Perception Subsystem: Target Identification and Tracking

# Motivation

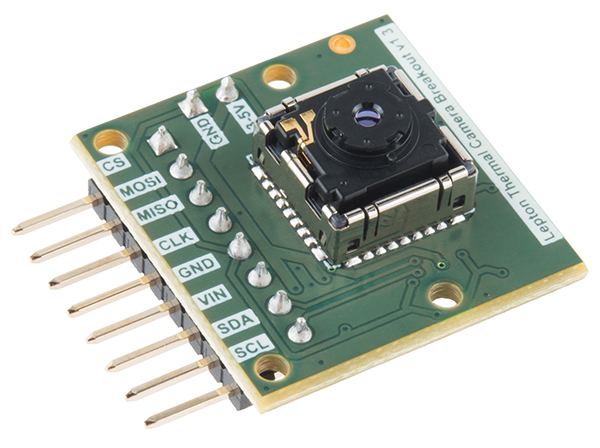
The primary purpose of the FlyNet drone is to perform autonomous search and rescue operations. This mission immediately implies some very broad requirements. These the requirements include the capablility of autonomous mapping and navigation as well as target identification and tracking. After fleshing these requirements out, there are a handful of requirements that pertain specifically to target identification and tracking. These are provided in the following table.

Table 1. List of requirements relevant to target identification and tracking

|  |  |
| --- | --- |
| Requirement Label | **Level 3 Requirement** |
| 1.1.1 | The system shall have visual sensor(s) |
| 1.1.2 | The system shall be capable of processing imagery |
| 1.1.3 | The system shall locate individual targets in images with 95% probability of detection |
| 1.1.5 | The system shall identify human targets within 5 meters |
| 2.1.1 | The PPU shall estimate target position to within 0.5 meters |
| 2.1.2 | Each PPU shall estimate at least 2 visually identified target tracks |
| 2.1.3 | Each PPU shall maintain estimates of all confirmed target tracks |

# Hardware

Requirements 1.1.1, 1.1.3, and 1.1.5 heavily influence the types of sensors that must be used to accomplish the target identification and tracking. It specifically dictates that some sort of visual sensor must be used that has enough resolution to detect a human target at up to 5 meters away with a 95% probability of detection. An infrared camera was determined to provide the greatest contrast between human targets and background due to the relatively large difference between indoor ambient temperature and human body temperature. This sort of infrared camera would provide a much more ‘binary’ image than a visual spectrum sensor, which would have to rely heavily on features of the target to identify it.

Figure 1. FLiR Lepton on a SparkFun breakout board.1

Resolution requirements of the camera had be be determind after the spectrum was decided. Fine details of a scene are not required when using an infrared camera to detect humans since they have a distinct body temperature that does not match the temperature of an indoor environment. The FLiR Lepton Long Wave Infrared (LWIR) camera was chosen as the preferred for numerous reasons. First and foremost, it operates in the desired LWIR spectrum and would make it easy to detect the targets. It is also the smallest, lightest, and cheapest commercially available sensor in its class. This makes it an ideal choice of an airborn platform that has limited power and thrust. Finally, the FLiR Lepton is available on a breakout board and has significant documentation for development. Basic specifications for the FLiR Lepton are below.

* LWIR sensor: 8-14 μm
* 51-deg HFOV, 63.5-deg diagonal FOV
* 80 x 60 pixels
* Thermal sensitivity <50 mK
* MIPI and SPI video interfaces
* I2C control interface
* Power draw <160 mW
* Price: $260

The FLiR Lepton requires two different communication protocols. It uses I2C for configuration and SPI for telemetry. The Odroid XU4 that has been chosen for use on the FlyNet quadcopter does provide these interfaces through an XU4 Shifter Shield.

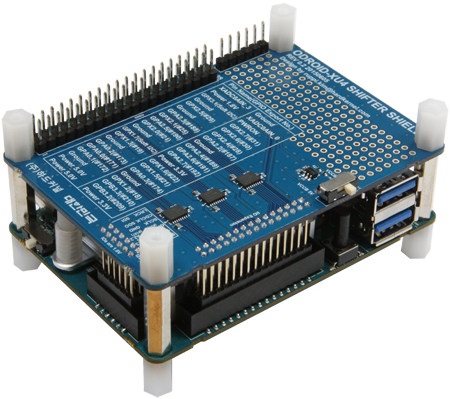
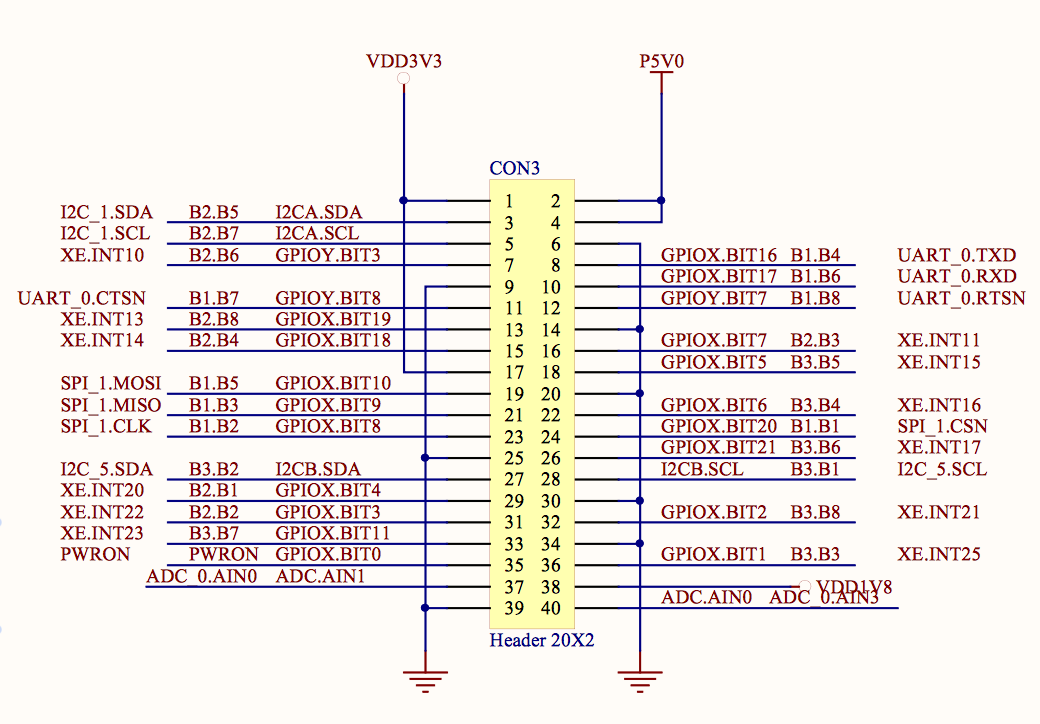
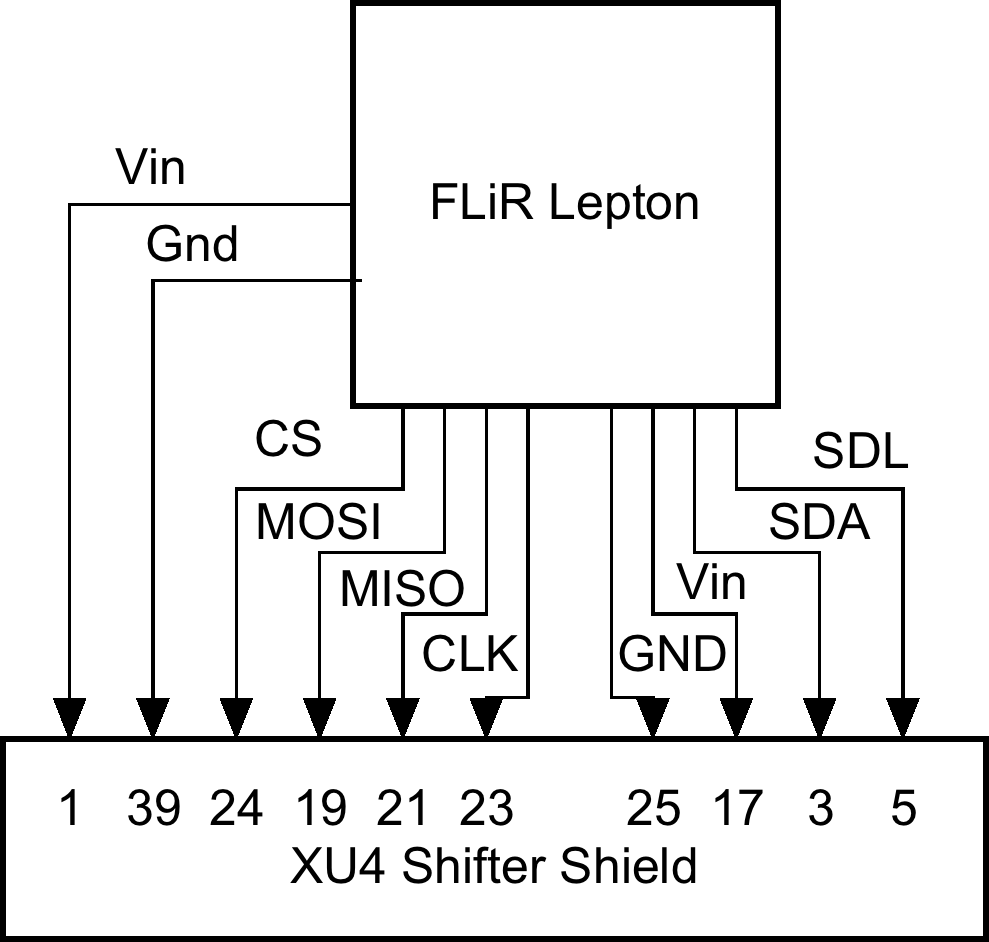
Figure 2. Odroid XU4 with shifter shield.2

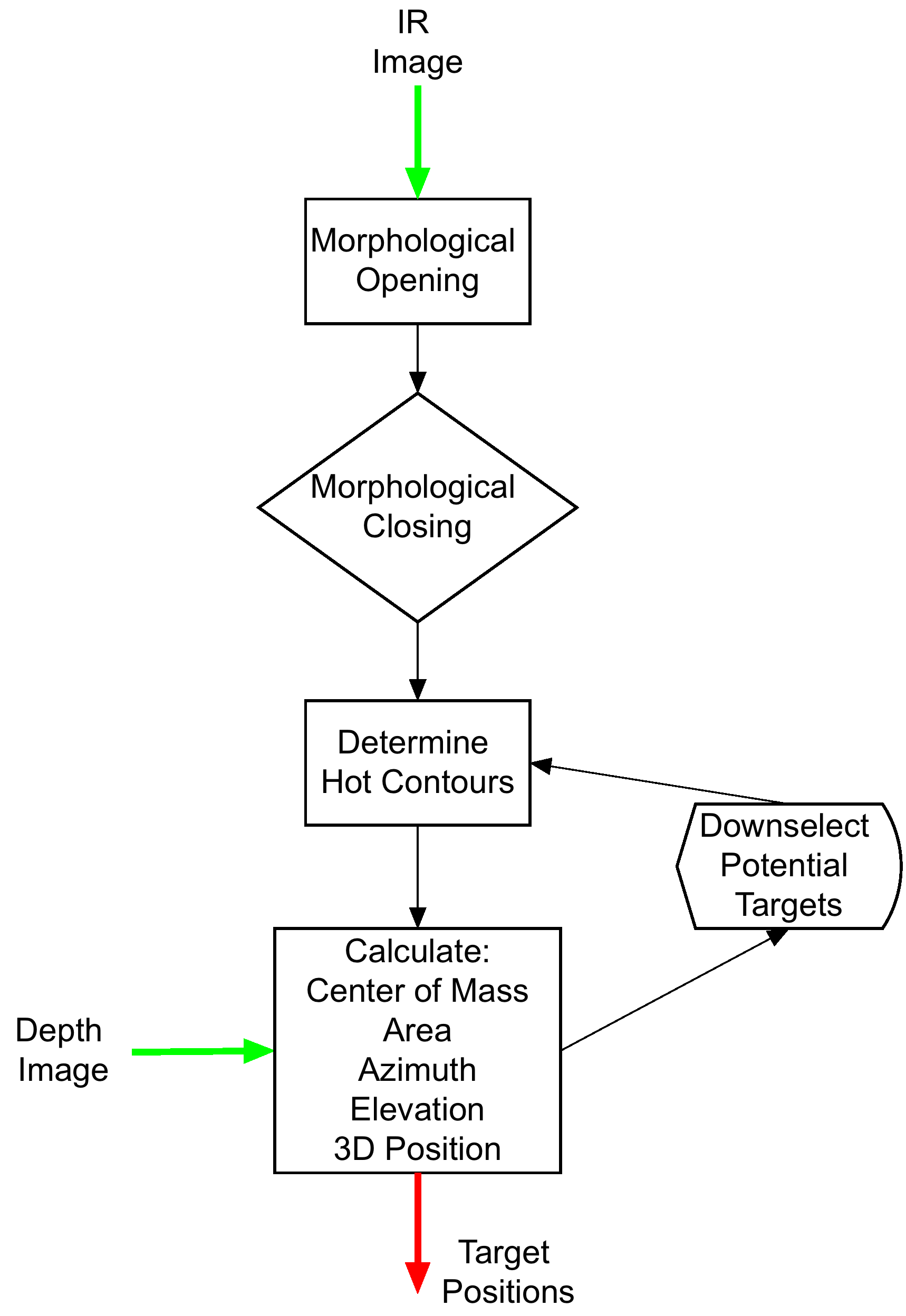
Figure 3 provides a pinout diagram of the XU4 Shifter Shield pins. Figure 4 shows the connection between the FLiR Lepton breakout board and the XU4 Shifter Shield.

Figure 3. XU4 Shifter Shield pinout3

Figure 4. Connection between FLiR Lepton breakout board and XU4 Shifter Shield

# Algorithm Overview

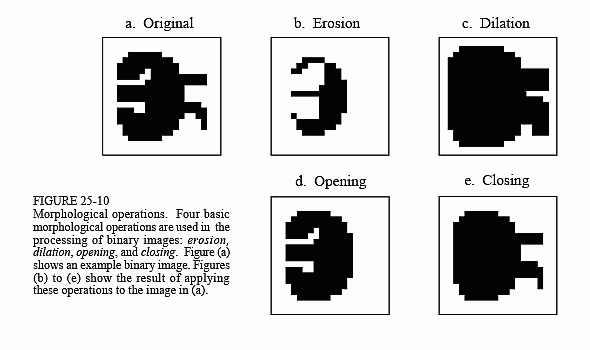
The tracking and image processing algorithms required to fulfill the requirements in Table 1 must run quickly and efficiently on a relatively low power microcomputer. Simplistic yet highly effective methods of image processing and tracking were used to keep processing overhead low and meet the 10Hz requirement. A flowchart of the algorithm is provided in Figure 5.

Figure 4. Algorith flowchart

The target tracking and identification (TTI) algorithm takes an 80x60 LWIR image from the FLiR Lepton and a 320x240 depth map image from the Guidance sensor as inputs. The LWIR image is immediately masked to return a binary image that sets objects that are close to human body temperature to white, and everything else to black. This binary image now highlights all objects that may be targets. Next, the new binary image goes through a process called morphological opening and morphological closing. This process puts the image through a series of dialations and erosions that removes some of the noise from the image and allows large objects to come across as clean white blobs. Figure 5 shows examples of morphological processing.

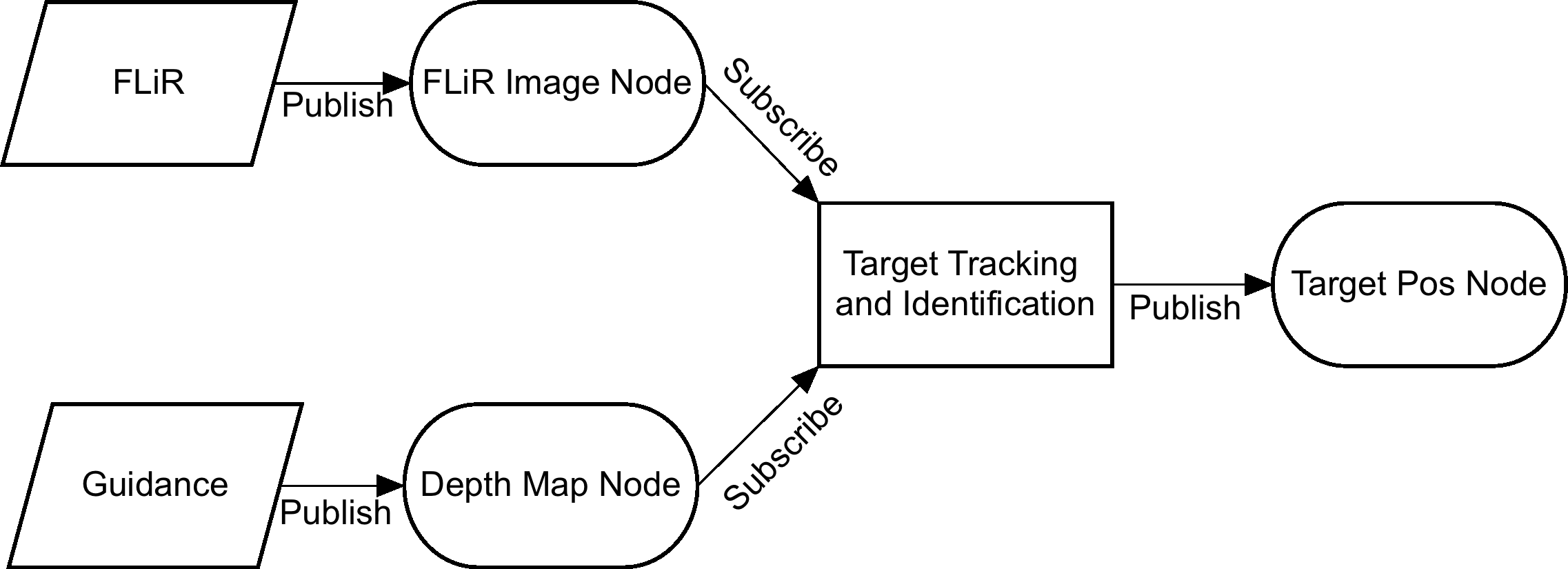
The contours of the binary image are calculated using the openCV contour function after the binary image has been opened and closed. This provides a list of all the different potential targets in the image. Many of the contours will not correspond to actual human targets, so they must be pruned to include only the contours that most likely correspond to a human target. Since the average size of a human is known and the TTI is required to detect a human at up to 5 meters away, a human sized object is expected to have a minimum area. This area along with aspect ratio of the object can be used to downselect the potential targets.

The center of mass of the contours is used to define the location of the contour in the image plane. This location corresponds to an azimuth and elevation angle. The depth map image provides a range for that specified azimuth and elevation. The three-dimensional location of the target is determined by using azimuth, elevation, and range.

Figure 5. Morphological processes4

# Software Architecture

The TTI software interfaces with ROS to obtain and disperse information. The FLiR camera will be controlled by C++ code that publishes the images to a ROS image\_transport node. Similarly, Guidance will publish its depth map to a ROS image\_transport node. The TTI software will subscribe to these two nodes in order to obtain the images. The final target positions will be transformed by using the ROS transformation functionality and they will be published to a target position ROS node for use by other code.

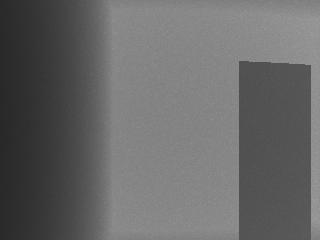
Figure 6. Software architecture overview

# Preliminary Results

Preliminary testing of the FLiR camera has shown that the images are not very noisy and are capable of discerning a human being from a distance of at least 5 meters. Figure 7 shows the first image obtained from the FLiR Lepton sensor. These images are of sufficient quality that the TTI algorithms have no problems performing adequately well.

Figure 7. First image from FLiR Lepton

Matlab code was written to create simulated FLiR images and depth map images that could be used to characterize the TTI code. Each of these simulated images is an image of the inside of a room with a target that has realistic contrast with the environment. These are then fed into the TTI algorithm to test how noise and target location in the image affects the three dimentional position solution. The tests showed that the TTI algorithms position estimate errors are on the order of the noise in the depth map images. Location in the frame does not affect the final position estimate as long as the whole target is in frame. Figure 8 and 9 show examples of the simulated FLiR and depth images, respectively.

Figure 9. Simulated depth image

Macintosh HD:Users:taylordean:Documents:GradProjects:trackingImages:flir.jpgFigure 8. Simulated FLiR image

# References

1*”*FLIR Lepton Hookup Guide,” *Sparkfun*, : <https://cdn.sparkfun.com/assets/learn_tutorials/3/5/9/13233-011.jpg> [Accessed 6 December 2015]

2“XU4 Shifter Shield”, *Hardkernel Co. LTD.*, <http://www.hardkernel.com/main/products/prdt_info.php?g_code=G143556253995> [Accessed 6 December 2015]

3“TXS0108E LEVEL TRANSLATOR,” *Hardkernel Co. LTD.*, 30 June 2015, <http://dn.odroid.com/homebackup/XU4_SHIFTER_SHIELD_REV0.1.pdf> [Accessed 6 December 2015]

4Smith, Steven W. “Chapter 25: Special Imaging Techniques,” *The Scientist and Engineer’s Guide to Digital Signal Processing,* <http://www.dspguide.com/graphics/F_25_10.gif> [Accessed 6 December 2015]