**An Iterative Image Registration Technique with an Application to Stereo Vision**

Abstract

The image registration techniques till now tend to be costly so a new image registration technique is found with fewer potential matches making use of spatial intensity gradient using newton-raphson iteration.This can be generalized to handle rotations,scaling and shearing.

Introduction

Definition of image registration-It is transforming different sets of data into one co-ordinate system.It is necessary in order to compare or integrate data obtained from different measurements.

Stereo vision-Extraction of 3-d information from digital images.

This technique which uses spatial intensity gradient takes more information from the images and is able to find the best match between the images with fewer comparisons unlike other techniques that examine the possible positions of registration in some fixed order.Our technique takes the advantage that the two images are already in approximate registration.This technique can also be generalized to deal with linear distortions of the image including rotation.

**The registration problem**

Translational image registration technique has following characteristics:-

F(x) and G(x) are the functions which give the pixel values at x in 1st and 2nd images respectively. We have to find vector h that minimizes the measure of difference between F(x+h) and G(x) at some x.

Typical measures of difference between F(x+h) and G(x) are:-

* L1 norm=∑x€R│F(x+h)-G(x)│
* L2 norm=(∑x€R(F(x+h)-G(x))2)1/2
* Negative of normalized correlation

= (─∑x€R F(x+h)G(x))÷((∑x€RF(x+h)2)1/2(∑x€RG(x)2)1/2)

**Existing techniques**

The brute force approach is that if the size of image G(x) is NxN and possible values of h is of size MxM then it takes O(M2N2) time to compute.

So to reduce the time complexity we use hill climbing technique .This technique takes an initial estimate h0. To calculate hk+1 from hk we evaluate the difference function at all points in a small 3x3 neighbourhood of hk and finds hk+1.

Sequential similarity detection algorithm(SSDA) estimates error for each difference vector h.One stops estimating the error for h if it is not likely to give the best match.

This algorithm takes an initial estimate h and uses spatial intensity gradient to obtain next h so this process is repeated in newton-raphson method so if the iteration converges it will do it in O(M2logN) steps.

**The registration algorithm**

One dimensional case

We have to find the disparity h between the two curves F(x) and G(x) so for minimum difference vector h G(x)=F(x+h).

We know that F′(x)=(F(x+h)-F(x))/h.

Now F′(x)=(G(x)-F(x))/h.

So h=(G(x)-F(x))/ F′(x).

Approximation to h depends on x.Easy way to find correct h is to simply average h for various values of x.

F’’(x)= (G’(x)-F’(x))/h

This average can be improved using a weighting function

w(x)= 1/(| G’(x)-F’(x)|)

The average with weighting is

h=(∑x(w(x)(G(x)-F(x))/ F′(x))/∑xw(x)

We can repeat this procedure of finding the estimate of h using newton-raphson iteration which is expressed by h0=0,

hk+1=hk+(∑x(w(x)(G(x)-F(x+hk))/ F′(x+hk))/∑xw(x)

**Performance of new registration technique:**

For finding hk+1 we need F’G , F’F, (F’)2. We cannot calculate F’(x) exactly. So we estimate it by F’(x)=(F(x+∆x)-F(x))/ ∆x

And similarly for G’(x) we choose ∆x small eg:one pixel

**Generalization to multiple dimensions**

The technique can be generalized to two or more dimensions using the L2 norm of error E=∑x€R(F(x+h)-G(x))2 where x and h are n dimensional row vectors

F(x+h)≈F(x)+h(∂F(x)/ ∂x)

where ∂/∂x is gradient operator with respect to x as column vector ∂/∂x=[(∂/∂x1) (∂/∂x2)…. (∂/∂xn)]T

Using this approximation to minimize E we get

h≈[∑x(∂F/∂x)T[G(x)-F(x)]][∑x(∂F/∂x)T(∂F/∂x)]-1

which has much the same form as one dimensional version

**Further Generalizations**

This technique can be extended to registration between two images related by rotation, scaling and shearing. This is expressed by G(x)=F(xA+h) where A is a matrix expressing the linear spatial transformation between F(x) and G(x). The error to be minimized here is E=∑x[F(xA+h)-G(x)]2 .To determine amount ∆A and ∆h, we use linear approximation

F(x(A+∆A)+(h+∆h))≈F(xA+h)+(x∆A+∆h) ∂F(x)/∂x

When we use this approximation error expression becomes quadratic which can be solved by differentiating with respect to the quantities in the expression and solving a set of linear equations.

This is useful in stereo vision where the two different views of the object are different because of difference in view points of cameras. If we model this difference as linear transformation we have F(x)=αG(x)+β

Where α is contrast adjustment and β is brightness adjustment

For stereo vision we obtain E=∑x[F(xA+h)-(αG(x)+β)]2 as quantity to minimize with respect to α,β,A and h.

**Application to stereo vision**

**The Stereo problem:**

The problem of extracting depth information from a stereo pair has in principle four components:

Finding objects in pictures, Matching objects in the two views, Determining the camera parameters, Determining the distances from the camera to the objects.

Stereo vision systems that work with features at pixel levels can use our registration technique which has a fast method of doing pixel level matching.

**A mathematical characterization**

Let c be the vector of camera parameters that describe the orientation and position of camera2 with respect to camera1’s coordinate system. Let x denote position of image in camera1. Suppose the object is at distance z from camera1. Given x and distance z of the object we can find p(x,z) i.e.; position of object in 3-space. The object would appear on camera2’s film plane at a position q(p,c) i.e.; dependent on p and camera parameters c. Let G(x) be intensity value of pixel x in picture1. Let F(q) be the intensity value of pixel q in picture2. The goal of stereo vision system is to find a relationship described above and solve for c and z using x, F and G.

**Applying the registration Algorithm**

First lets find distance z considering that we know exact parameters c. The linear approximation that we use here is

F(z+∆z)≈F(z)+ ∆z ∂F/∂z

Where ∂F/∂z=(∂p/∂z)(∂q/∂p)(∂F/∂q)

To update our estimate of z we find ∆z that satisfies

0=∂E/∂∆z ≈ ∂∑x[F+∆z∂F/∂z-G]2 /∂∆z

Solving for ∆z we obtain

∆z=∑x∂F/∂z[G-F]/ ∑x(∂F/∂z)2

Similarly considering that we know the distances zi i=1,2,…,n of each of n objects from camera1 we can find camera parameter c.

Using linear approximation F(c+∆z)=F(c)+ ∆c(∂q/∂c) (∂F/∂q)

and on solving we obtain ∆c=

[∑x((∂q/∂c) (∂F/∂q))T[G-F][∑x((∂q/∂c) (∂F/∂q))T(∂q/∂c) (∂F/∂q))]-1

**An implementation:**

This technique was implemented in a system that function well under human supervision. This algorithm will converge to the correct distances and camera parameters when initial estimates of zi’s and c are sufficiently accurate that we know the position in camera2 film plane of each object to within a distance on the order if the size of object