CS365

Artificial Intelligence

Project Report

Motion Tracking In Aldebaran Nao

Group Members

Sarvesh Kumar Singh

Harshal Agrawal

Abstract

Our work is aimed at trying to learn the motion flow by observing the Temporal-Informational Correlations between the sensors (visual) and actuators of Aldebaran Nao. We will also try to learn to follow a particular motion by learning the motion flow.
Our work is primarily inspired by the work of Olsson and Nehaniv[1], but we are using the different method to do the same.

Introduction

One of the amazing capabilities of many sensory systems is the ability to adapt to the current environment and to the events occurring in the current environment. In our work we are trying to make Nao able to adapt to the visual events in the environment. By adapting to the visual events we mean being able to detect the direction flow by calculation the direction of the optical flow in the visual data.

In our method for the motion flow detection we are generating the motion flow by changing the angle of the head of Aldebaran Nao by very little amount over time. This causes a relative motion flow between the environment and the Nao's visual sensors, then we are trying to detect the direction of this motion flow.

Input Used

We are taking the images from Nao with different head angles as input for our work, then we are using our method to find the direction of motion as observed by Nao using these images.

Our Method

First we are taking the two images with two different head angles and two different time stamps (say Image Im1 at time t and Image Im2 at time $t+t_0$).
Now we are taking 200 random pixels (Im1cpx1, Im2cpx2, Im3cpx 3 …) in the image Im1 and correspoong pixels (Im2cpx1, Im2cpx2, …) in the image Im2.
Then with each of these 200 points in image Im1 as center we are taking a block of pixels size 81X81 and then calculating the SSD of this "center-pixel" (let's say Im1cpx1) with each pixel in the corresponding 81X81 block in the image Im2 over a 7X7 SSD window.

Now, we calculate the distance [dX, dY] of the pixel with minimum SSD's of the 81X81 pixels from the corresponding "center-pixel" (Im2cpx1 in this case) in image Im2. In this way we have 200 [dX, dY] values each corresponding to one of the 200 random pixels seleted. Now we are taking the mean of these 200 [dX, dY] values to obtain a single [dX, dY] value. This value gives us the direction of motion flow as observed by Nao.

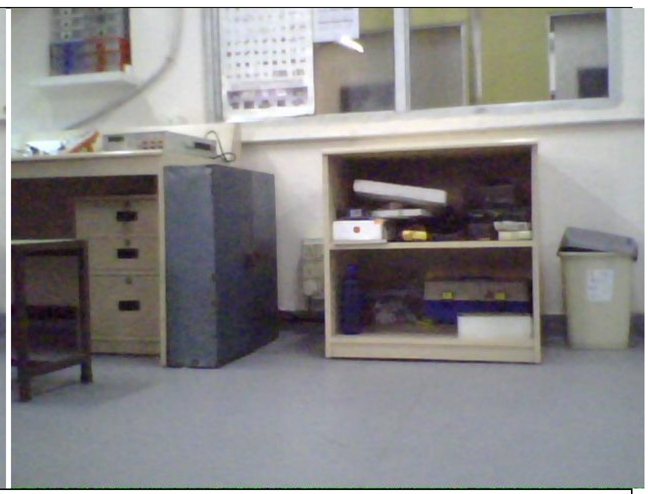
For example : [dX, dY] = [15, 0] means the Right direction flow.
[dX, dY] =of [-15, 0] means the Left direction flow.
[dX, dY] = [0,15] means the Down direction flow.

[dX, dY] = [0, -15] means the Up direction flow.

Results

The results for the various image sets are given below along with the images.

Image Set 1

	
Head Angle = 0.159494042397 rad	Head Angle = 0.177902042866 rad

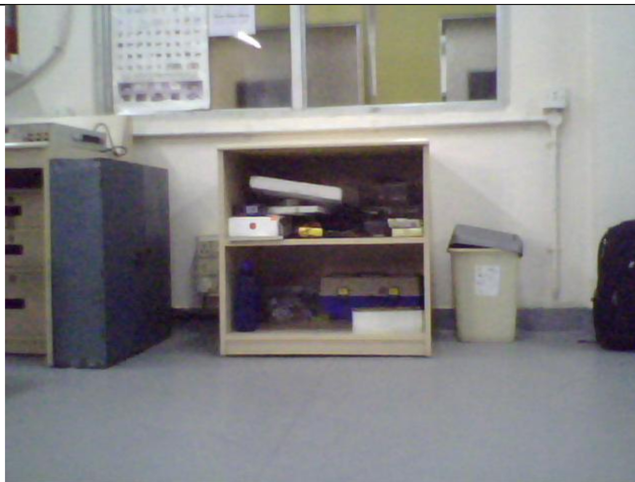
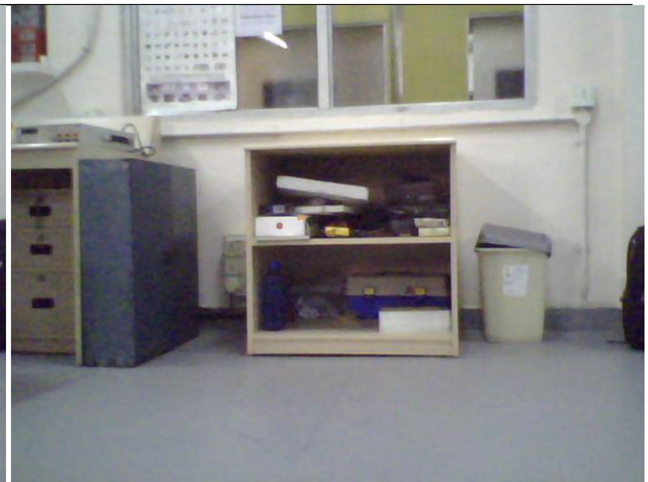
Anlgle difference = 0.019408 radians

Value = [15, 0]   Concluded Direction of flow : Right.

Original Direction of flow : Right.

The result was correct.

Image Set 2

Head Angle = 0.0521140396595 rad	Head Angle = 0.0781920403242 rad

Anlgle difference = 0.026078 radians

Value = [20, 0]   Concluded Direction of flow : Right.

Original Direction of flow : Right.

The result was correct.

Image Set 3


Head Angle = 0.00762803852558 rad	Head Angle = 0.0352400392294 rad

Anlgle difference = 0.027612 radians

Value = [21, 1]   Concluded Direction of flow : Right-Down.

Original Direction of flow : Right.

The result was nearly correct, but this result is showing a little downward motion too. This is due to noise in the image and can be reduced by smoothing the image in the first step.

Image Set 3 – with Gaussian Smoothing

Anlgle difference = 0.027612 radians

Value = [21, 0]   Concluded Direction of flow : Right.

Original Direction of flow : Right.

Extra Finding

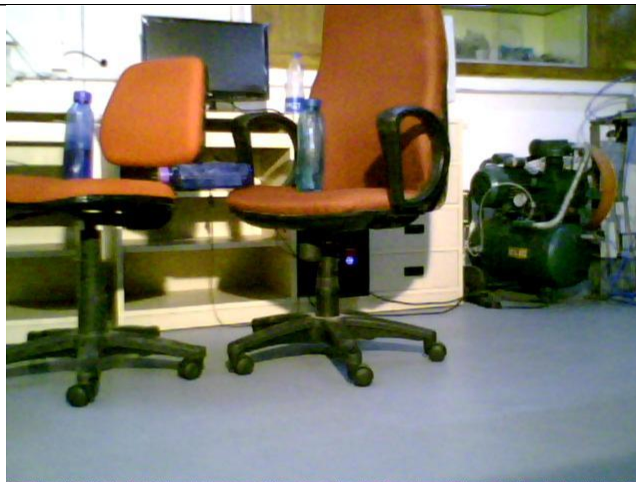
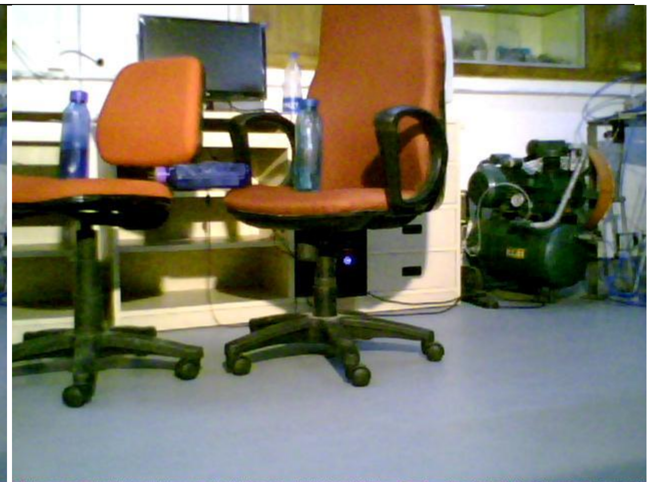
Here one important thing to notice is that if the head angle difference is more the [dX, dY] values also increases or in other words, the [dX, dY] values are also increasing with increasing difference in the images.

In this way we can also tell which image correspond to a larger motion by comparing the [dX, dY] values.

Angle Difference	**[dX, dY]**
0.019408 radians	[15, 0]
0.026078 radians	[20, 0]
0.027612 radians	[21, 0]

Now for set of Images from another scene.

Image Set 4

	
Head Angle : 0.012230038642 rad	Head Angle : 0.00149203836918 rad

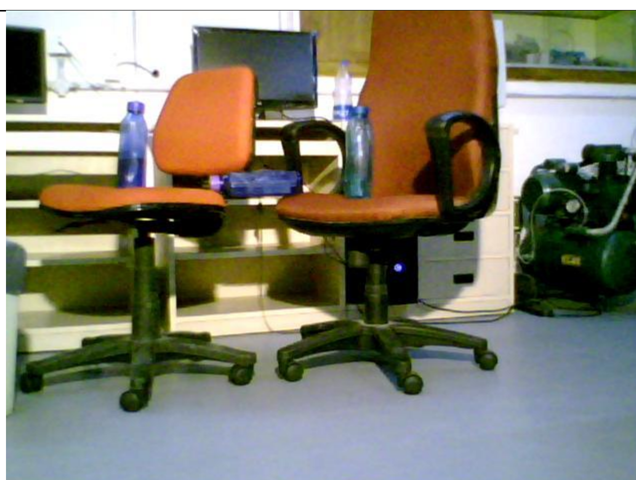
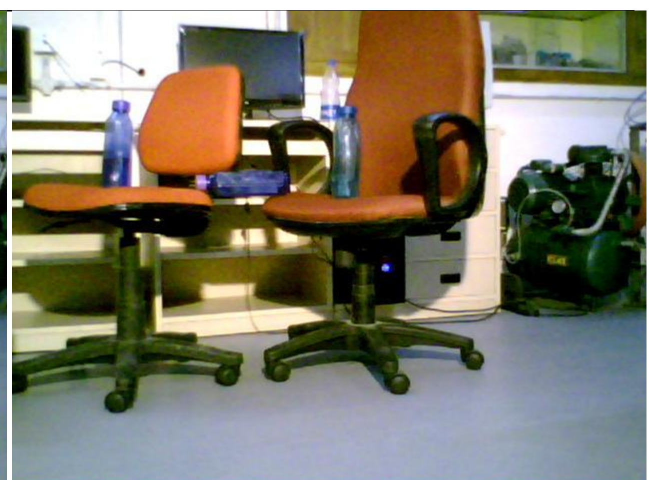
Anlgle difference = - 0.0107380002737 radians

Value = [-9, 0]   Concluded Direction of flow : Left.

Original Direction of flow : Left.

The result was correct.

Image Set 5

	
Head Angle: 0.0735900402069 rad	Head Angle : 0.0490460395813 rad

Anlgle difference = - 0.0245440006256 radians

Value = [-20, 0]   Concluded Direction of flow : Left.

Original Direction of flow : Left.

The result was correct.

We can see that the larger head angle difference is producing more negative dX value implying larger left motion as compared to the image set with smaller head angle difference.

Conclusion

Unlike the Olsson's algorithm in which they were first shrinking the image and then were doing the adaptive binning on the image and then were using this binning data to construct the sensorytropic map of the sensors, Our algorithm is giving better results on the original image (with Gaussian Smoothing) as compared to shrank images.

Even with the original images (without any Gaussian smoothing) we are obtaining almost accurate results.

We checked the algorithm for a variety of scenes (texture variation) and it was performing very well in all the situations.

Advantage over Ollson's algorithm : By comparing the [dX, dY] values for two sets of images from the same scene we can that which set corresponds to larger motion of the scene, which was not possible with the Olsson's algorithm.

Refrences

[1] Olsson, L., Nehaniv, C. L., and Polani, D. 2005. Discovering motion flow by temporal-informational correlations in sensors. Proceedings of the Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems, pp. 117-120. Lund University Cognitive Studies, 2005.