Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Real-Time Augmentation of a Newspaper Using Computer Vision

By

Claire Beatty-Orr

Supervisor: Dr. Kenneth Dawson-Howe

A dissertation submitted to the University of Dublin, in partial fulfilment of the requirements for the degree of M.A.I

Submitted to the University of Dublin, Trinity College, May, 2016

Declaration

I, Claire Beatty-Orr, declare that the following dissertation, except where otherwise stated, is entirely my own work; that is has not previously been submitted as an exercise for a degree, either in Trinity College Dublin, or in any other University; and that the library may lend or copy it or any part thereof on request.

Signature:

Date:

ACKNOWLEDGEMENTS

Firstly, I would like to thank my supervisor Dr. Kenneth Dawson-Howe for his time and guidance throughout this project.

I would also like to thank my parents, Robert and Eleanor, and my brother, Ian, for their support and encouragement throughout the project.

ABSTRACT

The following dissertation aims to augment an adapted newspaper using the combination of a projector and computer vision methods. The final aim is for video footage to be projected onto the newspaper and to appear as if it is printed on the newspaper, as seen on the Daily Prophet in the Harry Potter films. The projection should be correctly transformed to correct distortions.

The technical side of this dissertation is mostly related to augmented workstations, desks and textbooks which were introduced in the early 1990s. Most of these prototypes aimed to improve physical notes by digitizing them, which has now been accomplished through tablets. This dissertation is justified as an interesting application of computer science or computer vision. It can be used as a fun prototype to increase interest in the field or just to serve as entertainment for others.

The adapted newspaper is fitted with fiduciary marks for the system to detect. These marks must be identified and labelled as the system requires information for each of the four corners to transform the projection. Two different identification methods were completed each with their own limitations, namely the assumption method and the colour method. The colour method was designed to overcome limitations of the assumption method and is more complex, this has resulted in limitations of its own.

Both methods have the ability to correctly transform the projection into to the right location on the newspaper once the testing environment is ideal (especially for the colour method as it is very sensitive to the environment’s lighting). Both methods also experience a lag when updating the projection which is presumably related to costly computation times of the required algorithms or a lag in the projector. Aside from this the application can still be used in a number of environments and can still offer entertainment for those using it.

PREAMBLE

A video demonstration of this dissertation and two of its methods can be found at https://youtu.be/-MIsrQoiMN8, this footage compilation displays in order:

1. The user's perspective of the assumption method.
2. Monitor capture of the assumption method, which also shows projector/camera calibration.
3. The user's perspective of the multiple colour method.
4. Monitor capture of the multiple colour method, which also shows projector/camera calibration and colour calibration.

As this is a highly visual dissertation the accompanying video best displays the aim and end product of the prototype.

TABLE OF FIGURES

TABLE OF CONTENTS

1 INTRODUCTION

This chapter will define what the main objective of this project is and its motivation and justification, as well as a discussion on what is to come in the succeeding chapters.

1.1 AIM

The aim of this dissertation is to replicate the newspaper known as the Daily Prophet from the Harry Potter films. In the Harry Potter films, paintings and photographs are displayed as moving images opposed to still images. The Daily Prophet is the main newspaper in the Harry Potter world and any photographs printed onto it appear as moving images.

Figure 1-1 Snippets of pages from the Daily Prophet [1]

This dissertation will aim to create a real-life version of this newspaper through computer vision methods and a projector. On completion there is potential that the prototype will be demonstrated as part of an exhibition within the Science Gallery at Trinity College Dublin. This demonstration will take live footage of a scene (perhaps footage of the user as they read the newspaper at a desk), and this footage will appear on the newspaper from an overhead projector. This should create the illusion of the moving image being printed on the newspaper.

An overhead camera will capture images of the newspaper from a bird's eye view as the user picks it up from a desk and reads it. Using computer vision methods these frames will be analysed to locate and track the newspaper's location and orientation within each frame. The aim of this is to determine the location and orientation of the blank space on the newspaper designated for the projection, which normally would contain an image on an ordinary newspaper. This information should then allow the program to perspective transform the footage of the reader accordingly, this footage is then to be projected down onto the newspaper in the hopes that the illusion of the footage being printed on the newspaper holds. The projector will also be placed above a desk at a 90° angle so that it projects directly down onto the newspaper and desk. This overhead camera will also capture footage of the newspaper as the projector shines onto it.

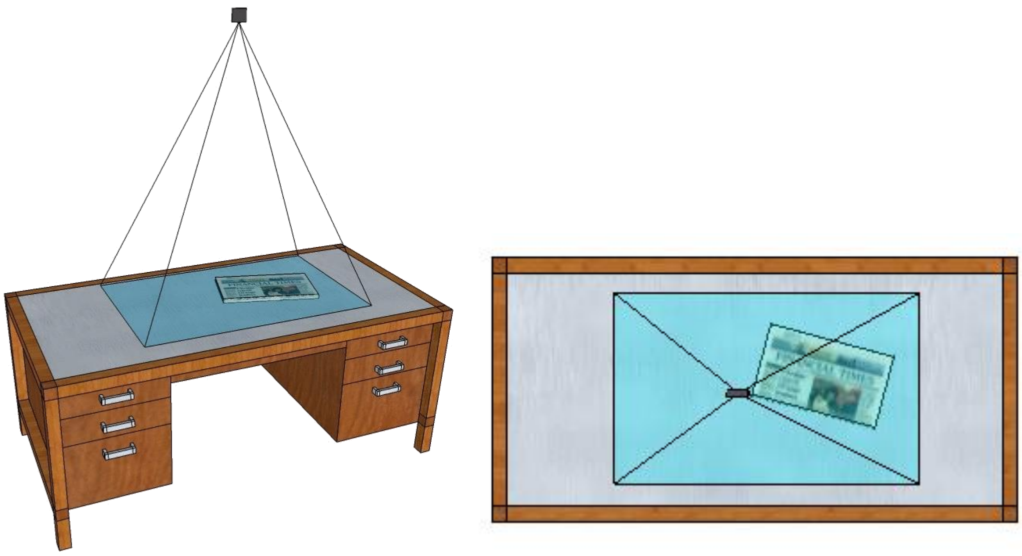


Figure 1-2 Above Left shows the isometric view of the prototype. Above right shows the bird's eye view of the prototype.

1.2 Motivation and Potential Impact

There is no definitive purpose of this dissertation however there are a few ways it can be justified. Firstly, the objective somewhat resembles smart desks (augmented desks), which popularly used computer vision and projectors to digitize physical documents. This prototype can also be justified as an interesting and fun application of computer vision or computer science and can offer entertainment for its users.

1.2.1 Smart Desks

One domain with which the dissertation has a large potential impact and has gained motivation from would be smart desks. The term smart desk can often be used very loosely and essentially can be described as an interactive or augmented work station, often designed to optimise and enhance the user's workflow. Not all smart desks will involve computer vision however many of the past have. Smart desks have been made widely redundant with the introduction of tablets to the market and therefore this research has other forms of purpose, but smart desks will not be ignored.

1.2.2 Interest in Computer Science and Entertainment

A more dominant motivation for this dissertation could be described as entertainment. This project can play the role of being a fun application of computer vision. Along with entertainment there is the hopes that it could also gain interest in the computer science and/or the computer vision field. At the beginning of this research there was potential for the dissertation to be demonstrated as part of an exhibition, 'SEEING', in the Science Gallery at Trinity College Dublin. This exhibition hopes to showcase "*different perspectives on enhanced vision, augmented ways of seeing, artificial eyes, and alternatives to vision*"[2]. This exhibition would be a good opportunity to see the response towards the final prototype from different age groups and backgrounds. It would be expected that Harry Potter fans would be highly interested with playing with the prototype, however would they be interested in the technology behind it? This exhibit could therefore potentially motivate people to get more involved with the computer science field if the project was successful enough.

1.3 Report Outline

This report will discuss related literature spreading over augmented desks and textbooks as well as other interesting applications of computer vision and projectors. This literature will vary in motivation from improving a workstation to educational aids to interesting and fun applications.

The report will then discuss the required equipment and its ideal set-up alongside the actual implemented set-up, details on this equipment will also be listed. This equipment section will also discuss the design of the adapted newspaper and the fiduciary markings design and inspiration. The calibration steps required before testing will also be discussed.

The methodology section will follow which will discuss the detailed theory behind the algorithms implemented to achieve the aim of the project, primarily mark detection, mark identification and perspective transformation. A number of methods for mark identification will be described. This section will also discuss proposed non-implemented features and the expected algorithms behind these features.

A chapter on analysis of the system will follow, this chapter discusses the end appearance of two of the fully implemented method and compares their results and features. This chapter will also discuss implementation issues encountered during development and why they occurred and how they were overcome where relevant.

The final chapter will be the conclusion which will discuss the success of the dissertation and its implemented methods and will also discuss the potential future work that could be conducted to continue the dissertation.

2 LITERATURE REVIEW

The following is a discussion of literature that related to the dissertation through similar motivation, justification and/or technical processes.

2.1 INTERACTIVE WORK STATIONS

As previously stated the work carried out within this dissertation can be adapted for use in augmented and/or interactive desks, sometimes referred to as smart desks. A wide variety of methods have been implemented to develop these smart desk prototypes, some involving computer vision, some involving projectors all similar to this proposed thesis.

2.1.1 Early Examples

Some prime examples for interactive desks were prototyped in the 1990s.

2.1.1.1 DigitalDesk

One of the earliest examples is that of DigitalDesk [3] from 1993. DigitalDesk argues the trade-offs between physical and digital documents. It aims to tackle the downsides of digital documents, mainly the lack of flexibility compared to physical documents (keeping in mind it is 1993). The writers argue that the loss of intuitiveness from physical to digital documents is one of the largest trade-offs, while they consider the main advantages of digital documents over physical ones to be spell checking, numerical calculation and keyword searching, at that point in time.

It's these differences that lead to the proposal of DigitalDesk which aims to combine the advantages of both physical and digital documents, to prevent the user from having to choose. Instead of replacing paper with digital documents the writers aim to enhance physical documents with computers. The paper aims to create this by projecting a computer display onto an existing physical desk. Cameras will also be directed at this desk to feed an image processing system (see below).

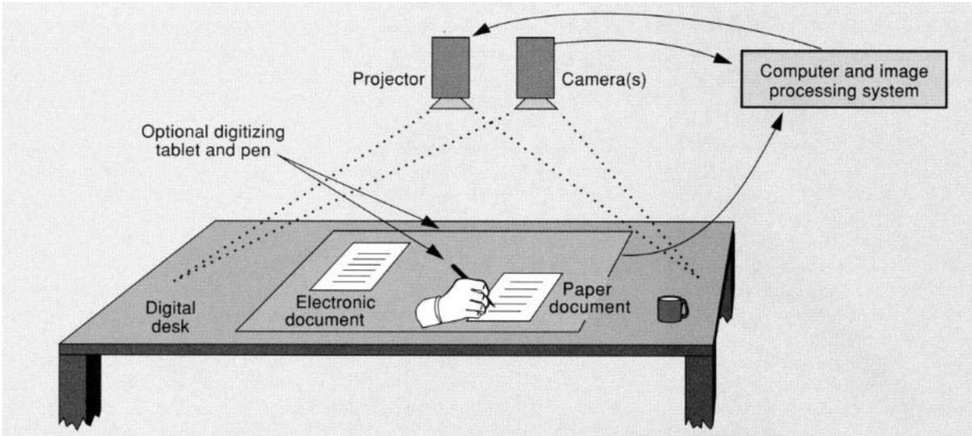


Figure 2-1 DigitalDesk's apparatus set-up [3]

One of the key relevant areas that DigitalDesk touches on is the implementation of its projected display, likewise to this dissertation, DigitalDesk plans to use projections to enhance a physical document. The writers highlight issues such as shadows (cast by the user as they lean over the document) and the affect the brightness of the room has on the projections. These issues are also relevant to this project.

Another relevant discussion of the paper is its image processing with regards to reading the sheet of paper. DigitalDesk has divided this process into three stages of image capture, image thresholding and character recognition. The main issue with image capture is the resolution of the camera, as the characters are difficult to read. One suggestion in the paper is to pre-scan pages and to use the low resolution camera when identifying pages later, however this is not suitable for scalability or interactive purposes (if the document is updated). To combat this DigitalDesk proposes a second camera which focuses on a specified area (highlighted by a projected box) where the second camera is zoomed in to a higher resolution. If the user wishes for a document to be read, they must place the document in this designated box. In the long run they state that multiple cameras could be added to cover the entire desk to eliminate this box. They also state that advances in high-definition cameras could also combat this, remembering again it's 1993.

This captured image is in grayscale and must be converted to binary so it can be used for character recognition. The writers conclude that global thresholding is unsuitable due to variances in the lighting in an office and state that adaptive thresholding would return better results however it is too slow to support user interaction. To avoid having to pass an image multiple times like with adaptive thresholding, they sacrifice results (but not as much as global thresholding) by using an estimate of the background illumination based on a moving average of local pixel intensities to calculate a threshold value at each point [4].

A final piece of relevance in DigitalDesk is the need for the system to be calibrated. DigitalDesk proposes four-point warping to overcome keystone effects of the projection (when projection is not at a right angle) [5]. The system uses cross shapes to locate key points on the display to calculate the four-point warping.

2.1.1.2 EnhancedDesk

EnhancedDesk [6] is another augmented desk which tries to combine the benefits of both physical and digital documents. Like DigitalDesk, EnhancedDesk also argues the trade-offs between paper and digital media, some of their arguments for paper have been outdated (remembering it is 1998), stating that paper offers a higher resolution and is much easier to carry.

EnhancedDesk offers a similar physical set up to that of DigitalDesk using a camera to capture footage of the desk and a projector to augment the desk. EnhancedDesk utilises a matrix code [7] for the computer to locate in the image. Corresponding information from a database can then be retrieved from this matrix code. After identification of this matrix code the computer then calculates the size and orientation of the papers so that projections can be projected in a correct location, this information is obtained from the matrix.

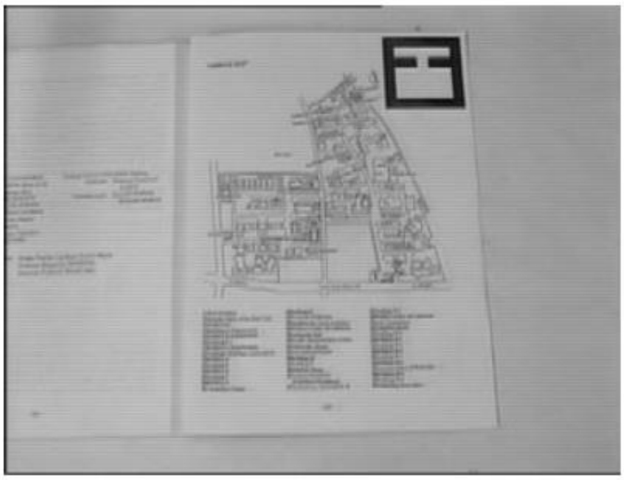


Figure 2-2 EnhancedDesk's documents are adapted with a matrix code [6]

2.1.2 Educationally Focused Examples

Smart desks and text books can be homed in to different domains, one such umbrella domain would be education.

2.1.2.1 A-book

A-book [8] is an example of an education focused project aimed at research biologists, it is a proposed interactive tool that can assist and enhance a researcher's notetaking. It is comprised of a pocket sized device which gives the user the ability to highlight and annotate and physical notes through the device and to make interlinks between notes, in the aim of digitizing the physical notes. This is not a smart desk or smart textbook but more of a tool that shares similar motives to some of the previously discussed works.

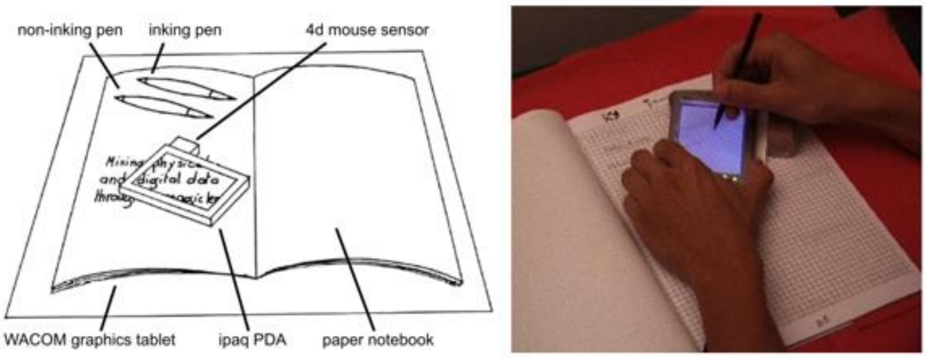


Figure 2-3 Diagram and picture of A-book prototype

2.1.2.2 SESIL

SESIL [9], is an augmented reality environment created for students to improve their learning. One feature of SESIL is the ability to recognise books and their individual pages, accomplished through computer vision. It works with a known textbook so that it has a PDF copy of the book, the system must work to recognise the page that the book is open on and find it from its database. SESIL also works to recognise and isolate the user's writing as they go, such as recognising when a sentence has been completed or a word is erased. By doing this SESIL hopes to enhance the ability to track a student's learning process and consequentially improve their learning.

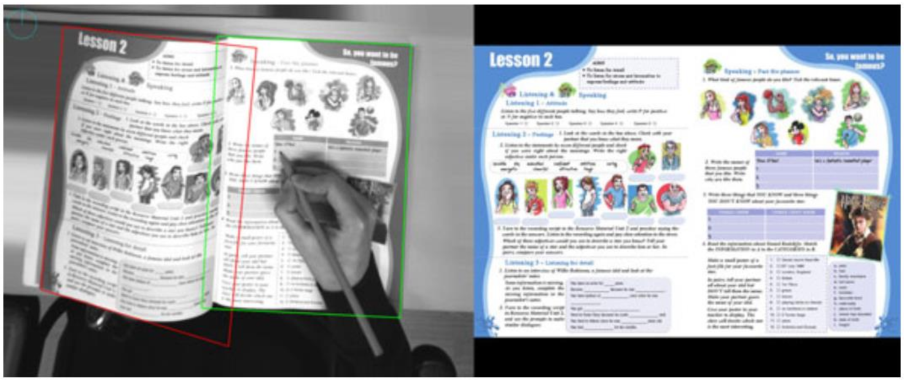


Figure 2-4 SESIL matches pages in database [9]

2.1.2.3 Augmenting Physical Books Towards Education Enhancement

A much more modern example of an interactive desk is the prototype described in "*augmenting physical books towards education enhancement*" [10]. Similar to DigitalDesk and EnhancedDesk, this prototype aims to enhance physical documents, however this prototype is also heavily motivated by enhancing education. This prototype is also comprised of a similar hardware set-up to previous examples, where the page is augmented with overhead projections.

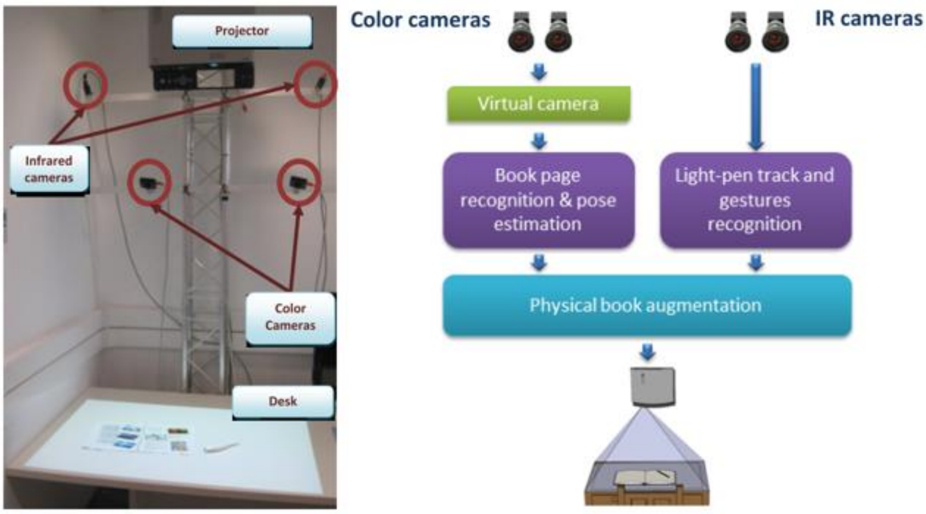


Figure 2-5 (Left) Hardware Set-Up of prototype, (Right) Prototype architecture [10]

This example makes use of IR cameras, for stylus tracking, which is often seen amongst augmented desks. IR cameras are a reoccurring feature in the discussed literature however this project has no immediate plans for its implementation. What is relevant is the user of colour cameras to recognise a document on the desk, which is more in line with the aims of this project.

This system utilises combined cameras to create a 'virtual camera', comprised of the cameras being synced together through calibration. This is to achieve a panoramic image of the desk surface. Checkerboards are placed on the desk as guides for the calibration and to align the left view with the right view as they are warped together (see below).

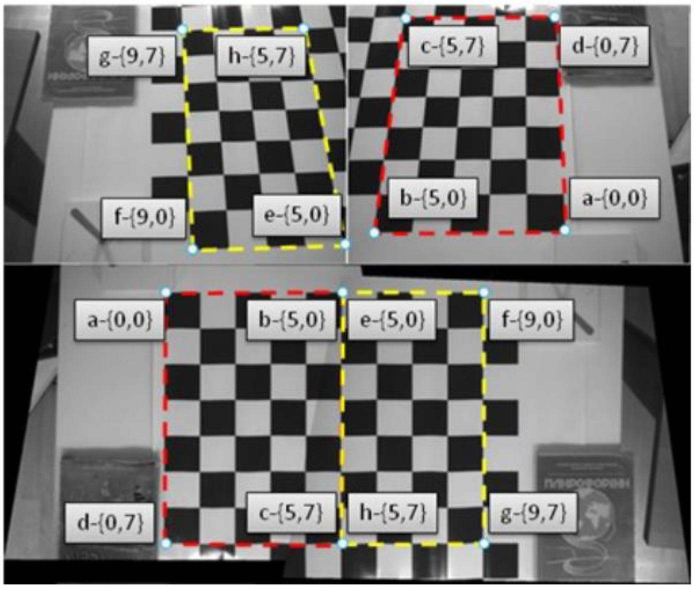


Figure 2-6 Virtual camera calibration and operation [10]

These checkerboards are used as references to physical coordinates on the desk which can now be related to virtual coordinates for image processing. The same transformations are then applied to the footage captured during runtime of the system.

Similar to EnhancedDesk, this system recognises pages and matches them to stored scans in its database. It uses the same recognition system as SESIL [9] which also provides it with the location and orientation of the book through initially detecting SIFT features [11]. This prototype uses this information to estimate the boundary of the book.

2.1.2.4 PaperLens

PaperLens [12] is an interactive table top application that allow the user to virtually interact with the 3D space above a surface such as a table. Primarily the user will be interacting with 2D and 3D information spaces. The system is aimed at researchers, for example medical researchers working in a medical lab.

The proposed system tracks a sheet of paper's navigation through these information spaces. Note below that when the paper is moved upwards and downwards on the z-axis from the table surface different information is displayed, such as the skeleton, the muscles and the nerves.

Figure 2-7 At difference heights, different information is displayed [12]

The system proposes other information spaces different to the volumetric information space displayed above. One such example is an information space where the sheet of paper initiates a zoom on an image. Note below that the projection can also be back projected to avoid a shadow being cast on the original image.

Figure 2-8 Zoom information space [12]

2.1.3 Other Examples

2.1.3.1 BendDesk

Similar to previous examples BendDesk [13] combines computer vision methods and projections. Opposed to directly tackling the differences between physical and digital documents, BendDesk focuses on combining the benefits of vertical and horizontal displays, aiming to enhance the entire workstation. BendDesk considers the main preference for a horizontal desk to be notetaking whereas a vertical display is more beneficial for reading. The user interaction between the vertical and horizontal surfaces should be as transparent and as seamless as possible for the user when they switch between both.

BendDesk combines both cameras and projectors to display images as well as detect a user's touch. Mirrors are used to deflect the projections on to specified areas. The surfaces of the desk are constructed so that when a user touches the surface a spot appears in the camera's image for it to detect a user's gesture.

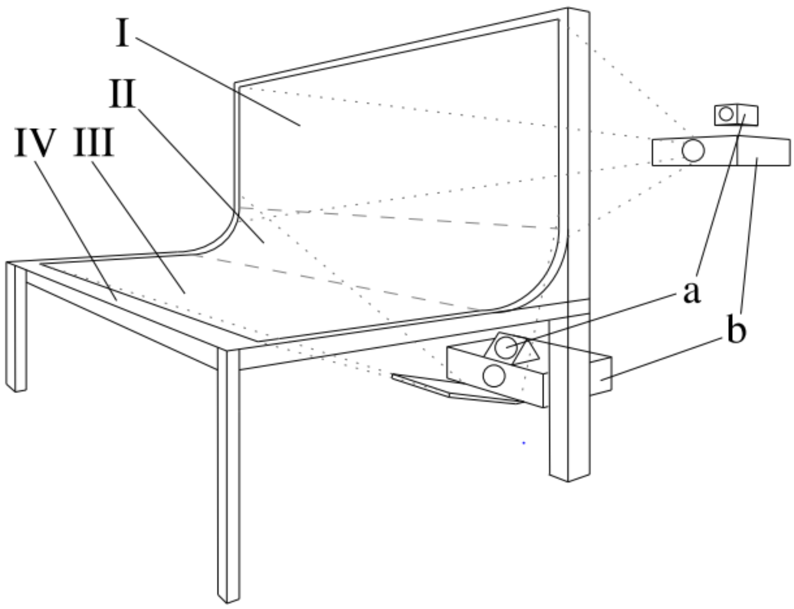


Figure 2-9 Hardware set-up of BendDesk. (I) Wall, (II) Curve, (III) Table top, (IV) non-interactive strip, (a) Infrared Cameras, (b) Projectors [13]

2.1.3.2 PaperWindows

It's clear that there's a thin line between augmented desks and augmented books/documents as a majority of the desks focus on augmenting a single page at a time. PaperWindows [14] is similar to previous works as it also considers the trade-offs between paper and digital documents, however PaperWindows differs in that its primary motivation is that physical paper is literally more flexible than digital documents (it can be physically bent and moved to form different shapes). Also similar to previous examples the writers state that physical sheets of paper are much easier to pass around and distribute in a meeting (it was 2005 and tablets were not yet on the market).

Figure 2-10 PaperWindows prototype using three pages [14]

PaperWindows uses an Infra-Red motion capturing system to track the documents, Vicon Motion Capturing System [15], which tracks Infra-Red reflective markers on the page. Similar to this dissertation PaperWindows uses an overhead projector to augment the documents, the projected image is also skewed to account for distortion, using information from the Vicon system. They found it beneficial to scale down their projections to enhance visibility as they positioned their projector 1.5m above the projection surface. Another interesting aspect of PaperWindows is that the writers found from a study [16], that users tend to position paper pages at angles between twenty-five and forty-five degrees relative to the surface. A calibration process was also required between the output of the projector and the initial position of the paper, at a few different skews.

2.2 ENTERTAINMENT

Entertainment was also discussed as motivation for this dissertation. Some other fun applications of computer vision and/or projectors are discussed below.

2.2.1 Game Applications

2.2.1.1 MagicBook

Keeping with the augmented document theme, MagicBook [17] is an interactive and augmented storybook aimed at children. MagicBook creates an interactive experience for the user and their storybook allowing them to move between a real world and a virtual world by using a handheld augmented reality display to read a physical book. The user may opt to view 3D images on their pages or immerse themselves into a 3D virtual scene.

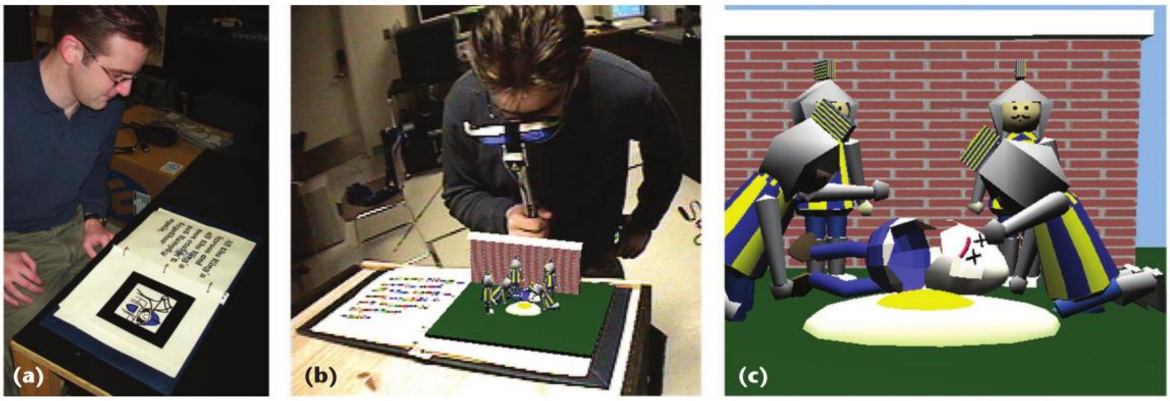


Figure 2-11 Above shows the three perspectives a user can opt to see with MagicBook [17]. (a) reality, (b) augmented reality, (c) immersive virtual reality

MagicBook uses a thick black border as a tracking mark for its computer vision based head tracking system. The black border can be seen in the above figure (a) and the handheld device with which the user watches through can be seen in (b).

2.2.1.2 *Perceptive Workbench*

The aim of the Perceptive Workbench [18] is to create a computer vision based user interface for virtual reality applications. An object placed on the system's display surface is recognized and tracked using computer vision methods. The main domain of interest is that of gesture recognition.

The user will cast a shadow over the bench and this shadow will be analysed by the vision system.

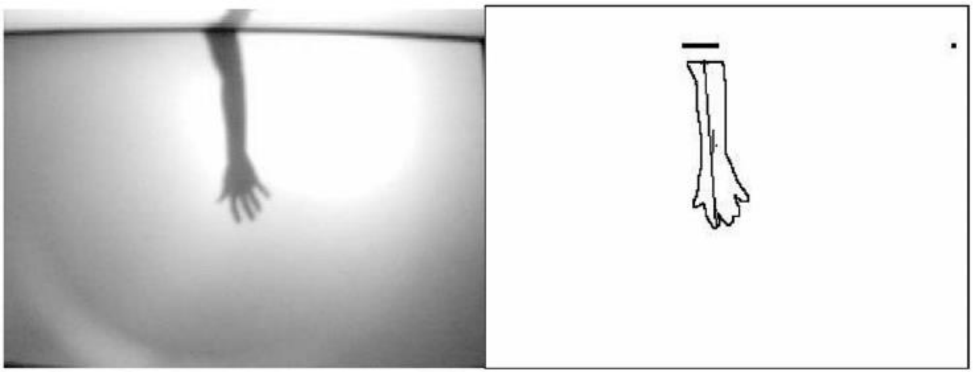


Figure 2-12 (Left) shadow cast by user's arm. (Right) Returned contour and arm direction [18]

Infra-Red lighting is used as it is invisible to the human eye so it will not interfere with the user's eyesight, hence the above shadow is captured only by the cameras as the shadow is cast from the IR lights.

2.2.2 Other Applications

2.2.2.1 *LightSpace*

LightSpace [19] identifies and tracks its users but also outputs projections to create a 2-way interactive room of inputs and outputs. The room allows users to project images onto different but flat surfaces (opposed to using a projector screen) such as desks or walls. What's fun with LightSpace is that users can then use their own bodies as tools to manipulate these projections and move them from one surface to another. If users hold hands the system will identify them as one mesh and they can then move projections over greater distances. A user simply has to place their hand on the projection and their other hand on the new surface to transform the image across. LightSpace uses depth sensing cameras to form these meshes of the users and the walls.

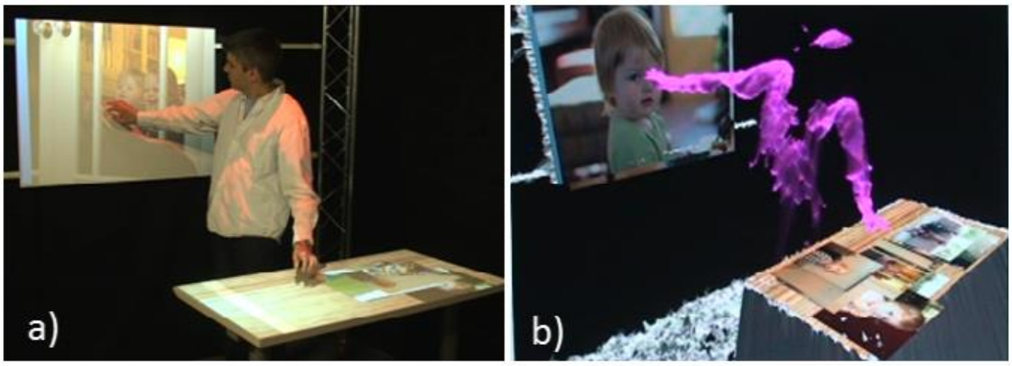


Figure 2-13 Above shows a user manipulating projections using their body as a tool with LightSpace [19]

2.2.2.2 Flexpad

Flexpad [20] is an interactive system for projecting media onto uneven sheets of paper or foam. The system combines a depth camera and a projector to achieve this.

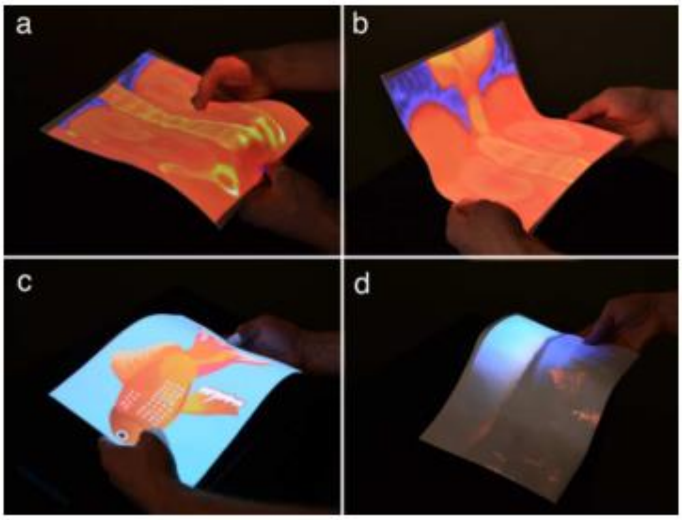


Figure 2-14 Application examples of Flexpad [20]

Above depicts numerous applications of Flexpad. (a) and (b) shows exploration of volumetric information, where curving the paper changes the displayed information. (c) shows an animation which Flexpad presents as an application for children, the above fish is animated by moving and deforming the paper. (d) shows Flexpad's application of allowing users to slice through time on a video by deforming it.

3 Equipment and Testing

The following chapter outlines the ideal set-up for the final prototype of the dissertation, the actual set-up that was implemented during development and testing. The chapter also discusses the design of the adapted newspaper and its marks.

3.1 Proposed Equipment and Set-Up

The following outlines the main pieces of apparatus used for the dissertation and their ideal physical set-up.

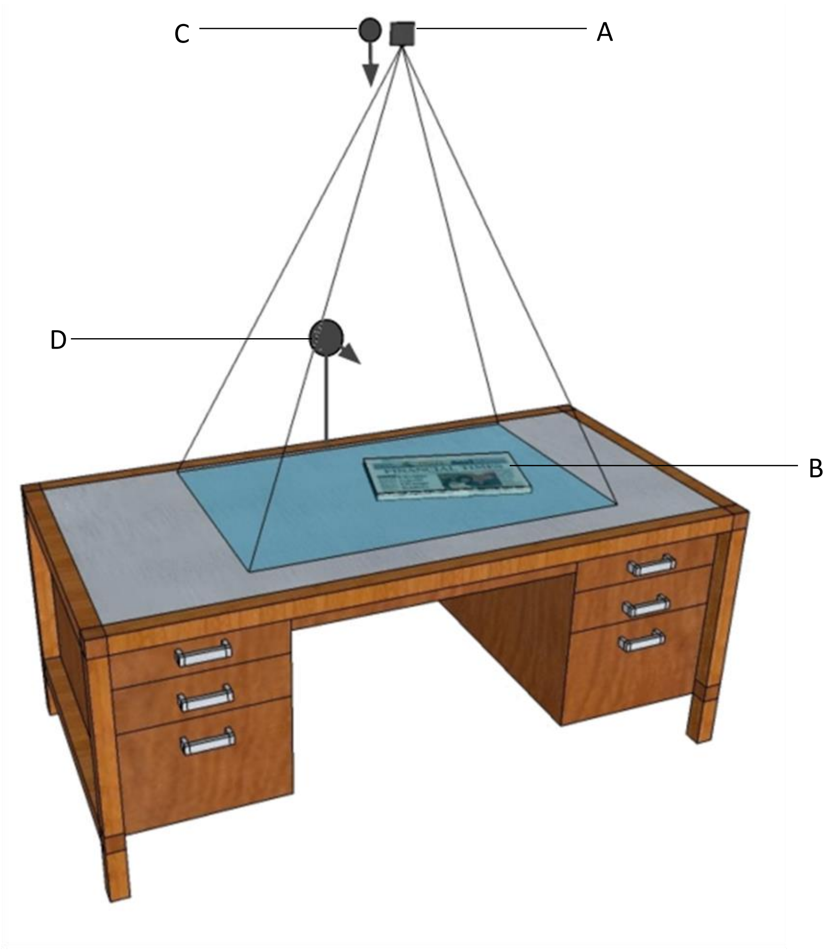


Figure 3-1 3D model of final prototype.

3.1.1 Apparatus

- A – *Overhead Projector*
- B – *Adapted Newspaper*
- C – *Input Feed*
- D – *Output Feed*

The ideal final environment for use of the prototype is depicted above. The input feed captures the footage that feeds the vision system (this is the footage that will be analysed). The input feed camera should be aligned as close as possible to the projector without interference. In an ideal situation this camera would be as high resolution as possible, without hindering the processing speed of the system. Similar can be said for the output feed, this captures the footage that will be transformed and fed out through the projector. The transformations can reduce the image resolution hence a higher resolution from the beginning may retain as much quality as possible, opposed to a mid-range resolution.

3.2 Newspaper Design

This section discusses the design of the adapted newspaper that was used in the final implementation of the system and why the design was chosen.

3.2.1 The Need for Marks

One of the first steps required for the system is its ability to locate the newspaper in a frame. A number of methods were brainstormed as to how to attempt this. The stronger ideas were to use the font and stylings of the newspaper (as this would maintain user transparency) as an identifier or to disguise fiduciary markings within a border around the content of the newspaper. The plan was decided to use fiduciary markings which could later be disguised into a border. It was decided that using the text would not be feasible as the text cannot be seen well enough from a distance, due to the camera resolution.

3.2.2 QR Code Design

The final design for the markings was adapted from finder patterns commonly found on QR codes. QR codes contain three finder patterns located in the top left corner, the top right corner and the bottom left corner. In the bottom right corner sits an alignment pattern which is later used to determine the orientation of the QR code.

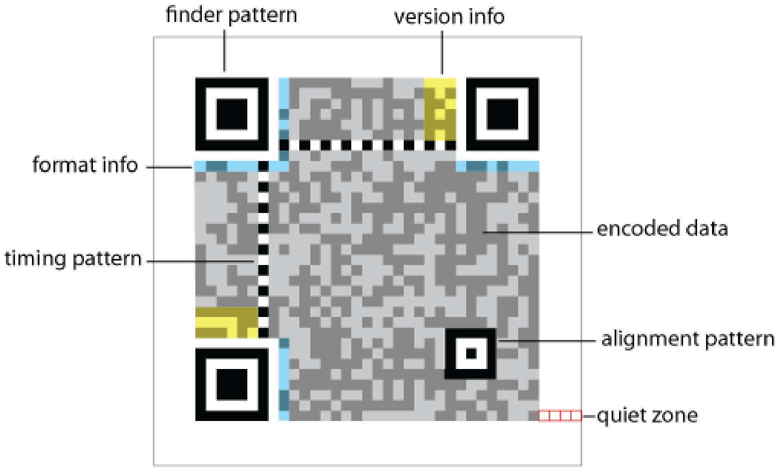


Figure 3-2 Schematic of a QR code depicting vital components [21]

Note the design of the above finder patterns and how they are essentially a series of concentric monochromatic squares. The finder patterns are the components of this above design that are of great interest to this dissertation. It is also important to note that the squares are also designed with specific ratios to one another.

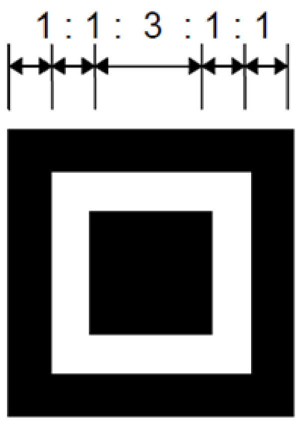


Figure 3-3 QR Code Finder Pattern Ratios [22]

The need for these ratios will become more evident when mark detection is discussed in 4.1. The incorporation of similar ratios into the final design is required.

3.2.3 The Final Mark Design

The previously described pattern above was adapted, using the same ratios, into a circular pattern. The reason for a circular design was for the instances where the newspaper is tilted at severe angles to the input feed. The circles are less abstracted than the squares at these angles.

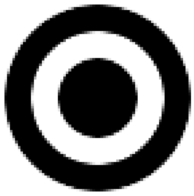


Figure 3-4 Finalised Monochromatic Mark

3.2.4 Coloured Marks

The system must be able to determine which mark is which (which is top left, which is bottom right for example) for when the output image is transformed. The process of how these marks are identified will be further discussed in a later chapter, 4.2, one of the potential methods for this uses colours.

The system will want to identify the average hue of the mark (or of the inner circle of the mark). Each mark must have a different hue value to decipher between them. These marks would then be adapted from the one shown previously.

Figure 3-5 Adapted Red Mark

The above is just an example of one colour. For the identification stage more varying colours will need to be produced. Care is to be taken when picking colours for the marks as the hue values need to be distinctive enough from one another to be identified by the system.

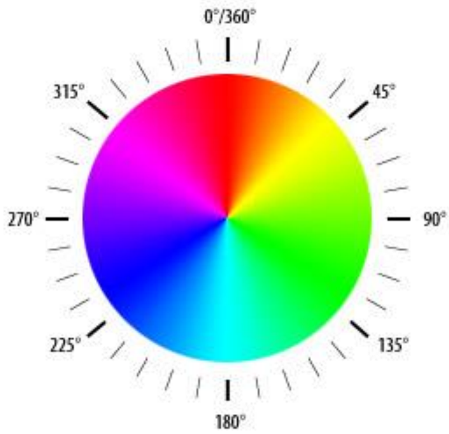


Figure 3-6 Hue Colour Wheel [23]

Note the following four marks that were implemented on the newspaper have equidistant hue values according to the above colour wheel.

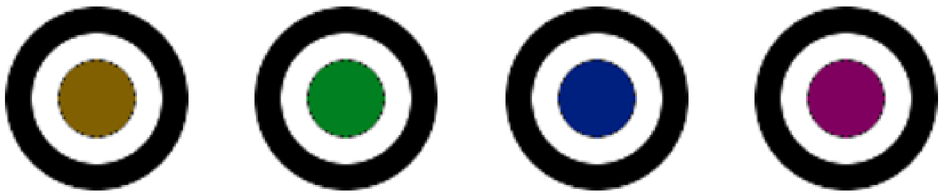


Figure 3-7 Marks with respective hue values of 45, 135, 225 & 315

Note that the above examples are all also of dark shades. This is highly necessary as if the colour in the centre is not dark enough the system won't be able to detect the mark when applying a binary threshold. This binary thresholding will also be further discussed in a later chapter, 4.1.3.

3.2.5 Mark Placement on the Newspaper

The marks are going to be used as the four points required for a perspective transformation later on. The footage that is being transformed is by default 480x640 and to keep default the marks should be placed in the same ratio of roughly 0.75:1. This also avoids further distortion due to different ratios, and eliminates a step that would be required to fix this distortion.

Science Gallery News

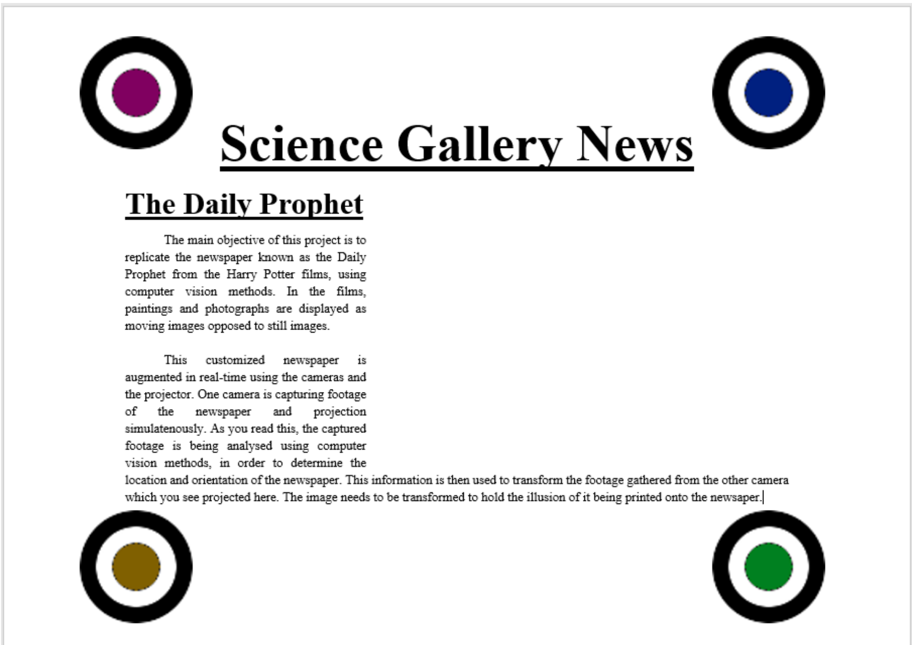
The Daily Prophet

The main objective of this project is to replicate the newspaper known as the Daily Prophet from the Harry Potter films, using computer vision methods. In the films, paintings and photographs are displayed as moving images opposed to still images.

This customized newspaper is augmented in real-time using the cameras and the projector. One camera is capturing footage of the newspaper and projection simulatenously. As you read this, the captured footage is being analysed using computer vision methods, in order to determine the location and orientation of the newspaper. This information is then used to transform the footage gathered from the other camera which you see projected here. The image needs to be transformed to hold the illusion of it being printed onto the newsaper.

Figure 3-8 Newspaper adapted with marks

A different adapted newspaper will be required for a different identification method. These identification methods will be further discussed in 4.2, the proposal of using colours to identify individual marks has been briefly mentioned within this chapter.



Science Gallery News

The Daily Prophet

The main objective of this project is to replicate the newspaper known as the Daily Prophet from the Harry Potter films, using computer vision methods. In the films, paintings and photographs are displayed as moving images opposed to still images.

This customized newspaper is augmented in real-time using the cameras and the projector. One camera is capturing footage of the newspaper and projection simulatenously. As you read this, the captured footage is being analysed using computer vision methods, in order to determine the location and orientation of the newspaper. This information is then used to transform the footage gathered from the other camera which you see projected here. The image needs to be transformed to hold the illusion of it being printed onto the newsaper.

Figure 3-9 Newspaper with adapted coloured marks

The markings are arranged in the same ratios as the previous adapted newspaper of 640x480. Note how each mark is a different colour as each colour will be representable of a location such as top left, top right, bottom left or bottom right.

3.3 Testing Environment

The following is a description of the actual equipment setup for development of the dissertation.

Due to limitations on space and other reasons, the projector and input camera were never set up overhead. Instead for development they faced a blank wall, rotated at 90^{o} angles from their proposed orientations. The newspaper would be held against the wall, to simulate a desk surface. All software should work the same if the apparatus was rotated to the original proposed set-up.

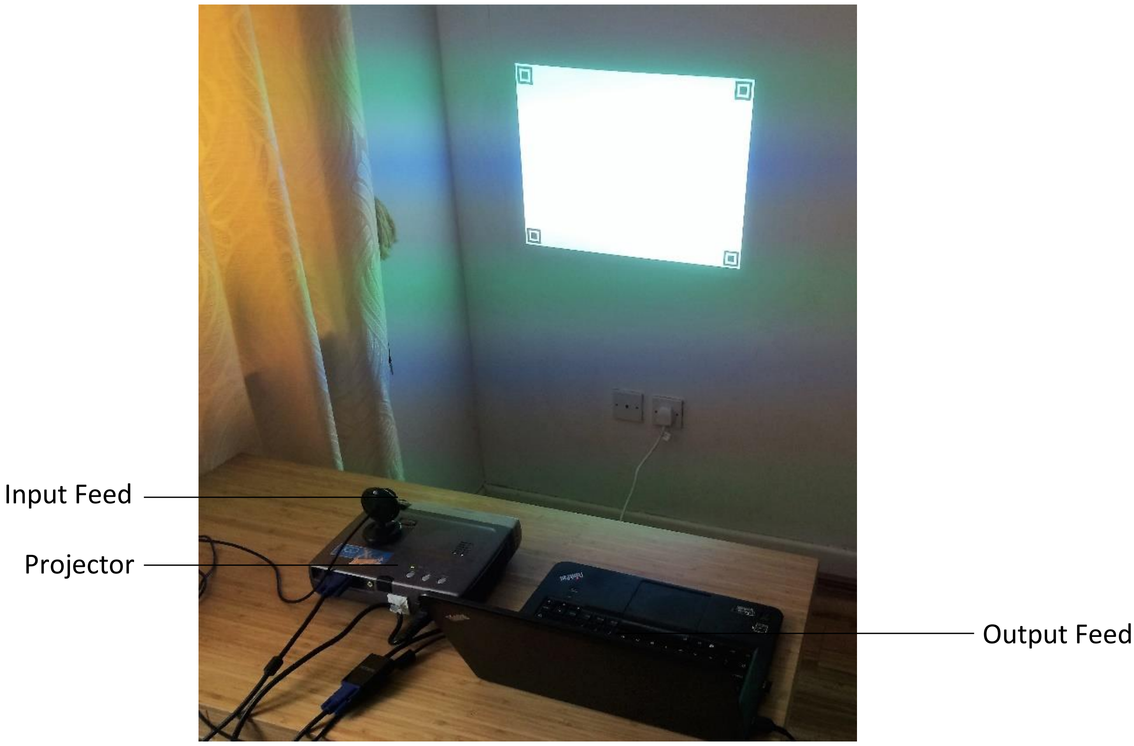


Figure 3-10 Development Environment during calibration

3.3.1 Actual Equipment

The following was the actual equipment used for development and testing:

- Input Camera: 1.3 Megapixel (1280 x 1024) Resolution
- Output Camera: 720p Low Light Resolution
- Projector: 1028x768 Resolution
- The system was built on the Visual Studio 2013 platform with the OpenCV 3.0 library
- The newspaper was printed on A4 paper

3.4 CALIBRATION

Two calibrations can be required (depending on which mark identification process is used) before the system may be tested or used. The following will outline the processes undertaken for these calibrations and why they are required.

3.4.1 Camera/Projector Calibration

This calibration process is always required. It must be ensured that the input feed can see the projection before searching for the newspaper. The projector will output the below image:

Figure 3-11 Camera/Projector Calibration Image

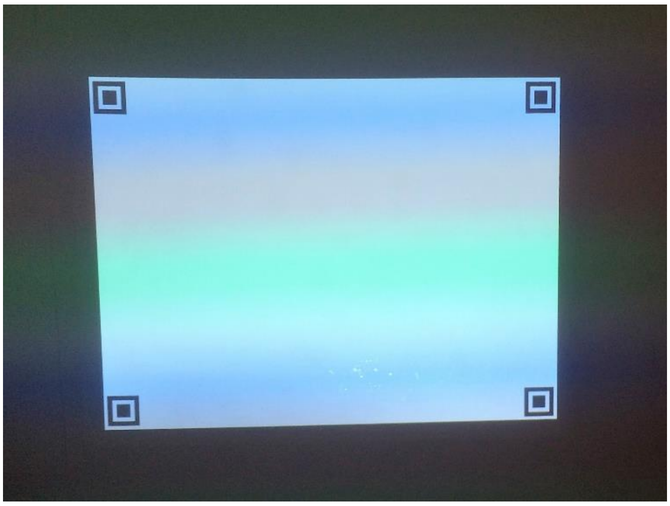


Figure 3-12 Calibration projection on the wall

This calibration image is purposely similar to the previous newspaper designs. It too is 480x640 and the squares are of the same design as the finder patterns found in QR codes. If the system is able to find four marks in a single frame (mark detection will be discussed later in 4.1) then it will identify which mark is the top left, top right, bottom right and bottom left using the assumption method (this method is described in further detail in 4.2.1). In short the assumption method assumes that which mark is top left in the image is the top left mark on the newspaper, as the projector and the camera are always aligned as close as possible the limitations of this method do not count it can always be safely assumed that top left is top left.

These marks are important as the system will need to know the location of the projection's boundaries for when it performs the perspective transformations, discussed later in 4.3. The system needs to know where these boundaries sit within the input feed.

3.4.2 Colour Calibration

Colour calibration is only required for the identification processes that use colour. The colour identification process will be discussed further in 4.2.2, but in short each mark is identified by the unique hue of its centre. Theoretically the centre of each circle has a defined hue value as seen in 3.2.4, but in practice these hue values deviate from their theoretical values. It is because of this deviation that calibration is required, as a range of minimum hue and maximum hue needs to be established for each colour. If a mark's hue sits within this range it may then be classified as this colour.

Figure 3-13 Colour calibration cards

This process requires user cooperation where the system requests to be shown a colour, the user may use the colour calibration cards seen in the figure above. The system will take 100 samples of the colour to find its average hue in the current environment. It will then use this value as a midpoint for the range. A predefined value such as 10 for example will be used as the displacement from the mean for the maximum and minimum possible hues.

For example, if colour A obtained an average hue of 70 during calibration, then this colour's range would be 60-70, if an offset of 10 was used. Then in the identification stage if a mark obtained a hue within said range it would be classified as colour A.

The process continues for the remaining colours, where the system asks the user to show it a specific colour each time. The system is reliant on the user here, for example, if the system asks to see pink the user must show it the pink mark as the system is designed to associate pink with a specific corner of the newspaper.

4 METHODOLOGY

This chapter details the theory behind how the newspaper was found and identified within a frame, it outlines multiple identification methods. The chapter also describes how the system transforms the projection. Details of theory behind future work is also included.

4.1 MARK DETECTION

This section will outline how a mark is located within a single frame and how it is confirmed to be a mark.

4.1.1 Mark Detection Summary

i. *Convert RGB image to grayscale image*
ii. *Convert grayscale image to binary threshold image*
iii. *Find connected components within the binary image*
iv. *Discard found components that are too small or too large*
v. *Sort components into ascending order by area*
vi. *Find pairs (components that wholly sit inside a larger component)*
vii. *Discard pairs if the ratio is unreasonable*
viii. *Identify a centre point*

4.1.2 Obtaining a Grayscale Image

The first key component of locating marks within a frame is to obtain a binary image of that frame, before this step is taken the image must be converted into grayscale. A grayscale image is much less complex than an RGB image as each pixel only contains a single value, which is required for the binary thresholding stage.

The input footage is captured in the RGB colour space and therefore the system must convert it to grayscale. This is accomplished by the following equation [24]:

$$Y = 0.299R + 0.587G + 0.11$$

Y is the output value for a single pixel in the grayscale image. *R* is the red value of a pixel in the RGB image and *G* is the green value of a pixel in the RGB image.

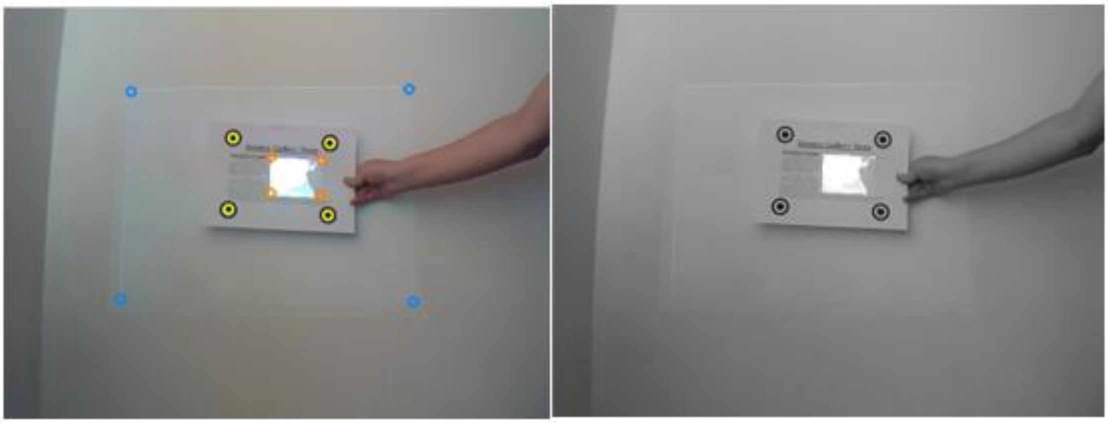


Figure 4-1 RGB to grayscale conversion (the markings on the left hand side RGB image are for tracking purposes)

4.1.3 Obtaining a Binary Image

A grayscale image contains a single value per pixel, to make this image binary (all pixels purely white or black) a threshold will be required. The side of the threshold that the pixel's value sits will determine if its new value will be 255 or 0.

Identifying an appropriate threshold can be difficult in this situation as the development area is not consistent with its lighting. This is why a single predefined threshold value is not acceptable, other algorithms are more appropriate where the system can dynamically calculate a threshold.

4.1.3.1 Adaptive Threshold

Adaptive thresholding is an appropriate method when the light distribution in the image is highly uneven, or if this is a large shadow cast on a portion of the image.

Adaptive thresholding [24] divides the image up into sub-images. The current pixel is compared to the local average of the pixels in the corresponding sub-image (the sub-image that the current pixel resides in). If this pixel's value is greater than or less than this average its binary value will be 255 or 0 respectively.

For this application adaptive thresholding can be costly in terms of processing time and the system requires real-time reactions, alternative thresholding options were explored.

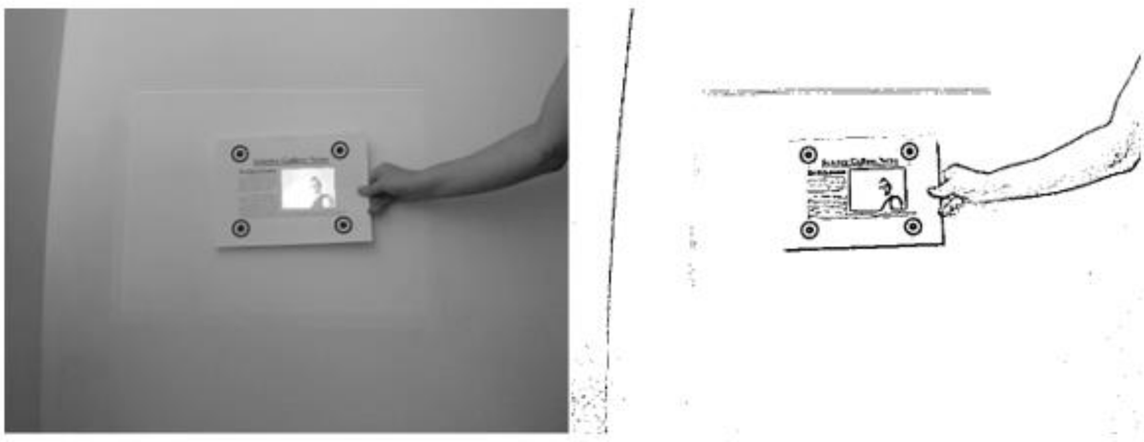


Figure 4-2 Grayscale to binary using adaptive thresholding

4.1.3.2 Otsu Threshold

Otsu Thresholding is an algorithm [25] used to automatically select an optimal global threshold (a single threshold value for the whole image) for an image. All possible thresholds will be considered, and for each threshold the algorithm will calculate the spread of pixels either side of this threshold. The threshold value with the lowest sum of foreground and background spreads will be selected.

First the foreground and background variances (measure of spread) must be calculated for the current threshold. Where $\mu_f(T)$ and $\mu_b(T)$ are the means of the foreground and background greyscale values respectively, $w_f(T)$ and $w_b(T)$ are the portions of points which are foreground or background respectively, and $\sigma_f^2(T)$ and $\sigma_b^2(T)$ are the variances of the foreground and background grayscale values respectively:

$$\mu_f(T) = \sum_{g=T}^{255} p(g) \, . \, g \Big/ w_f(T) \qquad w_f(T) = \sum_{g=T}^{255} p(g)$$

$$\sigma_f^2(T) = \sum_{g=T}^{255} p(g) . (g - \mu_f(T))^2 \Big/ w_f(T)$$

$$\mu_b(T) = \sum_{g=0}^{T-1} p(g) \cdot g \Big/ w_b(T) \qquad w_b(T) = \sum_{g=0}^{T-1} p(g)$$

$$\sigma_b^2(T) = \sum_{g=0}^{T-1} p(g).(g - \mu_b(T))^2 \Big/ w_b(T)$$

Then the within class variance $\sigma_W^2(T)$ can be calculated, the threshold that returns the smallest within class variance is the desired threshold:

$$\sigma_W^2 = w_f(T)\sigma_f^2(T) + w_b(T)\sigma_b^2(T)$$

The above can be computationally expensive, the threshold with the smallest within class variance is also the threshold with the largest between class variance. The between class variance $\sigma_B^2(T)$ is easier to compute:

$$\sigma_B^2(T) = w_f(T)w_b(T)(\mu_f(T) - \mu_b(T))^2$$

The above equations were obtained from [24].

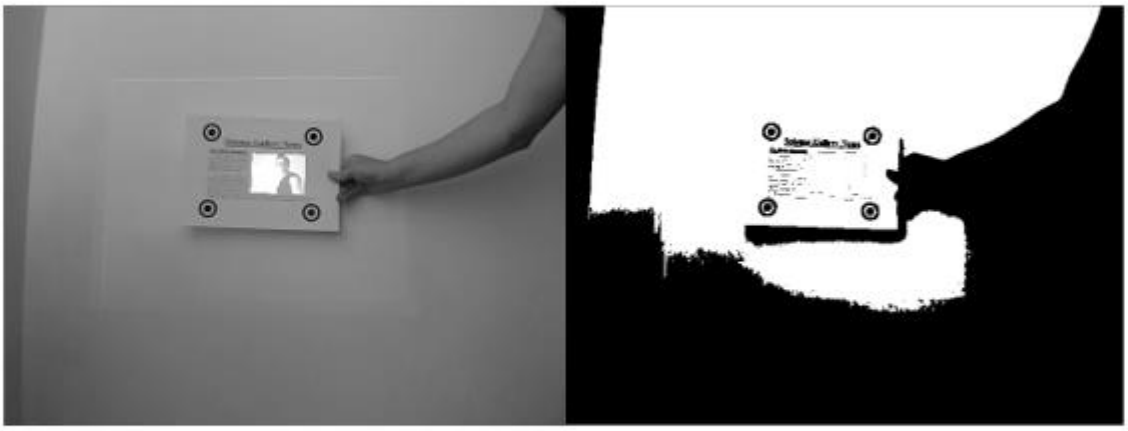


Figure 4-3 Grayscale to binary using Otsu thresholding

4.1.4 Finding Connected Components

Connected component analysis is vital for identifying the marks in a frame, as seen from the design of the marks in the previous chapter it would be expected that the marks should not be distorted by binary thresholding.

The aim of the connected components in this case is to locate the circle and surrounding ring as components. These can be identified by identifying the borders between foreground and background pixels after the image has been made binary.

The algorithm [26] works to follow the boundaries between foreground and background pixels identifying contours of components. When a foreground pixel (a non-zero value pixel) has been identified, it will be given a new label. Now the algorithm will inspect the 8 neighbourhood pixels of the current pixel looking for another foreground pixel. Any found neighbouring foreground pixels will be assigned the same label as their parent. This process continues for all neighbouring pixels until there are no more unlabelled neighbouring foreground pixels. This process repeats for the next component, assigning the next foreground pixel with a different label (note this will be a unlabelled foreground pixel that has no labelled neighbours).

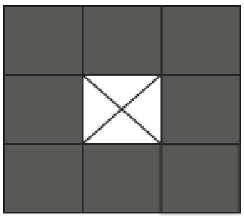


Figure 4-4 8-adjacency, the shaded pixels are the neighbourhood of the centre pixel [24]

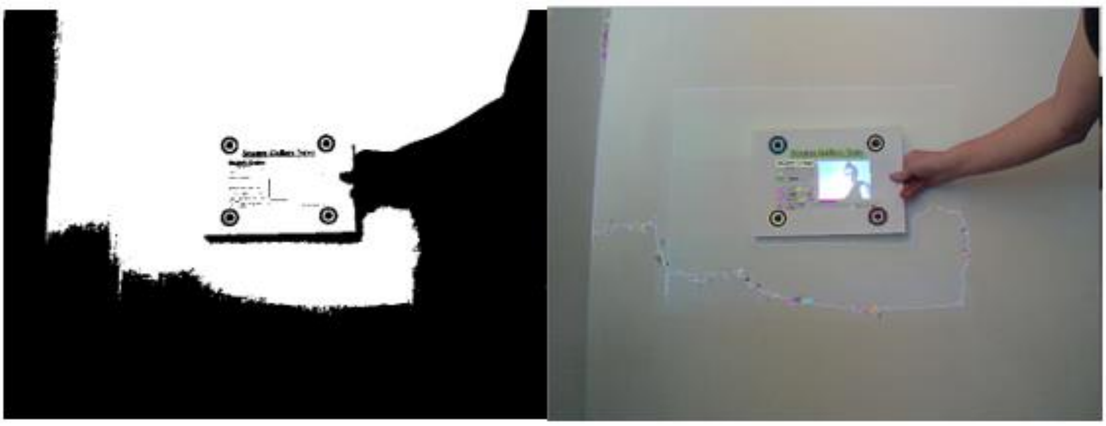


Figure 4-5 Contours of components found from binary image

4.1.5 Identifying Components as Marks

To identify the marks, the system must find a component or contour that wholly sits within another component or contour. These two components must also not have an unreasonable ratio to one another, for example the boundary of the newspaper and the boundary mark would not be identified as a mark as their ratio is extreme.

4.1.5.1 Arranging the Components

The area of all components is to be calculated by summing their pixels. To do some quick elimination of the contours, anything that has an area less than 8 pixels or greater than a fifth of the image can be discarded. This is to reduce the number of contours to compare and also because nothing worthwhile can be done with the marks if they are beyond those thresholds.

The order of the contours within their array must be sorted ascendingly by area [27]. This rearrangement leaves the smaller areas at the beginning of the array, increasing in size towards the end.

These steps ease the process of checking if components are concentric pairs.

4.1.5.2 Pairing Components

A pair is deemed a pair if every pixel of one component is inside the border of the other component. The previously taken steps are beneficial in improving the efficiency of this process.

The smallest component (component with the smallest area / component at the beginning of the array) is taken, this will be called component A. The next component in the array is then taken, this will be called component B.

A check is then done to determine if each and every pixel of A sits within the boundary contour of B [28]. This is done pixel by pixel, if a pixel dissatisfies the check the system will replace B with the next contour within the array. This process continues until there are no more available contours, this is when A will be replaced with the next ascending available contour.

If it has been determined that every pixel of A is within B, then these contours are stored as a pair. This pair is also sorted ascendingly, similar to before. It is also noted using flags that these two contours may not be checked to be parts of another pair.

The system then continues by taking the next smallest contour that is available (the next smallest contour that has not been arranged into a pair) as A. B is subsequently the next ascending contour that is available (also not arranged into a pair) from this contour A.

The final step of confirmation is to eliminate pairs that are unreasonable ratios to one another and therefore cannot be deemed a mark. If A is the same size as or greater than B, the pair is discarded. If B is 10 times larger than A, the pair is discarded. The pairs are also checked to ensure that there are two and only two contours in a pair before the ratios are checked. These ratios can be changed to be stricter if needs be as the marks have been designed to have specific ratios to one another.

The marks have now been confirmed as marks. The system still does not know which mark is which it only knows that it has four marks for example. Further processing must be completed to identify which mark is which.

4.1.6 Obtaining a Centre Point

Now that the system has determined the components to be marks due to their coupling, it must work to identify a centre point to note where the mark sits within the image. To achieve this the system will enclose the larger component in the smallest possible circle.

A component is stored as an array of points. The radius of the minimum enclosing circle will be half of the maximum distance between two points, and the centre point will be the midpoint of this distance. An iterative process will be used to find this maximum distance.

4.2 Mark Identification

This section describes multiple methods available for determining the role of a found mark, for example by identifying if the found mark is the top left mark on the newspaper. This is necessary for correctly transforming the image in a later stage.

4.2.1 Assumption Method

This is the most straightforward method however it subsequently has the most limitations. The assumption method does not depend on the colouring of the marks. It does require four marks and also requires the newspaper to be more or less in an upright position.

This method works off the assumption that the top left mark will also be top left in the frame, similarly for the remaining three marks.

If four marks are detected within a single frame, the system will initially search for the top left mark. It identifies this mark by finding the mark with the smallest sum of x and y coordinates. The axis' on the frame begin in the upper left coordinate at (0,0). This is why the mark with the small x + y is deemed as the top left mark.

A similar approach is adopted to identify which mark is the bottom right mark, instead of searching for the smallest value the largest x + y is what identifies this mark as the bottom right.

To identify distinction between the remainder marks, the x coordinate alone is looked at. The mark with the lowest value x coordinate is taken to be the lower left mark.

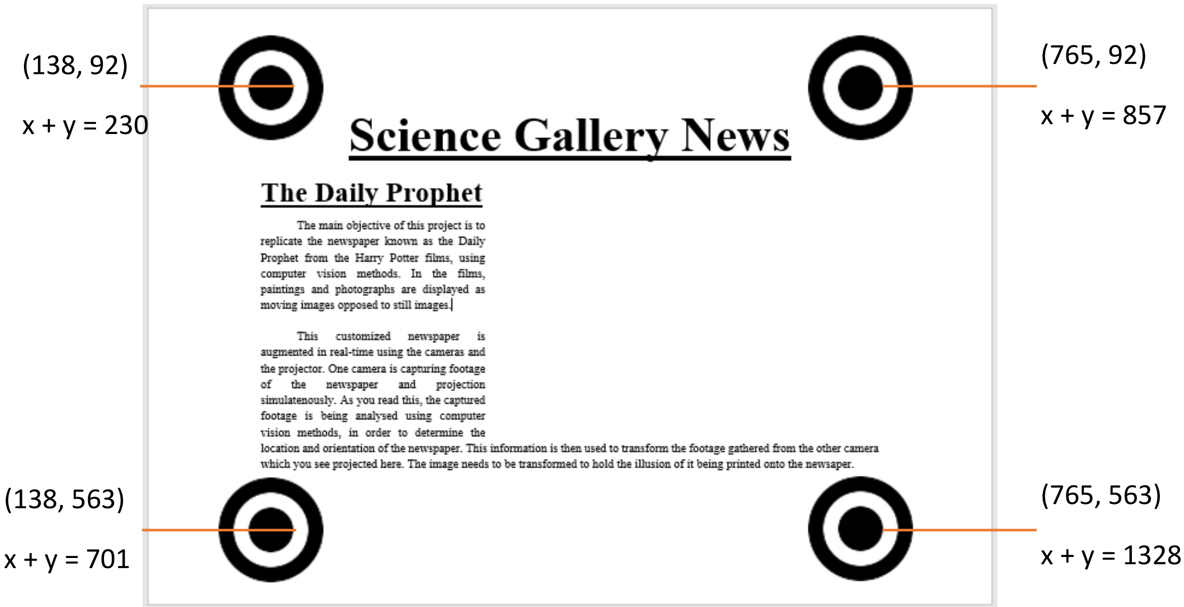


Figure 4-6 Sample coordinates and calculations for assumption method

4.2.2 Colour Method

The assumption method has its obvious limitations where the newspaper cannot be rotated beyond a limit. The colour method was developed to overcome these limitations by using colour to identify the marks. The use of colour to identify the marks has been briefly mentioned when discussing design of the Newspaper in a previous chapter.

Each colour can be associated with a mark. If the system can identify a colour it can then find this mark's associative role, if it is the top left mark for example. By identifying the marks with colour like this, the assumption method's rotation limitation is eliminated, the newspaper can be fully rotated upside down. If the top left mark is red the system will always know the red mark is the top left mark, irrelevant of where it is found within the frame, or its coordinates.

4.2.2.1 Identifying the Average Hue

Each mark differs by a differently coloured centre. These centres are created with equidistant hue values from one another so they can be distinct as previously discussed in a previous chapter. It is highly important that the centre of the marks is dark enough to be identified as a background pixel when binary thresholding is applied.

The same process follows for detecting the marks as previously stated. The system currently has all the pairs stored in ascending order, hence the system knows that the first component in a pair is the contour of the inner circle (the coloured circle in this case). By finding the minimal bounding rectangle [29] around this contour the system can easily calculate average values of the pixels.

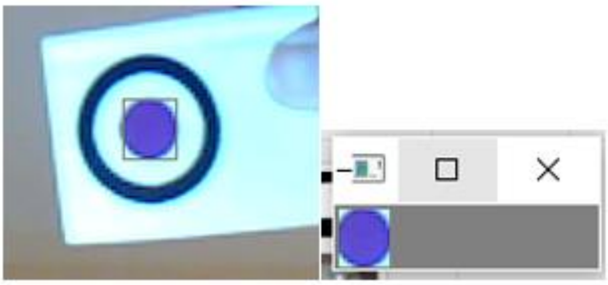


Figure 4-7 Bounding box around inner circle of mark

Note that the bounding box was cropped out of the original image and stored on its own, this makes it easier to count the average values of the pixels in the image. Also note that the bounding box has been draw on the original image.

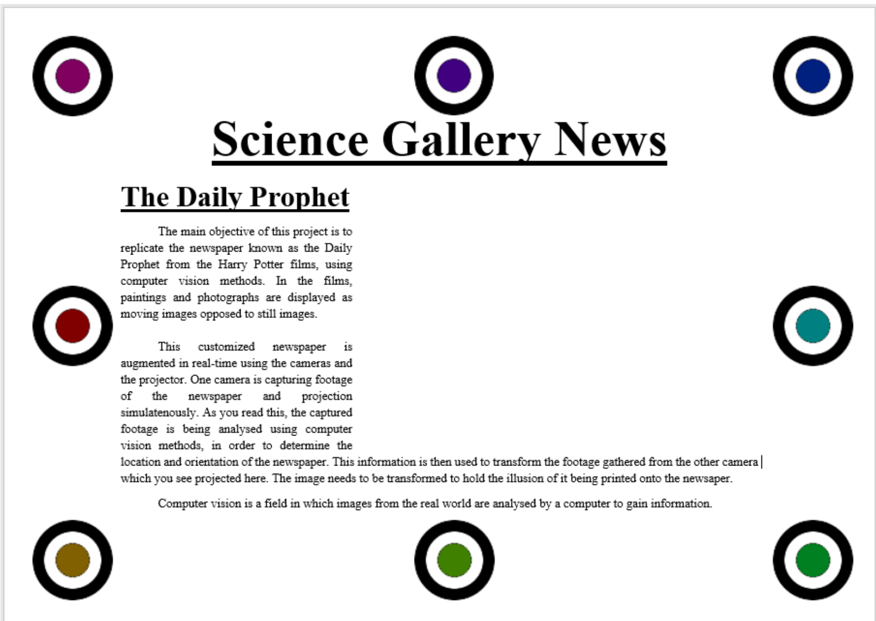
The image is currently in the RGB colour space. The system is attempting to use the average hue value of the image to classify it, hence this image needs to be converted to a different colour space. The image can be converted to the HLS colour space [30]. This is the hue, luminance and saturation space. The system can find the average hue of the cropped image by summing the hue value of each pixel in the image and dividing it by the total number of pixels in the image.

This average huge value can then be checked against thresholds (defined in the colour calibration stage previously discussed). If the value sits within the minimum and maximum thresholds for a colour, then this mark can be identified as that colour and the colour's coordinates are updated accordingly.

Limitations can be seen from looking at the cropped image alone and that it has some white pixels on the corner, these can distort the average hue of the centre of the circle and can bring all the separate hue values closer together. If more time was available finding the average values within a bounding circle might assist in eliminating this.

4.2.2.2 Using 8 Marks

A large limitation to the assumption method is that it requires all four marks to be found within a single frame. A large piece of motivation for using colours was to allow for more than four marks on the newspaper. This way, estimates could be made if the four corners were not found in a single frame, by extrapolating information from whatever combination of marks was available.



Science Gallery News

The Daily Prophet

The main objective of this project is to replicate the newspaper known as the Daily Prophet from the Harry Potter films, using computer vision methods. In the films, paintings and photographs are displayed as moving images opposed to still images.

This customized newspaper is augmented in real-time using the cameras and the projector. One camera is capturing footage of the newspaper and projection simulatenously. As you read this, the captured footage is being analysed using computer vision methods, in order to determine the location and orientation of the newspaper. This information is then used to transform the footage gathered from the other camera which you see projected here. The image needs to be transformed to hold the illusion of it being printed onto the newsaper.

Computer vision is a field in which images from the real world are analysed by a computer to gain information.

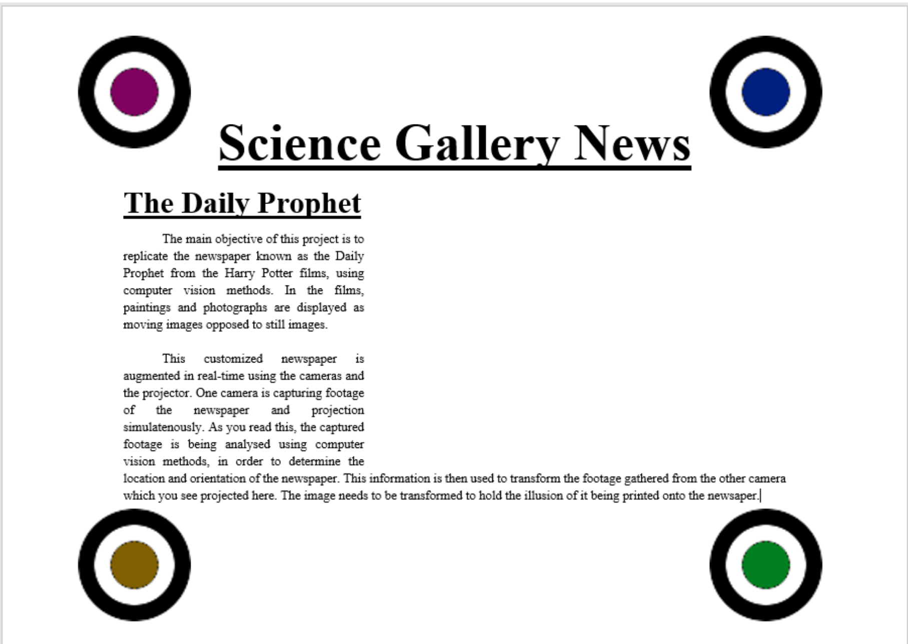
Figure 4-8 Adapted newspaper with x8 coloured marks

The above newspaper was developed. The hue values of the marks are all equidistant from one another (spread evenly on the hue colour wheel seen in 3.2.4) so that they have the best chance at being identified from one another.

This process was not further developed when a limitation was found where the coloured marks were not distinctive enough, due to a lack of time the project had to move on. A solution to this was to reduce the number of coloured marks.

4.2.2.3 *Using 4 Marks*

As previously stated by using 8 marks, the hue values were too close together to be distinct enough for the system to recognise. This resulted in reverting back to using 4 marks with greater differences between their hue values.



Science Gallery News

The Daily Prophet

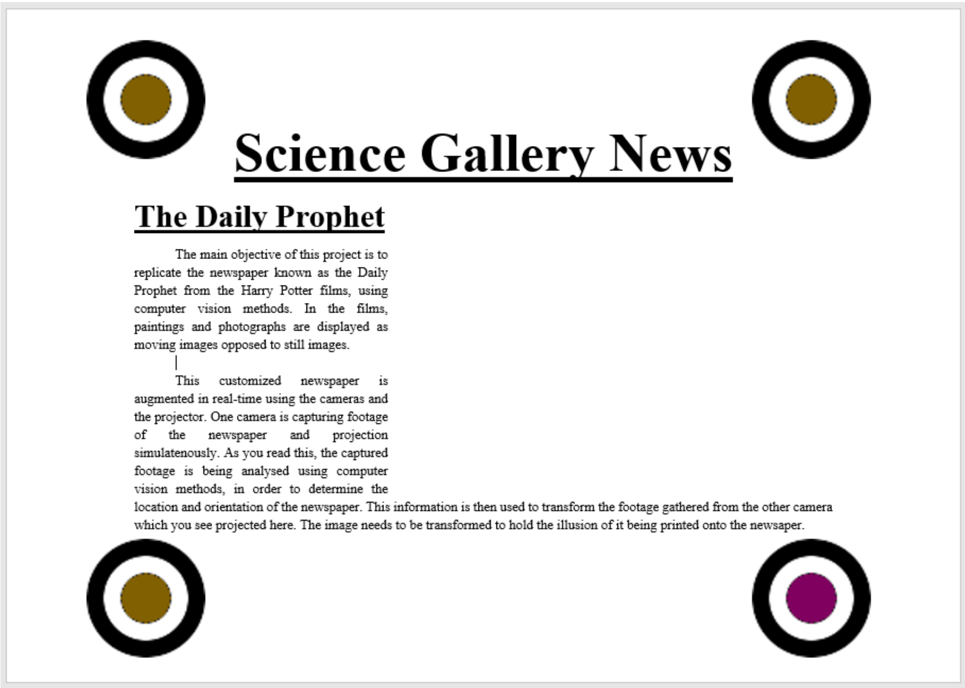
The main objective of this project is to replicate the newspaper known as the Daily Prophet from the Harry Potter films, using computer vision methods. In the films, paintings and photographs are displayed as moving images opposed to still images.

This customized newspaper is augmented in real-time using the cameras and the projector. One camera is capturing footage of the newspaper and projection simulatenously. As you read this, the captured footage is being analysed using computer vision methods, in order to determine the location and orientation of the newspaper. This information is then used to transform the footage gathered from the other camera which you see projected here. The image needs to be transformed to hold the illusion of it being printed onto the newsaper.

Figure 4-9 Adapted newspaper with x4 coloured marks

4.2.3 Two Colour Method

A large limitation with the multiple colour method is the system's difficulty in certain lighting to identify between the colours, this was seen when 8 marks were attempted. Reducing to 4 different marks still had similar limitations. This method works to use just two different colours which sit on opposite sides of the hue colour wheel.



Science Gallery News

The Daily Prophet

The main objective of this project is to replicate the newspaper known as the Daily Prophet from the Harry Potter films, using computer vision methods. In the films, paintings and photographs are displayed as moving images opposed to still images.

This customized newspaper is augmented in real-time using the cameras and the projector. One camera is capturing footage of the newspaper and projection simulatenously. As you read this, the captured footage is being analysed using computer vision methods, in order to determine the location and orientation of the newspaper. This information is then used to transform the footage gathered from the other camera which you see projected here. The image needs to be transformed to hold the illusion of it being printed onto the newsaper.

Figure 4-10 Adapted newspaper with x2 different colours

This method would use a similar algorithm to how a QR code identifies its orientation. First challenge is to identify the bottom right mark (as this is the only different colour), this is achieved using the same process as the previous colour method and the mark is classified according to thresholds defined in the calibration stage. This mark is to be treated similar to the alignment pattern found on QR codes.

The top left mark can be identified next. The top right and bottom left marks (or the diagonal pair they can be referred to) can be identified as the marks with the longest distance between them. From simply using the distance between two points formula [31]:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

The pair with the max distance is the diagonal pair and the remaining mark can be identified as the top left mark.

The final challenge is to distinguish between the diagonal pair, using the position of the bottom right mark relative to the other marks. This method has not yet been resolved due to lack of time.

4.2.4 Limitations and Potential Limitations

Lack of time with completion of this method is a limitation in its own as it could not be implemented and tested effectively. Another large limitation is attempting to use the QR code's assumptions as the newspaper is not the same layout as a QR code. All the marks in a QR code are equidistant within a square, whereas the above example is a rectangle. This can lead to miss classification of marks, as the system could mistakenly identify the top left mark and the top right mark as the diagonal pair if the newspaper was tilted at such an angle where this distance is the maximum, potentially seen below.

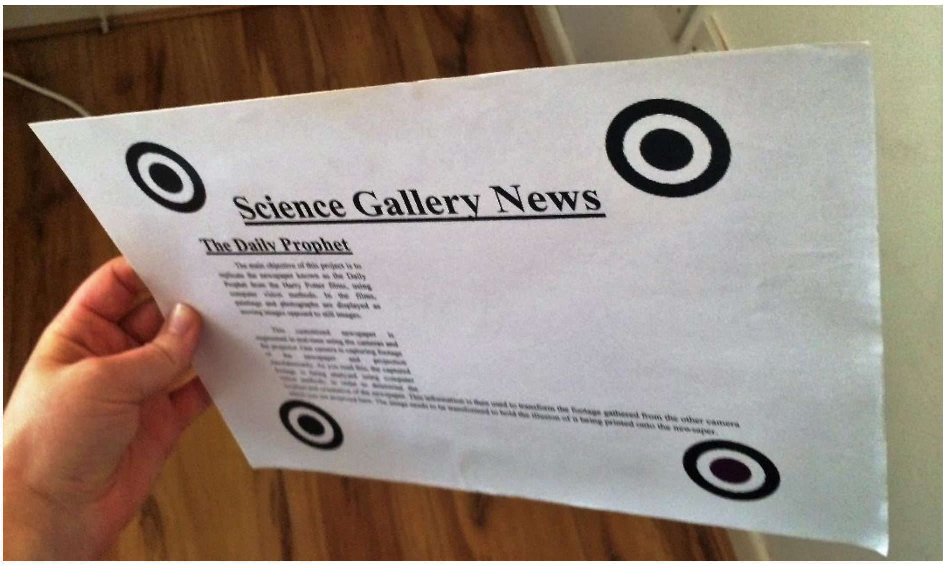


Figure 4-11 The diagonal pair should have the largest distance however severe angles may distort this

4.3 Perspective Transformations

This section describes the transformations applied to the projected output image captured by the second camera and any intermediary processes required.

The higher level goal is to transform the image from its original size (which would be the entire projection) to the size and location of the blank space on the newspaper.

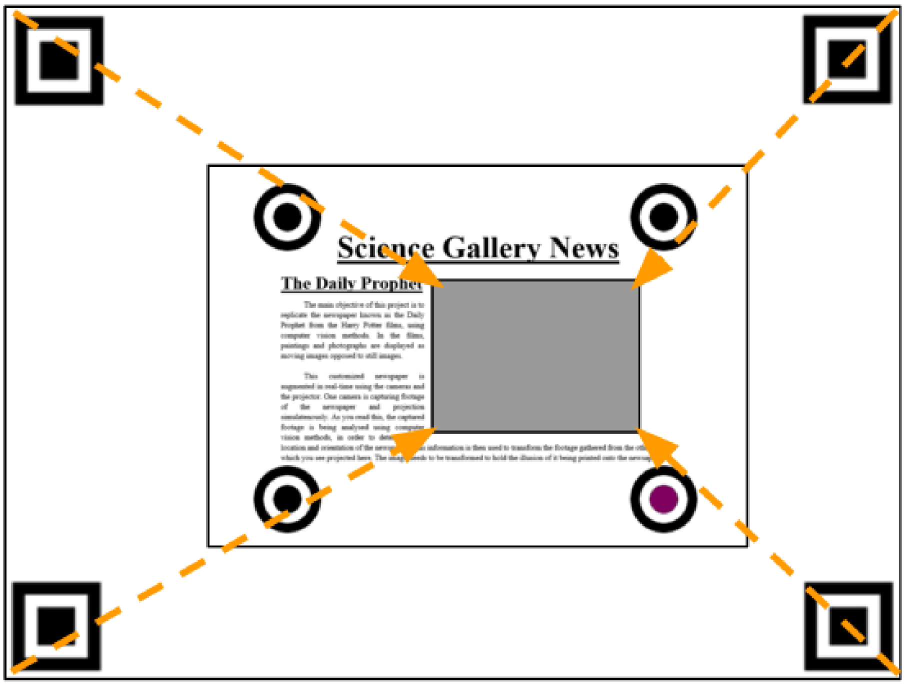


Figure 4-12 Graphic showing the desired perspective transformation

Above depicts the higher level goal of the perspective transformation. That is to transform the extrapolated outer calibration points to the interpolated inner newspaper points of the blank space.

4.3.1 Point Approximation

There are two sets of points which need to be determined before the perspective transform can be performed, the outer calibration points located on the corners of the projection and the inner newspaper points which are the corners of the blank space on the newspaper. Essentially the endpoints of the arrows depicted on the above figure. Both sets are estimated using vectors and predefined fractions from known points.

4.3.1.1 *Inner Newspaper Points*

These are the points which define the blank space located on the newspaper that is designated for the projection. They will be estimated from the marks on the newspaper.

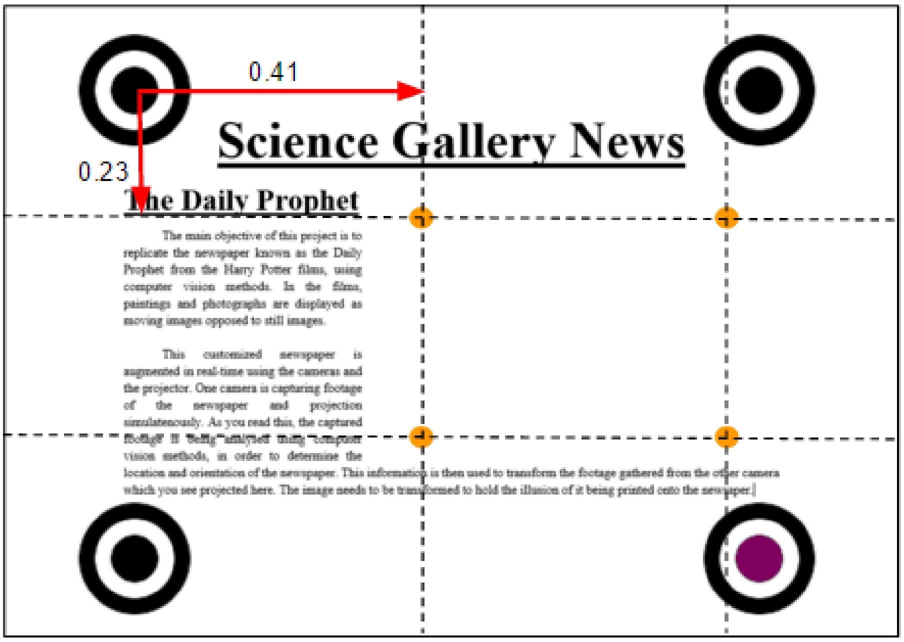


Figure 4-13 Interpolation of the inner newspaper points

To estimate the top left interpolated point shown above, it must be moved along the axis towards the top right mark and it must also be moved along the axis towards the bottom left mark. The system needs to know the direction to move the point in and the fraction to move it by.

The direction is calculated by subtracting the top left mark's coordinates from the top right mark's coordinates, and similarly said for the downward direction by subbing in the bottom left mark for the top right mark. 41/100 is the predefined fraction the interpolated mark sits from the top left mark to the top right mark. These fractions were pre-calculated as it is assumed that the newspaper cannot be moved at an angle that will miss calculate these approximations beyond use, however this assumption still has its limitations to be discussed shortly.

By multiplying the direction to move in with the fraction to move along that direction, the new interpolated point can be estimated. The below equations depict the calculations to interpolate the new top left mark:

$$dir_{TopLeftTopRight} = p_{TopRight} - p_{TopLeft}$$

$$dir_{TopLeftBottomLeft} = p_{BottomLeft} - p_{TopLeft}$$

$$InterpolatedPt_{TopLeft} = p_{TopLeft} + \left(0.41 * dir_{TopLeftTopRight}\right) + (0.23 * dir_{TopLeftBottomLeft})$$

For all the interpolated points the fractions were found to be as follows:

Direction	**Fraction**
Top Right – Top Left	*0.41*
Bottom Left – Top Left	*0.23*
Top Left – Top Right	*0.08*
Bottom Right – Top Right	*0.23*
Bottom Left – Bottom Right	*0.08*
Top right – Bottom Right	*0.27*
Bottom Right – Bottom Left	*0.41*
Top Left – Bottom Left	*0.27*

4.3.1.2 Outer Calibration Points

The calibration points which are found during the calibration stage must be extrapolated to estimate the outer corners of the projection. This are done using the same method as described above, only with different calculate fractions and subtraction is used instead of addition when translating the original point to its extrapolated point.

$$dir_{TopLeftTopRight} = p_{TopRight} - p_{TopLeft}$$

$$dir_{TopLeftBottomLeft} = p_{BottomLeft} - p_{TopLeft}$$

$$ExtrapolatedPt_{TopLeft} = p_{TopLeft} - \left(0.04 * dir_{TopLeftTopRight}\right) - (0.05 * dir_{TopLeftBottomLeft})$$

The fractions remain uniform throughout. Any left or right directions are a fraction of 0.04 and any upwards or downwards directions are a fraction of 0.05.

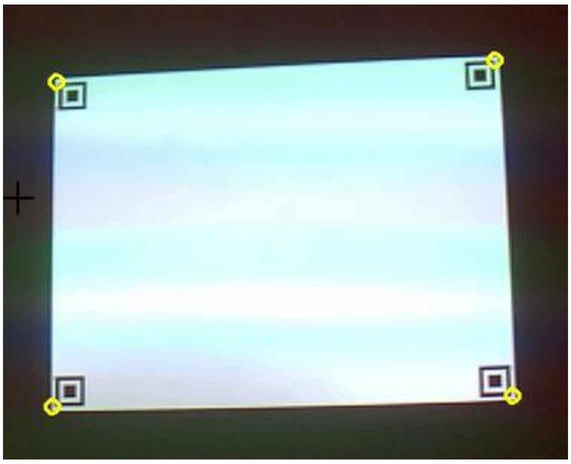


Figure 4-14 Extrapolated calibration points estimate the outer corners of the projection

4.3.2 The Mapping Matrix

The system must find the matrix which maps the extrapolated calibration points to the interpolated newspaper points. This calculated homography will then be applied to the output footage before it is projected. This mapping matrix can be seen below [32]:

$$\begin{bmatrix} t_i x_i' \\ t_i y_i' \\ t_i \end{bmatrix} = mappingMatrix \cdot \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

Where:

$$(x_i', y_i') = The\ transformed\ coordinates$$

$$(x_i, y_i) = The\ original\ coordinates$$

$$i = 0,1,2,3$$

It is important to note that even though the old coordinates and the new coordinates are numerically different, their directions remain equivalent. Hence the cross product between the coordinates will be zero [33]. This information is useful when isolating the homography before it is applied to the output feed. Each coordinate in the output image is updated by this mapping matrix [34]:

$$destination(x, y) = source\left(\frac{M_{11}x + M_{12}y + M_{13}}{M_{31}x + M_{32}y + M_{33}}, \frac{M_{21}x + M_{22}y + M_{23}}{M_{31}x + M_{32}y + M_{33}}\right)$$

4.3.3 Total Transformations

Three transformations are required to correctly relocate the projection. The main body of the transformations is the previously discussed one from the outer approximated calibration points to the inner approximated newspaper points. The mapping matrix is found and applied to the output image for each transformation with the same method as previously described.

1. *Actual Coordinates -> Extrapolated Calibration Coordinates*
2. *Extrapolated Calibration Coordinates -> Interpolated Newspaper Coordinates*
3. *Extrapolated Calibration Coordinates -> Actual Coordinates*

The below figures graphically depict the transformations respectively. The output and input footage is of the same dimensions however it is important to note for the first and third transformations, the actual coordinates of the output footage is used. The below diagrams just depict the transformations that will happen to the output footage and not the input footage (even though the below figures use the input footage to depict this). Also note that projection is not in the correct space below as the below figures were taken before the projection had enough time to reach its designated space.

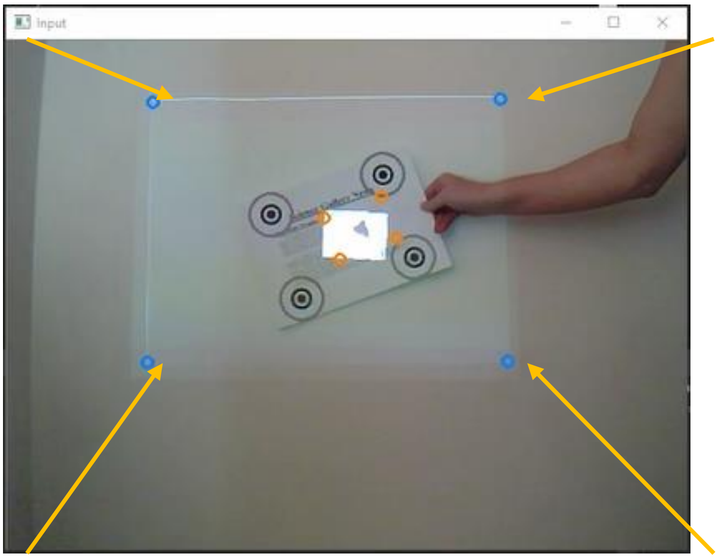


Figure 4-15 Transformation 1

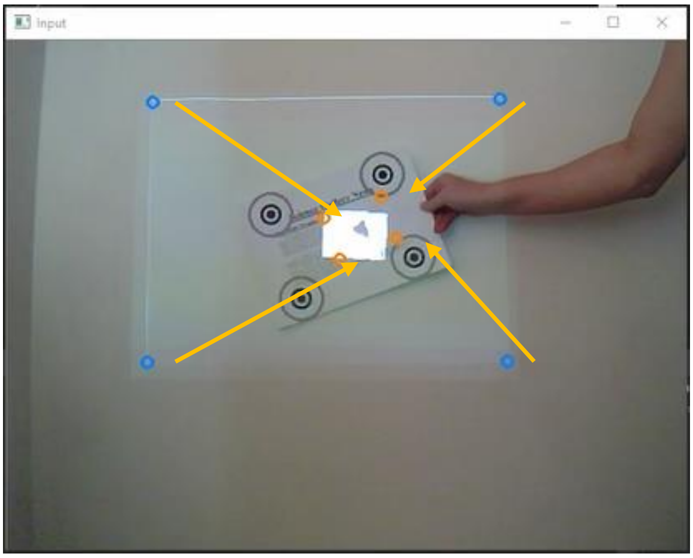


Figure 4-16 Transformation 2

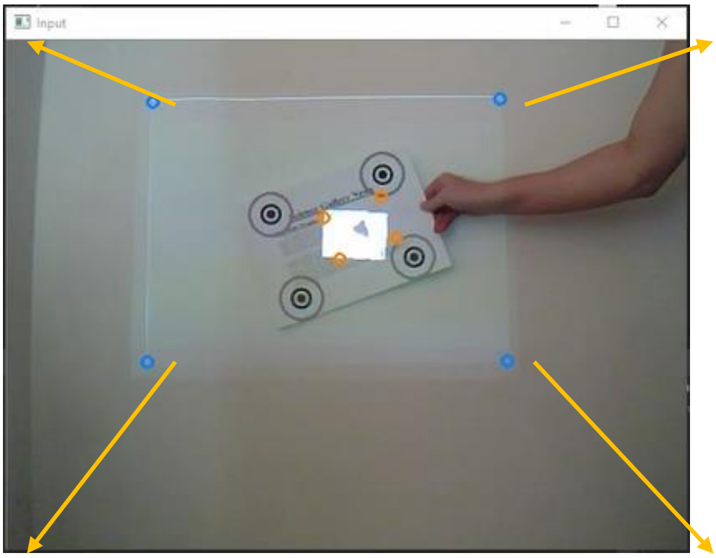


Figure 4-17 Transformation 3

The actual coordinates can be described as the coordinates which identify the four outer corners of the projected image. The extrapolated calibration points can be described as the apparent coordinates for the projected image, relevant to the input feed, where the actual coordinates of the projected image are the coordinates known to the system, (Note that the 'Total No. of Columns' or 'Total No. of Rows' indicate the maximum the width and height of the image according to pixels):

- *Top Left = (0,0)*
- *Top Right = (Total No. of Columns - 1, 0)*
- *Bottom Right = (Total No. of Columns – 1, Total No. of Rows – 1, 0)*
- *Bottom Left = (0, Total No. of Rows – 1)*

The second transformation is from the extrapolated calibration points to the interpolated newspaper points. This transformation cannot be applied to the output image alone. This is because the system believes the output image's coordinates to be the boundary coordinates described above and not the extrapolated calibration points.

The image's actual coordinates must first be mapped to the calibration (apparent) coordinates. Therefore, when the mapping matrix calculated from transformation no. 2 is applied to the output image, the initial coordinates of the output image before this transformation will now be the calibration coordinates, opposed to the actual coordinates.

Subsequently, when the second transformation is completed this mapping must be undone so that the image does not appear shrunken, this is the third transformation. It would appear shrunken as the image would have been transformed from the actual coordinates to a smaller size according to the calibration coordinates, it would be desired to undo this shrinking. Without transformations 1 and 3, transformation 2 would be applied to an image with the actual coordinates. This is inconsistent and would result in the image being projected in an incorrect location.

4.4 Future Methodology

The following describes the theory behind methods that were never implemented in the system due to limiting factors such as time. Below is the algorithmic approach that would be taken to add in these features however they have not been tested so their success cannot be guaranteed.

4.4.1 Prediction of Coordinates

For the system to work transparently, the projection should be updated as frequently as possible. If the direction and velocity of the newspaper's marks could be recorded, the projection could be estimated for frames where the newspaper has been unable to identify four marks. These estimated coordinates could then be used temporarily as substitutes for actual detected coordinates for the perspective transformations.

To achieve the above, it would be proposed for the system to store the coordinates of the previous instance where all four marks were found in a single frame. Essentially the last instance where the projection was updated. The system would also need to store the previous instance to this where the above is true. The third measure the system would need is the difference in frames between the two stored instances of the projection being updated.

The number of frames in between the two instances and the distance between each coordinate within the two instances could be used to estimate the velocity of the newspaper, to find a per frame distance travelled. The subtraction of the older instance from the fresher instance would indicate the direction of the newspaper.

$$direction_{12} = p_2 - p_1$$

$$distance_{12} = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2}$$

$$differenceInFrames_{12} = frameNumber_2 - frameNumber_1$$

$$distancePerFrame = \frac{dist_{12}}{difference\ in\ frames_{12}}$$

$$differenceInFrames_{23} = frameNumber_3 - frameNumber_2$$

$$estimatedDistance_3 = (distance\ per\ frame) * (difference\ in\ frames_{23})$$

$$estimatedPoint\ (p_3) = p_2 + (estimatedDistance_3 * direction_{12})$$

The above equations describe the steps that could be taken to obtain the location of an estimated point, where p_1 would be a single point on the first stored instance and p_2 would be the corresponding point on the second (fresher) stored instance. The system would need to repeat this process three more times for each corner.

This step would only be expected for short time periods as it is an estimate. The use of a logical clock such as the difference between frames opposed to time could lead to errors. Only straight line distances are accounted for which also could lead to errors, hence it is proposed that this would only be implemented for short periods. It would be proposed that the system would contain a limit to prevent the estimated projection from being updated if the newspaper has not been detected for a prolonged period. The system would also be unable to predict locations if there were not two previously stored instances, hence this feature would not be implemented until this was true.

This implementation would be highly useful for the colour method as this method rarely manages to identify four marks within a frame.

4.4.2 Smoothing of Transitions

Another feature which would aid user experience would be to smooth out the transformations opposed to them jumping from one position to the next. This can occur if the user moves the newspaper at a high pace.

This would involve interpolating intermediary positions to transform the projection as it moves towards the final destination. This feature could involve similar procedures discussed previously with coordinate prediction in 4.4.1, where the distance and direction between the old coordinates and the new coordinates would need to be calculated. The system could then decide for there to be two intermediary points for example. Another future aspect of this feature could be the system's ability to decided how many intermediary points would be used according to distance, if the distance was large then more intermediary coordinates would be required. With two intermediary points the system would project a transform at one third the distance of the total distance and then a transform at two thirds the distance of the total distance, before the transform sits in its final destination.

It would be expected for this method to be costly in terms of processing time with all the intermediary estimates and transformations being calculate, meaning it might not be suited to actually be implemented if time was available.

This would also be useful for the colour method as often the method is unable to identify four marks for a couple of frames and when it finally does the image jumps across.

4.4.3 Refined Search Area

Reducing the search area could potentially improve processing speed of the system. Using the previous knowledge of the mark's location, the surrounding area could be isolated for the system to only search this area. This could be repeated for each mark.

When detecting marks, the system already computes the area of the component, in 4.1.5.The system could use the knowledge of the centre point and this area to search an area that is ten times this area as an example. This could potentially save time compared to processing the entire image looking for all the marks again. If the system failed to find the mark within the system the search area could be increased accordingly, else the system could move onto the next frame and search the entire image in the case of a failed search.

5 Analysis

This chapter discusses the final appeal of the two methods and compares their performance as well as highlighting implementation issues that were encountered during the development process.

5.1 Assumption Method

Once the system can find all four marks the assumption method will be updated, so once the lighting is good in the testing environment the assumption method works very well as it is frequently updated. It only requires one calibration stage of calibrating the camera and projector, so its set-up is easier than the colour method.

This method does experience a lag with the projection, if the newspaper is moved rapidly the projection will lag behind before it reaches the newspaper's current position (if the newspaper is held roughly in the same position). This method generally takes between 180ms and 205ms per frame to compute, however that is when marks are found. If four marks are not found it takes between 40ms and 60ms and even less if no marks are in the frame (10-20ms). The lag could also be related to the projector's reaction time.

The projection also appears very pixelated and of poor quality, however its transformations are correctly adjusted.

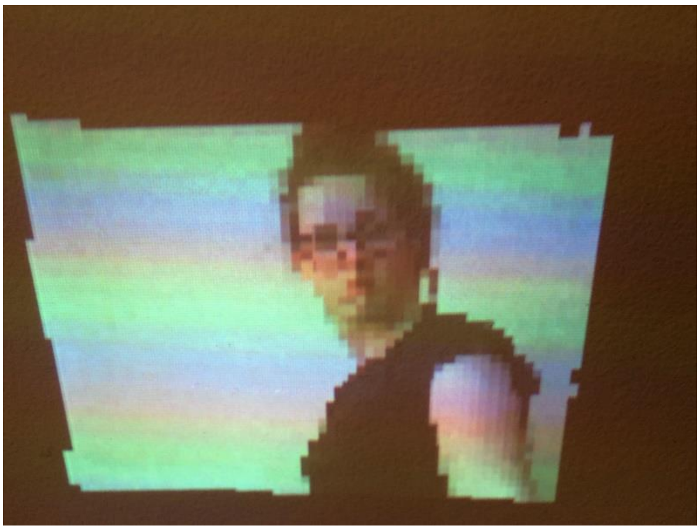


Figure 5-1 Highly pixelated projection

Also as previously stated this method does have limitations where the user must not rotate the newspaper too severely as the projection will then be incorrect. These limitations are predictable and avoidable unlike the limitations of the colour method.

5.2 MULTIPLE COLOUR METHOD

The colour method requires all four marks to be found but also to be identified as a colour before it updates, hence when compared to the assumption method the projection is updated much less frequently as this identification process is not completed as often. This also already depends on the lighting being natural and not poor artificial lighting which is a huge limitation to this method.

Unlike the assumption method the user is not limited to how far they want to rotate the newspaper. This is the main upside to this method however it has had costing effects as previously stated with regards to the projection being updated frequently. The original motivation for colours was to overcome the possibility of the user obstructing a mark, it was proposed to have more than four marks available. The system could make estimates from the found marks.

Errors do occur with the colour method and these are primarily related to colour miss-classification. They will be further discussed in 5.3. This error is generally unpredictable when compared to the assumption method as the assumption method's error (rotational limits) can be avoided. This unpredictability with the errors lowers the usability of the colour method.

Similar to the assumption method the colour method's projection is equally as pixelated and this method also experiences a lag when the projection is updated. The computation time of the colour method is minimally greater by roughly 10ms, however if the system cannot match all four marks it generally takes 70ms to 90ms per frame. It is difficult to more accurately measure the colour method as it rarely updates due to being unable to identify all marks. Again this lag could also be related to the projector's reaction time.

5.3 IMPLEMENTATION ISSUES

The following outlines large issues that interrupted the development process and how they were overcome where relevant.

5.3.1 Newspaper Design Issues

The design of the newspaper was a highly iterative process. Many adaptations were tested through trial and error. There were also numerous adaptations of the marks and the marks' placements on the newspaper.

A large concern was that the user might obstruct the newspaper with their hands.

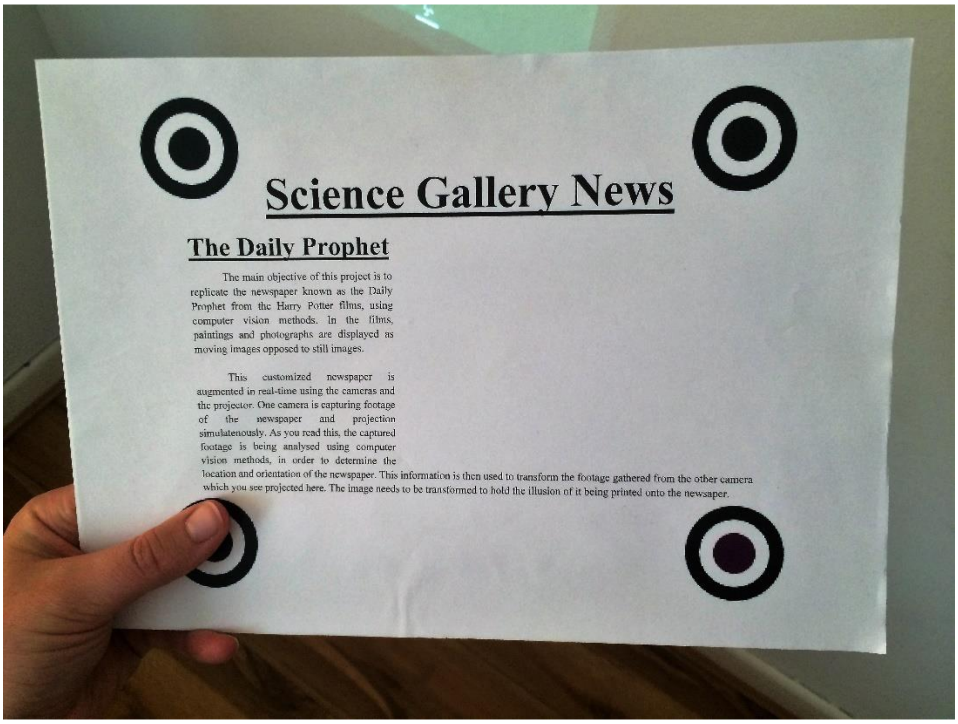


Figure 5-2 User obstructing bottom left mark

This was taken into account when designing some newspapers. The main thought was to use multiple marks so that if the user obstructed one the system could estimate the four corners from the remaining marks, this ran into problems of its own which is discussed in a bit with colour method issues in 5.3.4. Mark placement on the page was important, however a large limitation was the requirement for the marks to be large enough to be detected this might be resolved by using a higher resolution camera.

Different marks were sampled and tested through trial and error. There were initial difficulties in obtaining coloured marks that could be detected when a binary threshold was applied.

Figure 5-3 Evolution of Marks

The above figure shows the evolution of the marks during the design process, after it was decided to use QR code finder patterns as a guideline. The marks were converted to circles and then testing with colour began. Note the three coloured marks on the right are all of the same hue, 45. The first coloured mark (the centre figure) failed to be detected after a binary threshold was applied, hence experimentation was done with the black border. The best solution to this was to darken the shade while keeping the same hue which is how the mark on the far right was drafted.

5.3.2 Thresholding Issues

Otsu thresholding was used over adaptive thresholding. Theoretically adaptive thresholding should return much better results, especially in the situation of this project with varying light. Through practice adaptive thresholding proved to be processing too slowly for real-time implementation.

Poor thresholding was one of the only areas that would potentially cause errors with mark detection on rare occasions. The system requires four marks to be found to apply transformations to the projection, therefore if the binary image is of poor quality with obscured undetectable marks the system cannot continue. Most frequently the marks were obscured due to uneven lighting or the projector shining onto a mark.

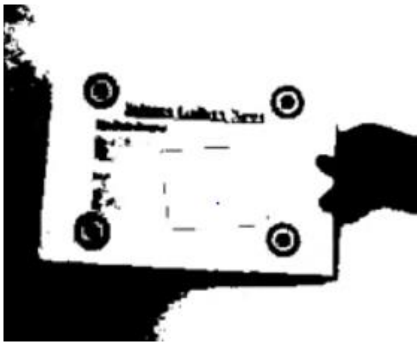


Figure 5-4 A poor quality binary image, note how the uneven lighting effects the left marks

5.3.3 Assumption Method Issues

It was stated that the assumption method was the most straightforward to implement however it had the largest limitations in terms of flexibility. Once the newspaper is rotated at too severe an angle the assumptions are redundant. Remembering that this method assumes which ever mark is closest to top left in the image must be the top left mark on the newspaper.

Below is a good example of when the newspaper has been rotated beyond its limits. The top right and bottom left marks would be incorrectly identified, recalling that whichever of these two contains the lower x value is the bottom left.

Figure 5-5 Newspaper's angle is too severe for the assumption method

In practice it was found also when rotated too much the system would assume the top right mark to be the top left and the below distortion would happen as well as obviously failing to work upside down, also seen below:

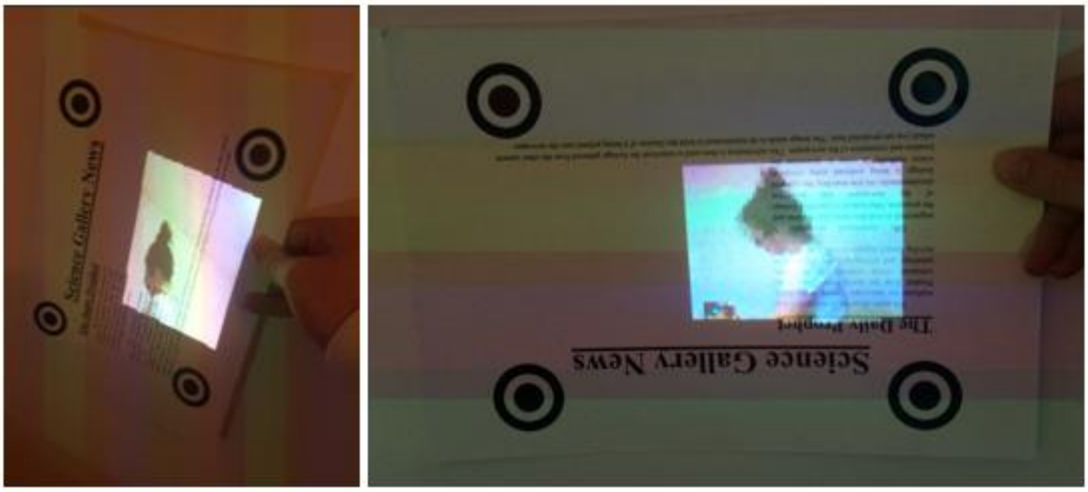


Figure 5-6 Assumption method errors

5.3.4 Colour Method Issues

It was briefly stated that the main limitations to the colour method was the system's difficulty in distinguishing between different hue values. Recalling that the hue value of the centre of each mark can be calculated by the system.

The aim of the colour method was to solve the limitations associated with the assumption method where all four marks were required, and the user could not rotate the newspaper beyond a limit. Hence the colour method initially aimed to use more than four marks.

Figure 5-7 Newspaper prototypes for colour method

The above image on the left depicts the desired design using eight different hues. These colours were created in an image editor where the hue values could be selected and were all equidistant from one another to give the best results, their hues ranged over the interval [45 360].

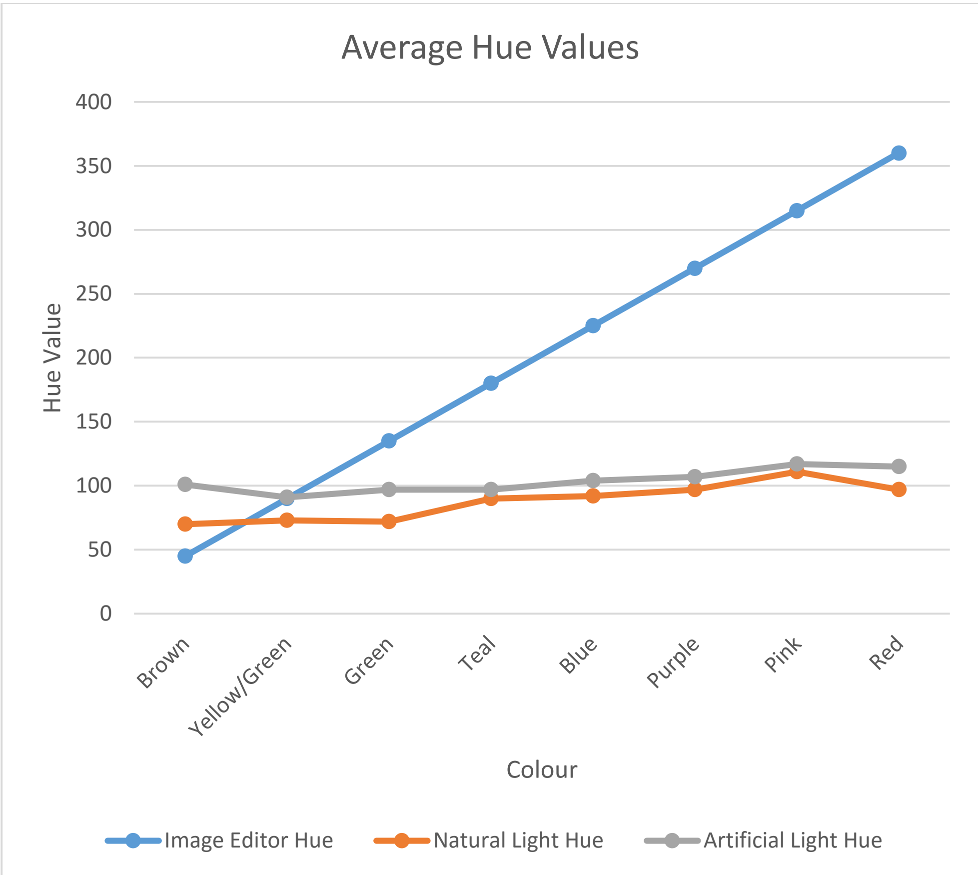
During colour calibration it became clear from the hue values that the values were not distinct enough. The software reads the hue values over the interval [0 179] opposed to the theoretical interval [0 360], however this is not considered to be the problem. The readings varied greatly and often would overlap.

An immediate solution to this was to reduce the number of colours, hence they would be more distinguishable to the system, which explains the image on the right in the above figure. This did not create an immediate solution however as the hue values were still returning incredibly close values to one another.

Potential areas where this error is occurring could be due to how the system is finding the average hue. It places a bounding box around the centre of the mark and calculates the average hue of this box, this has its evident flaws as stray whitish pixels are located in the corners of the box disturbing results. There could also be potential flaws with the creation of the images in the first place where black pixels have made it into the border of the colour also disturbing results.

The testing environment had the greatest effect on these values. Large differences were noticed between the calibration values when testing was conducted under natural light compared to artificial light. Two sample calibrations were made in the same day, one in the day time and the other in the evening after dark, and these calibrations returned the following averages:

COLOUR	IMAGE EDITOR HUE	NATURAL LIGHT HUE	ARTIFICIAL LIGHT HUE
BROWN	45	70	101
YELLOW/GREEN	90	73	91
GREEN	135	72	97
TEAL	180	90	97
BLUE	225	92	104
PURPLE	270	97	107
PINK	315	111	117
RED	360	97	115



Theoretically the method would work flawlessly if the image editor hue values were the real values, however this is not the case. There is a slight improvement in the values read under natural daylight compared to those under artificial light. The daylight values just about work when only four marks are used using brown, green, blue and pink. The above test case showed green and blue to be very close so these calibration values may not be sufficient and could cause miss classification for the identification stage of the process. This miss classification of a colour results in an incorrectly transformed image seen below.

Figure 5-8 Result of colour miss classification

5.3.5 Point Approximation Issues

Recalling that the outer calibration points and the inner newspaper points were estimated using predefined fractions. As the newspaper can be tilted at severe angles, these predefined fractions can lose their integrity, see below

Figure 5-9 Left shows interpolated points working well. Right shows limitations of these approximations.

The above left image shows the orange interpolated points working well as the newspaper is almost directly facing the input camera, note how they don’t overlap with the text, this is an ideal situation for the point approximation. The picture on the right shows the described limitation, this occurs when the newspaper is tilted at more severe angles. Note how the top left interpolated mark in the picture on the right is position slightly over the text. This error potentially occurs due to the change in perspective, hence the predefined fractions would no longer be as valid as they were in the left hand side image. There is also potential that the newspaper is not perfectly planar as warping can be seen on the image on the right which could also be the reason the approximations are distorted.

5.3.6 Transformation Issues

When outlining the perspective transformations in 4.3.3 it was stated that multiple transformations were required as errors occurred without the first and third transformation. These transformations were required to map the actual coordinates of the projector’s image to the apparent coordinates of the projector’s image relevant to the input feed, also described as the calibration coordinates (which have been extrapolated to estimate the outer corners of the projection).

A reminder of the three transformations:

1. *Actual Coordinates -> Extrapolated Calibration Coordinates*
2. *Extrapolated Calibration Coordinates -> Interpolated Newspaper Coordinates*
3. *Extrapolated Calibration Coordinates -> Actual Coordinates*

Below depicts the error that occurs when transformations 1 and 3 are removed. As the newspaper is moved further from (0,0) (the top left corner of the projection) along the x axis the error increases.

Figure 5-10 Perspective Transformation Error

The above image also shows the interpolated newspaper marks in orange, these are interpolated from the yellow marks. These marks highlight that the system is correctly identifying their location however the projection is still in the incorrect location.

A solution was found by first mapping the actual coordinates to the extrapolated calibration coordinates before the second transform was applied, as the mapping matrix for the second transform is from these calibration coordinates to the interpolated coordinates and this helps hold consistency.

5.3.7 Equipment and Environment Issues

Equipment was a limiting factor to the dissertation, it has been previously stated that higher resolution cameras could yield better results in a number of locations. If the cameras were of higher resolution:

- The input feed camera could be set up further away from desk/wall surface
- The marks could be smaller and more discreet on the newspaper
- The projected image would be of higher quality on the sheet of paper

However, if higher resolution cameras were used, it would be expected that the processing speed of the system would decrease and hinder usage. Unless of course a computer with better processing power was sourced.

The environment was also a limiting factor. Due to light fluctuations (especially the difference between daylight and artificial light) no consistent testing could be conducted. The system also required large clear space to be set up and tested, such space is difficult to source without obstructions and shadows.

6 CONCLUSION

This dissertation set out to replicate the Harry Potter newspaper known as the Daily Prophet through a means of combining computer vision and projections. This involved adapting a prototype newspaper with fiduciary markings for the vision system to look for. A lot of time was taken before coming to this conclusion to use fiduciary markings as the project is quite unique, and locating similar projects and research was very difficult, especially when looking for more recent and relevant pieces of work.

A number of designs were prototyped on a trial and error basis. A vital component to using the markings was for the system to be able to identify which marking was which (identify the top left, top right, bottom left and bottom right). A number of methods were designed for this task, namely the assumption method and the multiple colour method. The two colour method was briefly outlined and development would have been further ensued if more time was available.

Each subsequent method was developed to form solutions to limitations of its predecessor, the colour method would allow for the projection to work when the newspaper was rotated beyond the limits of the assumption method, the two colour method would allow for the colour method to be used in poor or artificial lighting.

The assumption method functioned well if the user used it as expected. A lag could be seen with the projection presumably due to computation time or slow reaction time of the projector. The multiple colour method restricts the user less as it allows them to use the system at greater angles (upside down/rotated 180^{o}). The multiple colour system requires an extra calibration step during set-up, and this calibration would frequently need to be updated as the colour method is highly sensitive to the surrounding environment and its lighting. A lag could also be seen with the colour method again presumably to computation time or slow reaction time of the projector.

Final testing (including the projector) of the colour method was conducted quite late as a large amount of time was lost to image editing and finalising the newspaper prototype, and designing all the previous components of the system. A lot of trial and error with the coloured prototype newspaper was conducted, very often a single coloured mark would not be detected after a binary threshold was applied. This would result in a redesign and a reprint of the adapted newspaper. This justifies why the third proposed method of the two colour method never completed development due to limiting time.

The colour method proved to work better under natural daylight opposed to artificial light and this contributed to its difficulty in testing. Due to the short days of Irish winter months there was a limited time frame with which the system could be tested each day. Also as the system required a large amount of space to test with the projector, it was not portable and could not be tested on campus. A single camera could be brought in to conduct minor testing during a college day however testing with the projector was much less flexible.

Overall the system successfully transformed the projection into the correct location on the newspaper for both the assumption method and colour method. Granted both methods have certain limitations to their environment set-up, the assumption method must have the camera and newspaper in an upright position relative to one another, the colour method requires natural lighting and his highly unsuccessful under artificial lighting. If conclusions were made quicker about the lighting disrupting the colour methods and more time was available, development could have been completed on the two colour method or other solutions could have been explored. The system remains a fun application of computer science or computer vision and it can be concluded that it would not be feasible as an augmented desk as the reaction time is too slow and there is little need for one in today's society due to the popularity of tablets. There are still areas with which this piece of work could continue to improve and will be further discussed shortly.

6.1 Future Work

For future work it is already clear that development would continue on the two colour method if more time was available. Another large area that would need improvement is the computation time of the algorithms, this could involve a large refactoring of the implemented code. Currently when using the system there is a lag for the system to update the projection's location, granted when update this projection is in the correct location however it is delayed. It would be hoped that refactoring the code and improving the processing time this delay could be reduced. If more time was available this lag would also be further explored as there is potential that it could be a result of the physical projector.

Other proposed features were mentioned in the methodology section such as the addition of pre-processing which would involve predicting where the newspaper would be next. This pre-processing could arguably also reduce the projection's lag when updating its position. Another proposed feature would be to reduce the search area for the next frame by using the assumption that the mark will be in the surrounding radius, there is potential that the user could move the newspaper too fast for this to be feasible. It could also solve the colour method's issue when it is unable to frequently identify all four colours within a single frame.

The projection is only ever in the wrong location if it is lagging (or moving to the correct location), if the newspaper is rotated past its limitations during the assumption method or if the colours are miss identified in the colour method. It was considered to implement a form of error detection by having the projection also project finder patterns however when converted to a binary image it is not possible to identify anything within the projection as it often identified entirely as foreground pixels.

There was also discussion about potentially smoothing out the transitions of the projections, this could be implemented further down the line if the lag in the projection was reduced and the projections appeared evidently much more erratic. The projections do sometimes appear jumpy when the colour method has not identified marks for a large number of frames, and the projection jumps from one location to another suddenly.

6.2 Reflection on Potential Impact

Large research was also conducted into interactive desks and textbooks as it was the most familiar previous works to this project. Little work has been done on interactive work stations using projectors in recent years, presumably largely impacted by the introduction of tablets. A large motivator for a majority of these previous works was the ability to digitize physical notes and this is a feature which tablets enable.

When identifying the potential impact of this project it was identified that it was primarily for entertainment and could be used as a fun application which could gain interest in the computer vision and/or computer science field. Interactive desks were explored as previously stated however this goal would not gain much traction due to the popularity of tablets. At the beginning there was potential for the project to be exhibited at the Science Gallery at Trinity College Dublin on completion, this would have been a good outlet to see public's reaction to the application and evidentially see if it gained interest in its respective field.

References

1. *Daily Prophet*. Harry Potter Wiki [Webpage]; Available from: http://harrypotter.wikia.com/wiki/Daily_Prophet.
2. *Science Gallery Dublin*. SEEING; Available from: https://dublin.sciencegallery.com/2016.
3. Wellner, P., *Interacting with paper on the DigitalDesk.* Communications of the ACM, 1993. **36**(7): p. 87-96.
4. Wellner, P.D., *Adaptive Thresholding for the DigitalDesk*, in *EuroPARC Technical Report EPC-93-110*. 1993.
5. Wellner, P.D., *Self Calibration for the Digital Desk*, in *EuroPARC Technical Report EPC-93-109*. 1993.
6. Kobayashi, M. and H. Koike. *EnhancedDesk: integrating paper documents and digital documents*. in *Computer Human Interaction, 1998. Proceedings. 3rd Asia Pacific*. 1998. IEEE.
7. Rekimoto, J. *Matrix: A realtime object identification and registration method for augmented reality*. in *Computer Human Interaction, 1998. Proceedings. 3rd Asia Pacific*. 1998. IEEE.
8. Mackay, W.E., et al. *The missing link: augmenting biology laboratory notebooks*. in *Proceedings of the 15th annual ACM symposium on User interface software and technology*. 2002. ACM.
9. Margetis, G., et al., *Augmented interaction with physical books in an Ambient Intelligence learning environment.* Multimedia tools and applications, 2013. **67**(2): p. 473-495.
10. Margetis, G., et al. *Augmenting physical books towards education enhancement*. in *User-Centered Computer Vision (UCCV), 2013 1st IEEE Workshop on*. 2013. IEEE.
11. Lowe, D.G., *Distinctive image features from scale-invariant keypoints.* International journal of computer vision, 2004. **60**(2): p. 91-110.
12. Spindler, M. and R. Dachselt. *PaperLens: advanced magic lens interaction above the tabletop*. in *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*. 2009. ACM.
13. Weiss, M., S. Voelker, and J. Borchers, *Benddesk: Seamless integration of horizontal and vertical multi-touch surfaces in desk environments.* Adjunct Proceedings ITS, 2009. **9**.
14. Holman, D., et al. *Paper windows: interaction techniques for digital paper*. in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2005. ACM.
15. Vicon Motion Capturing System; Available from: www.vicon.com.
16. O'hara, K. and A. Sellen. *A comparison of reading paper and on-line documents*. in *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*. 1997. ACM.
17. Billinghurst, M., H. Kato, and I. Poupyrev, *The magicbook-moving seamlessly between reality and virtuality.* Computer Graphics and Applications, IEEE, 2001. **21**(3): p. 6-8.
18. Starner, T., et al., *The perceptive workbench: Computer-vision-based gesture tracking, object tracking, and 3D reconstruction for augmented desks.* Machine Vision and Applications, 2003. **14**(1): p. 59-71.
19. Wilson, A.D. and H. Benko. *Combining multiple depth cameras and projectors for interactions on, above and between surfaces*. in *Proceedings of the 23nd annual ACM symposium on User interface software and technology*. 2010. ACM.
20. Steimle, J., A. Jordt, and P. Maes. *Flexpad: highly flexible bending interactions for projected handheld displays*. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2013. ACM.
21. *Quick Response (QR) Codes*. FUSE Technology; Available from: http://www.l2soft.com/FUSE_T-QR_Codes.aspx.

22. *The Functions of QR Code Function Patterns*. Keep Automation; Available from: http://www.keepautomation.com/tips/qr_code/functions_of_qr_code_function_patterns.html.
23. *The HSB/HLS Color Model*. Technical Guides; Available from: http://dba.med.sc.edu/price/irf/Adobe_tg/models/hsb.html.
24. Dawson-Howe, K., *A practical introduction to computer vision with OpenCV*. 2014: John Wiley & Sons.
25. Otsu, N., *A threshold selection method from gray-level histograms.* Automatica, 1975. **11**(285-296): p. 23-27.
26. Suzuki, S., *Topological structural analysis of digitized binary images by border following.* Computer Vision, Graphics, and Image Processing, 1985. **30**(1): p. 32-46.
27. *Operations on Arrays - Sort*. OpenCV Documentation; Available from: http://docs.opencv.org/2.4/modules/core/doc/operations_on_arrays.html#sort.
28. *Stuctural Analysis and Shape Descriptors - Point Polygon Test*. OpenCV Documentation; Available from: http://docs.opencv.org/2.4/modules/imgproc/doc/structural_analysis_and_shape_descriptors.html?highlight=pointpolygontest#pointpolygontest.
29. *Structural Analysis and Shape Descriptors - boundingRect*. OpenCV Documentation; Available from: http://docs.opencv.org/2.4/modules/imgproc/doc/structural_analysis_and_shape_descriptors.html?highlight=boundingrect#boundingrect.
30. *cvtColor*. OpenCV Documentation; Available from: http://docs.opencv.org/2.4/modules/imgproc/doc/miscellaneous_transformations.html#cvtcolor.
31. *The Distance Formula*. Purple Math; Available from: http://www.purplemath.com/modules/distform.htm.
32. *Geometric Transform - Perspective Transform*. OpenCV Documentation; Available from: http://docs.opencv.org/2.4/modules/imgproc/doc/geometric_transformations.html#getperspectivetransform.
33. *How to compute a homography*. Projective geometry, camera models and calibration; Available from: http://www.cse.iitd.ernet.in/~suban/vision/geometry/node24.html.
34. *Geometric Transformations - warpPerspective*. OpenCV Documentation; Available from: http://docs.opencv.org/2.4/modules/imgproc/doc/geometric_transformations.html#warpperspective.