MODEL BASED GLOBAL IMAGE REGISTRATION

Niloofar Gheissari, Mostafa Kamali, Parisa Mirshams and Zohreh Sharafi
Computer Vision Group, School of Mathematics
Institute for studies in Theoritical Physics and Mathematics, Tehran, Iran

Keywords: Image Registration, Model Selection, Panorama.

Abstract: In this paper, we propose a model-based image registration method capable of detecting the true transformation model between two images. We incorporate a statistical model selection criterion to choose the true underlying transformation model. Therefore, the proposed algorithm is robust to degeneracy as any degeneracy is detected by the model selection component. In addition, the algorithm is robust to noise and outliers since any corresponding pair that does not undergo the chosen model is rejected by a robust fitting method adapted from the literature. Another important contribution of this paper is evaluating a number of different model selection criteria for image registration task. We evaluated all different criteria based on different levels of noise. We conclude that CAIC and GBIC slightly outperform other criteria for this application. The next choices are GIC, SSD and MDL. Finally, we create panorama images using our registration algorithm. The panorama images show the success of this algorithm.

1 INTRODUCTION

Image registration refers to the process by which two or several images (taken from different view points) are transformed and integrated into a single coordinate system. This means image registration involves estimating transform parameters such as rotation, scaling and translation. Generally, there are two main approaches to image registration: local registration and global registration, each of which has its own advantages and disadvantages. Local methods find small patches (blocks) or interest points, match them and register two images based on the parameters estimated from the corresponding points or blocks. Global methods minimize a global energy term, which describes the error generated by aligning two images. This error might be the sum of squared differences of intensity values or a more complicated measure.

Local methods are able to deal with local deformations and distortions while they might generate undesirable results along patch boundaries. In contrast, global methods are robust. However, global methods are unable to deal with local motions. For a more elaborate survey on image registration methods we refer the reader to (Zitova & Flusser, 2003).

In this paper, we introduce a new image registration method that automatically selects the true underlying transform model from a library of candidate models. This model library consists of 2D transformation models that might describe the motion model between two images. We use 2D models (rather than 3D ones) because our application is global registration for making panoramic images. Using the correct model that can describe the true camera motion model is adventurous. For example, the algorithm is more robust to noise and outlier since any corresponding pair that does not undergo the chosen model is rejected. In addition, this method best suits virtual reality applications where a virtual object is to be placed in a real background. Since the true transformation model and the model parameters are computed, they can be applied to the virtual object so that its motion is consistent with the rest of the image (the real part) and generate more realistic images. As mentioned before, our algorithm chooses the true transformation model from a model library which includes pure translation, Euclidean, similarity, affine, and projective models as shown in Table 1. The model library includes all possible models from the most complex model (projective) to the most degenerated model (pure translation). The nested property of this library allows us to use an

Gheissari N., Kamali M., Mirshams P. and Sharafi Z. (2008).
MODEL BASED GLOBAL IMAGE REGISTRATION.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 440-445
DOI: 10.5220/0001074804400445

information theoretic criterion to select the true model.

Table 1: The model library used for image registration (Szeliski, 2006).

Model	*Transformation*	*DoF*
Translation	$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \end{bmatrix}$	2
Euclidean	$\begin{bmatrix} cos\theta & -sin\theta & t_x \\ sin\theta & cos\theta & t_y \end{bmatrix}$	3
Similarity	$\begin{bmatrix} 1+a & -b & t_x \\ b & 1+a & t_y \end{bmatrix}$	4
Affine	$\begin{bmatrix} 1+a_{00} & 1+a_{01} & t_x \\ 1+a_{10} & 1+a_{11} & t_y \end{bmatrix}$	6
Projective	$\begin{bmatrix} 1+h_{00} & h_{01} & h_{02} \\ h_{10} & 1+h_{11} & h_{12} \\ h_{20} & h_{21} & 1 \end{bmatrix}$	8

This paper is organized as follows. We first discuss some of statistical model selection criteria for computer vision applications in section 1.1. Next, we describe our model-based image registration method in section 2. An important component of the proposed registration method is model selection. We evaluate a number of different model selection criteria for image registration application in section 3. We show that CAIC and GBIC outperform the other statistical criteria. Section 4 is dedicated to making panoramic images, and in section 5 we present our conclusion.

1.1 Model Selection Criteria and their Use in Image Registration

Model selection criteria allow choosing the true model by establishing a trade-off between “fidelity” and the “complexity” of that model. Because the most complex (highest order) model always fit the data better than any other model, it has more degrees of freedom.

In this paper, we propose to use a model selection criterion to detect the true transformation model for registering a pair of images. If we use a more general model than the true model (over-fit), we allow noise and outliers to affect parameters estimation more severely. This is because having more degrees of freedom gives enough flexibility to the model to bend and twist itself and consequently fits to noise and outliers. In contrast, having a less general model (than the correct model), will result in under fitting. Under fitting has the danger of rejecting inliers as being outlier and so disregarding important information.

These model selection criteria score a model based on two terms. That is the accuracy of the fit (fidelity) that is usually the logarithmic likelihood of the estimated parameters of the model. This likelihood is equal to the scaled sum of squared residuals, providing noise is Gaussian. The term scoring the complexity is a penalty term for higher order models so that the criterion always avoids choosing the most general model.

Akaike perhaps was the first to introduce a model selection criterion known as AIC (Akaike, 1974). The main idea behind AIC is the fact that the correct model can sufficiently fit any future data with the same distribution as the current data. AIC has been modified in many ways. For example, many model selection criteria including CAIC (Bozdogan, Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions, 1987), GAIC (Kanatani, Model selection for geometric inference, 2002), and GIC (Torr, 1999) are derived from AIC.

Later, in 1978, Rissanen introduced MDL (Rissanen, Modeling by shortest data description, 1978). The underlying logic of MDL is that the simplest model that sufficiently describes the data is the best model. Kanatani derived GMDL (Kanatani, Model selection for geometric inference, 2002), which has a very similar logic to MDL, specifically for geometric fitting.

Another group of model selection criteria is based on Bayesian rules such as GBIC (Chickering & Heckerman, 1997). They choose the model that maximizes the conditional probability of describing a data set by a model.

Cost functions of the aforementioned criteria and two other model selection criteria Mallow CP (Mallows, 1973) and SSD (Rissanen, Universal coding, information, prediction, and estimation, 1984) are shown in Table 2. A more complete survey on different available model selection criteria can be found in (Gheissari & Bab-Hadiashar, Model Selection Criteria in Computer Vision: Are They Different?, 2003).

There have been a few papers; such as (Bhat, et al., 2006), in the image registration literature concerned about choosing the true transformation model between two images. However, they use a heuristic approach to decide whether the transformation model is a simple homography or the fundamental matrix.

Table 2: Different model selection criteria cost functions. N is the number of correspondences, p is the degree of freedom of a model, r is the residual, δ is the scale of noise of the highest order model in the library, and d is the dimension of our data. We set λ_1 and λ_2 to be 2 and 4 respectively.

Model Selection	***The scoring function***
MDL	$\sum_{i=1}^{N} r_i^2 + (P/2)\log(N)\delta^2$
GMDL	$\sum_{i=1}^{N} r_i^2 - (Nd+P)\delta^2 \log(\delta/L)^2$
GBIC	$\sum_{i=1}^{N} r_i^2 + (Nd \log(4) + P\log(4N))\delta^2$
Mallow CP	$\sum_{i=1}^{N} r_i^2 + (-N+2P)\delta^2$
GAIC	$\sum_{i=1}^{N} r_i^2 + 2(dN+P)\delta^2$
SSD	$\sum_{i=1}^{N} r_i^2 + (P\log(N+2)/24 + 2\log(p+1))\delta^2$
AIC	$\sum_{i=1}^{N} r_i^2 + 2P\delta^2$
GIC	$\sum_{i=1}^{N} r_i^2 + \lambda_1 dN\delta^2 + \lambda_2 P\delta^2$
CAIC	$\sum_{i=1}^{N} r_i^2 + P(\log N + 1)\delta^2$

To the best of our knowledge, the only image registration method that incorporates a statistical model selection criterion to choose the true model is the method proposed by Yang et al. (Yang et al. 2006). They start from the lowest order model and iteratively apply a modified version of AIC to choose between a model and its immediate higher order model. The process terminates if a model is chosen over its immediate higher order model or the most general model is selected.

Our model selection step differs from Yang et al. in different ways. This is firstly because they only apply a model selection criterion to select between only two models and not a library of models. In addition, they only consider similarity, affine and projective transformations while we also consider pure translation and Euclidean transformations for detecting camera motions. More importantly, the none-iterative nature of our method makes it faster than the method of Yang et al. Finally, since we apply the outlier rejection phase before and separate from the model selection task, we do not violate the general assumption (made by statistical criteria) that the data is outlier free.

2 THE PROPOSED IMAGE REGISTRATION METHOD

In the following subsections, we describe the details of the proposed image registration algorithm. Different elements of our algorithm include feature detection and matching, robust fitting, and model selection. An outline of the proposed algorithm is as follows:

1. Finding a set of correspondences between two images.
2. Finding an initial outlier-free (inliers) set of correspondences. This is achieved by robustly fitting a translation model to the correspondences.
3. Fitting all models in the model library to the above set of inliers.
4. Applying a model selection criterion to choose the true underlying model of the above set.
5. Finding the final set of inliers by applying the chosen model to the dataset and re-estimating the scale of noise.
6. Re-estimating the parameters of the chosen model.

We fit a translation model to the dataset (Step 2) only to find an initial set of outliers. In Step 3, a model selection criterion is applied to this initial set of outliers and the true model is chosen. We need to find this initial set of inliers since almost all model selection criteria available in the literature assume the data set to be outlier-free. In Step 2, we choose to use the lowest order model in the model library to ensure this initial dataset only include inliers.

Our algorithm is implemented by MATLAB and registers two images of size 350×530 in 5.5 second on an Intel 1.86 GHz platform. Most of this time is spent for feature detection and matching (about 5 seconds). We used Scale-Invariant Feature Transform (SIFT) (Lowe, 2004) for this purpose. To match feature points of two corresponding images, as suggested by Low (Lowe, 2004), we used the nearest neighbor ratio matching strategy. If we use the Speeded Up Robust Features (SURF) (Bay, Tuytelaars, & Van Gool, 2006) and implement the algorithm in C++, we expect to achieve real-time performance.

2.1 Parameter Estimation

Consider we have a number of corresponding points between two images. Let $\mathbf{x}$ be a point in the first image and $\mathbf{x}'$ its corresponding point in the second image, i.e.

$$x=(x, y) \leftrightarrow (x', y') = x$$

We are investigating a transformation T under which

$$X' = Tx \quad (1)$$

We estimate the initial parameters for each transformation model using a least square method. Projective transformation is estimated using Normalized Direct Linear Transformation (Hartley & Zisserman, 2004). For Euclidean transformation the procedure is different. For this transformation, we have the following system of equations:

$$\begin{bmatrix} -y & x & 1 & 0 \\ x & y & 0 & 1 \end{bmatrix} \begin{bmatrix} \sin\theta \\ \cos\theta \\ t_x \\ t_y \end{bmatrix} = \begin{bmatrix} x' \\ y' \end{bmatrix} \quad (2)$$

Also the following equation should be satisfied

$$(\sin\theta)^2 + (\cos\theta)^2 = 1 \quad (3)$$

We solve this non-linear system of equations by using the solution of the unconstrained equation system as an initial point.

2.2 Robust Fitting and Outlier Rejection

Finding correspondences between two images involves errors introduced by noisy features, erroneous feature descriptors, and inaccurate distance measurements. Hence, due to noise and outliers instead of Equation 1 we have

$$r = Tx - x' \quad (4)$$

where r is the residual (error). To reject false correspondences between two images as outliers, we apply the following procedure as suggested by (Bab-Hadiashar & Suter, 1999). We sort the residuals and compute the scale of noise for $m = p + 1, \ldots, n$ according to

$$\partial_m^2 = \frac{\sum_{i=1}^{m} r_i^2}{m - p_h} \quad (5)$$

Then, we iteratively increase m until $r^2_{m+1} > T^2\partial^2_m$ or $m=n$. Here, T is a constant factor obtained from Gaussian distribution table. We set $T=2.5$ in our experiments. The smallest outlier is where

$$r^2_{m+1} > T^2\partial^2_m \quad (6)$$

Use of this formula for the scale of noise can be justified by the fact that, if the model is correct and the error, r, is subjected to a normal distribution of zero mean, then the statement

$$\frac{\sum_{i=1}^{m} r_i^2}{\partial^2} \quad (7)$$

is subjected to a χ^2 distribution with m-p degrees of freedom (Kanatani, Model selection criteria for geometric inference, 2000). All points with residuals more than r_{m+1} are rejected as being outliers. We iteratively carry out the above process until no more outlier is removed. Having omitted all outliers, we compute the final transformation robustly.

2.3 Model Selection

After removing outliers, we fit each model in the model library to the remaining data. Then we compute the residuals for each model and compute the scores for a model selection criterion (according to Table 1). Our experiments (discussed in Section 3) suggest that CAIC and GBIC are the preferred model selection criteria among those we evaluated.
We use the scale of noise of the higher order model for model selection task. As explained by Kanatani (Kanatani, Model selection criteria for geometric inference, 2000), the scale of noise of the correct model and the scale of noise of the higher order models (higher than the correct model) must be close enough for the fit to be meaningful. Therefore, it is the most accurate estimation of the true scale of noise available at this stage. We compute the scale of noise according to

$$\partial^2 = \frac{\sum_{i=1}^{N} r_i^2}{N - p_h} \quad (8)$$

where N is the number of inliers and p_h is the number of parameters of the highest model in the library (here set to be 8).

3 EVALUATING MODEL SELECTION CRITERIA FOR IMAGE REGISTRATION

We evaluated nine different model selection criteria for image registration. In our experiments a model selection criterion was expected to be able to identify (from the model library shown in Table 2), the true underlying transformation model between a set of corresponding points. To achieve this, we gathered 40 challenging images from five different environments:

1. External views of different buildings.
2. Internal views of different building.
3. Natural scenes.
4. Views of roads and streets.
5. Views of stadiums and football field.

After applying different synthetic transformations to the above database, we fit each model in the model library to the remaining data. Then we compute the residuals for each model and compute the scores for all different model selection criteria listed in Table 1. For each criterion, we choose the model that minimizes the score of it.

Our objective was also to examine the effect of noise level on the success rate of each criterion. After finding correspondences, we add Normal noise of different variances (with zero mean) to the pixel coordinates of corresponding feature points.

This is because our objective was to examine the performance of different criteria independently from the particular matching algorithm used. However, the normal noise added to pixel coordinates simulates normally distributed errors generated by the feature matching process.

The performances of each criterion for different noise levels are shown in Figure 3. As can be seen from this figure, all criteria perform about 90% until the noise level reaches to 3% of noise. This means for example, if a pixel coordinate is 100, then up to 3 pixels matching error is well tolerated. As a result, we used CAIC in our experiments. The success rate (correct prediction) of each criterion in accurately recovering the underlying transformation model between all corresponding images is shown. As can be seen from this figure, GBIC and CAIC perform very similarly and have better performance than other criteria. The next choices are GIC, SSD and MDL.

4 MAKING PANORAMA

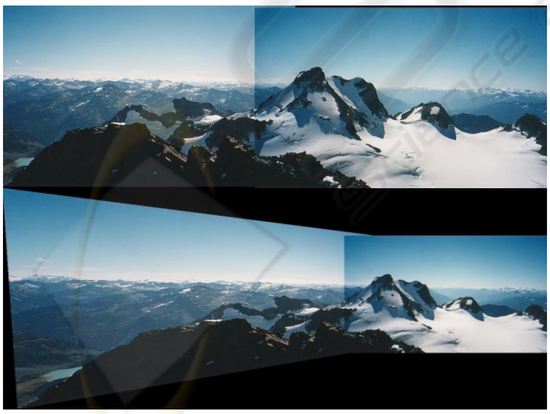

Figure 1: (top) The panorama image using the true model chosen by CAIC (translation model)- (bottom) the panorama image using a wrong model (projective). The difference between these two panorama images shows the importance of model selection.

On account of registering images for building wide view panorama images, we used the transformation computed by the robust model based method. In figure 1, we show the importance of model selection by applying the right model (translation here) and a wrong model (projective) to two images taken at different viewpoints. Fig 2 is a sample of panorama image with five images.

Figure 2: Another sample of the panorama images we created.

5 CONCLUSIONS

We proposed a model-based image registration method capable of detecting the true transformation model. We used a robust method to detect the scale of noise and reject false correspondences. The proposed algorithm is robust to degeneracy as any degeneracy is detected by the model selection component. Another contribution of this paper is the evaluation of nine different model selection criteria for image registration based on different levels of noise. We conclude that CAIC, GBIC slightly outperform other criteria. The next choices are GIC, SSD and MDL. Finally, we made panorama images, which show the success of this algorithm.

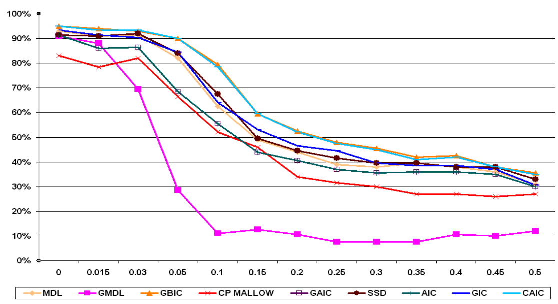

Figure 3: Percentage of success of different model selection criteria versus different noise levels.

REFERENCES

Akaike, H. 1974. "A New Look at the Statistical Model Identification." IEEE Transactions on Automatic Control AC-19(6):716-723.

Bab-Hadiashar, A. and D. Suter. 1999. "Robust Segmentation of Visual Data Using Ranked Unbiased Scale Estimate." ROBOTICA, International Journal of Information, Education and Research in Robotics and Artificial Intelligence 17:649-660.

Bay, Herbert, Tinne Tuytelaars and Van Luc Gool. 2006. "SURF: Speeded Up Robust Features." In Proceedings of the 9th European Conference on Computer Vision (ECCV06).

Bozdogan, H. 1987. "Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extentions." Psychometrica 52:345-370.

Chickering, D. and D. Heckerman. 1997. "Efficient Approximation for the Marginal Likelihood of Bayesian Networks with Hidden Variables." Machine Learning 29(2-3):181-212.

Gheissari, Niloofar, Bab-Hadiashar,Alireza. Dec. 2003. "Model Selection Criteria in Computer Vision: Are They Different?" In Proceedings of Digital Image Computing Techniques and Applications(DICTA 2003). Sydney,Australia.

Kanatani, K. 2000. "Model Selection Criteria for Geometric Inference." In Data Segmentation and Model Selection for Computer Vision: A statistical Approach, ed. A. and Suter Bab-Hadiashar, D.: Springer-verlag.

Kanatani, K. Jan. 2002. "Model Selection for Geometric Inference." In The 5th Asian Conference on Computer Vision. Melbourne, Australia.

Lowe, David. 2004. "Distinctive image features from scale-invariant keypoints, cascade filtering approach." International Journal of Computer Vision 60:91 - 110.

Mallows, C. L. 1973 "Some Comments on CP." Technometrics 15(4):661-675

Rissanen, J. 1978. "Modeling by Shortest Data Description Automata." 14:465 - 471.

Rissanen, J. 1984. "Universal Coding, Information, Prediction and Estimation." IEEE Transactions on Information Theory 30(4):629-636.

Szeliski, Richard. September 2004. "Direct (pixel-based) alignment." In Image Alignment and stitching. Priliminary draft. Edition. Microsoft Research.

Torr, P.H.S. 1998. "Model Selection for Two View Geometry: A Review." In Model Selection for Two View Geometry: A Review. Microsoft Research, USA: Microsoft Research, USA.

Yang, Gehua, Charles V. Stewart, Michal Sofka and Chia-Ling Tsai. 2006. "Automatic Robust Image Registration System: Initialization, Estimation and Decision." In Proceedings of the Forth IEEE International Conference on Computer Vision Systems (ICVS 2006).

Zitova, B. Flusser J., 2003. "Image Registration methods: A Survey, Image and Vision Computing." 21:97

Hartley, R. Zisserman, A. 2004. “Multiple View Geometry in Computer Vsion “(2nd Edition ed.). Cambridge University Press.