# Basic cost matrix

Table 1 cost matrix of detections and tracks with row and column headers

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **Detection 1** | **Detection 2** | **Detection 3** | **…** | **Detection n** |
| **Track 1** | Distance 1, 1 | Distance 1, 2 | Distance 1, 3 | Distance 1, j | Distance 1, n |
| **Track 2** | Distance 2, 1 | Distance 2, 2 | Distance 2, 3 | Distance 2, j | Distance 2, n |
| **Track 3** | Distance 3, 1 | Distance 3, 2 | Distance 3, 3 | Distance 3, j | Distance 3, n |
| **…** | Distance i, 1 | Distance i, 2 | Distance i, 3 | Distance i, j | Distance i, n |
| **Track n** | Distance n, 1 | Distance n, 2 | Distance n, 3 | Distance n, j | Distance n, n |

The basic cost matrix has detections as columns and tracks as rows. It is a square matrix which represents detections to be assigned to tracks. represents the cost to assign the detection to the track, calculated based on the Euclidean distance between the detection and the estimated position of the track in the frame.

In the case of uneven detections and tracks, i.e. detections to be assigned to tracks, the cost matrix still has to be a square matrix where . A common strategy to convert the rectangular matrix into a square matrix is to fill the missing rows or columns with zeros, where . An example with , m is shown below. In this way, the Munkre’s assignment algorithm can still produce assignments which will match the best track detection pairs.

Table 2 cost matrix of 4 detections and 2 tracks with row and column headers

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Detection 1** | **Detection 2** | **Detection 3** | **Detection 4** |
| **Track 1** | Distance 1, 1 | Distance 1, 2 | Distance 1, 3 | Distance 1, 4 |
| **Track 2** | Distance 2, 1 | Distance 2, 2 | Distance 2, 3 | Distance 2, 4 |
| **-** | 0 | 0 | 0 | 0 |
| **-** | 0 | 0 | 0 | 0 |

The problem with this approach is that the best track detection pairs may not always be optimal for the problem. Consider the case where the best track detection pairs are still a considerable distance from each other due to the tracked object being occluded in the frame resulting in its track being matched to noise. Clearly this is undesirable, especially in the case where the noise was not generated by the object as it would introduce errors into our predictions. In order to account for this deficiency, two changes need to be made. Firstly, additional dummy tracks and detections need to be added in to allow for non-assignment, where assignment of actual tracks or detections to these dummy tracks and detections would represent an unassigned detection and track respectively. Secondly, to avoid all tracks and detections being matched to dummy tracks and detections, a threshold value needs to be introduced, representing the cost of non-assignment as opposed to matching distant tracks and detections. This cost represents the threshold distance and is similarly measured in the same units as the cost of assignment.

# Modified cost matrix

The modified cost matrix is based off the assignDetectionsToTracks function from MATLAB and is where the idea of using a padded cost matrix combined with separate costs for unassigned detections and tracks came from.

To keep things simple, let us consider the case where , m shown below.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Detection 1** | **Dummy Detection** | **Dummy**  **Detection** |
| **Track 1** | Distance 1, 1 | Cost of unassigned track | Cost of unassigned track |
| **Track 2** | Distance 2, 1 | Cost of unassigned track | Cost of unassigned track |
| **Dummy Track** | Cost of unassigned detection | 0 | 0 |

First off, note that the size of the matrix is . The matrix has square to allow the application of Munkre’s assignment algorithm and has maximum dimensions of in order to accommodate the case where all detections and tracks are located outside the threshold distance from each other. In this case, all actual tracks and detections need to be assigned to dummy tracks and detections, requiring dummy tracks and dummy detections. This results in dimensions of which can be rewritten as

Secondly, the modified cost matrix is divided into four sections. Going clockwise, they are the valid assignments, unassigned tracks, dummy pairs and unassigned detections. To understand how these sections work, let us consider the following cases.

When the first detections are made, there are no tracks created yet. In the case of , m, all detections can only be made to the unassigned detection section.

|  |  |  |
| --- | --- | --- |
|  | **Detection 1** | **Detection 2** |
| **Dummy Track** | Cost of unassigned detection | Cost of unassigned detection |
| **Dummy Track** | Cost of unassigned detection | Cost of unassigned detection |

When the number of objects in the scene does not change, the ideal case would be that all detections would have come from tracked objects leading to a case where and where for all detections and tracks is lower than the cost of assignment. In the case of , m, assignments will only be made to the valid assignment and dummy pairs sections.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Detection 1** | **Detection 2** | **Dummy Detection** | **Dummy**  **Detection** |
| **Track 1** | Distance 1, 1 | Distance 1, 2 | Cost of unassigned track | Cost of unassigned track |
| **Track 2** | Distance 2, 1 | Distance 2, 2 | Cost of unassigned track | Cost of unassigned track |
| **Dummy Track** | Cost of unassigned detection | Cost of unassigned detection | 0 | 0 |
| **Dummy Track** | Cost of unassigned detection | Cost of unassigned detection | 0 | 0 |

Most likely however, noise and occlusions will result in a mismatch in the number of detections and tracks. Let us first consider the case of noise, when extra detections further than the threshold distance from any track are generated. In the case where , m, assignments will first be made to match the detection generated by the tracked object and the dummy pair. The remaining assignment can thus only be made between the remaining detection and the remaining dummy track.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Detection 1** | **Detection 2** | **Dummy Detection** |
| **Track 1** | Distance 1, 1 | Distance 1, 2 | Cost of unassigned track |
| **Dummy Track** | Cost of unassigned detection | Cost of unassigned detection | 0 |
| **Dummy Track** | Cost of unassigned detection | Cost of unassigned detection | 0 |

In the case where a tracked object has been occluded, it is likely that the track will be located further than the threshold value from the other detections. In the case where , m, assignments will once again be made to the valid assignments and dummy pairs first, before matching the remaining track with the remaining dummy detection.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Detection 1** | **Dummy Detection** | **Dummy Detection** |
| **Track 1** | Distance 1, 1 | Cost of unassigned track | Cost of unassigned track |
| **Track 2** | Distance 2, 1 | Cost of unassigned track | Cost of unassigned track |
| **Dummy Track** | Cost of unassigned detection | 0 | 0 |

New objects appearing in the frame behave similarly to noise in the context of a single frame, in that they result in extra detections far away from existing tracks. Objects leaving the frame behave similarly to occlusions in that they result in extra tracks located far away from existing tracks.

Finally let us consider a case where both noise and occlusions occur simultaneously. In such a case, there will be both tracks located far from detections and detections located far from tracks. While the number of detections could differ from the number of tracks, for simplicity let us consider a case where , m, all detections arise from noise and all tracks are occluded. Assuming the threshold distances are appropriately chosen, they will be lower than the distances within the valid assignments section. In this case, the assignment algorithm is likely to chose to match tracks with dummy detections and detections with dummy tracks, with no assignments made to the valid assignments or dummy pairs sections.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Detection 1** | **Detection 2** | **Dummy Detection** | **Dummy**  **Detection** |
| **Track 1** | Distance 1, 1 | Distance 1, 2 | Cost of unassigned track | Cost of unassigned track |
| **Track 2** | Distance 2, 1 | Distance 2, 2 | Cost of unassigned track | Cost of unassigned track |
| **Dummy Track** | Cost of unassigned detection | Cost of unassigned detection | 0 | 0 |
| **Dummy Track** | Cost of unassigned detection | Cost of unassigned detection | 0 | 0 |

Of course, with a mix of detections arising from tracks, noise and occlusions, it is likely that in practice assignments will be made to all four sections. In fact, this is the desired behaviour, where the assignment algorithm can properly categorise the detections and tracks into these distinct sections. However, depending on the threshold distance chosen as well as how detections are generated, it is entirely possible that incorrect assignments may occur.

Therefore, other filtering techniques such as pre-processing to remove as much noise as possible and making tracked objects more easily detectable as well as time dependent techniques such as monitoring the