



The University of Hong Kong

Department of Computer Science

Identifying and Replacing Moving Image in Different Viewpoints

Chan Yiu Keung Joseph(2011565354)

Tam Siu Sing Stanley(3035010353)

Abstract

This proposal proposes an efficient Augmented Reality (AR) Cloud System. Practically, in order to enhance the efficiency of cloud image matching, we will build up a cloud database server that allows multiple matching in parallel. In addition, Speeded Up Robust Features (SURF) will be used as images similarity calculation. Our group hopes that our system can facilitate the use of AR among public and provide user a better experience.

Table of Content

Background.....	3
Introduction.....	4
Objective and Task	5
Methodology	6
Process Flow of Uploading Image Descriptor.....	6
Process Flow of Matching Image.....	7
Image Matching Algorithm (SURF)	8
Image Weighting System.....	9
Demo Application	10
Source of Idea.....	10
Process Flow	11
Major Milestones	12
Division of Labor	14
Schedule	15
Budget Plan	16

Background

Augmented reality (AR), as one of the most exciting technologies nowadays, has been widely adopted on mobile devices over the last few years. Lots of mobile applications have implemented AR technology into various aspects, such as educational, business, as well as gaming. Examples include Word Lens, Vuforia, and Google Glass.

Along with the development of AR on mobile devices, number of problems and limitations were found. The major concern is that AR requires a sufficiently large database for storing image descriptors which are essential for finding the region of interest on a image. Since the storage of majority of mobile devices are relatively small, it is not feasible to put the entire database on the mobile at this moment of time.

Another concern of AR is the processing power and energy consumption of mobile devices. Since AR requires large amount of calculations, the processing power becomes one of the important factors that affect the efficiency of AR performance. Moreover, large amount of calculations also imply large energy consumption. Regarding to these problems, people are putting lots of effort on transferring the database and massive calculations from mobile to web server.

Introduction

The main purpose of this project is to set up an online cloud service for massive images matching within a reasonable time. On the client sides, images would be captured and sent to the server side, and OpenCV / Vuforia would be used for images pre-processing if necessary. After receiving the image, server will then apply some searching strategies to reduce the total number of comparisons during images matching. The similarity of two images will be calculated by using SURF(Speeded Up Robust Feature) algorithm.

After finding out the matched sub image on server, the server would return the matched image descriptor(sub image) to the client. The client will then use the image descriptor to locate the target using Vuforia / OpenCV. In order to reduce the usage of networking bandwidth, some caching techniques would also be used on the mobile side. In short, we are going to implement relatively efficient AR functionalities using online cloud.

For the demo app, it will use the services(API) provided by the AR cloud server. The main idea of the app is comment on everywhere. Through this app, users can put/view comment on different images, eg: company logos, banners , products label, or even news pictures etc.

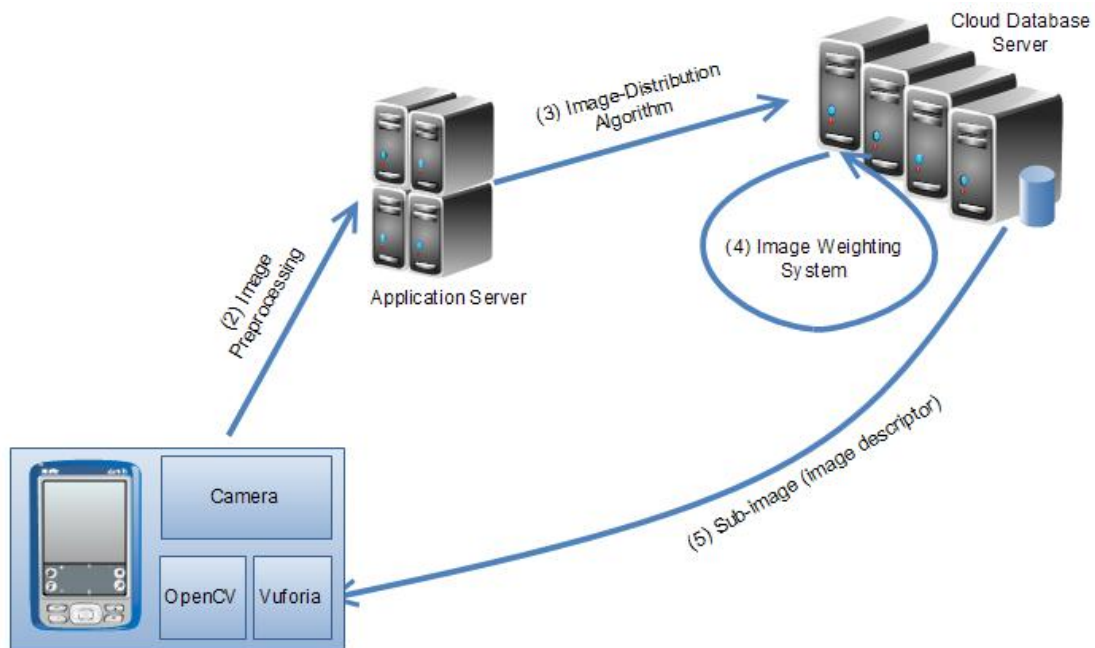
Objective and Task

The objective of this project is to provide an efficient AR Engine which target to minimized the processing time and the accuracy of image-matching. It can therefore encourage more people to use and increase the extendibility of AR application. To achieve the purpose, our team will implement the following component:

1. SURF algorithm for image matching
2. Cloud Database Server for storing image and parallel image matching
3. Efficient back-end System architecture for shortening the time of image matching
4. User-Friendly application for demonstrating the AR cloud services

Methodology

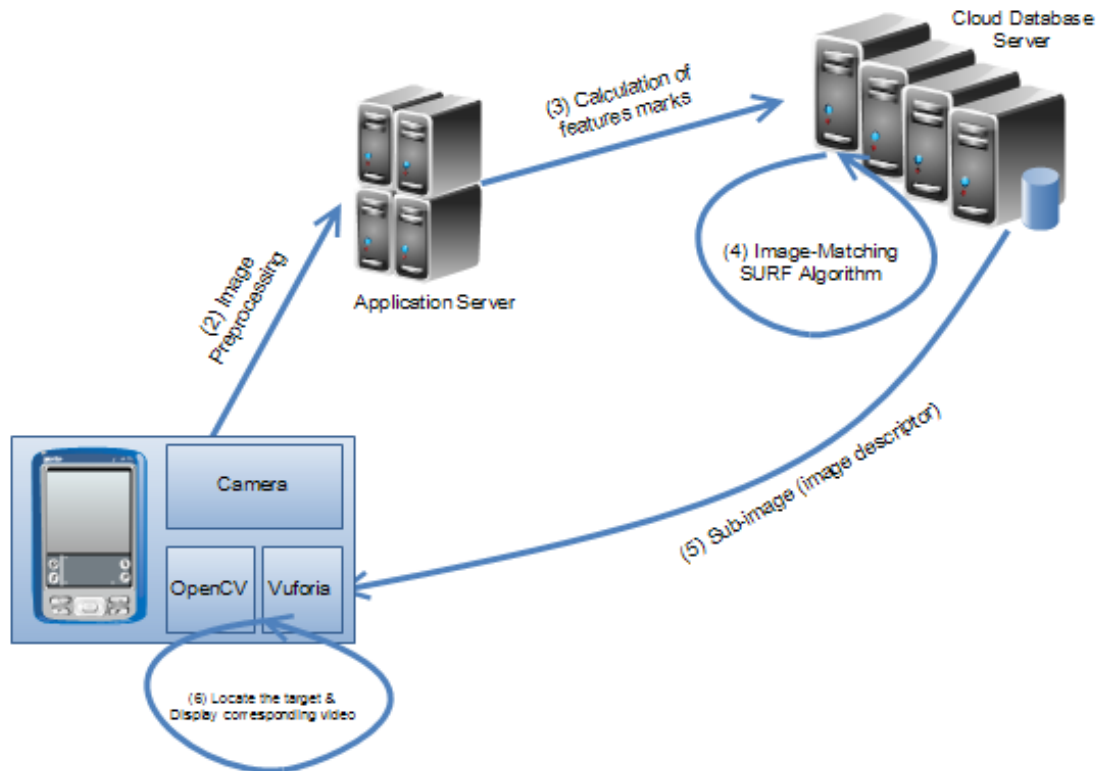
Process Flow of Uploading Image Descriptor



Steps:

- 1) Clients capture an image descriptor, and upload it to the server using TCP/IP
- 2) Image descriptor undergoes pre-processing using OpenCV / Vuforia if necessary.
- 3) Application Server applies Image-Distribution Algorithm to distribute images in different database based on the feature of the database
- 4) Database Server applies Image Weighting System to assign marks for each image's features(eg: number of SURF features points). The marks of each features will be used as weighting to facilitate image matching, such that reducing the number of unnecessary matching.
- 5) Database Server will return the image descriptor(sub image) to the client

Process Flow of Matching Image



Steps:

- 1) Clients capture the image, and upload it to the server using TCP/IP
- 2) Image undergoes pre-processing using OpenCV / Vuforia if necessary.
- 3) Database System calculates different features marks of the image, and apply the marks as the weightings to trim down the searching space.
- 4) SURF(Speeded Up Robust Feature) algorithm will be used to calculate the similarity of the two images.
- 5) Database Server will return the matched image descriptor(sub image) to the client
- 6) Client will use Vuforia / OpenCV to locate the center of interest by using the image descriptor provided by server.
- 7) Some hot-topic image descriptor will be cached in the client for reducing the usage of networking and bandwidth

Image Matching Algorithm (SURF)

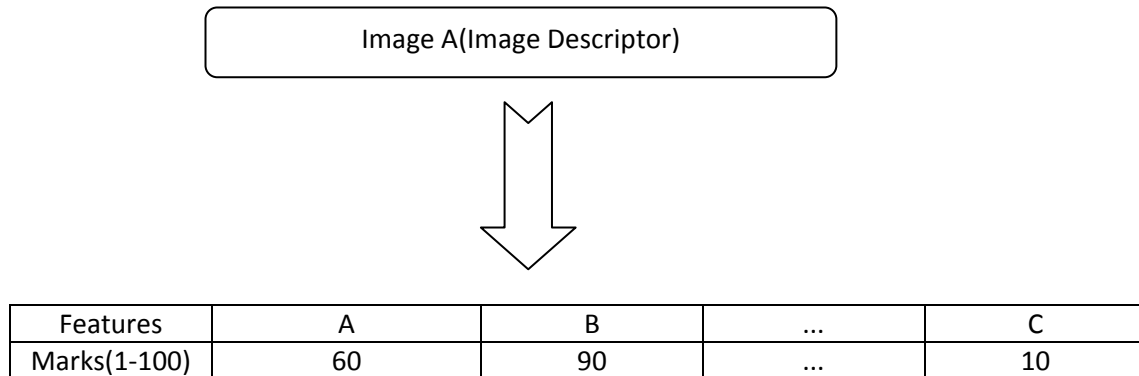


SURF (Speeded Up Robust Features) is a robust local feature detector, first presented by Herbert Bay et al. in 2006, that can be used in computer vision tasks like object recognition or 3D reconstruction. It is partly inspired by the SIFT descriptor. The standard version of SURF is several times faster than SIFT and claimed by its authors to be more robust against different image transformations than SIFT. SURF is based on sums of 2D Haar wavelet responses and makes an efficient use of integral images.

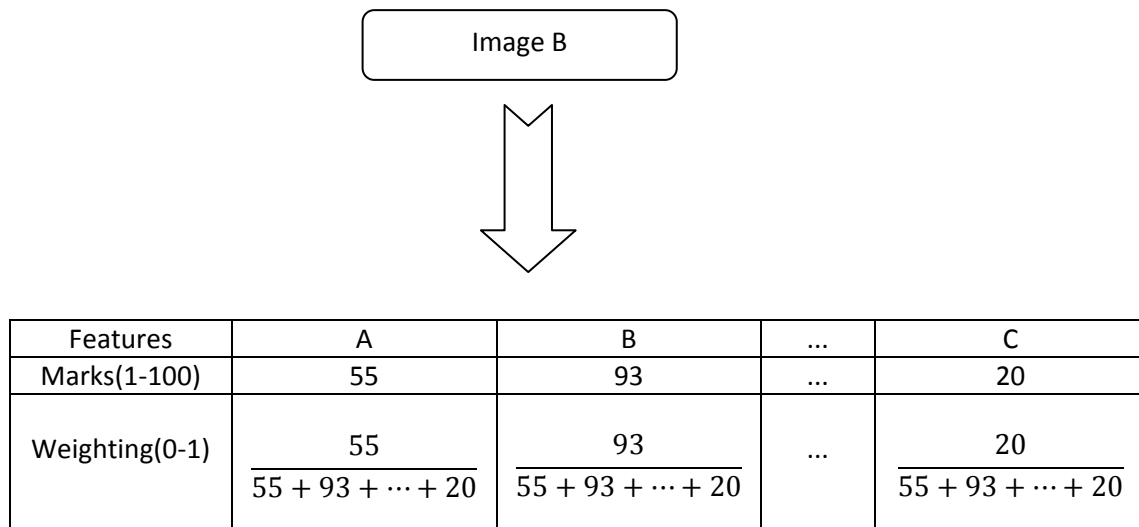
It uses an integer approximation to the determinant of Hessian blob detector, which can be computed extremely quickly with an integral image (3 integer operations). For features, it uses the sum of the Haar wavelet response around the point of interest. Again, these can be computed with the aid of the integral image.

Image Weighting System

(1) Calculating Features Mark (When adding new image to database)



(2) Calculating the weighting (When searching sub image)



For each image descriptor, the mark is calculated as followed:

$$A's \text{ Mark} = 60 \times \frac{55}{55 + 93 + \dots + 20} + 90 \times \frac{93}{55 + 93 + \dots + 20} + \dots + 10 \times \frac{20}{55 + 93 + \dots + 20}$$

The Marks will be sorted in decreasing order. Image with higher marks will be searched first!

Demo Application

Source of Idea

Nowadays, people always search the comments for organization before they decide to enjoy the service or the product provided by company in every aspects. One of the dominant example is the OpenRice forum, people will use the comment of the customer as a reference to decide whether they will go to the restaurant.

However, in present, there is still no central forum that can collect the comment from customer in different aspects, and industries. Moreover, It is inconvenient for customer to view the comment when they are urgent to make a decision. For example, in the current process, they need to first start a browser, or open a mobile application, and then they need to type in the name of the organization, and then search for comment. Sometimes, customer is unfortunate that they may need to enter the long name of the companies several times before they succeed. Our team consider that these steps are unnecessary and time-consuming.

And so, our team decides to design an AR mobile application that can return the comments of the company in only one step and in a very short time. User only need to open the camera and focus on the logo of the company and the application will then return the comment to them automatically. And the application also allow user to add comment on it. And our system will cover all the industry. This will be achieved with the help of user to identify the logo with the company name at the first trial.

Therefore, Our team guarantee that the application can facilitate users daily life, to view the comment as soon as possible, and discover the "covered treasure " in the street when they user the camera focusing on the logo when they are free.

Process Flow

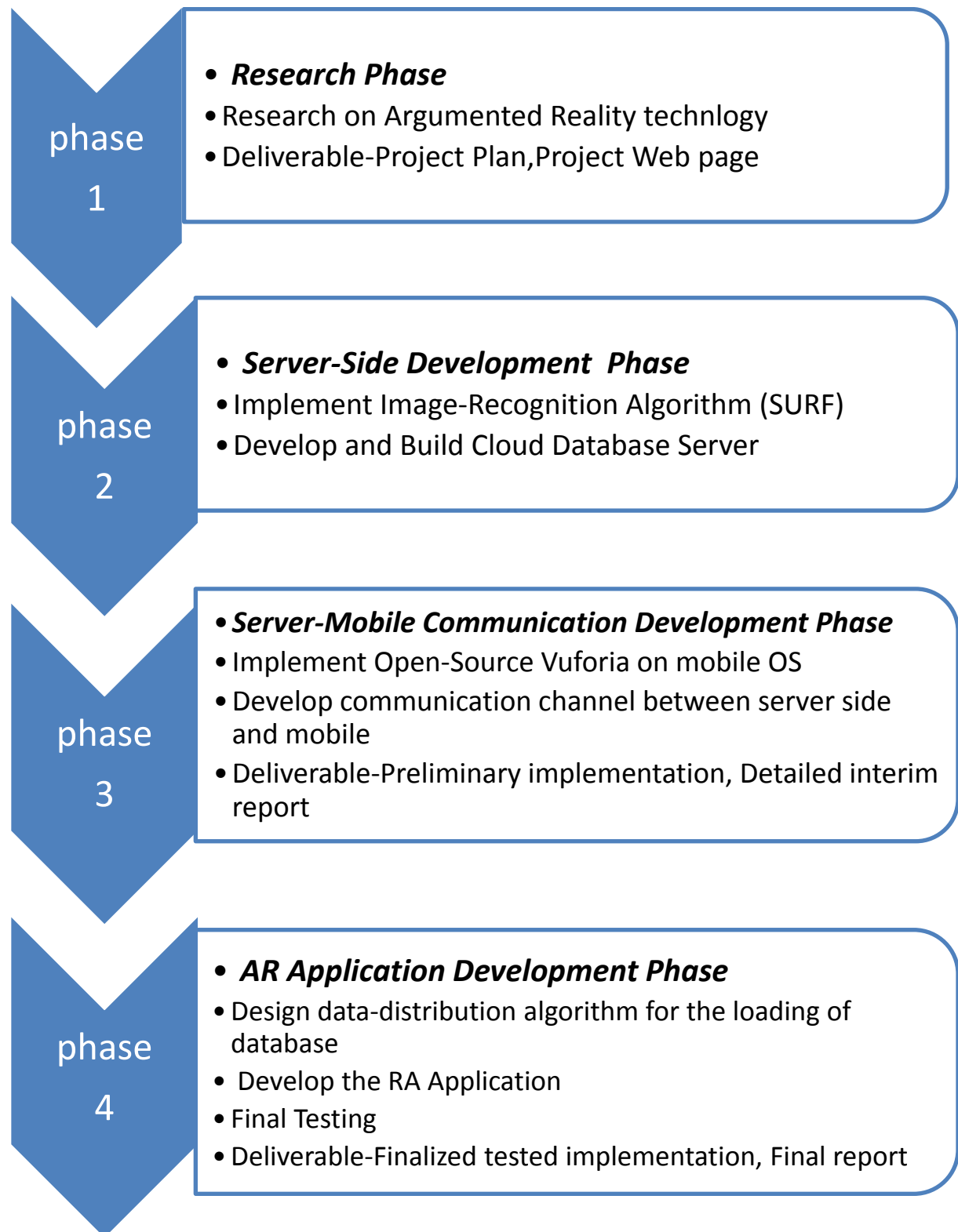


Steps:

1. Users open the camera application.
2. User focus on the logo of the organization(for example, HKU logo)
3. The application search the image in the database and gather the comments.
4. The application return a "+" button.
5. User press the "+" button to see the comments from other users
6. If the application cannot identify the logo(that means other users had not been focused on it before), the application invite users to help us to identify the logo by typing the name of the company.
7. User can add the comment if they want.

Major Milestones

The following marked some of the major milestones of the project.



Phase 1

Main Objective- to get a basic understand of the behind technology of Augmented Reality

1. Research on the development of the Augmented Reality
2. Research on the present successful AR application
 - Find out the key of successes
 - Find out the things needed to be improved

Phase 2

Main Objective- to develop the server side architecture

1. Implement Image-Recognition Algorithm(SURF)
 - Image-Recognition Algorithm is the most important algorithm that employed in the server. This algorithm determines the accuracy of the image-matching as well as efficiency, and so our team decided to develop it first.
2. Develop and build Cloud Database Server
 - Online searching property is our key focus, we will make use of machines to build a server using TCP/IP for accepting and responding requests

Phase 3

Main Objective- to develop communication between server-side and mobile OS

1. Implement Open-Source Vuforia on mobile OS
 - Vuforia is an open source AR application that develop in iOS and Android, we will make use of its functionality to play videos in the matched images.
2. Develop communication channel between server side and mobile
 - Before develop a real AR application, we would like to test the connectivity between the server-side and mobile device. Our team will develop a communication channel to guarantee high efficiency of data transmission

Phase 4

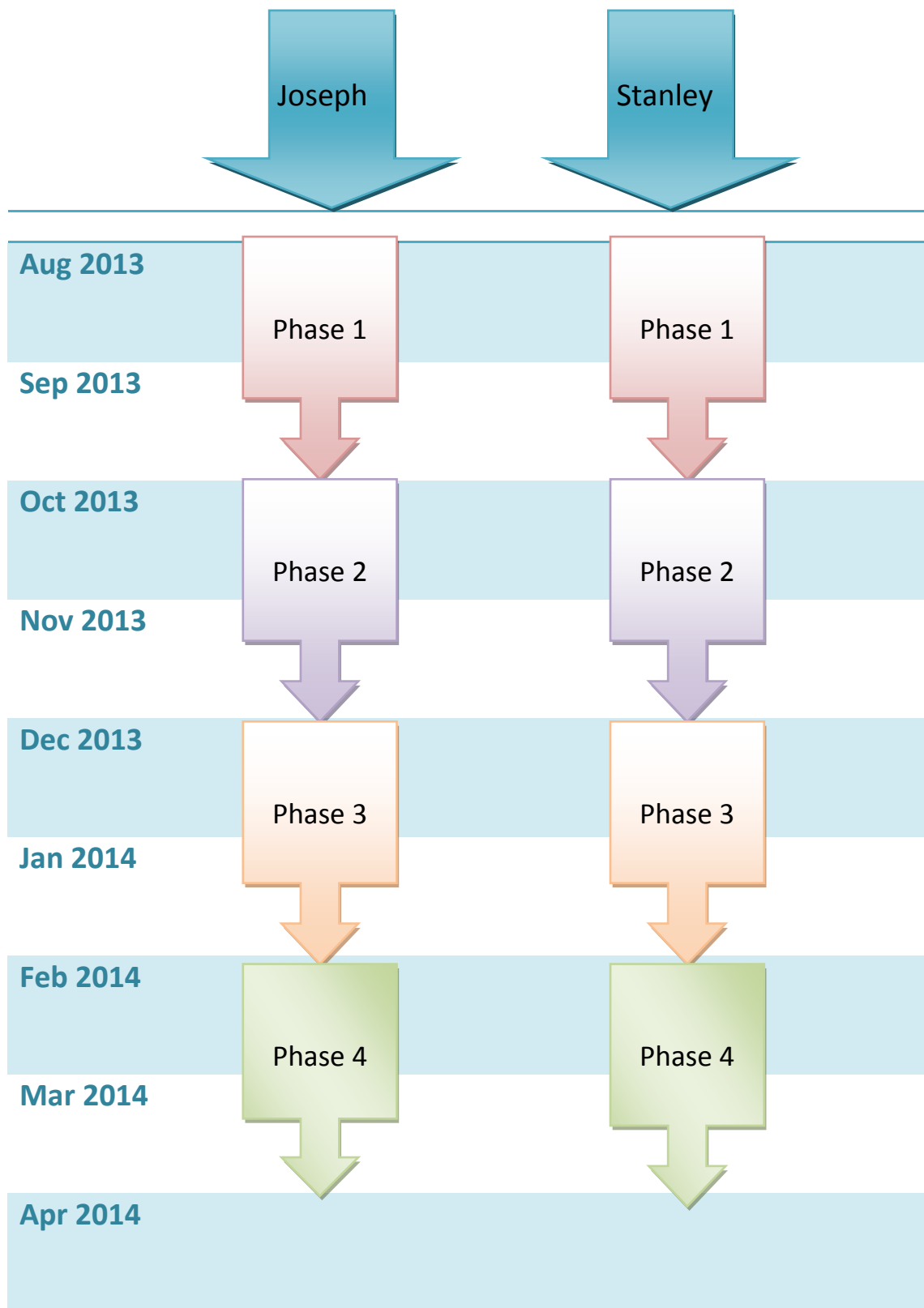
Main Objective- to develop AR application for demonstration

1. Design data-distribution algorithm for the loading of database
 - To prevent database being overloaded, we will develop an algorithm to distribute the workload to the database server evenly and so as to prevent broken down of server.
2. Develop RA Application
 - Our team will develop a useful AR application to demonstration the power of our engine.

Division of Labor

	Phase1	Phase2	Phase3	Phase4
Joseph	Research on the development of the Augmented Reality	Develop and build Cloud Database Server	Implement Open-Source Vuforia on mobile OS	Design data-distribution algorithm for the loading of database
Stanley	Research on the present successful AR application	Implement Image-Recognition Algorithm(SURF)	Develop communication channel between server side and mobile	Develop RA Application
All	Deliverable-Project Plan, Project Web page	N/A	Deliverable-Preliminary implementation, Detailed interim report	Deliverable-Finalized tested implementation, Final report

Schedule



Budget Plan

In this project, each students have been granted \$1000 from the Department of Computer Science, so our team have a total of \$2000 as our total budget. Throughout the development process, our team will spend the money on

	Unit Price	Quantity	Amount	Total
Income				
Budget from CS Department	\$1000	2	\$2000	\$2000
Less: Expenditure				
Server Rental	\$700	1	\$700	
100 Printing Unit	\$30	2	\$60	(\$760)
Total				\$1240