UNIVERSITY OF GOTHENBURG

Exploring Technical and Behavioral Aspects for Public Motion Controlled Applications' Development Using 3D TOF Camera

Abdul Waheed
Supervisor: Erik Borglund

Report no: 2010- 075
SSN: 1651-4769

Master's Thesis in Software Engineering and Management

Table of Contents

Exploring Technical and Behavioral Aspects for Public Motion Controlled Applications' Development Using 3D TOF Camera

Abdul Waheed
IT University Gothenburg Sweden
+46704780174
waheed@student.chalmers.se

Abstract

Context: Remarkable advances in multimedia games can be seen during recent years. In this thesis a novel way of playing games, merely using human body parts, is under consideration just like XBOX Natal concept [2]. To implement this idea, a newly made 3D Time-Of-Flight (TOF) camera is used to capture the motions i.e. gestures and postures of moving objects.

Objective: In this thesis two major research questions are addressed. One is to find out the possible usages and limitations of three dimensional (3D) motion controlled interfaces for Public Interactive Application (PIAs). Second one more technological question is to find out possible and feasible solutions and programming techniques to capture motion using 3D TOF camera.

Method: Qualitative research method along with Contextual Inquiry is used to gather and analyze data. Literature study, interviews and some explorations on relevant technology blogs (and forums) were made. Some inspirations were taken from Grounded theory method to analyze collected data.

Results: This thesis explores possible usages of TOF camera to develop motion controlled PIAs along with limitations with respect to the potential users. A detailed solution for 3D motion capturing is suggested, which consists of the Visual C++.NET, Win32 Application and OpenCV libraries synchronized by EmguCV wrapper.

Conclusion: This report provides a detailed study about the possibilities and limitations of using motion controlled interfaces in public areas. Valuable ideas providing extensive usage and future fertility of this nascent technology have also been described. Findings in this report can provide constructive principles for development of public interactive applications.

KEYWORDS

PIA, Motion Controlled, Time-Of-Flight

Definitions

Public Environment

Places and environments which are in access of common public e.g. airports, universities, city centers.

PIA

Public interactive applications, motion controlled applications to be installed in public environment

3D Motion

Three dimensional motions of objects which will be placed in front of camera

Motion Controlled Games

The games which would be played without using any controls e.g. mouse or keyboard, but to be played by capturing gestures and postures human body. Just like XBOX Natal games.

3D TOF Camera

A special camera capable of capturing 3D motions, this camera continuously emit the light which returns back after striking some opaque object. By calculating the time of light's returning back to cam the distance of object from camera is computed, which is known as time of flight (TOF). In this way this camera is competent enough to tell the range of object and pixel depths in the form of third dimension.

SV Systems

System developed by a number of 2Dimensional cameras to view and capture the motion.

LRS

Laser range scanners, 3 dimensional cameras equipped with laser beams provide 3D images with low intensity.

Avatar

The cartoon sketch of body or object which reflects physical features of real object in a specific morphological way.

Motion Capture

To capture motion of moving objects with the help of cameras. A computer vision term.

1. INTRODUCTION

This report is about the master's thesis project which has two-folded purpose, one is to find out the possibilities and limitations of bringing (introducing) the multimedia interactive applications into public environment e.g. airports,

city centers etc., and the second one to explore and devise some enduring technical (programming) solution to develop motion controlled multimedia applications by using Time-of-Flight (TOF) camera. The proposal of bringing multimedia applications into public places is quite new. Moreover the development of 3D public interactive applications using the TOF camera is also a novel and innovative paradigm. To make research and development (R&D) upon these unique and fresh ideas comprises the whole thesis work. This report contains functional and practical results of this dual natured research. These results assure that the usage of TOF camera to make public interactive applications can bring a new and easy way of interaction between human and multimedia applications for entertainment and information purposes.

Moreover in this chapter theoretical background has been discussed along with vision of whole project and research questions.

1.0 Vision and Purpose

This section contains vision, scope and research questions (RQs) of the project. If the scope of project is considered in the broader sense, it has the room to contain the implementation of quite diverse ideas of future applications. As it is discussed earlier that the idea of developing public interactive applications is fairly new, likewise the idea of using TOF camera in development of PIAs is ultimately novel. So the successful implementation of such modern notions can bring new concepts of public entertainment as well as new ways of interacting with multimedia applications. Although the title of thesis is not fairly self-explanatory as the key idea of motion control by 3D TOF camera is surrounded by numerous diverse notions i.e. using augmented reality (AR) and mixed reality concepts to make motion controlled interaction with applications more real. Likewise construction of multimedia physical games and informative applications containing 3D interaction are also amongst the chief goals of this challenging initiative and research. Still no complete formulated programming or software solution is available to use this technology for development purposes. Hence primary goal of this work is to research upon the future of above mentioned technologies to be used in public areas and to implement the concept with the help of newly constructed TOF (Time-of-Flight) 3D camera [5].

The idea is to install 3D public interactive applications (PIAs) i.e. games using the above said technologies. The first goal of the project is to find out the user behavioral aspects about the usage of this new technology in the form of entertaining games in public areas. And in first part following questions are addressed

a. What are the limitations and possibilities of using motion controlled interfaces and applications in public areas?
b. For what concepts are motion controlled technologies best used for?

Besides that to develop and implement PIAs technology wise research has also been made. It has been explored that which software and technical solutions are appropriate to develop such applications. While finding and devising optimal development solutions, the initially faced problems have also been described. For example lack of relevant material, unavailability of camera at the start, unclear and uncertain scope of project and less technical support were the remarkable barriers.

To implement the concept following question are addressed

c. What software tools and technologies should be used to process and use 3D data?

d. Using what software architecture and platform performance can be optimized?

e. What software process model should be adapted (or devised) to develop PIAs using the PIAs? (optional)

To take the input (3 dimensional data) from TOF 3D camera efficient system is required which may capture and transform data into useful information so that it might be used in developing applications, is the key purpose of questions c) and d). The data should be processed in a way so that the third dimension (depth of pixels, which is an additional benefit of TOF camera) in a video stream should be used for useful purposes. By using the third dimension, efficient interactive applications should be developed. After efficiently processing 3D data it should be translated into a comprehensive mode. To achieve these purposes adequate software techniques which must be compatible with the TOF camera are required.

Somehow the first task is to track different object placed in front of TOF camera by using some efficient algorithms. Further using the camera input by means of software we have to draw the avatar of the captured object.

The last research question addresses that working on such a research project (which requires plenty of R&D and have not been developed before) some special software process model should be adopted which may assure that the research made for one application should be used for development of other applications of same kind. To address this RQ a suitable software process model has been devised

which may promise not only the reusability of developed common components but also assure to document all the research made for reusability purposes (whether it was implemented or not).

1.1 Background

During the last five years we have seen an outbreak of different interaction technologies. The recent advances in multimedia applications have proved that the mouse and keyboard are no longer the only ways to interact with digital experiences such as games and information applications etc. Console games have been driving the development with the Playstation EyeToy concept [1] and the Wii technology [3]. These applications are however mostly interacted and played in an environment where the player can explore the technology in a safe and controlled way. In such environments these technologies have to use different kinds of peripheral devices to interact with the system, the Playstation and Wii devices have been shown in the pictures from left to right respectively.

Fig 1.1 PlayStation 'Move' **Fig1.2 Wii Remote**

Although a number of fascinating 3D games have been attaining significant attraction amongst different age groups, but are played in indoor environments. Soon these technologies will reach the masses and will be interacted in public areas as well.

Previously the Nintindo has introduced very good idea of 3D motion controlled games by using some specific controls. Now this is the time to bring the games into more real world scenarios by providing users a feel of playing and interacting with a real world. For this purpose the idea of (motion controlled games) playing by using body parts is being introduced. The idea is being implemented by capturing gestures and postures of user's body. To make the users free from all sort of peripheral controls it is necessary to make a well-built system which may capture and translate the body gestures and postures into digital signals in a consistent way. Previously and even in recent for 3D games 3D visions and scenes of real world are being created and captured with the help of Projections and secondly by creating the stereo depth perceptions by using multi-camera systems [6], whereas this thesis report explains the usage of Time-of-Flight camera (to capture 3D data) to develop public interactive applications.

1.2 Thesis Outline

Chapter2 contains the detailed research context of project. Chapter3 elaborates the related research work as well as explains who can be the readers of this report. Then in Chapter4 detailed research design is elaborated. Whereas in Chapter5 it is explained that for which concepts this technology can be used and who can be the potential users of PIAs, besides that all the limitations as well as possibilities of usages of PIAs has been described. Likewise in Chapter6 the results for programming solution and formulated software architecture are given in detail. Chapter6 describes that what integrated development environment and programming languages are best suitable to develop applications under consideration. Moreover in Chapter7 the proposed and adopted software process model for this project is discussed. Conclusively Chapter8 contains the suggestions, future work and conclusion of whole research.

2. RESEARCH CONTEXT

2.0. Principle Idea

In this chapter the whole idea of the project has been described. Besides that some of the obstacles in fulfilling the idea (in completion of project) have also been given in this chapter. The key idea is to install motion controlled public games at Skellefteå airport, Sweden like the XBOX natal concept by using the 3D TOF camera (equipped with cheap CMOS chip) for motion capture which is keeping this notion one step ahead as compare to other technologies of this family. Currently the idea of installing and to transform that 3D data (which is provided along with great much noise due to nascent technology of camera) to make it part of an augmented reality scenario and to draw the avatars of objects placed in front of the camera is the overall scope of the project. Being part of this project, my responsibilities are to find out the technical hurdles and barriers which may occur in processing the 3D data input given by TOF camera while capturing the motion. To translate and to process the data in comprehensive way requires adequate software techniques which should be compatible with the TOF camera, is the challenge of this thesis. Furthermore to find out the limitations and possibilities of using motion capture interfaces in public areas is also part of the thesis research work.

2.1. Technology: Better and Cheap but Nascent and Immature

As Houseman et al [5] explain in principle there are 3 methods used to acquire 3D information: Stereo Vision (SV) Systems, laser range scanners (LRS) and time-of-flight (TOF) cameras. These cameras are being development by using CMOS chips along with PMD (photonic mixer devices) technology.

The SV systems have flaw of limited field of view (FoV) and allocation problem as well (as mentioned in [5] quoted in [7]). Likewise the LRS are not capable of providing 2D intensity images and range data at same time (as mentioned in [5] quoted in [8]).

The core premise of TOF camera is that the camera continuously emits a light of certain frequency, with the help of this beam of light the range and distance of object is calculated by computing that how long the light photons took in returning back to camera after reflecting from the object. This camera provides the intensity images and range data with the pixel depth at the same time. TOF camera can provide 2D images of good quality and in addition can provide good vision on scenes in 3D sense along with motion tracking information which is the major challenge of this project.

2.2. Considerable Barriers

- **No Possibilities of Usability Testing**

At the moment it is quite difficult to bring scenario in the comprehension of users to get their opinion and feedback. It is also no easy to develop some simulator or prototype of users.

Some observations and some surveys to get the information qualitatively at the airport area will be made. 3D model of airport is being developed in house by the designers in company by discussing the ideas collaboratively, and system installation design is being developed in the form of 3D conceptual model.

Fig 2.1: 3D conceptual model of installed system

- **Undeveloped Camera Technology**

The TOF camera is quite novel device which provides the 2D image with range (pixel depth) in the form of 3rd dimension which tells that how far is some object. This will be used to capture the 3D motion of objects. This camera provides movies not in traditional AVI format but provides arrays of frames in PGM format. Soon we will be provided the camera which is capable of providing the 160X120 resolution along with 3rd dimension in the form of distance of object from the focal point of camera. First barrier in this regard is to filter the data by removing the noise (currently a lot of unwanted pixels are there with desired data).

TOF camera continuously emits a light of certain frequency, with the help of this beam of light the range and distance of object is calculated by computing that how long light photons took to return back to camera after reflection. This process is not very much mature yet and it causes to save a lot of noise pixels. Currently no specific algorithm or computer language APIs support to filter this data efficiently.

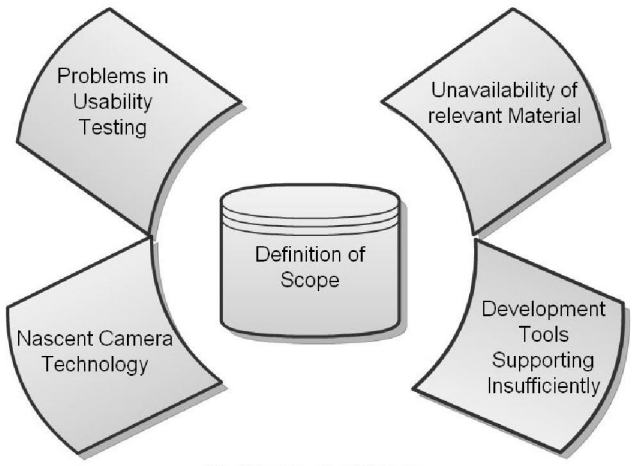
- **Unavailability of Relevant Material**

Standard literature in the form of research papers is not available specifically for motion control by TOF camera. However some literature can be found about the manufacturing and mechanical aspects of camera. Literature is available about motion capture by traditional ways of 3D vision is also available.

Although some websites and blogs are quite helpful in bringing out the better ways towards solutions, i.e. [4] is good source of relevant information.

- **Development Tools Support Insufficiently**

The tools and algorithm routines available for the image processing and motion capture are not purely designed for data input of TOF camera. OpenCV supports the 3D vision which is developed by 2D cams using the Stereo systems techniques.



Challenges at a Glance

Fig2.2. Major Barriers

All these barriers were supposed to bring a grinding halt to thesis in some stage, however the findings and results till now are still optimizing.

3. RELATED RESEARCH

3.0. Relevant Work

This chapter includes the summary of previous related work along with some relevant theoretical details. For questions (a) and (b) which are about finding the feasibilities of motions controlled interactive applications in public areas, no previous research or development work exists. However one gadget related to public gaming can be painfully seen which is "Urban Screens Festival and Conference 2008" in Melbourne [20]. This event was managed by developing some games to be played by mobile phones interacting with the screens installed in public places.

As for as the previous research for questions c), d) and e) is concerned, there is no accomplished programming solution or directly relative techniques to address these questions are available. Couple of completive i.e. Primesense [16] and Canesta [15] are working on these ideas but no final product or research results have been shown in public yet. Although at the end of previous month in April 2010 Canesta uploaded three articles about using the 3D Camera for technical purposes i.e. 3D data filtering, tracking and some of the background subtraction algorithms have been provided in detail in those articles. Brief summary regarding the those article is given following

- "A Time-Of-Flight Depth Sensor – System Description, Issues and Solutions" [21], this article explains some basic issue like noise and compatibility problems while handling the third dimension in captured data.
- "Real-time Foreground Segmentation via Range and Color Imaging" [22]. As this article is fairly explanatory by its name, explains how to extract the different segments out of a stream of frames and some ideas about embedding virtual reality in it.
- "3D Head Tracking Based on Recognition and Interpolation Using a Time-Of-Flight Depth Sensor" [23]. This article describes some of the head tracking algorithms to be used in 3D data.

Currently two other known companies are working with the TOF camera however no application has been development completely by any vendor.

- First one is Canesta [15] which is manufacturer of TOF camera as well, is making 3DTV interface by using TOF camera. Canesta uses the cheap single CMOS chip in TOF (Time of Flight) camera to transform light signals into electric pulses. The 3D TV interface will allow users to interact with the TV without remote control. And TV channels will be changed just by waving hands in air.
- Primesense [16] is another vendor who is working on 3DTV interface by using TOF camera this is the vendor who is providing the 3D sensors to Microsoft for project XBOX 360 Natal.

3.1. Who Should Read This Report

The user experience designers as well as developers who want to develop or design motion controlled applications using the 3D TOF camera can have plentiful practical knowledge after reading this report.

The designers who are looking for the users' behavior against motion controlled interfaces especially in public places can have good reasons to read this report.

Besides that the programmers and application developers (specially related to game programming and computer vision) can have beneficial information about capturing the 3D data using TOF camera. This report provides a vision about the future of interactive applications in public environments.

4. RESEARCH METHODOLOGY

To address the research questions explained in Chapter 1, the detailed research plan and research methodology has been described in this Chapter. Qualitative approaches for data gathering and analysis have been adopted i.e. Observations, Contextual Inquiry, Literature study, inspirations from Grounded Theory method and exploration of different forums and blogs are the considerable tools.

4.1 Research Design

To address the research questions (a) and (b) the observations and literature study is carried out. The Contextual Inquiry [13] was conducted at two public places. One place is airport Skellefteå (Sweden) and other place is Campus Skellefteå (Sweden). User centered design was taken into account for contextual inquiry and couple of interviews were also conducted.

Moreover to learn about the potential users, other public places (instead of airport) are also figured out to find the feasible places for interactive installations.

For research questions (c), (d) and (e) (which are related to software technology) literature review and interviews are conducted. There is no significant amount of relevant material available, which actually explains the development and implementation of such system. Therefore exploratory study is conducted as Stebbins et al [11] explain that when we do not have any clear idea what and how to find then exploratory study is done.

Furthermore to gather and analyze data about our inductive ideas, some inspirations were taken from grounded theory (GT) method of qualitative research because the ideas were novel. The way of implementing ideas was undefined and unclear that's why a firm theory about the usability and implementation of technology was fairly needed to be built.

According to Martin and Turner [12], grounded theory is "an inductive, theory discovery methodology that allows the researchers to develop a theoretical account of the general features of a topic while simultaneously grounding the account in empirical observations or data." [12]

Grounded theory approach implies 'continuous interplay between data collection and analysis' [10][12] Grounded theory seeks to get information that is grounded in collected data instead of existing theory [9]. Keeping in mind different coding schemes of GT the gathered data was analyzed to get results.

Furthermore to resolve the several technical problems, different Computer vision and programming forums and blogs were browsed and explored until the favorable outcomes were derived. Research on different issues is still going on.

5. POSSIBILITIES & LIMITATIONS OF MOTION CONTROLLED INTERFACES IN PUBILC

With the help of observations, literature study and Contextual Inquiry the research question a) and b) are addressed here. A couple of public places were studied and analyzed qualitatively to envisage the design constraints of future applications. A Contextual Inquiry was made at Skellefteå, Sweden airport as well as at University Campus of Skellefteå, Sweden. To better design the application this is very essential to study and observe the users in that specific environment where they are supposed to use the applications. David Woods, et al. [19] states "Designs are hypotheses about an envisioned world". For installment of motion controlled interface game is airport of Skelleftea, Sweden, airport was visited three times to gather the data following the contextual enquiry.

a. Once on weekend to examine the flight and passenger data and receivers at airport who can come to receive the other passengers.
b. Once airport was visited on long holidays start i.e. at start of Easter when people travel to move to their home towns.
c. Once airport was visited on a regular working day.

In all these three different kinds of days we could find different sort of people travelling. We found the receivers while waiting their guests were getting bored and they can be the potential users.

A number of new discoveries came into acknowledgment i.e. during working days mostly travelers were of age 30 to 60 years old and they were more interested in informative, news and serious applications instead of more thrilling games. Conclusively it can be stated that in working days the applications should be serious like some interactive news on 3D TV instead of games or entertaining ones.

Whereas on weekends and Special holidays i.e. at the start and end of Easter a number of students and less than ten years old kids were also observed at airport with big proportion and they seemed more interested in thrilling and adventurous games on big screen. Besides that on weekends and especially at starting and ending of long holidays (a day before starting of Easter holidays was observed) great many young students also travel. And for young travelers entertainment sort of games and applications can be attractive use of technology.

On the other hand some kind of CI was conducted in University campus and students were observed daily for a month on numerous places inside the campus. Following places to install motion controlled applications were analyzed to be suitable venues i.e.

a. Cafeteria inside the campus was considered to be a great place for adventurous as well as for some puzzle games to be interacted merely by body parts could be fun.
b. Small hall adjacent to the sports hall could be made an attractive place by installing motion controlled games to be played by intensive body kinematics, which can provide good exercises to students.
c. Lobby beside the library can be an attractive place for students by putting some informative applications. Some 3D TV screens equipped with motion controlled applications can provide a trouble-free search facility to students. Student can interact with such applications to discover

numerous categories and groupings of books with 3D maps.

Interviews were conducted with different users to ask about their opinion about the PIAs and the following things were considered during the interviews to learn about the users and environment.

a. Previous users' experience
b. Domain knowledge
c. Age
d. Would you like to play the multimedia games
e. The location within airport where users spend more time or passes by more frequently
f. Average time they spend on airport
g. What can engage them, or persuade them to stop
h. Where else they have to wait besides the airport (can help to find other areas where we can install this prototype in future)
i. Which places are least visited by users
j. Traveling alone or with someone (how different people spend their time when they are alone and when they are in group)
k. What are the key work tasks these people perform that you might want to support?

Consequently following ideas can be taken into account to use this technology which can be more enhanced while working on more practical lines:

a. Using the augmented reality in captured 3D data, it can be used in different public places like museums, and historical places to create conversation between the past and present. And people can interact with such mixed reality scenarios.
b. By using motion controlled interfaces i.e. using body parts to interact with some mixed reality scenario people can interact with the past and history. They can be provided to discover the historical things in real time by interacting with some mixed reality scenarios.
c. By developing digital games with motion controlled interface for physical exercises, taking into account usability (for all age groups) and technical perspectives quite prolific applications can be developed.

6. EXPLORATION OF TECHNICAL & PRORAMMING SOLUTIONS FOR APPLICATION DEVELOPMENT

This chapter explains in detail the explorations made to find appropriate software and programming solution which may provide optimal bases to develop motion controlled PIAs with TOF camera. The applications and concepts which are considered possibly good to implement (as discussed in the chapter 5) have been taken into account to develop in this chapters.

As far as motion capturing in 3D development with SV systems concerns, the concrete software solutions are available in literature with best coding techniques outfitted with excellent algorithms. However to capture motion with the 3D TOF camera ample amount of study material is not available, typically in context of software and programming solutions for reasonable learning of this young domain. Instead rarely available literature about TOF camera merely addresses the physical and architectural issues of TOF device with respect to its performance, electronic circuitry etc. For example Husmann et al [8] explains the performance improvement of TOF camera considering the sensors, extrinsic and intrinsic parameters and other hardware stuff. Likewise Bothe et al [14] throws light over the tangible features of TOF camera i.e. compactness, and in more detail explains the fringe projection techniques and elaborates its high gradient capturing capabilities while grabbing a 3D scene.

Therefore it is need of hour to go one step ahead to invent and publicize the programming and software solutions for motion capturing with TOF camera to extend this revolutionary idea in making software applications.

A step by step development of suitable software architecture has been described in this section. The construction of architecture flourished gradually by testing numerous tools and technologies. Modifications in architecture have been taking place. Once in the whole project whole software architecture was changed and redeveloped when a completely new and modified camera driver was released with new artifacts.

6.0. Outcomes: Working with Two Dimensional Camera

In first step the openCV with numerous algorithms for image processing and motion capturing in two dimensional environments. Ordinary Two dimensional cameras were used to take input. Secondly OpenCV was configured with the .NET framework using the win 32 application with Visual C++. Because during initial couple of months of thesis work TOF camera was not available, therefore work was progressed with the two dimensional camera and different computer vision algorithms for object detection were tested e.g. Canny edge detection, and contour detection etc.

using cvCanny and cvThreshold functions. with the help of OpenCV libraries [6]. These results were tested successfully by tracking the video streams as well as still images on color and grayscale both formats.

6.1. Outcomes: Initial Working with First Release of TOF Camera

Then we switched to TOF 3D camera. Its first version of artifacts explains that the camera should be registered and implemented with the .NET framework using the Windows Form Applications with Visual C++. To cover up the more features of motion capturing techniques OpenCV was configured with the Windows Form Application in Microsoft Visual Studio 2008. The architecture could have the following picture.

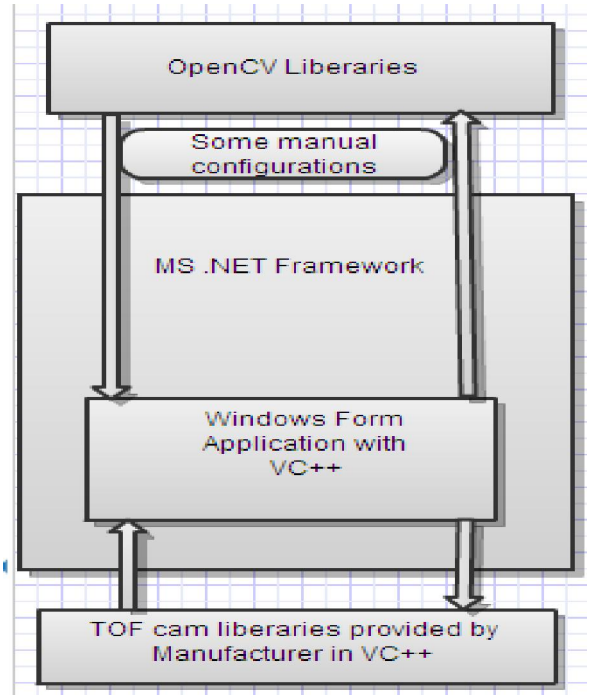


Fig 6.1 Integration development environment configuration diagram

While resolving the issues occurred in configuration of .NET Framework and OpenCV libraries, it was explored that a new version of a plug-in called EmguCV came into appearance (in October, 2009). This version of EmguCV was known as "the first .NET wrapper that is compatible with OpenCV 2.0". "Emgu CV is a cross platform .Net wrapper to the Intel OpenCV image processing library. Allowing OpenCV functions to be called from .NET compatible languages such as C#, VB, VC++, IronPython etc. The wrapper can be compiled in Mono and run on Linux / Mac OS X". [17] By using this developed platform simple computer vision applications were made to check out the efficiency of camera. It could not work better with different filter and image processing algorithms because there was a lot of noise data along with necessary data.

Some other wrappers are also available e.g. SharperCV and OpenCVDotNet etc. However EmguCV due to following capabilities was chosen proved to be more compatible.

A Comparison of OpenCV Wrappers

Name	Emgu CV	OpenCVDotNet	SharperCV	Code Project
Cross Platform (Mono)	√	X	X	X
OpenCV 2.1	√	X	X	X
Machine Learning	√	X	X	√
Exception Handling	√	X	X	X
Debugger Visualizer	√	√	X	X
Actively Maintained [1]	√	X	X	√
License	GPL or Commercial License with a small fee	Non-commercial GPL	Non-commercial Academic use	Non-commercial

[1] "Actively Maintained" refers to update within one year.

Fig 6.2 Comparison of EmguCV with other wrappers

By configuring EmguCV with OpenCV and .NET framework the whole development environment got the following shape.

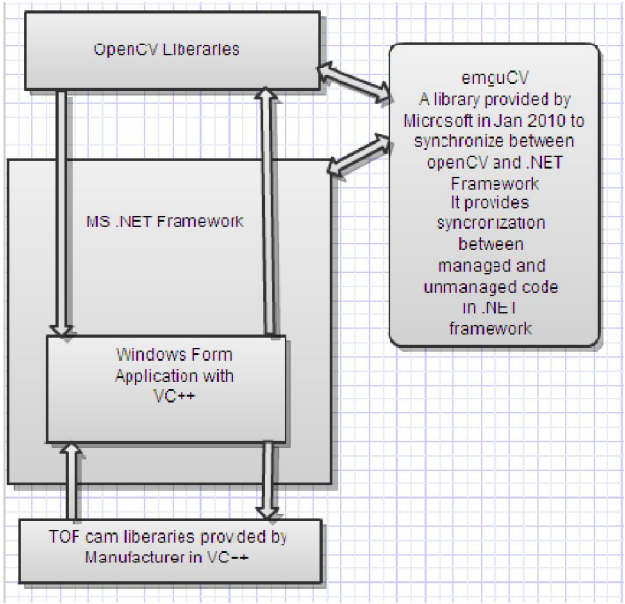


Fig 6.3 Integration development environment configuration diagram

In this well settled framework couple of image processing executions was done like changing converting managed bitmap data into unmanaged data and vice versa by utilizing OpenCV libraries. Background subtraction was done and other explorations were consideration with the 3D data provided by TOF camera when we got new version of camera driver with new programming solution.

6.2. Outcomes: Working with Second Release of TOF Camera

Working continued over a month upon the first release of TOF camera driver and application provided by Canesta&Optronic[15][18] . That application to handle the camera was developed based on the Windows Application Forms in .NET frame work using Visual C++.

Later on a new version of Camera driver with enhanced efficiency was supplied by the Fotonic a department of Optronic dealing with TOF camera.

This new release has been developed by Optronic [18] using the Win32 APIs in Visual C++ supported by the bulk of DirectX9 components to address the 3 dimensional issues. Since this entirely changed the structure of programming application for camera driver so it enriched the results i.e. frame rate was enhanced from 40 fps to 60 fps (frames per seconds), 3D data rendering is fast and noise has been removed up to some extent. Depth range was also enhanced. On the other hand a very complex data structure has been used; to make modifications in this programming solutions to apply different algorithms of image processing and 3D data filtering seems far difficult than doing on previous solution. As these executions may lead us towards realization of motion capturing in 3D environment therefore it is essential to make this application (which is developed in Visual C++ with DirectX9) compatible with the image processing and motion capturing algorithms' rich library i.e. OpenCV. Currently this application is getting 3D data from TOF camera as input in the form of matrices and drawing on the 2D monitor screen by making TrinangularLists and PointLists. Now currently we are making the application which may apply some image processing and data filtering functions upon matrices containing 3D data.

Following diagram shows both the tracks followed by us during exploring the First and Second releases of TOF camera driver with different programming solutions.

After successfully configuring and testing the mentioned platform the first main application yielded is background subtraction. Background subtraction was done using the TOF depth sensor. During separation of background and foreground in a color stream, the basic problem which was faced is segmentation of those portions of foreground in correct way which portions we want to keep [22]. The method of background subtraction using the depth sensor is given in literature provided by Canesta [15] who is manufacturer of TOF camera and working on different applications. The distance of scene from camera was defined as third dimension and a threshold was defined for the foreground pixels. All the pixels taken at certain depth (coming threshold area) were considered as foreground pixels and rest of them were considered as background pixels. With the help of depth when the pixels were identified as the foreground pixels, then by comparing the average depth of foreground pixels with the uncertain pixels it was decided which pixels should belong to foreground and which ones should be considered as back ground pixels. Mathematical probability was used (devised by R. Crabb et. al) to decide about the surrounding pixels to be included in foreground or not [22].

Then by using *erode* and *dilate* (cvErode and cvDilate) filters given in OpenCV library [6] the output was refined by removing the noise data.

Likewise using the OpenCV functions (e.g. cvCheckContourConvexity, cvConvexityDefect) for convexity recognition along with depth of pixels the hand tracking was done successfully.

Luca-Kenade, Laplace, Erosion algorithms will be used to filter the data using the APIs provided in OpenCV library supported by C/C++ [25,26].

The method given in [23] was used to develop the depth signatures of objects to track and capture the motion. With the help of Hand Tracking module numerous applications can be developed which can be interacted by waving hand in air without using any control.

Two simple projects were implemented during the thesis work. Which are very good examples of public interactive applications with 3D TOF camera.

One of them is “Island Demo” while the other is “Hand Fishing”. Author majorly participated in development and programming part of these projects.

In the “Island Demo” project an UFO (kind of alien ship) was flown over the animated island. The motion of flying UFO was captured by users by waving their hands in air in three dimensions, while standing in-front of big screen. This application was exhibited in public hall during the Creative Summit 2010 [27]. The demo video of project is available in this link [28].

The idea of "Hand Fishing" is to put a screen in public area. Screen will show some animated flowing water with fishes. User will move her/his hand (in air) in a specific region in front of screen (to catch the fish in animated screen). When the depth pixels (co-ordinates) will match with the scaled pixels of fishes in water, system will show that user has caught the fish. This application is titled as Hand Fishing.

A useful library of functions has been developed i.e. hand tracking, image format conversion, and background subtraction along with some filters. These functions are for specifically working with the 3D TOF camera. By using these functions a firm base for numerous interactive applications can be developed in quite less time.

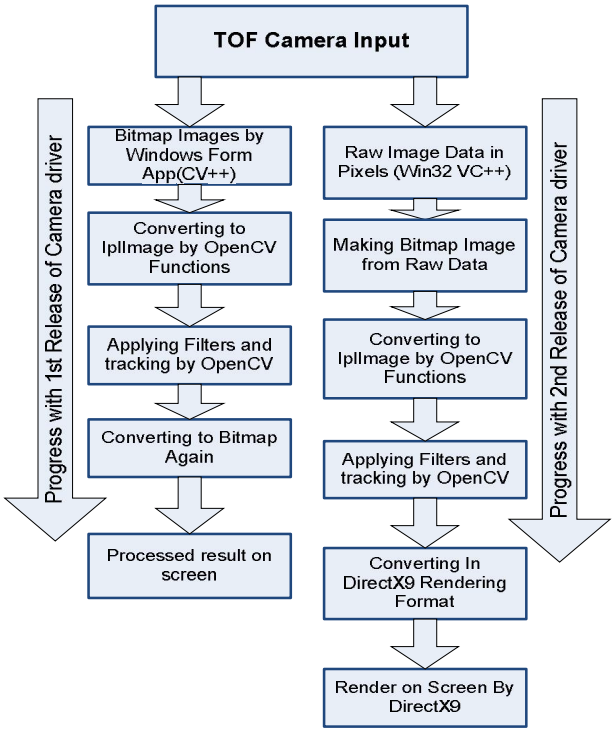


Fig. 6.4: Two R&D tracks followed with both the releases of Camera Drivers.

6.3. Summary of Technical Solutions

After configuring the integrated development environment which is compatible and suitable for working with the TOF camera a library of functions has been developed which would be used in a good number of interactive applications. Following functions have been developed.

a. Background subtraction using the depth sensor (using third dimension)
b. Hand tracking
c. Different filters, e.g. smoothing, edge detection and contour detection

Above mentioned functions can provide a good base to the development of any motion controlled interface (interactive application).

7. DEVELOPMENT OF PIAs: PROPOSED SOFTWARE PROCESS MODEL

Finally some suggestions have been delivered here about the software development life cycle of such and new domains which require loads of R&D. A software process model for the development of future interactive applications have been devised which helps in reusability of research results. By using the steps proposed here the rapidly developed applications can be yielded. This process model will help in development of such applications which need a complete cycle of research to find out possible programming solutions before starting development (and even during the development).

This development model is derived by making some modification (addition) in Software Product Line Engineering (SPLE) model. An amendment is suggested here that the Reference Architecture of Software Product Line Engineering [24] should have a capability for a research module. As SPLE captures 'Commonalities' and 'Variabilities' among different software products of same domain [24], so to make SPLE more functional for pure research based development projects a new phase of '*Research Results Engineering*' should be positioned as a sub phase of Domain Engineering Phase. During the development of one application of some specific domain following steps should be taken into account to perform this newly proposed phase.

a. This step should be performed after analyzing the common and variable requirements of numerous applications of same domain during the Domain Engineering Phase.
b. While finding out numerous solutions and suitable tools during research, it should be decided that which research results are best choices. All the found results should be categorized into two following classes by hit and trial method during development.
    i. Best found results
    ii. Other found results

    Then all the *best found results* should be documented thoroughly that how and for which applications (and scenarios) these are the best solutions to choose.
c. Besides best found results all *other found results* (solutions) should also be saved with appropriate documentation instead of ignoring them. Because *other found results* can be used in development of other applications of same family (probably in future).

d. It should be documented that for what causes *other found results* were not chosen in solution development.
e. It should be figured out that which other applications of same family can be developed by using the *other found results*. This can be attained by running following threads of analysis
    i. Re-analyze the common requirements and variable requirements of different applications belonging to same family.
    ii. Analyze the best found results and other found results while taking into account the common requirements of different applications of same domain or family.
    iii. Analyze the *best found results* and *other found results* while considering the variable requirements.
f. By comparing the requirements of different applications with other found results it should be figured out the in future which *other found results* can be appropriate for different applications of same domain. All the above said analysis should be documented. In this way most of the research results can be reused.

[***Example:*** for example during this thesis work two simple but different applications were developed which are "Island Demo" and "Hand Fishing" both belong to same domain which is "PIAs with TOF camera". For both applications hand motion was required to be captured. Different techniques and algorithms for hand motion detection were tested e.g. contour detection, edged detection, image erode, image dilation etc. Finally it was figured out that for island demo the noise reduction methods like *Erode* and *Dilate* were needed but for the other application "Hand fishing" these two methods were not required for hand detection. Consequently *erode* and *dilate* proved to be the *best found results* for "Island Demo application". Whereas the *edge detection* appeared as *other found results* for this application to use.
On the other hand edge detection method was very fruitful for hand detection in "Hand Fishing application" and proved to be the *best found result* for hand detection.]

Hence in this proposed phase it should be analyzed that which results (research) can be used for further development in future and how much research can vary according to variation in domain. In the following figure new phase is highlighted as green color box.

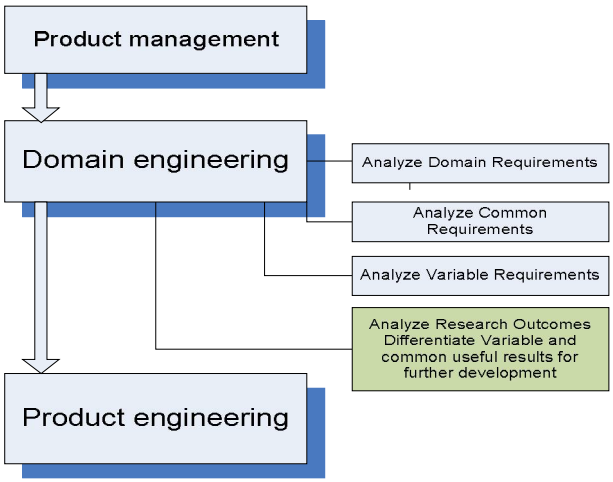


Fig 7.1 Putting Research Engineering Phase in SPLE

This adaptation will be suitable for the domains which have not been developed before. Like the advances in motion controlled applications with TOF camera are quite new so the research results for development of such applications should be made reusable. In this newly added phase (in SPLE) it will be analyzed that how much common research can be used for development of other applications of same or similar family.

8. CONCLUSIONS

All the research and development conducted during this project had led us to three following conclusions in accordance with our research questions.

8.0. Future of Technology Under Consideration

In this thesis report with the help of observations, contextual inquiry and interviews it has been concluded that the development and installation of PIA's (public interactive applications) supported by 3D TOF camera would be a fascinating idea which can provide a new way of interaction amongst people on public places. Not only motion controlled entertaining applications (games) would be appreciated rather the informative and news sort of motion controlled applications would also get significant magnetism on public places. Results have shown that such applications should be simple instead of being complex and time consuming. These results can provide effective and informative start towards research of new idea of interactive multimedia applications to be installed in public for information and entertainment purposes. The RQ a & b was answered in the study carried out in Skellefteå airport and Skellefteå Campus, Sweden. In this paper we cannot state that these findings could be general to all other public spaces, but they can be transferable to other public spaces. However there is a need for further research in the area of

identifying the specific requirements the motion capture applications in public spaces will have.

8.1. Technical Solutions

Numerous tools and software platforms were searched out and configured in a successful way to develop a multipurpose functional library. This set of functions would provide and common platform in further development of public interactive applications. OpenCV was configured with VC++ .NET in Windows Form Applications as well as Win32 Projects. Different well known algorithms were explored i.e. Luca Kenade tracker [22], and some filters algorithms were used to track the objects for motion controlled applications. A small library of functions was developed to address general computer vision and motion capturing problems related to 3D TOF camera. These functions can provide a base to develop further motion controlled applications. These findings can prove to be quite constructive principles for development of public interactive applications.

Still an extensive research in this technical area is required. A number of more generic functions can be developed by exploring and composing more ideas about motion capture applications. Though this research can prove to be a worthy starting point to develop public interactive applications, still an elongated research is required to find out more optimal technical solutions within the evolution of more ideas.

8.2. Process Model Supporting: Reusability of Research

By having and deriving a number of technical and user experience related outcomes after the research software process model is suggested (addressing the research question e). This process model suggests that the product line engineering mechanism should be modified and enhanced in a way so that it may be capable of providing reusability of research results in further (solely research based) development. This enhancement in PLE can provide a good way to reuse the research of some pure research and development project for further development. This can bring a prolonged discussion in software engineering area specific to SPLE. Further research can make the reference architecture more and more flexible and open to reuse as more as possible research results in future applications' development.

9. ACKNOWLEDGEMENTS

Great many thanks to Dr. Erik Borglund, for his continuous guidance in both academic and theoretical issues throughout the research of Master's thesis. I am also very thankful to David Eriksson for making it possible to perform the study at Northkingdom and for providing all the facilities. I am thankful to Muhammad Fahied who worked with me on same project as an interaction designer and provided me user experience as well as technical guidance.

10. REFERENCES

[1] EyeToy, http://uk.playstation.com/games-media/games/detail/item103380/EyeToy%C2%AE-Play-Hero/ 11:00, January 19, 2010

[2] Natale XBOX, http://www.xbox.com/en-US/live/projectnatal/ . 12:30, January 19, 2010

[3] Wii Technology, http://www.wiitechnology.com/ .14:00, January 20, 2010

[4] Project Natal http://www.popsci.com/gadgets/article/2010-01/exclusive-inside-microsofts-project-natal 12:30, January 25, 2010

[5]S. Husmann, T. Edeler, I2MTC (2009), International Intrumentation and Measurement Technology Confrenece, "Performance improvement of 3D-TOF PMD camera using a pseudo 4-phase shift algorithm".

[6] Bradski, G.R., Kaehler, A., (2008) "Learning OpenCV: Computer Vision with the OpenCV Library," ISBN-10: 0596516134

[7] O. Faugeras, (1993) "Three Dimensional Computer Vision: AGeometric Viewpoint". MIT press, Cambrigde Massechusets

[8] S. Husmann, T. Ringbeck and B. Hagebeuker, "A performance review of 3D TOF vision systems in comparison to stereo vision systems," in stereo visio , I-Tech Education and Publishing, Vienna, Austria.

[9] Gasson, S. (2004) Rigor in grounded theory research: "An interpretive perspective on generating theory from qualitative field studies", The handbook of information systems research, Idea Group Publishing

[10] Coleman, G. and O'Connnor, R. (2007) Using grounded theory to understand software process improvement: A study of Irish software product companies, Information and Software Technology, Vol. 49, pp. 654-667

[11] Stebbins, R.A. (2001) Exploratory Research in the Social Sciences, Sage Publications, ISBN-0761923993

[12] Martin, P.Y. and B.A. Turner. (1986) "Grounded Theory and Organizational Research," The Journal of Applied Behavioral Science, (22:2), pp. 141-157.

[13] Greffeny, G. (2004). Contextual Enquiry - A Primer. 12:30 PM 2nd April 2010, http://articles.sitepoint.com/article/contextual-enquiry-primer

[14] Bothe, T. Gesierich, A. Li, W. Kopylow, C.v. Kopp, N. Juptner, (2007). 3D-Camera for Scene Capturing and Augmented Reality Applications. 3DTV Conference.

ISBN: 978-1-4244-0722-4
[15] Canesta TOF camera manufacturer using the CMOS chip, http://www.canesta.com/ 8:00 PM 4 April 2010
[16] Primesense, Natural Interaction, http://www.primesense.com/ 8:00 PM 4 April 2010
[17] EmguCV, http://www.emgu.com/wiki/index.php/Main_Page 2:00 PM 4 April 2010
[18] Optronic, http://www.optronic.se/content/About-Optronic/Default.aspx 2:00 PM 4 April 2010
[19] Woods, D.D. and Dekker, S.W.A. (2001) "Anticipat-ing the Effects of Technology Change:A New Era of Dynamics for Human Factors," Theoretical Issues in Ergonomics Science, vol. 1, no. 3, 2001, pp. 272–282.
[20] Urban Screens Festival&Conference (2008) in Melbourne, http://www.urbanscreens08.net/
[21] S. Burak Gokturk, Hakan Yalcin, Cyrus Bamji Canesta Inc. "A Time-Of-Flight Depth Sensor – System Description, Issues and Solutions"
[22] Crabb, R., Tracey, C., Puranik, A., and Davis, J., Canesta Inc. "Real-time Foreground Segmentation via Range and Color Imaging"
[23] Göktürk, S. B. and C. Tomasi. "3D Head Tracking Based on Recognition and Interpolation Using a Time-Of-Flight Depth Sensor"
[24] Ziadi, T., J. Jean-Mark, and F. Fondement, IRISA, "Product Line Derivation with UML". Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France
[25] Lucas, B.D., and Kanade, T. (1981) "An iterative image registration technique with an application to stereo
vision",Proc. 7th IJCAI, 674-679.
[26] Shi, J. and Tomasi, C. (1994) Good Features To Track, CVPR , , 594-600.
[27] Creative Summit 2010, Sweden.
http://www.creativesummit.se/
[28] 3D interactive Island Demo with TOF camera.
http://www.vimeo.com/12659176