MULTIVARIATE ENTROPY DETECTOR BASED HYBRID IMAGE REGISTRATION

Pavan S. Sridhar G. Sridhar V.

Applied Research Group, Satyam Computer Services Limited
#14, Langford Avenue, Lalbagh Road, Bangalore, India 560 025
Email: {Pavan_S,Sridhar_Gangadharpalli, Sridhar}@satyam.com
Phone: +91 80 2223 1696 Fax: +91 80 2227 1882

ABSTRACT

Image registration is used to match two images for spatial alignment and intensity alignment. One of the possible applications of image registration is for the evaluation of printed image with respect to a given reference image. We propose a new hybrid image registration algorithm to identify the spatial or intensity variations between two color images. The proposed approach extracts salient descriptors from the two images using a multivariate entropy-based detector. The transformation parameters are obtained after establishing the correspondence between the salient descriptors of the two images, which yields better accuracy and lesser computational cost compared to the approaches present in the literature.

1. INTRODUCTION

Image registration has been studied in various contexts such as medical imaging, remote sensing, computer vision and pattern recognition. Image registration can be classified into the following four basic types: i) Multimodal registration (Registration of images of the same scene acquired from different sensors) ii) Template registration (Finding a match for a reference pattern in an image) iii) Viewpoint registration (Registration of images taken from different viewpoints) and iv) Temporal registration (Registration of images of same scene taken at different times or under different conditions). The registration methods generally fall into one of the following three categories: i) Intensity based registration works directly on the intensity values of the image. The drawback of this technique is that it is computationally expensive. In addition, the poses of the two images under comparison have to be close enough to attain a local optimum. ii) Feature based registration is based on extraction and comparison of features of the images in consideration. Control points such as corners, edges, contours etc are used to match the reference image with a template. The disadvantage of the feature-based methods is that they are adversely affected by scale changes and the correspondences are generally less accurate. iii) hybrid registration methods combine the merits of both feature and intensity based methods.

In this paper we present a novel hybrid registration technique for comparing a color image with a reference image. The objective of this registration is to find the specific differences between the images including the identification of the transformation parameters.

2. RELATED WORK

Image registration has been an active topic of research for a long time. Various variations of registration techniques appeared in the literature and a comprehensive survey on image registration is presented by Lisa in [1].

A variation of intensity based image registration is proposed in [4] wherein the images are aligned by maximizing the mutual information between the image intensities of the corresponding voxels present in the two images. On the other hand the efficiency of feature based registration to a large extent depends on finding proper correspondence between the identified features in both the images. Different techniques were proposed for finding the proper correspondence between the images being compared. For instance in [2][3], the accuracy of the feature-based method is improved by finding the correspondence between the two images using a region present in the local neighborhood of an image feature (such as corners present in the image).

Recently several hybrid methods have been proposed which combine the merits of both feature and intensity based methods. Timor *et.al* in [6] proposed a hybrid technique, which uses an entropy-based method to extract salient regions, which act as meaningful descriptors of the image. These descriptors are then used for finding the correspondences between the two images. A variation of hybrid registration method is presented in [7] where the entropy based method has been extended to cover the affine transformations by replacing the circular sampling window by an ellipse. Further to this, entropy-based method has been applied to the image registration problem to calculate the transformation parameters in [5]. This work introduces a region configural-matching step in addition to the previous approaches of region component matching based on the salient region technique. The region configural matching step is used to detect a joint correspondence between the feature pairs, which will contribute to the global image "alignedness" and reduce the probability of occurrence of false matches. The drawback of this technique is that the similarity is calculated for every pair of regions and for each sampled value of the set of transformation parameters, which will be computationally expensive. In addition the checking carried out to ascertain the parameters (which will contribute to the global image alignedness) has to be performed for each of the transformation parameter obtained.

In this paper, we propose an algorithm for establishing the correspondence between two color images and to identify the transformation parameters, which attempts to overcome the above-mentioned drawbacks.

3. ENTROPY BASED IMAGE REGISTRATION

The ultimate aim of any image-matching algorithm is to match the reliability and ease with which the human visual system extracts meaningful descriptions of the image and uses it for image comparison. As mentioned in earlier sections the main challenge in feature based approach is the extraction of features, which are unique and salient. In the case of grayscale images a univariate probability density function is used to calculate the entropy for the extraction of salient regions [5]. In this paper, we propose to use a multivariate entropy based feature detector to extract the unique features based on the pixel intensities. Details of the proposed salient region detection are explained below:

For a pixel 'Q' in the image, consider a square region of scale (side) 'L' in the neighborhood of 'Q' such that 'Q' is one of the vertices of the square 'Sq'. A square window is considered instead of a circular one to attain the objective of covering the entire image by a finite number of such regions with no overlapping windows. The local differential entropy of the region 'Sq' is given by:

$$H_{L,Q} = \int_{Sq} p_{L,Q}(v)\log_2\left(\frac{1}{p_{L,Q}(v)}\right)dv \tag{1}$$

where $p_{L,Q}(v)$ is the probability of finding a pixel of intensity vector 'v' within a square region 'Sq' of side 'L' with the top-left corner vertex situated at 'Q'.The best scale 'L_{opt}' for a pixel 'Q' is selected as the one which will maximize the entropy.

3.1. Multivariate density function

The probability density function (PDF) is calculated using a non-parametric multivariate density estimation method. The non-parametric density estimation methods are used to model the density function of the data without making any assumptions about the underlying distribution. The simplest form of a non-parametric density estimation method is the histogram. But the use of histograms has various drawbacks [9]. The most prominent among them is that the number of bins increases exponentially with the number of dimensions. Consider the case of a joint histogram of 3 dimensions (RGB). In this case there would be a bin for each combination of RGB and generally most of the bins would be empty. The general expression for the non-parametric density estimate is given by:

$$p_{L,Q}(v) = \frac{k}{NV} \tag{2}$$

where 'V' is the volume surrounding the pixel having intensity 'v', 'k' is the number of pixels inside 'V' and 'N' is the total number of pixels. Suppose the region 'Sq' that encloses the 'k' pixels is a hypercube of side 'h' and volume $V=h^D$(D is the number of dimensions) centered at the pixel having intensity 'v', then the number of points inside the hypercube is given by:

$$k = \sum_{n=1}^{N} K\left(\frac{v - v^n}{h}\right) \tag{3}$$

where K(u) is the Kernel function which can be modeled using a Gaussian density function:

$$K(u) = \frac{1}{(2\pi)^{D/2}}\exp\left(-\frac{1}{2}v^T v\right) \tag{4}$$

The multivariate PDF $p_{L,Q}(v)$ is given by:

$$p_{L,Q}(v) = \frac{1}{Nh^D}\sum_{n=1}^{N} K\left(\frac{v - v^n}{h}\right) \tag{5}$$

The parameter 'h' is also called the smoothing parameter or bandwidth. The optimum value of the bandwidth for a true Gaussian density distribution and a Gaussian kernel is given by [9]:

$$h_{opt} = 1.06\sigma N^{-\frac{1}{5}} \tag{6}$$

where σ is the pixel sample variance and N is the number of pixels.

3.2. Region Matching

In this step we establish the correspondence between the two images in the wavelet domain to make use of the geometric transformation invariance property of wavelets [8]. We have selected the Haar wavelet decomposition, since it is simple and the fastest to compute.

A signal $f(t)$ can be constructed from the mother wavelet $\psi(t)$ as illustrated by the following equation:

$$f(t) = \sum_{\phi}\sum_{k} a_{\phi,k} 2^{-k/2}\psi(2^{-k}t - \phi) \tag{7}$$

where k and Φ denotes the value in the scale and shift space respectively and $a_{\phi,k}$ is the corresponding coefficient. Haar wavelet coefficients are computed for both the reference image and test image. The similarity between the wavelet coefficient vectors for the two regions is calculated using the Euclidean distance. Let d_{ij} ($\forall$ i $\in$[1,M], j $\in$[1,N]) be the similarity value between the regions R_{1i} and R_{2j} in Im_1 and Im_2 respectively. For a region R_{11} in Im_1 its corresponding region in Im2 is R_{2k} when d_{1k}=min (d_{1j}). The detailed image segmentation and matching procedure has been explained in Fig.1.The transformation parameters required to align the two images are calculated as described in the Fig.2.

```
Let Im1 & Im2 be the reference & test images respectively.
Let (tx, ty, θ, s) represent the transformation parameters,
where tx & ty denote the translation in the x and y direction
respectively, θ denotes the rotation angle and
's' denotes the scale.

Regionselect()
{
Choose a pixel 'Q(x,y)' in the image.
Check whether 'Q' belongs to some other segment in the image.
If 'Q' is not part of any segment in the image then select 'Q'.
Grow a square region starting from 'Q' of dimension (Lx L)such
that 'Q' is the top-left vertex of the square 'Sq'.
Checkentropy(L,Q)
{
Calculate the multivariate probability density function using a
kernel function as given in the equation (5).
Calculate the entropy ( H_L,Q) of the region as given in
equation (1).
}
if H_L,Q< H_L-1,Q
{
Select region Sq(L-1×L-1).Mark all the pixels present in the region
i.e.mark(x,y)=1 ∀(x, y) ∈ Sq
Move to the next nearest pixel which is not included in any other segment
in the image and repeat the steps from the regionselect process onwards.
}
else
{
Grow the square region(L=L+1) with the same pixel 'Q' as the top-left
vertex and repeat the Checkentropy process.
}}
Repeat the above processes until all the pixels present in the image are
covered i.e. mark(x,y)=1∀(x, y) ∈ Im1.
Repeat the above mentioned steps to segment the test image Im2.
Let 'M' and 'N' be the number of regions present in Im1 and Im2
respectively.
for each region 'R1i' (i ∈ [1,M]) present in 'Im1'
{
   Let 'R2j' (j ∈ [1,N]) be any region present in Im2.
   Correspondence(R1i ,R2j)
   {
   Apply haar wavelet decomposition to both 'R1i' and 'R2j'.
   Let 'NR1i' and 'NR2j' represent the number of wavelet
   coefficients obtained from 'R1i' and 'R2i' respectively.
   Select the most significant coefficients from 'NR1i' and
   'NR2j' such that the number of coefficients is equal to the
    min(R1i ,R2j).
   Calculate the similarity between the two sets of wavelet
   coefficients using the euclidean distance.
   }
   Select the region 'R2j' which exhibits the maximum similarity
   w.r.t 'R1i'.
}
```

Fig 1.Entropy based Image segmentation and matching

```
Findtransformation(R1,R2)
{
Find the ratio of the sizes of the two regions-->[size].
Find the displacement by aligning the corner pixels
of both the regions-->[transx],[transy].
Sample the parameter space for rotation in the
range [-π , π ]->[rotate].
Calculate the modevalue(element with maximum frequency of
occurence) for [size],[[transx],[transy] & [rotate].

 Let scalemode,txmode,tymode and rotatemode represent the
modevalues of the vectors [size],[transx],[transy] & [rotate]
respectively.
For each of the unique elements in [size] find the value which
will contribute the most to the 'global image alignedness'.Assign
this value to 's'.
if rotatemode=0 & scalemode=1
   tx = txmode,   ty = tymode.
   θ= rotatemode,  s =scalemode.
else
Select the most salient region in Im1 and check the similarity
value w.r.t the corresponding region in Im2 for the rotation
sample space[-π , π ].  θ= the rotation angle which gives the
maximum similarity.
}
```

Fig 2:Calculation of Transformation Parameters

4. RESULTS AND DISCUSSION

In order to quantitatively validate the robustness and accuracy of the proposed algorithm we conducted some experiments using a pair of color images. The proposed algorithm was implemented mainly in Matlab.Standard color images (Caltech motorbike [available from http://www.robots.ox.ac.uk/~vgg/data/] and Lena) are used for the experiments as reference image. The test image is obtained by applying a known transform to the reference image. The five transformations which are applied are as listed below:

1) ***Addition of noise to the image*****:** Reference image is transformed by adding varying amounts of random noise. The proposed algorithm is then tested on these images to identify the specific parts of the image, which are corrupted by noise. Fig 3a illustrates the results, which indicate that the proposed approach effectively identifies the increasing difference between reference image and the transformed images. Fig.4 shows the reference image along with some of the segments.

2) ***Translation change between the reference image and test image****:* The translation parameter is varied in the range {-20:20 in steps of 1}. The negative shift values indicate right shift. The average normalized translation percentage error is around $\pm$ 0.5 as shown in fig.3b.

3) ***Rotation change between the reference image and test image:*** The rotation parameter is varied in the range {-180^0:180^0 in steps of 1^0}. The negative rotation values indicate clockwise rotation. The average normalized rotation percentage error is around 0 as shown in fig.3c.

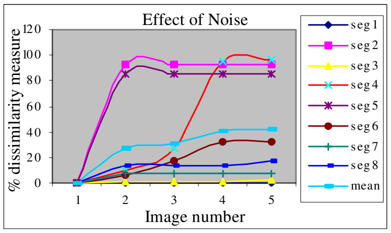

Fig.3a:Effect of noise on image segments

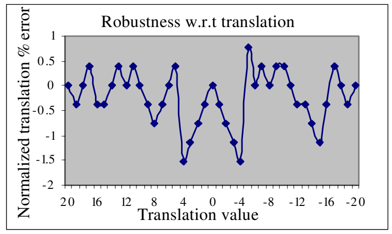

Fig.3b:Robustness of the algorithm w.r.t translation

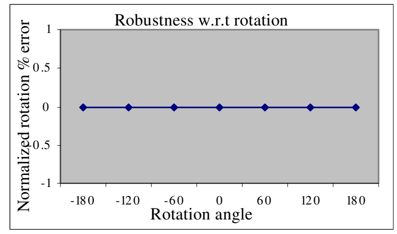

Fig.3c:Robustness of the algorithm w.r.t rotation

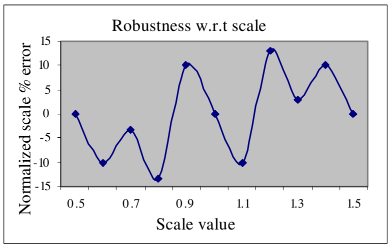

Fig.3d:Robustness of the algorithm w.r.t scale changes

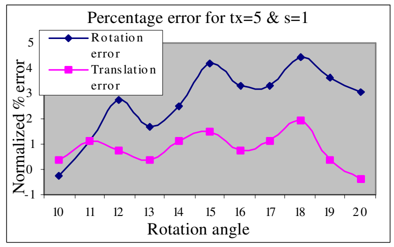

Fig.3e:Robustness of the algorithm w.r.t change in a combination of the transformation parameters(for t_x=5 & s=1).

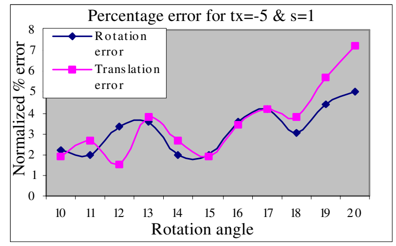

Fig.3f:Robustness of the algorithm w.r.t change in a combination of the transformation parameters(for t_x=-5 & s=1).

4) ***Scale change between the reference image and test image***: The scale parameter is varied in the range {0.5: 1.5 in steps of 0.1}. The average normalized scale percentage error is around ± 9 as shown in Fig.3d. Normalized error percentage is given by:

$$S_{err} = \left(\frac{S_{obt} - S_{app}}{var\_range} \right) * 100 \quad (8)$$

where S_{obt} , S_{app} and var_range are the obtained value, applied value and the variation range of the particular transformation respectively.

5) ***A combination of change in all the transformation parameters:*** The rotation parameter is varied in the range {10:20} for the translation parameter {-5,5} & scale parameter {1}. The average normalized percentage rotation and translation error is around +2 and +3 respectively as shown in fig.3e and fig.3f.

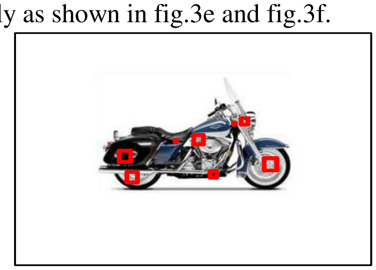

Fig.4: Reference Caltech motorbike image

5. CONCLUSIONS

We have presented a new hybrid image registration algorithm to identify the spatial or intensity variations between two color images. The proposed approach extracts salient descriptors from the images and identifies the transformation parameters after establishing the correspondence between the salient descriptors of the two images, which yields better accuracy and lesser computational cost compared to the existing approach.

6. REFERENCES

[1] Lisa Gottesfeld Brown,"A Survey of Image Registration techniques", *Appeared in ACM Computing Surveys,* vol.24, Issue.4, pp.325-376, December 1992.

[2] S.H.Srinivasan, Mohan Kankanhalli,"Wide Baseline Spectral Matching", *Appeared in Proc.International Conference on Multimedia & Expo*, July 2003.

[3] S.H.Srinivasan,"Spectral Matching of Bipartite Graphs", *Appeared* in *Proc.3rd International Conference on Hybrid Intelligence Systems*, December 2003.

[4] Frederik Maes, Andre Collignon, Dirk Vandermeulen, Guy Marchal, Paul Suetens,"Multimodality Image Registration by Maximization of Mutual Information", *Appeared in IEEE Transactions on Medical Imaging*, April 1997.

[5] Xiaolei Huang, Yiyong Sun, Dimitris Metaxas, Frank Sauer, Chenyang Xu,"Hybrid Image Registration based on Configural Matching of Scale-Invariant Salient Region Features", *Appeared in Proc.IEEE Workshop on Image and Video Registration*, July 2004.

[6] Timor Kadir, Michael Brady," Saliency, Scale and Image Description", *Appeared in International Journal of Computer Vision*, 2001.

[7] Timor Kadir, Andrew Zisserman, and Michael Brady," An Affine Invariant Salient Region Detector", *Appeared in Proc. of the 8th European Conference on Computer Vision*, 2004.

[8] Charles E. Jacobs, Adam Finkelstein, David H. Salesin,"Fast Multiresolution Image Querying", *Appeared in Proc.SIGGRAPH*, August 1995.

[9] L.Belanche, A.Nebot,"Density estimation", *A Lecture on Intelligent Data Analysis and Data Mining*, UPC, 2001.