ROBOGRAPHER

FACIAL EXPRESSION RECOGNITION USING SWARMS

CONCEPTUAL DESIGN REVIEW

SPONSOR: DR. KATIA SYCARA

TEAM G

JIMIT GANDHI

GAURI GANDHI

ROHIT DASHRATHI

SIDA WANG

TIFFANY MAY

Contents

1. Project description

What if photographers have a robotic assistants capturing happy moments in events such as weddings, birthdays or graduation ceremonies? ‘Robographer’ is the preliminary effort aimed at developing such autonomous assistants which not only click photos, but also recognize and capture human expressions accurately with equal competency. The principle of the project is facial expression recognition and accurate head pose tracking using a swarm of robots. Instead of working individually, a swarm of mobile robots will work collaboratively to accurately estimate the human expressions and click snaps accordingly.

By using multiple cameras, it is possible to improve the estimate of the facial expression and head pose through redundant noisy measurements and better handling of real-world issues such as occlusions. By adding mobility to the cameras, thus making each camera a dynamically actuated information source, it possible to further improve the estimate while also better tracking the human as they move through an uncontrolled environment.

The project Robographer has been conceptualized by common vision of a group of students who believe that robots can perform better when they work in collaboration with each other and also with humans while at the same time ensuring the safety and comfort of the humans in the environment. The goals of the project are as follows

1. The first component involves evaluating multiple cameras and applying computer vision and sensor fusion techniques to fuse the information from multiple static cameras using ROS and Intraface[1] to develop a more accurate facial pose and expression estimate. Intraface is a facial expression recognition software developed by human sensing lab at Carnegie Mellon.

2. The second component of the project will involve designing and building accurate pan-tilt units for the cameras and incorporating multi-camera head tracking into the developed sensor fusion software.

2. Use case

“Rohit Dashrathi, the luckiest man in the world, is going to have a big wedding today. And he has got a thought to introduce some new elements to this joyful ceremony.

One of his friends introduced him to the latest product from Gandhi Inc., a group of 7 robots which can take candid pictures of happy and smiling people collaboratively. This service can be rented from the company and is charged on an hourly basis.

Hosted in a 4000 square foot area of the Sida hall and attended by 1000 guests, this wedding ceremony is going to last more than three hours. The situation is mostly out of control for professional human photographers. But Rohit is not worried about it, as he has delegated the photography of this very special day in his life to a very very special army of robotic photographers, known as “Robographers”.

Throughout the event, Robographers have become a hot topic of discussion. These amusing robots wearing lovely dresses roaming around have added great charm to the wedding. Adjustable pan-tilt cameras on these units can capture photos of people of all heights, leaving no one unattended. The collision avoidance system keeps all the guests safe. Based on the processing speed greater than 20 Hz, Robographer can detected the human, recognize happy faces and take pictures at the very moment when they smile with joy. The most surprising part is, they are able to take candid pictures without drawing too much attention and all of the pictures are very natural and of high quality.

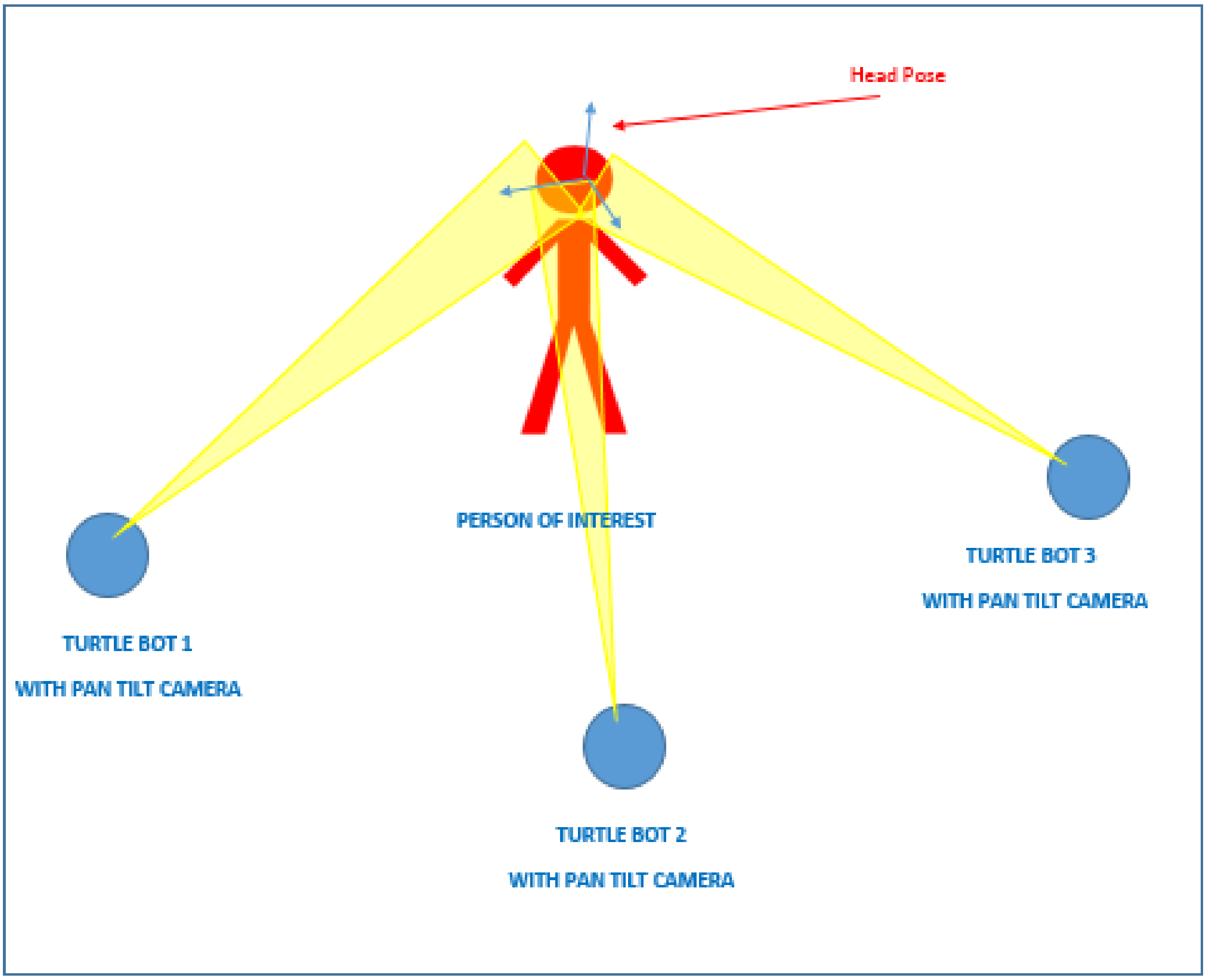

Figure 1 Project 'Robographer': Use Case

Besides recognizing smiling faces, these robots also remember people's faces and thus avoid taking pictures of the same guest too many times. Armed with the head pose tracking technology, the Robographers can tell if the picture is good enough or not by the angle of the faces in the image. The best part is they can communicate with each other and find all smiling faces together.

Eventually the robots cover the whole event on both time scale and space scale, which is a high standard where no human photographer can reach. All the guests get natural and smiling pictures in their e-mail boxes sent directly from the Robographer. To Rohit's surprise, the cost of hiring these autonomous photographers is half but the number and quality of pictures is doubled. This is the best wedding ever."

-Article published in Tiffany Times, May 2015

3. System level requirements

MANDATORY SYSTEM LEVEL REQUIREMENTS

FUNCTIONAL REQUIREMENTS:

Robots in the system shall:

M.F.1: Detect Human Figures
M.F.2: Detect and Follow Head Pose
M.F.3: Detects Faces
M.F.4: Detect Facial Features
M.F.5: Detect Smile Individually & Collaboratively
M.F.6: Communicate With Each Other
M.F.7: Drive Autonomously Between Multiple Locations
M.F.8: Avoid Obstacles
M.F.9: Localize Themselves
M.F.10: Navigate Collaboratively

NON-FUNCTIONAL REQUIREMENTS:

Robots in in the system shall:

M.N.1: Be Supported With Good Lighting Conditions (Fully Illuminated Face All the Time)
M.N.2: Have Wireless Communication mode
M.N.3: Have adjustable elevation
M.N.5: Be easy to operate
M.N.6: Should maintain physical stability (Robots should not topple)
M.N.7: Weigh not more than 11 kg
M.N.8: Should have minimum 2 robots in SWARM
M.N.9: Have minimum 2 Hrs. of battery time

PERFORMANCE REQUIREMENTS:

Robots in in the system shall:

M.P.1: Detect Faces At **0.5s**
M.P.2: Recognise **Smiling** Expression At **0.4s**
M.P.3: Drive Autonomously Between Multiple Locations at **15-20cm/S**
M.P.4: Detect Obstacles at **10 Cm** Minimum Height
M.P.5: Click Photos In **<1.5 S** Response Time after Expression Detection
M.P.6: Click Photos When Individual/Collaborative Smile Assessment **>50%**
M.P.7: Take Pictures within **6 Ft** Range
M.P.8: Click At Least **70 %** Smiling Photos (Measure of Overall Performance Requirement)

DESIRABLE SYSTEM LEVEL REQUIREMENTS

FUNCTIONAL REQUIREMENTS:

Robots in the system should:

D.F.1: Should identify multiple expressions such as Sad, Disgust, **Happy**, Scared, Surprise, Neutral, blinks
D.F.2: Should detect multiple faces
D.F.3: Search for human figures quickly
D.F.4: Navigate as quickly as possible
D.F.6: Identify human figures and facial features by same camera
D.F.7: Display photos
D.F.8: Be able to print photos
D.F.9: Work throughout an event (Avg. 4 Hrs.)

NON-FUNCTIONAL REQUIREMENTS:

Robots in the system should:

D.N.1: Have flash light for good lighting to capture photos
D.N.2: Communicate through Wi-Fi
D.N.3: Have a Graphical User Interface
D.N.4: Not find same person again and again
D.N.5: Have small setup time
D.N.6: Have automatic adjustable elevation
D.N.7: Incorporate 6 robots in swarm

PERFORMANCE REQUIREMENTS:

Robots in in the system shall:

D.P.1: Drive autonomously faster between multiple locations At **40 cm/s**
D.P.2: Detect Obstacles on the surface
D.P.3: Take Pictures within **20 Ft** Range
D.P.4: Click Smiling Photos only (100% Overall Performance measure)

4. Functional architecture

The functional architecture for Robographer is as shown in figure 2, which can be divided into 3 parts: detection, collaboration and planning. It is structured such that the inputs & outputs are on the left and its internal architecture is enclosed in the big blue box in the center.

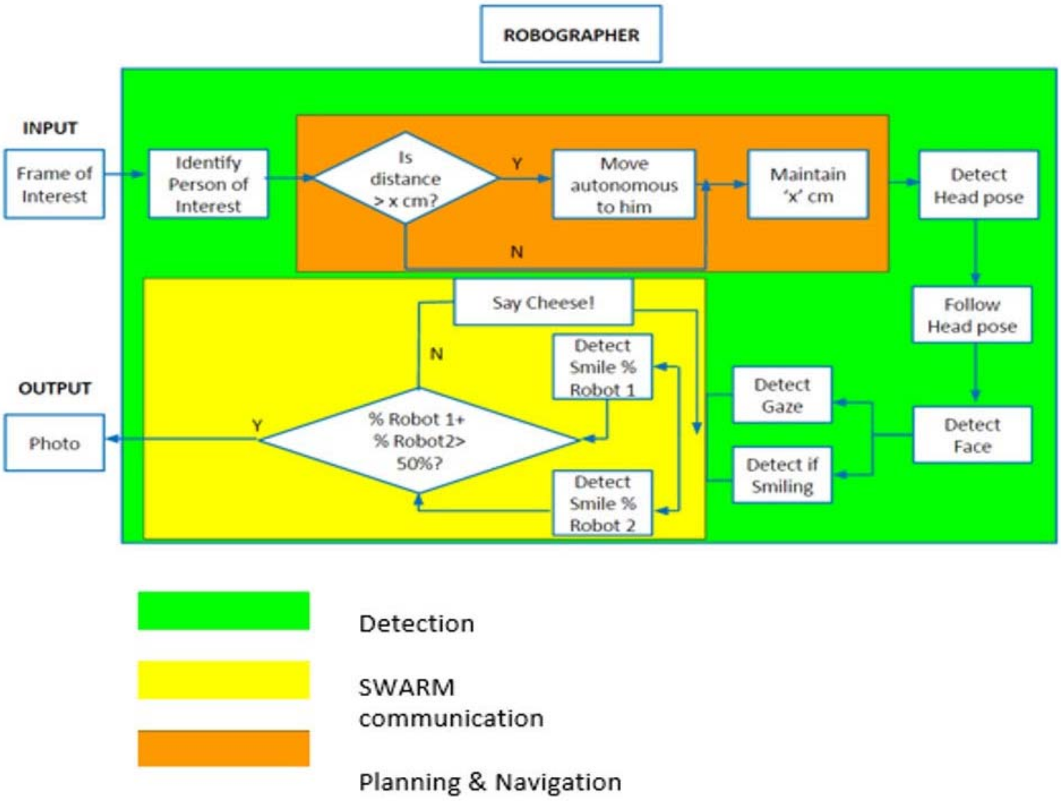

Figure 2 functional architecture

Work flow:

1. First, the swarm of turtlebots will go (with cameras units) into a room to detect where the person of interest is standing.

2. After human detection, the turtlebots will measure the distance between themselves and the person.

3. If the distance is less than specified measure, the turtlebots would move autonomously toward the person, until they reach the desired minimum distance. The robots will then maintain desired distance.

4. Next, the robots will detect the person's head pose. If there are 3 robots, for instance, one may stand straight in front of the person, while second may stand 30 degrees left, and the third robot may stand 30 degrees right. Collectively, the robots will then follow the head pose of the human, and would walk following the person.

5. The next step is to detect the face, which including two parts: detecting gate and lip. In this way, each robot may calculate the percentage of the person's smiling, and the swarm of robots will calculate the overall percentage of the person's smiling according to the position, angle and the smile percentage. If the overall percentage of smile is larger than 50%, the robots will confirm that the person is smiling. If not, the robots will send a voice signal to the human, such as "say cheese", asking him to smile more. After detecting collective 50%, smile percentage the robots will take a photo of the smiling person.

5. System level trade studies

So far we have performed four major system level trade studies. These are as follows:

5.1 Navigation device trade study

For selecting the best navigation and planning algorithm for our system, we weighed different criterion as per our requirements. For detection and recognition of facial expressions from the best possible angle and distance from the subject, accuracy and precision need to have the highest priority. Availability and feasibility in our concerned application are also important features to consider. Weight of hardware employed makes sense because the maximum payload capacity of turtlebots is 5kgs. Cost of hardware gets the lowest precedence as we can allocate sufficient funds from the given 4000$ for the navigation sub-system. Based on all these criterion, visual odometry using stereovision cameras gets the maximum score.

Criteria	Weight	Visual odometry using raspi camera module	Lidar[4] odometry	Visual[5] odometry using stereovision camera	April tags using global camera	Grid map feed	IMU+wheel odometry
Availability	25	20	20	20	20	20	25
Accuracy and precision	35	30	35	35	25	15	15
Feasibility in event photography	15	10	15	15	5	5	15
Weight of hardware	20	15	5	15	20	20	20
Cost of hardware	5	4	3	3	3	5	5
Total score	100	79	78	88	73	65	80

Table 1 navigation device trade study

5.2 Programming language trade study

We listed down all the languages and properties we may use in the project, and the features can be seen from the table below. We will use Intraface and openCV as our main programming tools, and both Intraface and opencv are originally written in C++. In this way, to make the project more robust, we find it is better to use C++ language. Furthermore, speed is very important for facial expression recognition, and C++ is one of the fastest languages among all the popular languages, hence, we choose C++ as our primary languages. We will use python or Matlab as our secondary language in our project to do some subsidiary work.

Hardware compatibility/language	C	C++	C#	Python	Java	Objective-c
ROS		Yes		Yes	Experimental	
Oculus rift		Yes	Through Unity3d			
Android Google glass	Through ndk (hard!)	Through ndk (hard!)			Yes	
Leap motion		Yes	Through Unity3d	Yes	Yes	Yes
Compiled Interpreted	Compiled	Compiled	JIT	Interpreted/ JIT	JIT	Compiled

Table 2 programming language trade study

5.3 Swarm communication trade study

Swarm communication is a very important and most unique subsystem in the system. Swarm communication is expected to involve the transfer of photos to master computer as well as decision signals from one robot to another. As the transfer of images requires very high data speed, the transfer speed will be the most crucial feature while choosing the method of communication. Based on the following the chart, we will use Wi-Fi or the Wi-Fi module of ZigBee (different from the original ZigBee) to connect turtlebots together.

	Weight	Wi-Fi			Bluetooth			ZigBee/Xbee		
			Score	W. Score		Score	W. Score		Score	W. Score
Transfer speed	10	54 Mb/s	9	90	1Mb/s	5	50	100 Mb/s	3	30
Range	8	100m	9	72	15m	5	40	10m(indoor)	3	24
Battery consumption	5	high	3	15	middle	5	25	low	10	50
Cost	3	high	3	9	middle	5	15	low	10	30
Total score		186			130			134		

Table 3 swarm communication trade study

5.4 Camera trade study

The score of camera for each parameter have been selected on a scale of 1-4 depending upon the desirability of the parameter. For example, we require high video and photo resolution camera. The camera with low resolution is scored as 1 and high resolution is scored 4. On the contrary if we require that the camera is light in weight then, lighter the weight higher is the score i.e. 1 for a heavy camera and 4 for a light camera.

The camera we require should have high video resolution, high photo resolution, high zoom ability, should be light in weight, small in size. Cost is not much of concern since the cost of most cameras do not differ much than 100-200$. The weights to each of the parameters have been assigned on the scale of 1-4 depending upon the degree of desirability of the parameter.

Accordingly, the weights have been assigned above and net score was calculated. The best score was achieved by soloshot[3] camera and quite close to that on second place Panasonic hcv110[6].

Criteria	Weightage of parameter(out of 4)	Foscam (out of 4)	Logitech c920 (out of 4)	Soloshot	Panasonic hcv110
Cost	2	2	1	3	4
Weight	3	2	4	3	3
Dimensions	2	2	4	3	3
Resolution for video	4	2	4	4	4
Zoom	4	2	1	4	4
Resolution for photo	4	1	4	4	3
Field of view	3	4	2	2	2
Total score		46	64	75	73

Table 4 camera trade study

6. Cyber-physical architecture

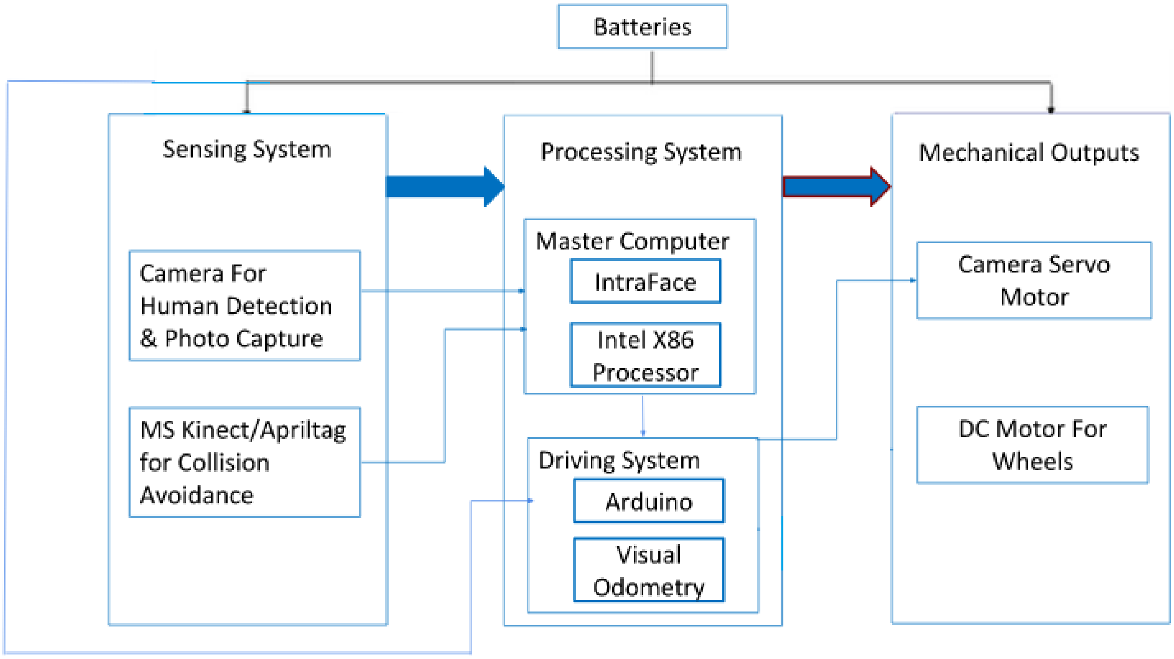

Figure 3 Cyber-Physical Architecture

To make it easier to get the whole idea, the cyber-physical architecture is divided into three different parts: sensing system, processing system and mechanical outputs. Generally speaking, the data will be transferred from sensing system to processing system, and then eventually to the mechanical outputs. Sensing system, mechanical outputs and the Arduino board in the processing system will be supplied power by batteries.

Now we will look at these parts one by one.1.) Collaboratively we have two different camera: (high resolution pan-tilt camera for human detection and photo capture cam), MS Kinect for collision avoidance. As mentioned previously, pan-tilt camera is a system combined with camera and two servo-motors, one for horizontal movement and the other for vertical movement. Kinect's build-in IR sensor will be used to detect the distance between the obstacles (human, wall) and prevent a turtlebot from colliding with them. April tag mechanism is similar to QR code and is primarily used for orientation and navigation. We may attach it to each robot to determine their relative distance. 2.) Under the processing system, we will have one master computer serving as base station to control a turtlebot. Each turtlebot will have its own master computer. With the help of Intraface, which is an API (application programming interface), we can detect the head pose, facial landmarks and expression within 7 ft. Visual odometry is a method to realize localization by computer vision. It can be implemented by the pictures attached on the wall or symbols on the floor. For better computing and processing speed, all the images obtained will be processed by x86 processor. In the driving system, the pan-tilt camera will be controlled by an Arduino board to find the person of interest. The dc motor is connected to the build-in board of turtlebot throughout the operation. All of the data will be transferred by Wi-Fi. The master computer works under Linux, with ROS and ROCON[2](a framework under ROS for multi-robot control), which can implement the swarm system.

To grasp the holistic picture, we are now going through the whole system again. On the first transfer route the image will be obtained by pan-tilt camera for human detection and photo capture, then processed in master computer using Intraface. After sending the the processed information about distance measurement to Arduino and hence the camera servo motor, the pan-tilt camera will adjust itself to match the height of target human. On the other hand, the MS Kinect will detect the distance between obstacles and robot to keep sending data to computer. When the distance is too close, the robot will make a detour or turn around. The data is transferred from MS Kinect to master computer to dc motor in this case. Remember there's still a third possible route. The pan-tilt camera obtains the image of a human figure, which is recognised and verified as a human by master computer. The master computer can also send the instructions to dc motor for wheels in the mechanical outputs part, to make the turtlebot move toward the human. After approaching the target, the API in the master computer will determine whether the person is smiling or not. The decision to or not to retrieve/ save a photo is determined with the master computer. In this third route, the information will not be sent to any motor.

7. Subsystem descriptions

The Robographer system has 3 different subsystems (refer functional architecture) as follows:

7.1 Detection sub-system

The task of detection can be divided into 3 parts:

7.1.1 Human detection

The swarm of robots will sense and search for human figures using multi cameras fitted on the turtlebots. It will form a collective estimation of the human figure when it is detected. After the human detection, image will be processed by the visual odometry system to calculate the distance between the swarm of robots and the standing human. This distance would act as an input to the planning and navigation subsystem for deciding the desired path choice to reach in front of the human figure.

7.1.2 Facial expression detection

Our main task is to do facial expression recognition with multi-cameras, from 3d direction, instead of the traditional 2d perspective. And it can be achieved by head pose detection, lip angle detection, and eye detection. The software Intraface and opencv will help us in this task. After the swarm of robots reaches in front of the human, next task would be to detect the person's head pose. Swarm of the robots will follow the head pose of the human, and would walk following the person in turn to detect the face and expressions. The robots in system will detect gate and lip of human face for accurate face detection using the camera. This face & expression detection part will be taken care by Intraface. In this way, each robot may estimate the percentage of the person's smile (given by Intraface). Furthermore, these readings will be sent to master computer which will calculate the overall average percentage of the person's smile according to the positions, angles and the individual smile percentages of images provided by each camera in swarm. If the overall percentage of smile is larger than 50%, the robots will confirm that the person is smiling and photo will be captured. If not, the robots may play a recorded voice signal to the human, such as "say cheese" for asking him to smile more. At last, the robots will take a photo of the smiling person.

7.1.3 Obstacle detection

Furthermore, the detection system will also be equipped with the sensors to detect obstacles, other entities in path (humans, tables, chairs, other robots etc), which also provides planning system with relevant information about the environment.

7.2 planning and navigation sub-system

The project involves the use of an accurate navigation sub-system so that the swarms can easily move through the indoor environment and can place themselves at accurate positions to perform facial recognition collaboratively. There are three basic sub-blocks of navigation-

7.2.1 Self- localization

The swarms first need to establish their position and orientation in the environment independently. By default, the turtlebots have wheels and IMU odometry for localization. This method does not give precise readings of the position of robots. For facial recognition and photography, a robust localization system is needed. We are implementing visual odometry using a stereo vision camera to accomplish this.

7.2.2 Planning

The path is to be planned between the present location of the turtlebot and the final location from where it will recognize the faces and click the photo. This will be an extension to the visual odometry algorithm. We will develop an algorithm for the robot to move towards the human and maintain a minimum distance from him and the other robots. It will also follow the head pose collaboratively.

7.2.3 map-building

We will use a pre-generated (using ROS) occupancy grid map of the environment.

Fallbacks of the navigation sub-system:

a. Robots collide among themselves

The system should insure that the swarms do not collide among themselves. For this, we intend to develop some algorithms for collision avoidance and test them in simulated environments.

b. Robots collide with humans or objects in the environment

To do this, we will employ some algorithms to avoid collisions with objects of minimum 20 cm height. In case, any object comes in the near zone of the robot, the robot will send out a signal and change its path.

c. Robots are lost

In this case, when the robots lose sense of their location, we will develop an algorithm so that they plan a path in the given map to reach a common docking station situated at one corner of the given environment.

7.3 Swarm collaboration sub-system

For the purposes of photography, robust facial expression recognition is a must. Multi robots can give much better idea of a person's expression compared to a single robot. Also, in a dynamic environment like a wedding or a party, people can move from a place, change their head pose or orientation within a few seconds. In order to capture photographs, the swarm robots should plan a strategy in such a way that even if the person of interest moves within a small area or changes the orientation of his head pose, they will still have a robust feedback of his expression without having to make major changes in their initial path.
This collaboration will be divided into two parts:

7.3.1 Identify a person of interest

The robots (turtlebots) will first move in a group sticking together but try to cover every angle using the pan tilt camera until they detect a human or person of interest. This person of interest will be identified based on his/her distance from the robot, his initial head pose and if possible his facial expression.

7.3.2 Collaborative approach and facial expression analysis

Once the robots identified the person of interest, they will break the formation and follow a strategy of approach depending on the head pose of the person of interest. While approaching the person they will continue to track the head pose of the person along with his or her expression, all in collaboration and determine the confidence of their analysis. The approach to the person of interest will be in such a way that even if he or she tilts, pans head, one or two of the robots will still have a facial expression reading consistent with each other. If not then the robots will check the head pose of the person of interest and re-plan the strategy of approach in order to get the robust expression reading. Once the person of interest is within a range within which a clear lively, smiling photograph can be taken, the robots will decide amongst themselves the best view or angle and click the photograph accordingly. This photograph will then be stored in the database and will be available at the user’s disposal.

Fallbacks of the swarm subsystem

a. Person of interest moves out of the range of detection of all robots:

Sometimes the person may move from one end of the room to another or turn around by 180 degrees while the swarm is approaching them. The turtlebots (robots) will not be able to cope up with the speed of the person and end up following the person all the way to the other end or lose track of him while in the approach mode. If none of the robots are unable to detect the head pose and expression of the person of interest, the robots will switch back to the ‘identify the person of interest mode’ and start all over again,

b. The robots identify and click photo of the same person again and again

One of the key performance requirements of the system is that the Robographer should cover the event clicking pictures of most of the guests. In order to avoid the robots clicking many pictures of a single person, they should be capable of storing distinct features of that person such as clothing, hairstyle, height etc. Along with the number of clicked pictures of that person. In this way when they identify that particular person as their person of interest again, they can know that picture of this person has already been clicked for ‘x’ times and move on to find another person of interest.

c. Robot stops functioning correctly or discharges

In order to make the system more robust to such scenarios, the system must be capable of operating. The advantage of using swarm is that even if one robot fails the other robots can collaborate and move further with the task completion without the dysfunctional robot. In this way the task gets completed even though the performance of the overall system might be affected. Also the user will be notified about the dysfunctional robot so that appropriate action can be taken to fix it. If it is the case of low battery, the effort will be that the robot will return to its charging base while its battery is low.

8. Project management

8.1 Work plan & tasks

The work plan for this project is detailed in the work breakdown structure flowchart shown in figure 4. The main categories of the work breakdown structure are based on the subsystem segregation discussed earlier. The major subsystems are further divided intro tasks to be accomplished. The overall plan is to develop and validate the subsystems and then integrate them together to form the main working system.

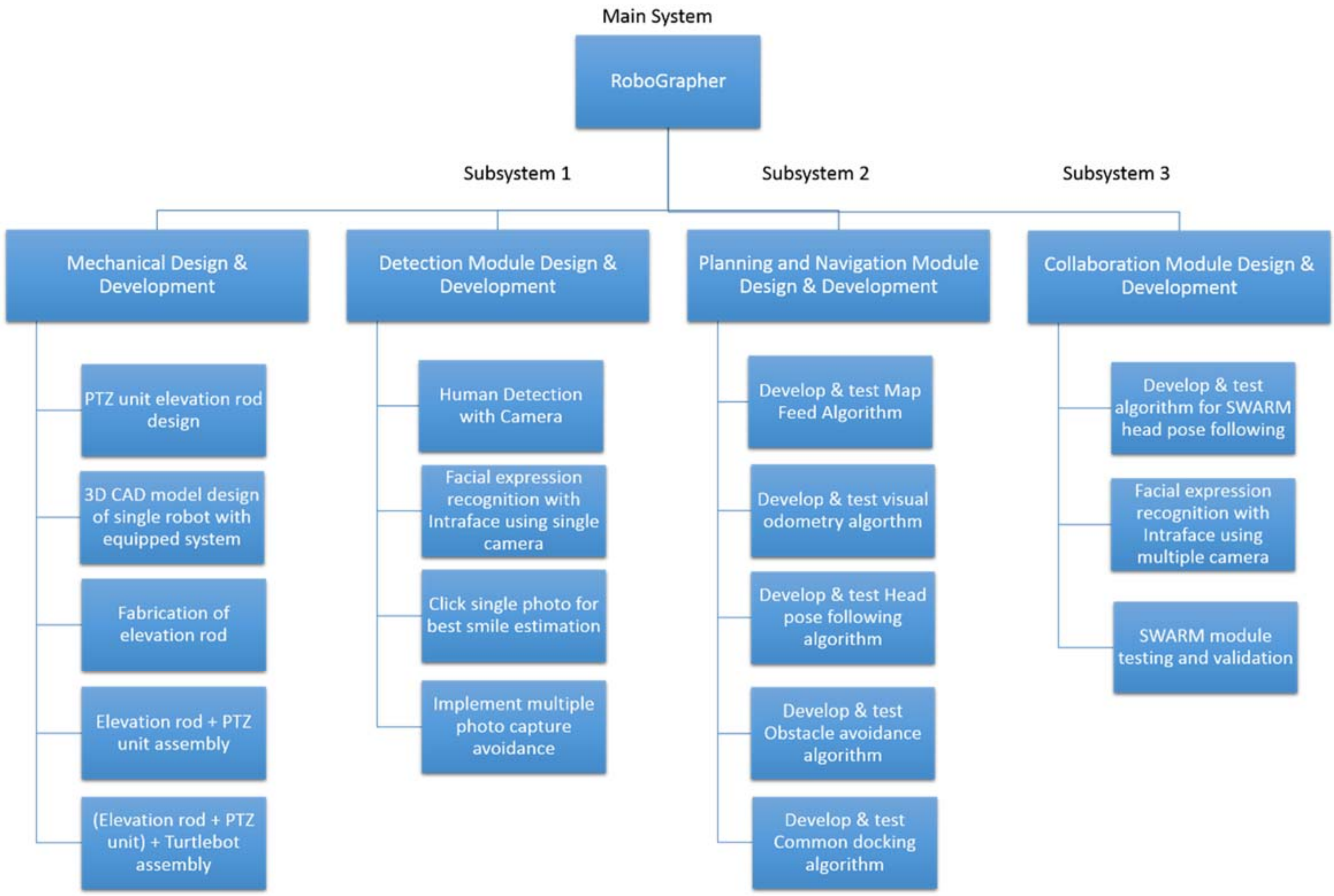

Figure 4 Work Breakdown Structure

8.2 Project schedule & key milestones

The following figures shows the parallel work schedule of fall 2015 as well as spring 2016, developed for the project and captured in the Gantt chart. This chart consists of the major tasks only; subordinate tasks grouped inside and not shown in following figure. The Gantt chart also shows the predicted time and deadline assigned for each task to be completed. Major deadlines and internal milestones are shown at the timeline on the top of the chart. This chart will be further broken down into subtasks for more detailing. Each major task is also assigned the name of responsible person leading that particular task.

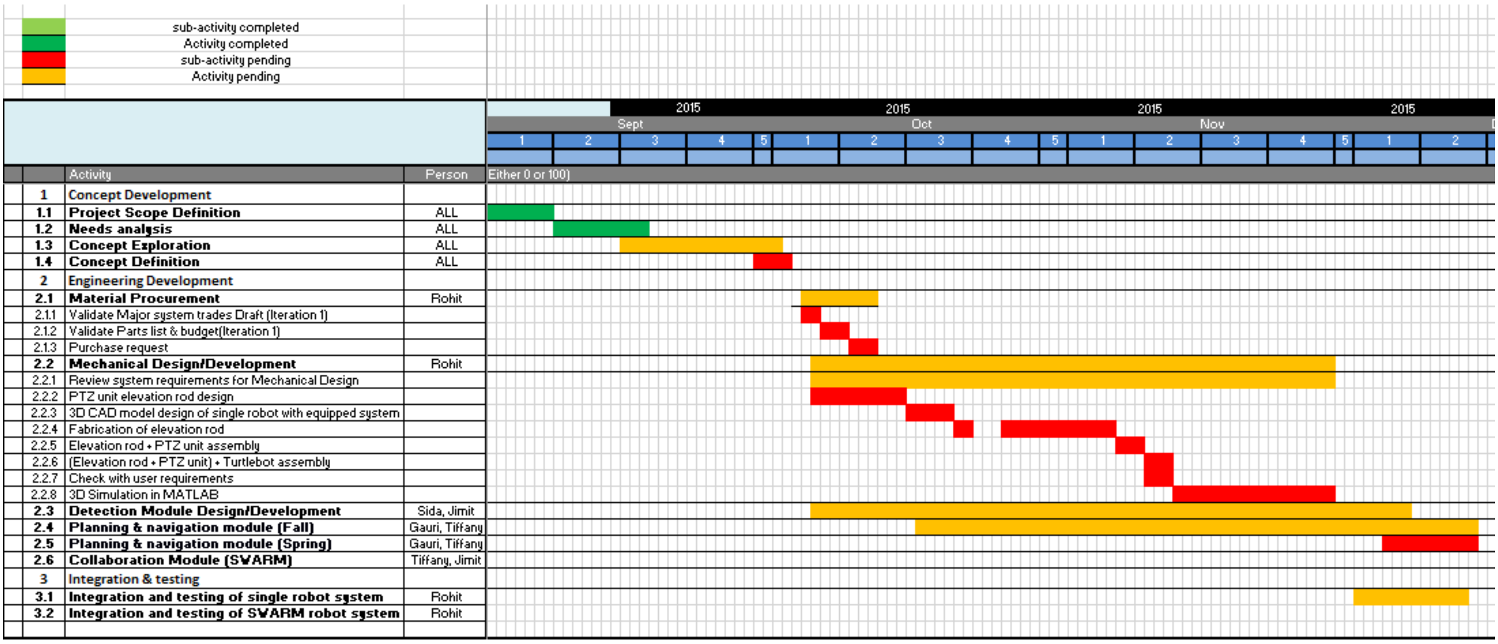

Figure 5 Project Schedule for Fall 2015

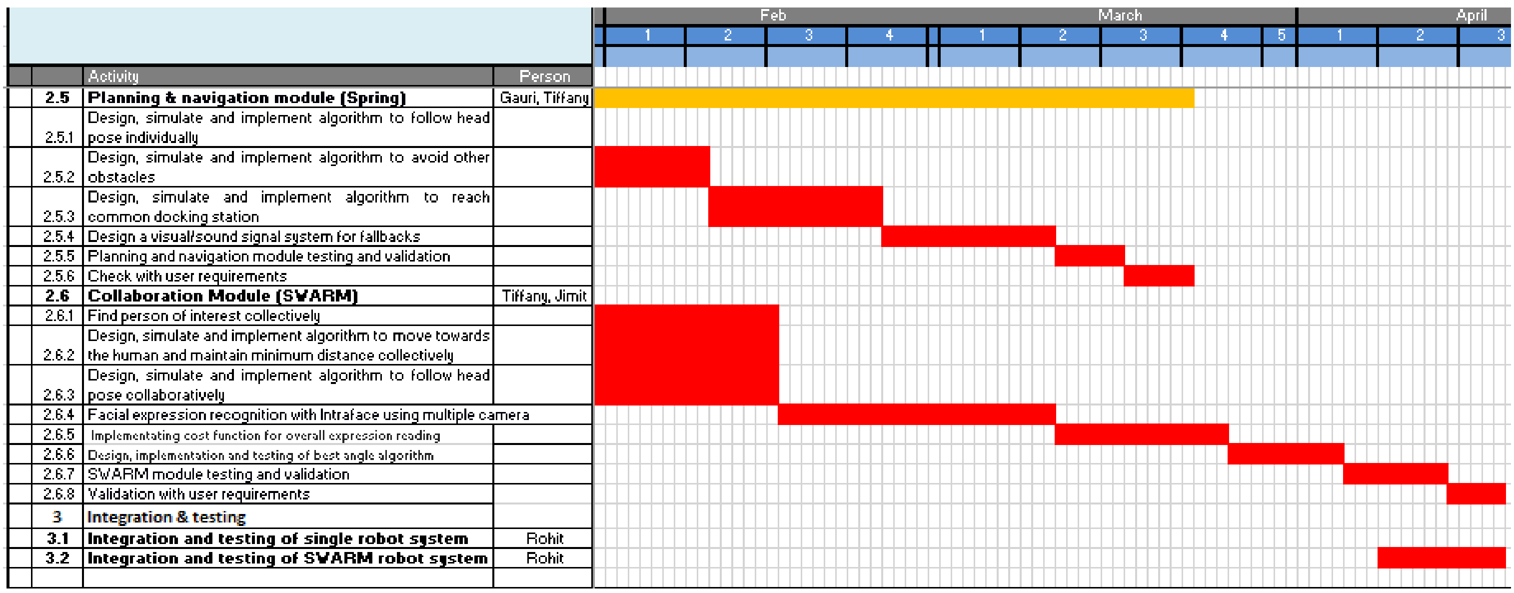

Figure 6 Project Schedule for Spring 2016

Following table also summarizes the key milestones from the Gantt chart project schedule:

Milestone	Comment	Date	Progress review
Pan tilt unit 3d model	Cad designing of PTZ unit is completed. It will be a representation of actual unit which will be attached to turtlebots	14-oct	
Grid map design completion	Preliminary design for grid map of site of demonstration completed	22-oct	Progress review 1
Pan tilt unit 3d model	Cad designing of PTZ unit is completed. It will be a representation of actual unit which will be attached to turtlebots	14-oct	
Facial expression recognition with Intraface using single camera	Camera is able to detect smiling face using Intraface	29-oct	Progress review 2
Build an algorithm for map-feeding	Program for feeding the map to Robographer system is accomplished	30-oct	
Fabrication of elevation rod for pan tilt unit completed	Elevation rod providing additional axis for height adjustment of Robographer system has been fabricated	05-nov	
Best pose estimation for smiling face in (-30,0,30) range (find best angle)	Algorithm for choosing best photo angle in the range (-30,0,30)	21-nov	
Click single photo for best smile estimation	Camera is able to take photo for best smile percentage estimated	24-nov	

Table 5 key milestones

8.3 System validation experiments

8.3.1 Fall validation experiment

1. Mechanical setup:
- One turtle bot shall be mounted with a pan tilt camera unit.
2. Detection of human:
- The single turtle bot shall be able to detect a human in its vicinity.
- Once detected it shall move towards it.
- The turtle bot shall now mark that human as a person of interest.
3. Detection of human face and head pose:
- The turtle bot shall then pan and tilt the camera as required in order to detect the head pose of the human.
- Once head pose is detected it shall try to recognize the various facial features of the human face such as lip width, distance between the eyes, width of the eye, hairstyle, and eye-wear.

4. Expression analysis:
- The turtle bot shall then analyze the facial features and give an estimate of the expression of the human face.
5. Follow the head pose
- If the human face changes orientation or human himself moves the turtle bot shall keep track of the head pose and if possible shall follow the human.
6. Click photograph
- Once the turtle bot gets a read of the facial expression, it shall click a photograph instantly if and only if the human expression is a smile.
7. Localization
- The turtle bot shall constantly localize itself all the while in a map provided by the user.

8.3.2 Fall final validation setup

1. Location - any rectangular room with good and uniform lighting. The surface of the room shall be smooth or carpeted. There should be no obstacles. Only one human should be present in the area of testing.
2. Equipment needed – turtle bot, pan tilt camera unit, battery (power supply).
3. Requirements to be fulfilled:
- Detect human
- Detect head pose
- Recognizing faces
- Recognizing expressions & click photos

8.3.3 Spring validation experiment

1. Collaborative detection of human in the environment
- A system of 4-6 turtle bot shall flock (stay together), searching in varied directions in an optimized manner and try to detect humans in the vicinity.
- Once detected, the turtle bots shall collaboratively decide a person of interest amongst all detected humans based on parameters such as distance, head pose and if possible then facial expressions.
2. Approach the person of interest
- The system shall break formation and plan a strategy to approach the person of interest from various directions such as to get a more robust expression reading.
- Each turtle bot shall be assigned the final goal coordinate and angle of approach. The goal coordinate will be decided on the basis of a certain minimum distance from the person from where photograph can be taken.
- The turtle bot as an individual shall plan a path to approach the person of interest in the assigned direction.
- The turtle bots shall individually approach the person of interest along the path that each bot planned while keeping track of the head pose.
3. Collaborative analysis of expression
- Each robot shall analyze the expressions individually from different angles.
- The system as a whole shall fuse the analysis of the expression and then determine the overall estimate of the expression of the person.
- If the expression is that of a smile the turtle bots shall determine the best view of them all and click a picture of the person of interest.

4. Move further to take another photo
- Once a person is photographed, the swarm shall again move in flock while trying to detect the next person of interest.

8.3.4 Spring final validation setup

1. Location - any rectangular room with good and uniform lighting. The surface of the room shall be smooth or carpeted. Obstacles present. A group of not more than 7-10 humans present in the room.
2. Equipment needed – 4-6 turtle bots with mounted pan tilt camera unit, battery (power supply).

Requirements to be fulfilled

- Recognizing expressions collaboratively
- Plan path
- Communicate within themselves
- Move autonomously from one location to another
- Avoid obstacles
- Click photos from best possible angles

8.3.5 Team member responsibility

Domain	Mechanical	Detection	Collaboration	Planning and navigation	Communication and interfacing
Primary	Rohit	Sida	Jimit	Gauri	Tiffany
Secondary	Gauri	Jimit	Tiffany	Sida	Rohit

Table 6 team member responsibilities

Each member of the team will be involved in two roles. One primary role and one secondary role. The roles are divided according to the domains of work breakdown structure. The member holding primary role for a domain shall be responsible for the respective domain.

8.3.6 Provisional list of parts and budget

Part name	Quantity	Approximate rate	Total price
Soloshot camera	6	350$	2100$
Servo motors	6	20$	120$
MS Kinect sensor	6	110$	660$
Raspberry pi	6+2	36$	288$
Stereo vision camera	6	150$	900$
Total			3768$

Table 7 provisional parts list and budget

8.3.6 Risk Management

Major risks are caused by (1) device factors, (2) resources factors, (3) project management. We have tried to come up with some ways in which they could be mitigated.

a. Device factors

1. Damage to device
2. Breakdown of turtlebots
3. Delay in shipping
4. Damage to master computers
5. Fail to choose the proper cameras

b. Resources factors

1. Long time to familiarize ourselves with ROS and other working environment
2. Failure of power supply
3. The power supply is too expensive
4. Fail to do human detection correctly
5. Fail to do human expression recognition correctly
6. Fail to find the correct algorithm to do human expression recognition with multiple cameras
7. Fail to navigate or localize the swarm of robots
8. Robots collide with each other
9. Robots bump into obstacles
10. Failure of corporation between swarm of robots

c. Project management

Honestly speaking, one of the most significant risk is falling behind schedule. This is a very realistic issue that we should seriously consider. The project of Robographer would take extensive time on it, and apart from this project, we will have 4 other courses to take. So it's easy to imagine that we will have many dues per week. Every student in MRSD program works with “heart at the work”, and we will try our best to take advantage of every moment in a day. In this way, it's very important to assign different tasks to different people, which will be more effective in project management. And every group meeting, it's important for time management, make plans for the next week and make summary for the past week.

In case of something goes wrong or something happens unexpected, we should try our best to stay ahead of schedule. In a project, everything may happen, and what we do is to do as perfect and fast as we can.

References

Project Report Format: https://sites.google.com/site/mrsdproject201415teamc/

1 http://www.humansensing.cs.cmu.edu/intraface/

2 http://wiki.ros.org/rocon

3 http://shop.soloshot.com/

4 https://en.wikipedia.org/wiki/Lidar

5 http://www.roborealm.com/forum/index.php?thread_id=2681

6 http://www.amazon.com/Panasonic-HC-V110-Weight-Digital-Camcorder/dp/B00AW54YWS