

# Integrating Occlusion Monitoring into Human Tracking for Robot Speed and Separation Monitoring

William Shackleford, Richard Norcross, Jeremy Marvel, Sandor Szabo\*  
National Institute of Standards and Technology (NIST)

100 Bureau Dr.

Gaithersburg MD, 20899

011 1 (301) 975-4286, 011 1 (301) 975-3440, 011 1 (301) 975-4592, \*retired

William.shackleford@nist.gov, Richard.norcross@nist.gov, Jeremy.marvel@nist.gov

## ABSTRACT

Collaborative robots are used in close proximity to humans to perform a variety of tasks, while more traditional industrial robots are required to be stopped whenever a human enters their work-volumes. Instead of relying on physical barriers or merely detecting when someone enters the area, the collaborative system must monitor the position of every person who enters the work space in time for the robot to react. The TC 184/SC 2/WG 3 Industrial Safety group within the International Organization for Standard(ISO) is developing the standards to help ensure collaborative robots operate safely. [1][2] Collaborative robots require sophisticated sensing technologies that must handle dynamic interactions between the robot and the human. One potential safety risk is the occlusion of a safety sensor's field of view due to placement of objects or the movement of people in front of a safety sensor. In this situation the robot could shut down as soon as even a single sensor was partially occluded. Unfortunately this could greatly diminish the extent to which the robot could work collaboratively. In this paper we examine how a human tracking system using multiple laser line scanners [3] was adapted to work with a robot Speed and Separation Monitoring (SSM) safety system and further modified to include occlusion monitoring.

### Categories and Subject Descriptors

C.1.2. Computing Methodologies / Artificial Intelligence / Robotics / Sensors

### General Terms

Algorithms, Measurement, Performance, Standardization, Verification.

### Keywords

Human Tracking, Laser Line Scanners, Robotics, Safety.

## 1. INTRODUCTION

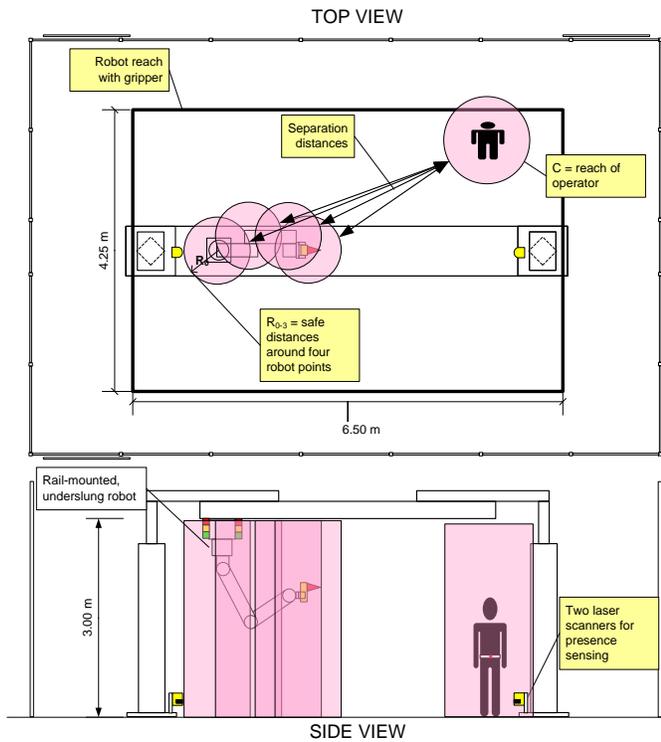
The Intelligent Systems Division (ISD) of the National Institute of Standards and Technology (NIST) is part of the team preparing the portion of ISO technical specification (TS) 15066 that deals with a form of collaborative robot safety termed speed and separation monitoring (SSM). SSM prevents contact between a moving robot and any person in the workcell by limiting robot speed and maintaining an adequate separation distance.[4] NIST has developed a prototype SSM safety system that uses laser range and detection scanners to measure the position and velocity of humans (or any moving objects) and computes the separation distance between the human and robot based on the robot's reported position and velocity. The system issues stop or slow signals depending on a minimum separation distance equation proposed in the ISO TS.

## 2. TESTBED

Our system consists of an under-slung robot mounted on an overhead rail (Figure 1 & Figure 2). The human tracking is done using two laser line scanners mounted horizontally and facing each other from opposite ends of the work volume. The system uses two laser scanners, one mounted horizontally to the base of each column that supports the under-slung robot rail (see Figure 2). The scanners are mounted at 0.39 m and 0.41 m above the floor facing each other on opposite sides of the robot work volume 5.05 m apart. This configuration detects the entire robot work area and reduces stationary and moving object occlusions. Also, placing the scanners below the robot's reach eliminates the need to discriminate between the robot and other objects that have entered or moved since the system was initialized. The system distinguishes between people and static objects such as the legs of a conveyor table and the rail support structure by subtracting a previously recorded background scan from regular scans during normal operation. For collaborative operation, the tracking system sends the position and velocity of each person to the SSM safety system. The safety system slows or stops the robot based on the relative distance between the robot and the nearest human. This allows the robot to move through one part of the work-volume while a person moves through another part of the volume.



Figure 1 Under-slung Robot under Rail



**Figure 2 Robot Test-Bed Setup**

### 3. SSM Controller

Equation (1) shows the collaborative form of the minimum separation distance equation.

$$S = K_H * T_R + T_B + K_R * T_R + B + C \quad (1)$$

Where:

- $K_H$  = Speed of human
- $K_R$  = Speed of robot
- $T_R$  = Reaction time to detect human and issue a stop – a parameter measured during timing test.
- $T_B$  = Brake time – see below.
- $B$  = Brake distance – see below
- $C$  =  $C_H + C_R$ , the region surrounding the human and robot respectively. For the testbed, this region includes the uncertainty in position and dimension of each

For the SSM testbed, the brake distance is:

$$B = (K_R^2) / 2A$$

$$T_B = K_R / A.$$

$A$  = Acceleration: worst-case deceleration level measured during stopping tests

The robot reports its own position and velocity ( $K_R$ ) while the human tracking system uses the laser scanners to report the positions and velocities ( $K_H$ ) of each person or unaccounted for object detected in the work-volume. The distance between the robot and each human is computed by the SSM controller. The SSM controller issues a stop whenever the distance to any human is less the minimum separation distance ( $S$ ).

## 4. HUMAN TRACKING

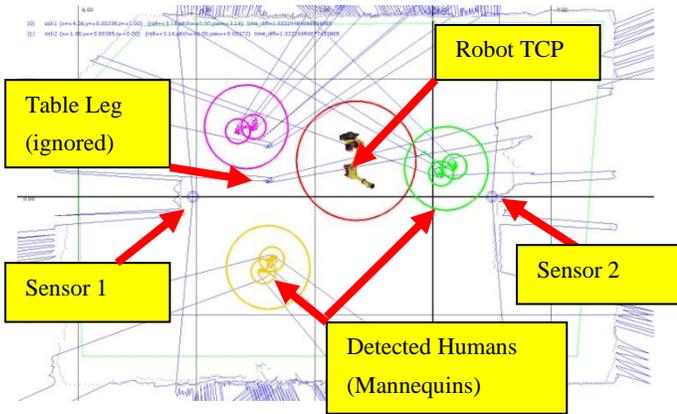
The human tracking system is an expanded version of a system we developed for inexpensive ground-truth measurement [3].

The tracker combines the range values into a single coordinate system. To accomplish this, the operator must first establish the position and orientation offset between the two sensors. This is done manually by visually aligning on a display the scans produced by each laser scanner. An object is placed in the Field of View (FOV) of both laser scanners. The operator drags the display of the object from one laser scanner over the display of the same object from the other laser scanner and rotates the object until the displays are aligned.

The background is recorded which contains all the static scanned objects in the FOV. Several frames of data are taken and combined to reduce sporadic noise. Objects seen during this background scan include the legs of a conveyor and the two columns supporting the robot. The tracker detects humans by detecting changes between the current range measurements and those recorded in the static background. Areas where background static objects exist are not processed by the tracker. This eliminates the problem where someone stands still in the robot work volume and eventually is considered part of the background. However, the operator needs to reestablish the background when static objects are moved. Otherwise a human could enter undetected through the previously occupied space. Future work will examine ways to automatically detect changes and automatically update the background.

The human tracking is calibrated to convert positions received from a coordinates system relative to each sensor to positions in the robot's coordinate system. The registration procedure uses a 10 cm (3.9in.) diameter x 91 cm (36 in) high tube placed in the robot's gripper facing down toward the floor. The robot is driven to three widely-spaced positions with the tube low enough to intersect the laser scanner plane. The robot's positions appear on the display along with the tracking system's measurement of the tubes. The operator uses display controls to manually align the robot position and the tube and software automatically calculates the transformation. All subsequent human tracker positions are transformed into the robot's coordinate system enabling the SSM controller to compute the correct separation distances.

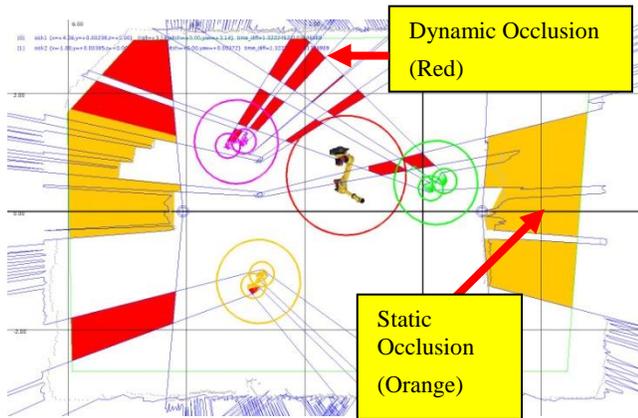
During SSM operation, the tracker groups range values into leg groups and human (center of two legs) groups, matches groups from previous groups, maintains a history of the group, and filters the position of each human using a Kalman filter. The filter assumes constant velocity will be maintained and can be tuned by setting the expected acceleration variance and measurement variance. The final position and velocity of the human sent to the SSM controller are taken from the estimated state of a Kalman filter. The results of this tracking are shown in Figure 3.



**Figure 3** Tracker display showing sensor location, range data, locations of moving objects and the location of the robot's tool center point.

## 5. OCCLUSIONS

One issue is occlusions due to multiple objects or people blocking the laser scanner FOV. These occlusions can mask the approach of other people thereby preventing the SSM from issuing a stop. We extended the tracker to detect occluded regions. The results of the occlusion detection algorithm are shown in **Figure 4**. The figure shows regions occluded by static objects (yellow) computed from the background range data and regions occluded by dynamic objects (red) computed from the tracking range data.

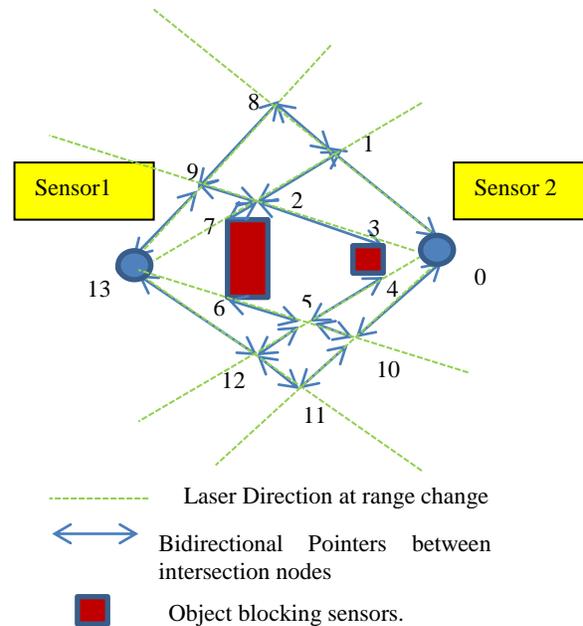


**Figure 4** Detecting regions the sensors can't see due to occlusions by fixed objects and by moving objects.

## 6. GRAPH SEARCH ALGORITHM

To find the occluded areas the tracker creates a bidirectional graph network. Each node in the graph contains the location where the laser was reflected and a node number obtained by incrementing a global count as each node is added. The node is connected to the sensor location for the first and last element in each sensor's range scan. The sensor locations are added as nodes so the graph can be traversed more easily. Points other than the first and last element are connected to the next and previous node. The size of the graph is reduced by combining consecutive nodes of approximately equal range from the sensor. The size of the graph is also reduced by combining all consecutive points outside a manually chosen protected area polygon. The system creates a graph for each sensor. The graphs are combined by searching for intersecting rays between nodes. At each intersection the connections between the original nodes are broken and all involved points are

connected to the new node at the intersection. The combined graph is searched to find all polygons. To find a polygon, begin at any node, and then traverse to any node connected to it. After the first move always choose the next connected node with the smallest possible angle to the previous node. Repeat until you return to the starting node. If you go to every node and apply this to every connection, you will have many polygons stored redundantly. For example, the polygon found starting at node 2 of 2,3,5,6,7 in Figure 5 would also be found by starting at 6 as 6,7,2,3,5. To eliminate these redundancies each polygon is normalized by starting the polygon at the minimum node number. The polygons can then be compared to eliminate the redundant ones. The outer polygon (in the example 0,1,8,9,13,12,11,10) will also be found in this way and is eliminated by testing any point not on the edge of the polygon to determine if it is inside the polygon. Each polygon in the list is labeled as occluded or not by testing one internal point to determine if the polygon is visible to at least one of the laser scanners. The internal point is computed by averaging three consecutive points in the polygon with an internal angle less than  $180^\circ$ . The point is tested by comparing its distance to each sensor with the range reported by that sensor at the appropriate angle.



**Figure 5.** Example Graph network for occlusion analysis.

## 7. Protected Area Polygon

The sensors can see areas on the other side of the fence that are not of concern for safety. To reduce the processing time needed to find obstacles and occlusions, a polygon drawn approximately just inside the fence line is added to the graph. Objects and obstacles outside this protected area are ignored.

## 8. Simulation

A simulator was developed to test large combinations of obstacle locations. The simulator places a given number of 0.3 m (1 ft.) diameter circular obstacles at random locations within the protected area polygon. Obstacle locations that would overlap a

laser scanner are regenerated. For each range measurement a laser scanner produced, the simulator calculates the distance to the outside edge of the closest obstacle and adds 2.5 cm (1 in) standard deviation Gaussian noise to the range measurement. The noise parameter was chosen from the laser scanner’s data sheet and the obstacle radius was chosen based on the approximate cross sections of our mannequins at the average height of the laser scanners. For each test the simulation generates one thousand combinations of obstacle locations.

## 9. Ground Truth Sampling

It is not really practical to use a high-precision range sensor to provide ground truth as to whether a given position should have been marked as occluded. Any displacement between the ground truth sensor and the laser scanner under test could make a position occluded for one sensor and not for the other. Instead, we use a simpler and more robust algorithm. This method works only at a single point in space. The distances to the point from the two laser scanners are compared against the range value provided by that scanner in the direction of the point. If any range measurement is greater than the distance to the point, the point is visible or else it is occluded. The area within the protected area polygon is randomly sampled and ground truth is only computed at those sample locations. Some points will be sampled within the radius  $C_H$  around a detected person or obstacle. Those points are ignored for purposes of occlusion ground truth since the robot would be required to stop as if there were a person there regardless of whether the point was occluded or not.

## 10. Performance Metrics

The following values were computed for each simulated or real-sensor data experiment as metrics for the effectiveness and/or efficiency of the system.

**Processing time** – average wall clock time measured as the system computes the occluded area. It does not include time for the robot to respond, nor for the raw data to be collected.<sup>1</sup>

**Percentage Occluded** – the percentage of the area as reported by the graph algorithm as occluded.

**Percentage False Occluded** – the percentage of sampled points labeled as occluded by the system under test but visible in the ground truth.

**Percentage False Visible** – the percentage of sampled points labeled as visible by the system under test but occluded in the ground truth.

**Percentage of unseen obstacles** - the percentage of obstacles that were more than  $C_H$  away from any detected person.

## 11. Simulation Results

The results of the first simulation set of tests are summarized in Table 1. One of the most disturbing results is the percentage of unseen obstacles with even two obstacles in the scene. The primary reason for this was that there was a blind area behind each laser scanner visible only to the laser scanner on the opposite side. Fortunately this area is not within the robot’s work volume. However people in these areas could be moving towards the robot work volume while their positions and velocities were not being reported to the robot due to the occlusion.

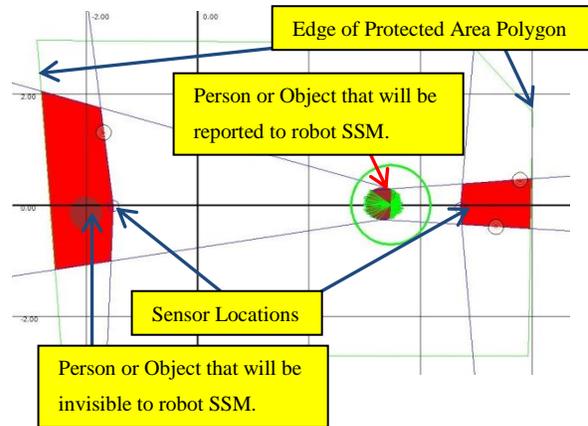


Figure 6. Image from simulation showing occluded areas behind each sensor and one unseen person/obstacle.

## 12. Real Laser Scanner Data Results

There are a number of problems with trying to reproduce simulated results with real sensor data. While it is easy to actually move people or mannequins around randomly, it is more difficult to ensure that their random positions were not biased to avoid or create occlusions. To ensure the positions were really chosen at random, the same computer program that generated obstacles for the simulation generated a set of obstacle positions to which the mannequins were then moved. The protected area polygon needed to be modified to eliminate areas that the mannequins could not be placed because the area was occupied by a conveyor table or a robot support. Figure 7 and Figure 8 show one snapshot taken from this data. Table 3 provides the cumulative averages for the entire set of tests, including 10 random position combinations and 100 frames of data collected for each combination.

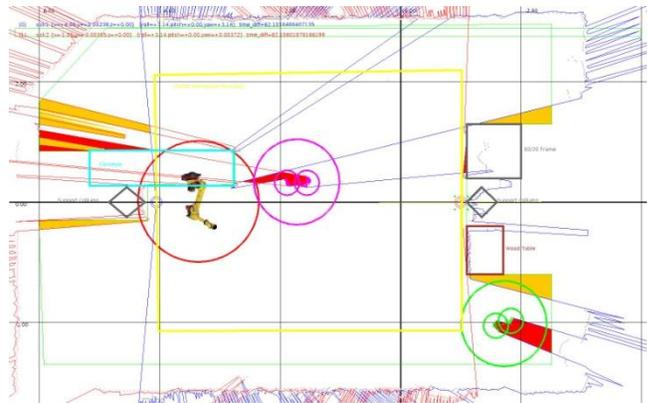


Figure 7. Snapshot of data collected from real laser scanners with mannequins placed at randomly generated obstacle positions.

<sup>1</sup> Tested on 2-core 2.1 GHz 32-bit laptop.



**Figure 8. Robot testbed with mannequins moved into places selected randomly by computer.**

### 13. Occlusion Mitigation Strategies

There are several possible strategies for reducing or eliminating the risks of the robot failing to stop or slow because the person was in an occluded area.

- 1) Shut down the robot whenever the number of detected people exceeds some maximum. This assumes that no person would be completely occluded unless at least some number of people is detected. The system could be tested and proven to handle at least that number of people. Since the number of people detected was an output from the existing human tracking system, occlusions do not have to be analyzed in real-time. Equipment being carried that hangs down below the height of the laser scanners could be considered an additional person. This might cause unnecessary and unexpected shutdowns. Additional sensors could be added to allow more people to be detected and allowed in the area or to reduce the chances of a person being occluded. This was tested for our testbed in simulation. (See Table 2 for the results.) The additional sensors reduced the size of the occluded regions and the probability that a person would be fully occluded. The additional sensors also increased the amount of processing required to compute the occluded polygons.
- 2) Use physical barriers to prevent people from standing in areas that would cause a large area to be occluded. The laser scanners could also be used to enforce a policy where some areas of the work volume could be used for a collaborative activity and other areas would result in an immediate shutdown upon detection of people.
- 3) Occlusion software could execute in real-time if there are sufficient computing capabilities. Either the occlusion monitoring

software or the SSM could use the list of occluded regions to compute the distance of the robot to the nearest occlusion and then compare the distance to the minimum separation distance given in Equation (1) as it does with the positions of people to determine when to shut down the robot. Since no estimate of a person's speed can be measured when they are occluded from the laser scanner, a constant maximum for  $K_H$  would have to be used. It may be necessary to use a less accurate, although faster, method of determining the occluded regions, such as sampling only the centers of grid squares.

### 14. Conclusions

Allowing humans to work in close proximity to robots will require an ability to detect people in and around the robot work volume. One technology already being used to protect people near robots is the laser line scanner. Although laser scanners are primarily being used only to shut down the robot, they can be adapted to provide real-time robot positions and velocities to allow the robot to adapt to the presence of people. One challenge in making this transition is accounting for the possibility that the laser scanners may be occluded. We presented a method for finding polygons of occluded areas and a way of testing such methods. This could be used either offline or online. Offline it could be used to show that the laser scanners are unlikely to be occluded for a region large enough to hide a person. Online the system could be used to stop or slow a robot before a person enters an occluded area.

### 15. REFERENCES

- [1] ISO 10218-1 Robots and robotic devices – Safety requirements – Part 1: Industrial robots, July 1,2011, [www.iso.org](http://www.iso.org)
- [2] ISO 10218-2 Robots and robotic devices – Safety requirements – Part 2: Industrial robot systems and integration, July 1,2011, [www.iso.org](http://www.iso.org)
- [3] W.P. Shackleford, T.H. Hong, T. Chang, “Inexpensive Ground Truth and Performance Evaluation for Human Tracking using multiple Laser Measurement Sensors.” *Proceedings of the 2010 Performance Metrics for Intelligent Systems (PerMIS) Workshop.* 2010. [http://www.nist.gov/manuscript-publication-search.cfm?pub\\_id=906630](http://www.nist.gov/manuscript-publication-search.cfm?pub_id=906630)
- [4] Testbed for Evaluation of Speed and Separation Monitoring in a Human Robot Collaborative Environment, October, 2011, NISTIR

**Table 1 Simulation results for two sensors with original layout**

Number of Obstacles	Processing time (ms)	Percentage Occluded	Percentage False Occluded	Percentage False Visible	Percentage of unseen obstacles
1	26	2.28	0.03	0.00	0.00
2	37	4.92	0.05	0.00	1.15
3	61	7.10	0.06	0.00	2.43
4	84	9.71	0.08	0.00	3.32
5	110	11.74	0.10	0.00	4.78
6	133	14.01	0.10	0.01	6.11
7	163	16.38	0.12	0.02	6.22
8	179	18.69	0.11	0.01	8.58
9	224	20.85	0.11	0.01	9.12
10	226	22.93	0.14	0.01	10.41

**Table 2 Simulation Results for four sensors.**

Number of Obstacles	Processing time (ms)	Percentage Occluded	Percentage False Occluded	Percentage False Visible	Percentage of unseen obstacles
1	231	0.33	0.00	0.00	0.00
2	691	0.72	0.00	0.00	0.00
3	1490	1.15	0.00	0.00	0.03
4	2555	1.69	0.00	0.00	0.00
5	4002	2.28	0.01	0.00	0.00
6	5622	3.02	0.01	0.00	0.07
7	7752	3.77	0.01	0.00	0.17
8	9616	4.66	0.02	0.00	0.26
9	11143	5.56	0.03	0.01	0.27
10	12558	6.49	0.03	0.01	0.52

**Table 3 Results using real sensors and mannequins**

Number of Obstacles	Processing time (ms)	Percentage Occluded	Percentage False Occluded	Percentage False Visible	Percentage of unseen obstacles
2	45	5.60	0.17	0.11	0.00