
in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE

THE UNIVERSITY OF TEXAS AT EL PASO
December 1987

ACKNOWLEDGEMENTS

I gratefully acknowledge the support (contract 385-50-761) of the U.S Department of Justice and Dr. Stephen Riter for granting me the required financial support and the use of their facilities during the preparation of this thesis.

I also wish to thank the members of my committee, especially Dr. Andrew Bernat, who has not only been a great source of inspiration for my academic performance, but also helped me see the world in a better perspective.

I wish to thank D. Po Wen Hu for helping me prepare the thesis in a proper manner.

ABSTRACT

The process of motion detection is the isolation of all objects that exhibit change in relative position in a sequence of time varying images. The technique discussed in this thesis is based on images acquired through a standard television camera, and digitized to yield a gray value image. The previous work on computer vision based motion detection is reviewed and classified into categories depending on technique used for matching and velocity estimation. It is established that there is a need for a motion detection technique that can efficiently segment an image and represent the data in a method to enable easy adaptation of the motion detection technique to conceptual motion analysis. An algorithm to efficiently segment an image and detect the moving regions is presented, together with results of its application to real world images. A modified algorithm for image segmentation and data representation is described, which is a variation of a simple technique based on the local properties of the pixels. The efficiency and robustness of this algorithm is established in the test results and timing analysis of its application to real world images. The use of this algorithm to applications such as target tracking, and motion detection from a moving camera are discussed.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Since time unknown humans have always tried to relieve themselves of their own work by imposing upon other weaker or less intelligent beings. Mechanization and automation gave a big boost to our efforts in this direction. But the revolutionary change came with the introduction of computers. Computers have been able to replace or aid workers in many low level jobs. The advent of artificial intelligence gave a whole new dimension to this problem. An increasing effort to use computers to aid scientists and engineers in achieving their every day goals has lead to the development of highly intelligent systems that try to emulate the human worker to a great extent. Such intelligent behavior has emphasized the pressing need for visual perception by the system. Computer vision has been used in many applications and has opened new areas of research. One of the most important issues of computer vision is dynamic scene analysis, since the simulation of human behavior requires a powerful visual system capable of understanding dynamic scenes. The use of dynamic scene analysis in solving some of problems has come a long way from analyzing satellite pictures for weather forecasting

to solving complex biological problems. Chapter 2 gives an overview of the varied applications of dynamic scene analysis.

Machine understanding of image sequences requires the development of computer internal representations for invariant entities in a depicted scene as well as discernable changes in appearance and configuration of such entities from frame to frame [Nag83]. Most of the earlier efforts to solve the problem of dynamic scene analysis has emerged out of application oriented problems which directed the research towards the development of low level detection systems. However, current research tends to shift towards modeling motion as a more general phenomena. This has lead to the development of conceptual motion understanding systems. The literature relating to motion analysis has been categorized into a hierarchy of abstractions for the description of motion. Details of this description are given in Chapter 3.

A review of the literature indicates that the several different solutions to dynamic scene analysis have some inherent problems. Understanding motion by a visual system is complex and requires key information about motion parameters. Most of the algorithms that have been

developed are either low level systems that detect the motion component in each pixel and do not concern themselves about an object concept, or they are high level conceptual motion analysis systems that are not concerned with the low level processes of matching and correspondence. This problem is established in Chapter 4. It is also shown that most of the current low level detection systems are either inappropriate for adaptation to conceptual description of motion, or they are very inefficient. Hence, this establishes a motivation for the development of a motion detection system at an intermediate representation level, that can be easily adapted to conceptual motion detection.

Chapter 5 gives a description of the algorithm presented in this thesis. The algorithm is based on region growing and correspondence. An efficient technique to segment an image based on local properties of the pixels is given in Section 5.1 of Chapter 5. The representation technique uses a pointer for each pixel of the image and gives random access to the parameters of the regions in the image. In order to make the algorithm efficient, the parameters used for correspondence are computed during the region growing procedure itself. Section 5.2 gives details about the parameters used and the method of

estimation for each of them. Two time varying images are subjected to this process of segmentation and parameter estimation. With this representation the problem of motion estimation reduces to region matching and displacement computation. The random access of the region parameters enables easy matching with limited search. Section 5.3 gives a description of the process of motion detection. Once the moving regions have been isolated, their position in succeeding frames can be computed and the regions can be tracked. Section 5.4 gives the procedure to track the regions from frame to frame.

The equipment used in the development and testing of the algorithm had the bare minimum capabilities. All the utilities and software were written by the author. Details of the equipment used is given in Chapter 6. The algorithm was tested on many real world images and Chapter 7 discusses the test results on applying the algorithm to such images. A detailed description of the operations required for the execution of this algorithm is given in Chapter 8. Since there are many parameters used in the algorithm, and it has been established that the technique would work with lesser number of parameters, the operations used to compute each parameter are given separately. This technique is then compared with other similar techniques in

terms of efficiency and robustness in Chapter 9. This technique can be adapted to many different applications including motion estimation from a moving camera. Chapter 10 establishes the conclusions and details about possible future work in the enhancement of this algorithm.

CHAPTER 2

APPLICATIONS OF DYNAMIC SCENE ANALYSIS

The analysis of dynamic scenes has been of importance in a wide variety of applications. Some of the early implementations in this area has been the analysis of cloud motion patterns. Satellite pictures of the earth taken with a sufficiently large time interval will indicate the approximate direction of motion of the clouds. This information has been used to estimate the wind speeds at these places. This application received wide interest and many different approaches have been tried [LNT70,MaA79].

Medical applications have used time varying image sequences in many ways. Different techniques to analyze the movement of the heart and other organs are in use today. Other applications include studies in metabolism and blood circulation by measuring the displacement of radioactively marked pharmaceutical introduced into the body. Some of the new applications of image sequence analysis has been in the analysis of human motion for purpose of athletic training and research on artificial limbs [Nag83].

Dynamic scene analysis has been used to study behavioral patterns. Quantitative information about the

movement of pedestrians in halls, corridors and high circulation areas has been used by architects for design purposes.

Security applications have attracted a lot of attention and much research has been done in this area. The U.S border patrol maintain computerized motion detection equipment to monitor any illegal activity along the borders, where such work must be done on a continuous basis and would become strenuous on any human officer. Military surveillance applications need to detect, identify and track a moving target as a matter of routine work. Motion detection can be used as a powerful alternative to the conventional security alarm systems.

Robotics and automated machine tools have begun to play an important role in manufacturing engineering technology. In the early days of robotics the primary function of a robot was to 'Pick and Place', but today robotics has grown beyond pressure transducers and optical sensors. A major constraint in the use of robots was the perfect alignment of parts on an assembly line. Visual aids with the use of dynamic scene analysis have helped remove this restriction and create more intelligent robots.

Current interests have been in applying a vision system to an autonomous vehicle. This has been a major

challenge, as such a system must be very general purpose, fast and must act on a domain independent basis.

CHAPTER 3

LITERATURE REVIEW

Dynamic scene analysis has grown with increasing importance, due to the wide variety of practical applications, such as robot vision, autonomous vehicles, security applications, traffic control, military surveillance and medical uses among others. Most algorithms developed for military applications use only binary images. Since the purpose of this theses is to establish a general purpose motion detection and tracking algorithm based on gray value or color images such techniques have been excluded from consideration.

The various approaches to motion detection have been classified in different ways by different authors [MaA78,JaJ82]. Martin and Aggarwal [MaA78] have classified the techniques based on a human perception model. They have three classes based on the extent of perception by the system. The first category is at the peripheral level where the system only identifies the area of motion activity with no attempt to analyze the type of motion or generate an object concept. The second category is at the attentive level. At this level the systems try to develop an object concept from the time varying images. The third

category is at the cognitive level where the kind of motion is identified with its domain dependant knowledge. Jain and Jayaraman [JaJ82] have given a classification based on the actual technique used in the algorithms. Some of the categories in their classification are differencing, correlation, motion for segmentation and optical flow.

There is an abundance of literature on motion detection and dynamic scene analysis, and many variants of the different approaches have emerged. It is very difficult to give a perfect classification of all the approaches, as researchers have tried many combinations of different techniques to solve the major problem of dynamic scene analysis. The classification adopted in this thesis is a variation of [JaJ82] and is not claimed to be perfect or in any way better than those already existing. The different techniques have been classified into five major groups based on the type of matching technique used and their applicability. The categories are differencing techniques, optical flow, cross correlation, feature matching or correspondence, motion for segmentation and domain dependant conceptual motion analysis. These broad classes can be further divided and are explained as separate sections. Due to the great extent of the published literature in this area, all the known techniques could not be considered for review. A representative set

of algorithms have been selected to establish the different approaches that have been adopted in estimating the motion parameters of an object in a time varying sequence of images.

3.1 DIFFERENCING TECHNIQUES

The basic idea in motion detection is to establish the spatial changes in a sequence of two or more time varying images. Perhaps the easiest way to do this would be taking the absolute difference between the gray values of corresponding pixels from two images. However this simple approach has many problems as has been demonstrated in the algorithms developed [JaN79,OnO80, Jai81]. Some of algorithms have incorporated this simple idea, only to help identify the general area of motion activity [Jai81], while others use a sequence of difference images to obtain an object concept of the images.

Jain [Jai81] has identified the different kinds of difference picture regions that are generated from the motion of different kinds of objects. This technique can be used only to identify major areas containing motion activity.

Oneo and Ohba [OnO73] have developed a technique based on differencing, which can be used only in specific

application of fixed camera positions and fixed picture background. The application discussed in their paper is traffic scene analysis. They obtain a reference frame of a street intersection with no moving objects and this image is stored in memory. Then a sequence of difference pictures are obtained using subtractive television, a concept where a signal level differencing is applied to the images. Since the differencing is done with a reference frame without the objects, all new objects in the scene appear as difference regions. A region growing is applied to a sequence of two such frames and the regions are matched and compared to obtain the magnitude of velocity. This technique has very limited applications. If this system were to be used during the day with the sunlight as the illumination source, then new reference frames have to obtained very often due to the changing position of the light source and consequent movement of the shadows in the image. Note that the reference frames must not have any moving objects in it and this places severe restrictions to the application of this technique.

Jain, Martin and Aggarwal [JMA79] have developed a technique that gives the motion in the object as a whole and not its component parts. They generate a difference picture using cross correlation with a window size of 4 pixels by 4 pixels. This difference picture is segmented

into difference regions. The difference regions are analyzed and classified based on the ratio CC/CP, where CC (CP) is the number of points which are both extreme points in the difference region and edge points in the Current frame (Previous frame). It has been shown in [JaN79] that the ratio CC/CP is less than one for difference picture regions that belong to the background and much greater than one for difference picture regions that belong to the object, and equal to one (ideal case) for regions belonging to both types. Once the type of the region is identified a region growing is applied to the background difference picture region in the first frame and the object region in the second frame. Region growing is done on a row by row basis, starting with the edge pixels on the difference picture regions. Once the object mask has been obtained the regions are refined, to fill holes in the region and to delete isolated noise pixels not belonging to the moving regions. No timing analysis is given for comparison with other techniques. However, it appears that the system would be computationally inefficient as we need a difference picture based on cross correlation, a region growing of the difference picture, an edge picture to compute the CC/CP ratios, then a region growing of the objects itself.

Jain and Nagel [JaN79] have developed a system that