MULTI-LAYERED APPROACH FOR TRACKING GENERIC MULTIPLE OBJECTS AND EXTRACTION OF VIDEO OBJECTS FOR PERFORMANCE ANALYSIS

Synopsis of the thesis Submitted to
Acharya Nagarjuna University
in partial fulfillment of the requirements for the Award of
DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE

By
Y.SURESH BABU, M.C.A.,

Under the Guidance of
Dr.G.S.PRASAD
Professor
Department of Computer Science and Engineering
R.V.R & J.C College of Engineering
Chowdavaram, Guntur.

ACHARYA NAGARJUNA UNIVERSITY
NAGARJUNA NAGAR
INDIA

May, 2012
This is a proposal for a thesis topic of “Multi-Layered Approach for Tracking Generic Multiple Objects and Extraction of Video Objects for Performance Analysis”. It is anticipated that this work will also provide and demonstrate a design framework for Tracking Multiple Generic objects in complex situations of a moving picture, prepare a content based video indexing for tracked objects, Segregate the objects from the content video index or crop the object, Segregated objects are applying for Super Zooming to increase its intensity for better identification. And also a comparative study of multiple objects tracking algorithm with the proposed and Various Zooming algorithm with the proposed.

Statement of Research:

According to the survey mentioned, various algorithms are written and developed to track multiple objects of the same class for a particular related video format only, but lack with different video format and different object class type. These algorithms have not been check for performance too among themselves.

Survey also states that traced video objects extraction is done using various methods, these methods used to extract object by segregation, segregation of objects are done only to particular class object type. But it lack with extraction of different objects of different class type.

The goal of work is

- To develop a generic algorithm to track generic multiple objects
- To segregate multiple generic objects
- Zooming objects with the proposed and comparison with the proposed
- Comparison of various tracking algorithms with the metric

To develop a generic algorithm to track generic multiple objects Multi-level chronological function algorithm for tracking multiple generic objects in a motion picture

Introduction

Tracking multiple objects simultaneously is key for many vision applications, such as visual navigation and object activity recognition. Even though each object can be tracked separately, tracking objects together is important for obtaining good results if objects have complex interactions [1]. I categorize object interactions into two classes. The first type of interaction constrains the object relative locations, i.e., objects tend to keep relative positions
or spatial layout during a short period of time. The second type of interaction is object mutual occlusion, i.e., an object in front occludes other objects in the same region.

Explicitly modeling interaction of objects enables tracking multiple objects more robustly, especially in cluttered environments. But, the search space also increases drastically compared to that of tracking objects separately. Naive exhaustive search becomes intractable. Efficient exhaustive searching schemes such as extended Generic programming function [1] are still too complex to be applied to problems with a medium number of observations and objects. I propose a chronological programming relaxation scheme for a specific class of multiple object tracking problems, in which the metric for inter-object position interaction term is convex while the intra object terms quantifying object state continuity along time may use any metric. The proposed scheme explores a large search space efficiently and almost always gives a global optimum because of the special structure of the formulation. Multiple object tracking has been studied intensively. For example, Kalman filtering has been a classic scheme for object tracking. Recently, particle filtering has been popular for tracking multiple objects such as ants [2] with complex interactions. Particle filtering has also been studied for tracking hockey players [3] in which object interaction is not explicitly modeled. Bayesian networks have been applied to optimizing trajectories of football players in video [5]. This approach does not consider track interaction among objects.

Generic programming function (GENERIC FUNCTION) is also widely applied to multiple object tracking. The single chain Viterbi algorithm can be extended [1] to optimize multiple tracks simultaneously. The computational complexity of extended generic function is $O(mk^2n)$, where $k$ is the number of observations at each stage, $n$ is the number of objects and $m$ is the length of the sequence. Extended generic Function is thus hard to apply to large scale problems. An efficient approximate generic programming function scheme [4] has been studied to find a single object’s path with heuristics used to determine the sequence of path assignments in a multiple-camera setting. While simple heuristics such as best-track-first assignment works well for multiple camera tracking, it does not always give correct solutions when objects have complex mutual occlusion patterns, especially for single camera applications. Chronological programming (Chronological Programming) is another approach that can be used for more efficient search in object tracking. Optimizing object tracks using 0-1 Integer Programming [6] has been studied for radar data association. This formulation is different from the proposed scheme in that a variable is defined for each feasible trajectory and object tracking is solved as a set packing problem. Other approximation methods for solving similar integer Chronological Programming formulations as [6] are studied in [7, 8], which turn out to be quite similar to the sequential generic method [4]. Unlike previous Chronological Programming methods, our proposed scheme is based on a multiple-level-
shortest-path model that tries to connect edges into paths and has much fewer variables. Belief Propagation (BP) [9] has also been used for optimizing tracking. Occlusion is explicitly modeled in this method. However, multiple object tracking results in a loopy graph structure making it difficult to guarantee convergence to a global optimum. Even though intensively studied, robust and efficient tracking of multiple objects with complex interactions remains unsolved. I propose a novel chronological scheme to optimize multiple object tracks simultaneously by explicitly modeling spatial layout constraints and mutual occlusion constraints and formulated object tracking as a multipath searching problem. Each path is composed of a sequence of states, e.g., locations and appearances, of an object through time represented by nodes in a graph. Different tracks are constrained so that objects cannot occupy the same spatial region. Convex penalty terms are included to constrain the consistent objects’ layout in space, i.e., the objects’ relative positions do not change abruptly from frame to frame. The state continuity metric term along time may use any metric. Based on the special structure of our formulation, a chronological programming relaxation approach effectively solves the path searching problem when paths overlap and objects occlude each other. As my results illustrate, the chronological program almost always yields integer solutions that globally optimize object tracks and has low order polynomial average complexity.

**Frame work used for Multi-level Object Tracking**

![Diagram showing network model for multiple Object tracking programming.](image)

The network model for multiple Object tracking programming.

\[
c(v_{n,m,i}, v_{n,m+1,j}) = \begin{cases} 
  g(s_n, v_{n,m+1,j}) + 
  \lambda_1 \cdot g(v_{n,m,i}, v_{n,m+1,j}) + 
  \lambda_2 \cdot d(v_{n,m,i}, v_{n,m+1,j}) & \text{if } v_{n,m,i} \text{ is occluded} \\
  c_{const} + g(s_n, v_{n,m+1,j}) & \text{otherwise}
\end{cases}
\]

- **Overlapped regions.**
- **Spatial layout consistency**

(a): Partially overlapped regions (b) Fully overlapped regions.
(a) The iteration process of a Pipe contour for multiple objects, (b) the enlarged region of the intersection of segments, (c) the splitting and connecting procedure.

Procedure for Evaluation of The problem

- Discrete optimization
- Multi-level function
- Energy flow MLP function
- Estimation of multi-level flow
- Improved energy flow
- Protocol used for splitting and connecting contours
- Determination of Intersection Segments
- Chronological Programming Relaxation

Proposed Algorithm

The Merging of multi-pipe and the nearest path for tracking multiple objects which are advanced.

Algorithm of the Boundary Detection of Multiple Objects

Algorithm for tracking

Logic State – The goal of the algorithms to track multiple generic objects, of different shapes and size and reduces collision. The basic work of proposed algorithm is to create a multi-layered pipe, which stores each objects trajectory in a pipe.
Each of the object track information stores in a pipe stack and its variation of motion, when the objects states motion leads to collision, merge split algorithms eliminates collision resolution problems.

The collision between objects tracking is minimal by using this algorithm; the collided state of objects is identified easily.

**Results:**

![Generic Object in a Motion Picture](image)

*Figure 3* (a) X-location curve (b) Y-location curve

**Example 1:** I apply generic function with best-track-first assigned heuristics to the same data. The energy function of generic function is the same as the proposed scheme except for the spatial layout consistency term. Approximate generic function not easily extended to include such regularization terms since it optimizes each track separately and then assigns tracks sequentially. Figure 2 shows the tracking result of approximate generic function.

In this example, generic function first picks the object 0 track as a better fit and determines the track for object 1 after removing assigned boxes for object 0. As shown in this example, greedy track assignment selected wrong labels at the first and third occlusion instances. Simply reducing the occlusion label cost will not solve the problem and it also causes many missed detections. At last finding the complexity of the proposed.

**Generic objects classifying and Segregate multiple objects motion picture**

Video object (VO) extraction is of great importance in multimedia processing. In recent years approaches have been proposed to deal with VO extraction as a classification problem. This method calls for state-of-the-art classifiers because the performance is directly related to the accuracy of classification. Promising results have been reported for single object extraction using Support Vector Machines (SVM) and its extensions. Multiple object extraction, on the other hand, still imposes great difficulty as multi-category classification is an ongoing research topic in machine learning. I introduce a new class of multi-category learning scheme for multiple VO extraction, and demonstrate its effectiveness and advantages by experiments.
Video object (VO) extraction, the process of segmenting and tracking semantic entities with pixel-wise accuracy [11], is an important yet challenging task for content-based video processing. For the purpose a great deal of approaches have been proposed [12]–[20], which provide satisfactory results for extracting VOs of homogeneous motion characteristics. Unfortunately, dealing with VOs with abrupt motions or occlusions remains a challenge. In recent years classification-based approaches have been proposed to meet the challenge by handling object tracking as a classification problem. Each VO is considered as a class, and VO extraction is achieved by classifying every pixel to one of the available classes. By doing so temporal associations of objects between frames are automatically maintained through correct classifications which is therefore motion-assumption free. As a result, the approaches are more robust to complicated motion fluctuations. What learning algorithm to use is key to the success of the classification-based approaches? By using powerful classifiers high classification accuracy can be achieved which leads to better performance for VO extraction. However, most of the results reported are limited to single object scenarios. In other words, only binary classification between the object and the background has been tackled. At the first glance, the extension from single object to multiple object extraction is straightforward since conceptually one only needs to replace the binary classifier with a multiclass classifier. Unfortunately, the implementation of such an extension is far more difficult than it appears because multi-category classification is still an ongoing and immature research topic itself in machine learning. Only recently have works emerged to offer new tools that can help tackle the multi-object problem. This work presents an attempt of such. Over the last decade, margin-based classification technologies for which the best known have drawn tremendous attention due to their theoretical merits and practical success. Instead of directly estimating the conditional probabilities, the margin-based classifiers focus on the decision boundary which, however, makes it difficult to generalize their applications from binary to multi-class scenario. “Single machine” and “error correcting” are two mainstreams for multi-class margin-based classification. As its name suggests, the “single machine” type of approaches attempts to construct a multi-class classifier by solving just a single optimization problem. On the contrary, the “error correcting” type of approaches works with a collection of binary classifiers, for which the primary goal is to determine what binary classifiers should be chosen to train and how to combine their classification results to make the final decision. Among all the methods published in the literature, “one-against-all”, “one-against-one” and directed acyclic graph (DAG) are most popular choices in solving real-world problems. A good overview of multi-class classification can be found. As a natural extension of binary large margin classification, the “single machine” type of approaches is intuitively appealing. It has drawn even more attention when certain formulations are reported to yield classifiers with consistency approaching the optimal Bayes error rate in the large sample limit. Multi-
class $\Psi$-learning is such a learning algorithm. Moreover, $\Psi$-learning aims directly at minimizing the generalization error (GE), which is the reason why its binary version has shown significant advantage over SVM in terms of generalization both theoretically and experimentally. The extended multi-class $\Psi$-learning retains the desirable properties of its binary counterpart. In addition a computational tool based on the recent advance in global optimization has been developed to reduce the time of training for the “single machine”. The purpose is to two fold. First, it introduces multi-category $\Psi$-learning to tackle the multiple VO extraction problem. Secondly, it reports the performance of the new learning algorithm on several MPGE4 standard video sequences instead of synthetic data which many multi-class learning algorithms are tested on. The rest of the work is organized as follows. First introduction of multi-class $\Psi$-learning. Then a multiple VO extraction method using this new learning methodology is explained. It provides the experimental results which is followed by conclusions.

MULTI-CATEGORY $\Psi$-LEARNING

Multi-Category $\Psi$-Learning

Seeking a function $f$ to minimize GE is the ultimate goal for any learning algorithm. Theoretical Advantage of Multi-Category $\Psi$-Learning

1) $\Psi$-learning estimates the Bayes classifier $\sim f$ as opposed to the conditional probability. However, the optimal classification performance of $\sim f$ is realized via the $\Psi(u)$ function which differs from $1$-$\text{sign}(u)$.

2) The convergence rate of learning decreases as the number of classes $M$ increases although the order remains the same for finite $M$.

3) Unlike the binary case, the optimal performance of linear learning may not be achieved at large $C$ for multi-category problems.

MULTI-OBJECT EXTRACTION USING MULTI-CATEGORY $\Psi$-LEARNING

Related Work

Filtering and association and representation and localization are two major techniques for object tracking. Rooted in the control theory, the former technique deals with the dynamics of the objects while the latter heavily relies on image processing technologies. The way these two techniques are combined and weighted is application dependent. For example, the filtering and association method prevails in the application of aerial video surveillance because the motion of the objects is the major concern. For content-based video processing, on the other hand, objects of interest are usually heterogeneous in spatial features, non-rigid in temporal domain, yet rich of visual information. For this reason representation and localization is the technique used most in VO extraction.
For VO extraction using the representation and localization technique, a reference model representing the object must first be created which can be done either in an automatic [12]–[17] or semiautomatic fashion [18]–[20]. A variety of models has been proposed including: 2-D, such as RGB[21], the histogram of colors, or the coefficients of the DCT transform of the block centering at the pixel [22]. Then in the second step a classifier is trained and the classification function is obtained to discriminate the pixels that belong to the object from those that belong to the background. Different classifiers have been attempted such as neural networks.

Ψ-learning [22] or even an ensemble of linear classifiers. The third step is to evaluate the classification function at every pixel in the subsequent frames. The final step is to generate the tracked object based on the classification results for which the way of implementation varies. For example, a so-called confidence map is first produced according to the classification results, and tracking is then realized by locating the object where the peak of the confidence map occurs. The output of the tracker, however, is a rectangle that tightly encircles the object of interest. For the task of VO extraction the fourth step can even be skipped since after the classification step we already know for every pixel if it belongs to the object. However, for efficiency purposes it is not necessary to do the classification pixel by pixel. By exploiting the spatial redundancy, we introduce the block-level classification instead and design a pyramid refining scheme to refine the boundary in an efficient and scalable manner.

Accuracy and complexity are two critical issues for VO extraction which have to be traded off in practice, and the major advantage of the classification-based methods is the potential to achieve both. The methods are accurate because powerful classifiers are designed for the purpose of object/background separation. Low complexity, on the other hand, is achieved through evaluating the classification function at each pixel which involves only simple calculations, e.g., \( w^T x + b \) for linear SVM while the time-consuming processes of object modeling, extracting, and searching are circumvented.

*Figure. An overview of the proposed approach for multiple VO extraction.*
Proposed Approach

As mentioned before, the choice of the learning algorithm is key to the success of the current approach because the performance of the algorithm is directly related to the classification accuracy. Considering the single VO extraction as an example, the background and the object are often not separable. As pointed out $\Psi$-learning aims at the minimization of GE and therefore has the advantages in non-separable cases. For this reason, a method for single VO extraction that employs binary $\Psi$-learning as the classifier is proposed in [22]. To tackle the challenging task of multi-object extraction, multi-category $\Psi$-learning has to be employed. As shown in Figure, the approach consists of two phases: the training phase and the tracking phase. For classification at the pixel level, individual pixels are represented by pixel-wise color or intensity information, which however would result in misclassifications due to the negligence of the spatial relationship among pixels. Another concern is the size of the training set. If every pixel is included, it would contain too many training samples to yield a quick training especially when the frame size is large. The same efficiency issue exists if we do the pixel-by-pixel classification in the tracking phase. Fortunately, in most video sequences there is abundant spatial correlation that we can take advantage of to make the approach more efficient. Let $p$ denote a pixel and $N(p; d)$ the set of pixels within a small distance $d$ from $p$. Due to the spatial correlation of images, the class labels as well as the feature vectors of $p$ and $N(p; d)$ tend to be similar to each other. Based on this observation, we introduce the concepts of object blocks and background blocks, and suggest the representation and classification both be done at the block level as follows. Suppose we have $M$ VOs of interest. The training phase begins with dividing the first frame, chosen as the training frame, into $(M+1)$ types of blocks (the number of different VOs plus background) depending on which object or background the pixel at the center of the block belongs to. The block size is empirically chosen as 9x9, and evidently the number of blocks determines the size of the training set. We use the same method as in [22] to represent each block as well as the centering pixels. Namely, Discrete Cosine Transform (DCT) is first applied to each block and then based on the DCT coefficients $c(i; j)$ the local and neighboring features are constructed for each block.

EXPERIMENTAL RESULTS

In all, applying the proposed multiple VO extraction method will be tested to four standard MPEG-4 test video sequences, which exhibit varieties of temporal and spatial characteristics.

These sequences are animals, students, Trevor, and Sun Flower Garden, respectively. Also the performance comparisons are made between multi-category $\Psi$-learning and three
popular multi-class algorithms, namely one-vs-all, one-vs-one and directed acyclic graph (DAG). All experiments are carried out on a Pentium IV 2.5-GHz PC.

VO extraction is of great importance for content-based video analysis, and a great deal of research has been performed for multiple object extraction. Unfortunately, multi-object scenario which is more realistic imposes a much greater challenge. Following the idea that handles VO extraction as a classification problem, this paper aims to tackle multiple object extraction by solving a multi-class classification problem and using multi-category \( \Psi \)-learning which is an emerged area in research.

**A frame work using hyper-based methods for image zooming and super resolution**

Image zooming is the process of enlarging the image to the desired magnification factor. But while enlarging an image there are few parameters that we have to be kept in mind. When the image is zoomed, artifacts like blurring, jagging and ghosting may arise. So the main focus is on the reduction of these artifacts.

The algorithm deals with the edges. It is basically designed to preserve the edges. It’s as adaptive zooming algorithm which focuses on preserving edges. The algorithm reduces the jagging. Blurring is reduced a lot in the algorithm.

To compare the algorithm with existing algorithms, I have taken few real world images and results are visually compared. And I have come to decision that the algorithm is better than the traditional methods. I have compared the images by four ways – MAE, MSE, CCC, and PSNR.

Zooming is the process of enlarging something only in appearance, not in physical size. This enlargement is quantified by a calculated number, called *magnification*. When this number is less than one it refers to a reduction in size, called *minification*. Image zooming [23] is among the fundamental image processing operations. Typically zooming is related to scaling up visuals or images to be able to see more detail, increasing resolution, using optics, printing techniques, or digital processing. In all cases, the zooming of the image does not change the perspective of the image. Applications are varied in different fields. In medical imaging, zooming can serve to improve the chances of diagnosing problems by highlighting any possible aberrations. Enhancing image details can also be useful for the purposes of identification, whether for improving the quality of an image interpreted by a biometric recognition system or trying to get a clearer view of the perpetrator of some crime. In entertainment, zooming can be used to resize a video frame to fit the field of view of a
projection device, which may help to reduce blurring. Finally, the most obvious application of image zooming is to simply allow one to enjoy a larger version of a favorite image obtained from any commercially available digital imaging device such as a camera, camcorder or scanner.

Traditional image zooming techniques use up-sampling by zero-insertion followed by linear filtering to interpolate the high-resolution samples. The main drawback of this approach is that the frequency content of the high-resolution image is the same as the low-resolution image. This is due to the fact that linear techniques are incapable of introducing new information into the image. The lack of new high frequency content results in a variety of undesirable image artifacts such as blocking, staircase edges and blurring.

Motivation

Studies on image zooming have been done in various ways. Instead of much work in this field there are very few image zooming software which provide adaptive zooming methods. Even the latest software uses the traditional linear magnification algorithm. However such software’s are recently brought into market. Many interpolation algorithms currently used in consumer products produce magnifications that include undesirable artifacts (flaws) like blurring, “jaggies”, and “ghosting” (Figure). The consumer, presented with the magnification, will be unsatisfied with the unrealistic image upon seeing the artifacts. As is known, sharpness of edges and freedom from artifacts are two critical factors in the perceived quality of the interpolated image [23]. More advanced magnification methods such as learning-based algorithms require specific training data, large running times, or extensive user input. Still there is a need for a more general interpolation technique that can minimize these flaws. The motivation of this research is to create a magnification algorithm that creates realistic real-world image magnifications that do not require a large amount of user input. A realistic real world image would be an image that is free from artifacts such as blurring and jaggies. The image must include smooth contours and also rapid edge transitions in areas of the image where sharp edges are found in the low resolution image. Our adaptive zooming algorithm can produce magnifications with these properties.

Work Description

The main thrust of the thesis is to bring out the actual performing of zooming of images while preserving the edges (adaptive zooming). As we all know that this is the field where a lot of work has already been done. But still my approach is to study all the work earlier done and find a better zooming algorithm. I have started from the literature survey of the earlier techniques. I realized that most of the algorithms focus on the traditional methods,
but we have changed our path to adaptive methods. The traditional methods have many artifacts, but I have tried to resolve them. Firstly, I start from the literature survey. Most of the methods use the interpolation based techniques which is quite useful but it is non-adaptive in most of the cases. I have described each and every interpolation based techniques with visual examples.

Secondly, the problem statement has been defined. Here, in depth my description of the problem is given. In next, the proposed algorithm is explained.

The algorithm is an adaptive based algorithm which basically focuses on the edges. Later, the result has been evaluated with the help of proper mathematical methods like cross correlation, PSNR, etc. Finally, the conclusion has been given with the benefits as well as disadvantages of the proposed algorithm.

Problem Statement

An image zooming is of interest in many applications like scientific visualization, multimedia applications and image analysis tasks. A generic image zooming algorithm takes as an input a digital image and provides an output a picture of required size preserving as much as possible the information context of original image. Several good zooming techniques are now days available. For a large class of zooming techniques this is achieved b mean of some kind of interpolation: replication, bilinear and bicubic are the most popular choices and the routinely implemented in commercial digital image processing software. Unfortunately, these techniques, while preserving while preserving the low frequencies content of the source image, are not equally able to enhance high frequencies in order to provide a picture whose visual sharpness matches the quality of the original image In most of the cases zooming leads to distortion of image. This distortion can be of many ways such like jagging - blocks are formed due to replication of pixels, burring- It is unclearity of the image, ghosting- it is the distortion of the image.

Current interpolation algorithms attempt to solve these artifacts in a number of ways. There are other factors have to be kept in mind while making an algorithm like speed (zooming should not take much time for enlargement), memory requirement (the algorithm should not take much memory space). I have taken into account information about discontinuities or sharp intensity variations while doubling the input image. The proposed technique is mentioned briefly below
Proposed Algorithm

Algorithm Basic Stages

This section describes the basics of image zooming.

![Figure: Basic Magnified Image](image_url)

In Figure: the image A is magnified 2 times to image B. The blank circles are the unknown pixel values and we have to find them. The above image A (2x2) is magnified to B (4x4). But in reality the original image (nxn) is magnified to (2n-1)x(2n-1).

Our algorithm includes 3 stages. All three stages have been discussed below.

Stage I

First step is to expand the original (n x n) image to (2n-1) x (2n-1). I can make it to (2n x 2n) by extending the initial image borders. The figure given below describes the above theory. In the figure the X (5 x 5) is the original image and Y (9 x 9) is the magnified image. X(i,j) is the pixel in the original image where i is the ith row and j is the jth column. In the same way in Y(m,n), m is the mth row and n is the nth.

![Figure: Zooming Phase](image_url)

The black dots represent the pixels in the original image (X). And those pixels are as it is copied in the image Y. These pixels can be mapped as:

\[ X( i , j ) = Y( 2i-1 , 2j-1 ) \]

The white dots represent the unknown pixels whose values we have to find. When an image is magnified 2 times, total number of pixels will be four times. Hence, we have now 1 known and 3 unknown pixel and we have to find 3 other with our algorithm.
**Stage II**

In this stage we have the centre pixel X. As we can see in the Figure below, the centre pixel X has to be found out. The centre pixel X is deduced with the help of the algorithm which is described in the section

The $H_1$ and $H_2$ are considered in case of the horizontal boundaries. $V_1$ and $V_2$ are considered in case of vertical boundaries. In all other case these four pixels are left vacant. The will be filled in the later stages. $H_1$ is considered in case of upper boundary of image then,

$$H_1 = \frac{(A+B)}{2}$$

$H_2$ is considered in case of lower boundary of image then,

$$H_2 = \frac{(C+D)}{2}$$

![Figure : Stage II Zooming](image)

V1 is considered in case of left boundary of image then,

$$V_1 = \frac{(A+C)}{2}$$

V2 is considered in case of right boundary of image then,

$$V_2 = \frac{(B+D)}{2}$$

Now, we have to repeat the above procedure for the complete image until all the unknown pixels of the boundaries and the centre pixels are found. After this we have to move to stage III. Eventually we also have to plot the values of the boundaries of the image.

**Stage III**

In this stage, we have to start from the beginning of the image and find the left over pixels as shown in figure below. In the figure (I) and (II), A & B are the pixels from original image. And X1 and X2 are the pixels derived from stage I.

![Figure](image)

All the pixels \{A, B, X1, X2\} are considered as in Stage I and computed in the same way. Finally the value of M is computed and put in the place of centre pixel. The (I) and (II) are computed simultaneously for the complete image. The M is computed using algorithm as shown. After all these steps we finally get the zoomed image.
Results

Visual Comparison

The first qualitative analysis of adaptive method is a series of image comparisons that will be presented. It is clear that visually the output of the algorithm is better than the other existing techniques if the input image has sharp edges. This gives brief image comparisons, and also tested to various types of images and labeling their corresponding interpolation algorithm later.

Figure: Boy Image comparison. (a) Pixel Replication (b) Bilinear Interpolation (c) Bicubic Interpolation (d) Proposed Algorithm
REFERENCES


