Localisation using Active Mirror Vision System

Luke Cole (u4014181) Supervised by Dr. David Austin

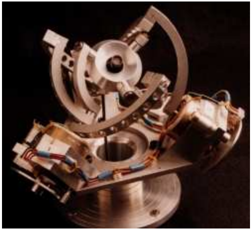
September 14, 2005

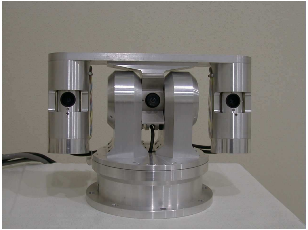
Localisation

Localisation consists of answering the question "Where am I?" from the robot's point of view.

- That is, a problem of estimating the robot's *pose* (position, orientation) relative to its enviroment.
- The robot's pose is typically the x and y coordinates and heading direction (orientation) of the robot in a global coordinate system.

Active Vision



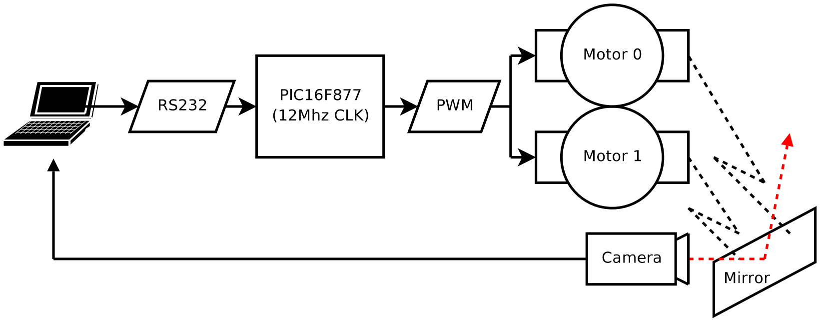


Approach

- **Novel Vision System**: Camera and motors mounted to fixed platform and camera view point changed via re-orienting a mirror.
- **View Selection algorithm**: Continuously re-orient vision system to most significant visual landmark. The most significant landmark is determined by considering:
    - *Visibility* of landmark.
    - *Orientation time* to landmark.
    - *Variance of probability distribution*.

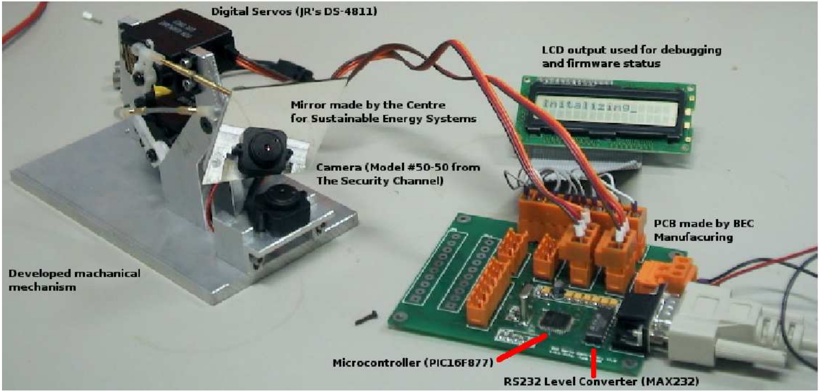
It was found the robot could best localise itself using a video frame rate of 1Hz.

Design and Architecture



Primary Design Requirements	
Field of view	60°
Range of motion (vertical and horizontal)	60°
Angular resolution	0.09°
Velocity	$600^\circ.s^{-1}$

System Overview

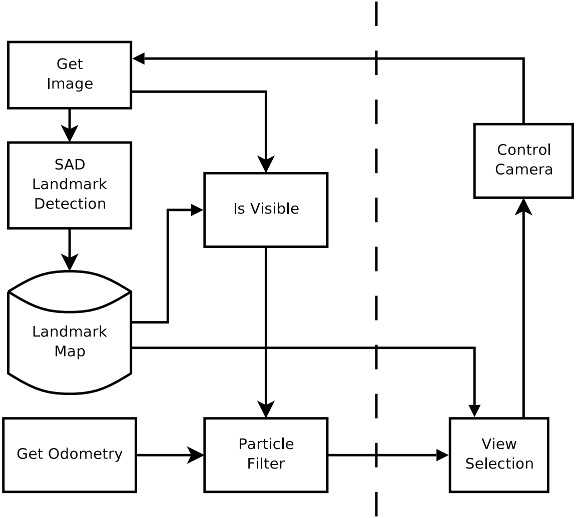


System Characteristics

Item	Qty	Item Cost (ea)
Digital RC Servo (JR DS8411)	2	150AUD
CMOS Pin-hole camera (Jaycar QC-3454)	1	90AUD
Mirror	1	30AUD
Machining (20 hours @ $40/h)	1	800AUD
Printed Circuit Board	1	100AUD
Electronic Components	1	60AUD
Total Cost		1380AUD

Specification	Unit	Measured Tilt	Measured Pan
Saccade Rate	Hz	3Hz	5Hz
Angular Resolution	$^\circ$	0.4	0.4
Angular Repeatability	$^\circ$	0.1	0.1
Max. Range	$^\circ$	90	45
Max. Velocity	$^\circ.s^{-1}$	666	666
Max. Acceleration	$^\circ.s^{-2}$	666	666

Localisation Algorithm



Visual Landmark Map

See window manager desktop (4).

Particle Filter

A robot's pose is represented by a probability distribution given by:

$$p(x_t \mid o_t, a_{t-1}, o_{t-1}, a_{t-2}, \ldots, a_0, o_0) \tag{1}$$

where, x denotes the robot *state* at time t, a denotes absolute position measurements and o denotes relative position measurements.

A particle filter algorithm represents equation (1) by a set of n weighted samples distributed according to equation (1), that is:

$$\{x^i, p^i\}_{i=1,\ldots,n} \tag{2}$$

where, x^i is a sample (particle) and p^i are called the importance factors, which sum up to one and determine the weight of each sample.

Using Bayes rule and Markov's assumption equation (1) can be put into recursive form known as *Bayes filter*:

$$\eta\rho \int \alpha p(x_{t-1} \mid o_{t-1}, a_{t-2}, \ldots, a_0, o_0) dx_{t-1} \tag{3}$$

where, η equals $p(o_t \mid a_{t-1}, d_{0\ldots t-1})^{-1}$, α equals $p(x_t \mid x_{t-1}, a_{t-1})$ and ρ equals $p(o_t \mid x_t)$.

The particle filter is an approximation of equation (3) and is generally performed as follows:

1. Robot moves. Move samples according to a_{t-1} using the motion model α.
2. Robot makes an observation, which yields the importance factors using the perceptual model ρ.
3. Normalise importance factors so they sum up to one.
4. Sample new particles according to the weights. Go to step (1).

IsVisible Algorithm for p^i

$$p^i = 1 - \frac{1}{n\sigma} \sum_{k=0}^{n} s_k \tag{4}$$

where p^i is the importance factor for the i^{th} particle, n is the number of landmarks, s_k is the score for the sum of absolute differences (SAD) between the k^{th} landmark and the new image, and σ is a constant defined by:

$$\sigma = Width \times Height \times BypesPerPixel \times MaxPixelIntensity \tag{5}$$

- If k^{th} landmark is not visible, $s_k = \sigma$.
- Landmark visibility determined by IsVisible algorithm, which maps the landmark global coordinates (in millimeters) to the image plane (in pixels), and if the coordinates exceed the image size, the landmark is not visible.

View Selection

Re-orient vision system to landmark k with maximum weight w.

$$w_k = \begin{cases} 0.0 & \text{if } BehindWall(p_{mean}, l_k) \\ 0.0 & \text{if } ExceedVisionLimits(p_{mean}, l_k) \\ \frac{v_k + t_k + p_k}{3} & \text{otherwise} \end{cases} \tag{6}$$

$$v_k = \frac{ABS(cos(AngleDiff(p_{mean}, l_k))) + \frac{l_{depth}}{Distance(p_{mean}, l_k)}}{2} \tag{7}$$

$$t_k = 1.0 - \frac{ReOrientationTime()}{t_{MAX}} \tag{8}$$

$$p_k = ABS(sin(AngleDiff(e, l_k))) \tag{9}$$

where, p_{mean} is the mean pose, l_k is the k^{th} landmark, t_{MAX} is the maximum orientation time, l_{depth} is the distance between the landmark and the camera when it was acquired for the map and e is the first eigenvector of the covariance matrix of the particles.

Results

See window manger desktop (4).

Conclusions

- Mirror based active vision system shows real potential as a solution to active vision. Developed system is cheap, fast and reliable.
- View selection worked as anticipated, adding efficiency to visual localisation and improving time to localise.
- 1Hz video frame rate best.

Future Work

- Mechanical modifications to mirror vision system to increase orientation angles.
- Faster microprocessor.
- Explore different materials such as plastic.
- Explore different methods to deriving the importance factors.
- Integration into simultaneous localisation and mapping (SLAM).

Acknowledgements

This work was supported by funding from National ICT Australia and the Australian National University. The Australian National University is funded by the Australian Government's Department of Education. National ICT Australia is funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council through Backing Australia's Ability and the ICT Centre of Excellence program.