

NONLINEAR SIMILARITY BASED IMAGE MATCHING

Muhammad Sirajul Islam
and Les Kitchen

*Department of Computer Science and Software Engineering ,
The University of Melbourne, VIC 3010, Australia.
Email: {msislam, ljk+iip}@csse.unimelb.edu.au*

Abstract Image matching is an inarguably important operation for many practical sophisticated systems in machine vision and medical diagnosis. Many gray-level image matching applications use the sum-of-squared-difference (*SSD*) or sum-of-absolute-differences (*SAD*), which are very sensitive to noise. Almost all images have some kind of noise, which causes the matching tasks significantly difficulty. In this paper we explore a new, less noise sensitive image-matching technique. It uses non linear similarity measure *min* or *median* on interest points to find a match. The algorithm has been tested using a range of images with different gaussian noise. The result shows a significant improvement over traditional Euclidean distance measure technique for image matching.

Keywords: Computer vision, Image processing, Interest points, Non maximum suppression, Feature points.

1. Introduction

Image matching is a common operation in many applications which include object tracking, motion estimation of objects in two successive frames, medical diagnosis[1], etc. The greatest challenge in matching two images lies in coping with the effects of noise. Noise may be caused by a wide variety of effects, e.g., detector sensitivity variations, transmission or quantization error, environmental variations, etc. Presence of noise in an image is very common due to the nature of image capturing devices.

There are various techniques for image matching. They can be categorized broadly into two classes: low-level (intensity level) image matching, and high-level matching techniques. Depending on how the features for matching are produced, the matching process can be divided into two types:

- Algorithms that deal with pixels directly for matching (low-level), for example cross-correlation techniques. In cross-correlation approaches two points are matched using some distance measure among the neigh-

bouring pixels of the two considered pixels. In an ideal situation this distance should be zero for a perfect match. This is not the case in the real world: as we mentioned earlier, the presence of noise is very common in digital images.

- The second category of matching uses description by extracted features, for example identifying edges and their relationships. Algorithms of this type highly depend on effective grouping of features and relationships among them.

The second type of matching relies on feature extraction and finding relationships among the extracted features. Finding relationships can be computationally expensive. One of the main disadvantages of cross-correlation or traditional Euclidean distance measures for image matching is that these techniques are significantly sensitive to noise and perturbations. Both of these two categories of matching algorithms may adopt some kind of probabilistic approach—like maximum likelihood image matching [2].

In this paper we describe a new technique for image matching which uses a nonlinear similarity based technique. Although our approach falls under the heading of low-level image matching technique, we try to improve the matching technique making it less sensitive to noise. Fuzzy logic has been an area of research in engineering since 1965. There are very few works, e.g., [3], using fuzzy information for image matching. Our algorithm has the flavour of fuzzy logic but does not strictly follow the steps of a fuzzy system, so we do not call it a complete fuzzy system based technique. A brief description of the steps is given in the next paragraph.

For matching two images it is very costly, in terms of time and memory, to match every point. Therefore, most of the matching techniques use a few hundred or so pixels, called *interest points (IP)*. We first extract IPs from two images and take the few hundred best points for pairing as matches. We use a 7×7 Moravec interest operator [4] for extracting IPs. The matching process treats the first image as a template which would be matched to the second. For similarity measures of an interest point the dot products of the normalized vectors, obtained by slicing a 3×3 neighborhood in the directions of NS, EW, NE, NW, are calculated. These four features are used to find the degree of similarity between interest points in two images. See Figure 1.

Our matching technique is based on the basic idea that the overall similarity of two points to be matched depends on the similarity of individual corresponding vectors. The experimental results show that our approach can detect 5 to 10% more correct matches than the traditional Euclidean distance transform measure.

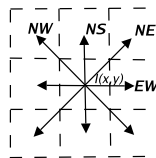


Figure 1. Vectors considered for feature extraction.

In Section 2 some related work is discussed. Section 3 describes our proposed technique; experimental results are shown in Section 4. Some final conclusions are drawn in Section 5.

2. Related Work

Various approaches have been adopted for image matching (or finding a patch inside an image). One of the techniques uses the Hausdorff Distance [5] for matching. In this approach, edge extraction is usually done with one of the many edge detectors known in the image processing literature, like the Canny edge detector [6], Laplacian, Sobel, etc. After applying some algorithm that minimizes the Hausdorff Distance between two images, the best match is taken. This approach considers the shape of the objects in an image but does not consider the intensity value and it is feasible to find matches only for objects that exhibit sharp edges.

Another low-level feature-based image matching technique is RIMA [7] which is an extension to distance transform (*chamfer*) matching. In this technique, edge points are extracted from digital images, converted to binary images, which are distance transformed, and then the distance transform is used for image matching. The matching is estimated by superimposing the distance transform of the template on the distance transform of the source image. RIMA needs to keep a distance-transform image as well as the binary edge map. An advantage of our technique is that it does not use a binary image like RIMA, hence it requires less memory.

Matching algorithms based on fuzzy information are an interesting area of research and some work has been done on this. One piece of work based on fuzzy features is [8]. We claim our approach is somewhat fuzzy, but the significant difference between [8] and our work is that our algorithm does not strictly follow the steps of fuzzy systems. In [8], the matching task is done by coarse-to-fine matching; Fuzzy information is used for fine matching using steps of fuzzy systems like defining membership functions, creating fuzzy rules, etc. We do not use any membership function or fuzzy rules, although it can be a future work as discussed in Section 5.



Figure 2. Different stereo views of corridor image.

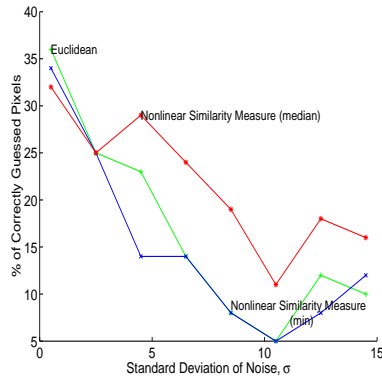


Figure 3. Comparison for corridor image.

3. Nonlinear Similarity based Matching

The general idea behind this approach is that two unit vectors are very similar if their dot product is very near to unity. This resembles to the fuzzy IF-THEN rule based formulation. In this section we describe our method of interest point extraction, how fuzzy features are extracted and what fuzzy operator is used and the reason behind that.

Interest Point Extraction

For extracting interest points we use a 7×7 Moravec operator [4], slightly modified to reduce directional bias, followed by 3×3 non-maximum suppression, with provision for resolving ties. We then choose the strongest few hundred points. However, to reduce the number of points that need to be solved, we first apply a conservative threshold to filter out the weaker interest points. For more details see [9].

Similarity Measure

In order to find a match we have to have some kind of similarity measure. To formulate such a measurement we consider the 3×3 neighborhood of the

point of interest and extract four vectors as shown in Fig 1. Let a, b be two unit vectors and θ be the angle between them. So the dot product of these two vectors is $a \cdot b = \cos \theta$. As $\cos \theta$ approaches 1 the degree of similarity between these two vectors increases to the maximum. For a mathematical formulation of our algorithm let us consider two interest points i and j from the first and second images respectively. Let $v_{rd}^i, v_{ld}^i, v_{hz}^i, v_{vt}^i$ be the normalized vectors along the NE (right diagonal), NW (left diagonal), EW (horizontal) and NS (vertical) directions, respectively, with respect to the center pixel i . Here, $v_{rd}^i \equiv [I(x+1, y-1), I(x, y), I(x-1, y+1)]$ and the definitions for the others are analogous. Now, four similarity measures $S_{rd}^{ij}, S_{ld}^{ij}, S_{hz}^{ij},$ and S_{vt}^{ij} are calculated taking the dot product of the pair of corresponding vectors of the two considered pixels i, j . For example, $S_{rd}^{ij} = v_{rd}^i \cdot v_{rd}^j$.

We are treating this similarity measure (zero to one) as a fuzzy grade of truth in the proposition that the corresponding vectors match. Using these values, we can take some decision about the degree of similarity of two considered pixels from the two noisy images. Because of noise, pixel values may change and hence the orientation in feature space of the four vectors will also change.

The Combination Operator

For estimating the similarity between two points from two images, we have to have some fuzzy operator. We can have fuzzy rules like:

$$(S_{rd}^{ij} \in \text{High}) \wedge (S_{ld}^{ij} \in \text{High}) \wedge (S_{hz}^{ij} \in \text{High}) \wedge (S_{vt}^{ij} \in \text{High}) \Rightarrow (\text{Pixels } i \text{ and } j \text{ are Highly Similar}).$$

We know that traditionally the *minimum* operator is used for evaluating this AND connective (and *maximum* is used for OR). But the *minimum* and *maximum* operators are not good estimators. Zimmermann and Zysno (1980) revealed through experiments that the *minimum* operator does not work well as a model of the *and* connective, producing too conservative (low) results. Therefore they proposed some compensator operator. Let us consider a scenario from our experiment. Assume that the similarity measure along the right diagonal, S_{rd}^{ij} , is small because of some noise in that direction. On the other hand, the remaining three of the similarity measures are very high, because the two pixels are in fact really similar. Then it would be unjust to the three high values if we chose the minimum by using the *minimum* operator. Hence, we choose the *median* operator (we take the average of the middle two values), which as its very meaning explains that it is not biased towards any particular directional intensity change (caused by some random noise). Experimental results also conform with our view of choosing the *median* operator, as will be evident in Section 4.



(a) Cameraman image.



(b) Pepper image.

Figure 4. Images used for experimental purpose.

4. Experiments

Our experimental methodology is as follows: We have two images to be matched. We extract the best one hundred interest points from each image. For each interest point in the first image, we find the best match among the interest points in the second image. The best match is determined in three ways: The first way by finding the point with the maximum similarity according to our nonlinear similarity based approach, as laid out in Section 3, using the median to combine the similarities from the four neighborhood slices. The second way is similar, but uses the minimum of the four neighborhood-slice similarities instead of the median. And the third is by finding best match as the point with minimum Euclidean distance, treating the 3×3 neighborhood around each point as vector in a 9-dimensional Euclidean space.

For each way we compute the percentage of correct matches. A match is deemed to be *correct* if it lies within the neighborhood used for computing the interest points, that is, in our case, within three pixel positions of the exact-match position, for our 7×7 Moravec operator. Obviously, computing the percentage of correct matches requires knowing which matches are correct, that is, knowing the “ground truth” for the matching.

Figure 2 shows a pair of synthetically generated images (a stereo pair), taken from [10], for which the true matches are known. Figure 3 shows the percentage of correct matches for each of the three measures, for different levels of added Gaussian noise.

It is difficult and time-consuming to determine the ground truth for many pairs of images. Therefore, as an expedient for obtaining more data for comparisons, we adopt a tactic of matching an image to itself. In this case the ground truth is known trivially: a point should match to itself. Of course, for two identical images, the matching task is far too easy to be a fair basis for evaluation of matching techniques. However, we can make the matching task sufficiently difficult by adding two *different* sequences of noise respectively to two copies of the same image. This approximates reasonably well two frames from a motion sequence for which the motion just happens to be zero. Since the matching algorithm searches for the best match according to the measure that it is using, and does not “know” the ground truth, this does represent a reasonable task for evaluating matching techniques.

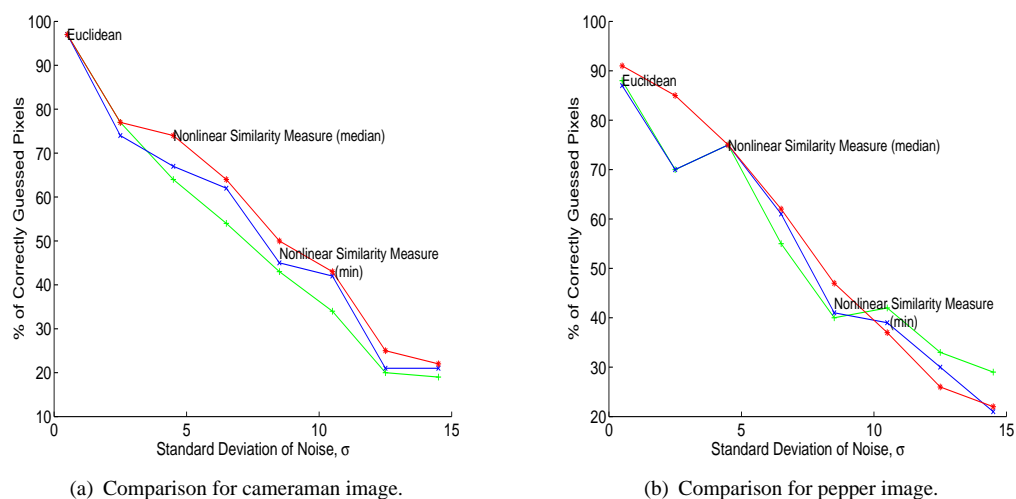


Figure 5. Comparison results.

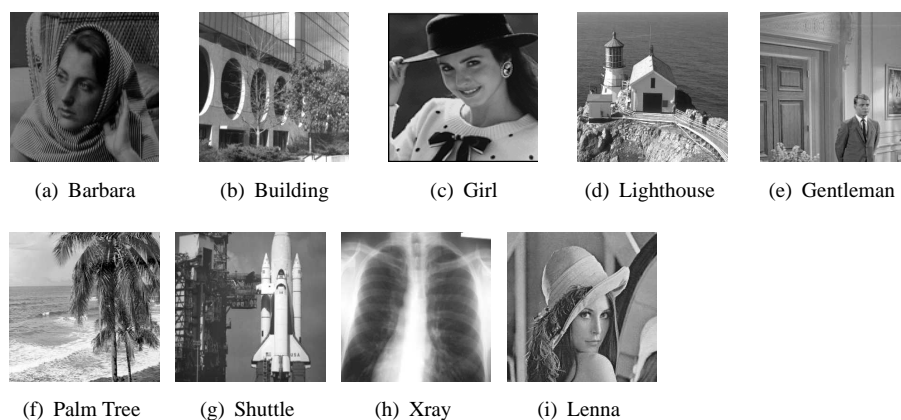
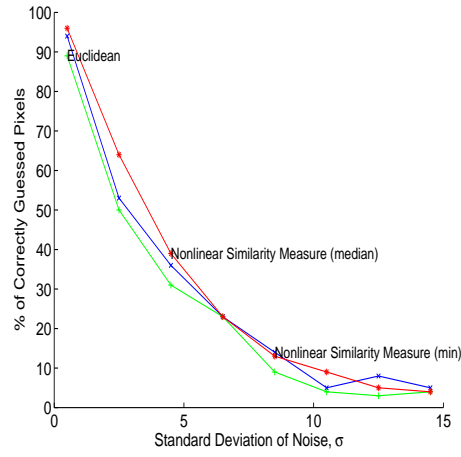


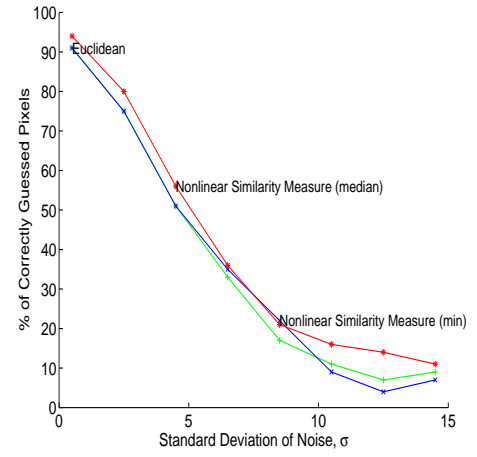
Figure 6. Images used in experimentation.

Figures 5(a) to 5(b) (among a number of experimental results) show the matching performance for the different measures for different levels of noise on the task of self matching for the images shown in Figure 4. It is re-assuring that the results are reasonably similar to those obtained for the stereo pair of images in Figure 2, giving support to the idea that self-matching under noise does provide a reasonable way of evaluating matching techniques.

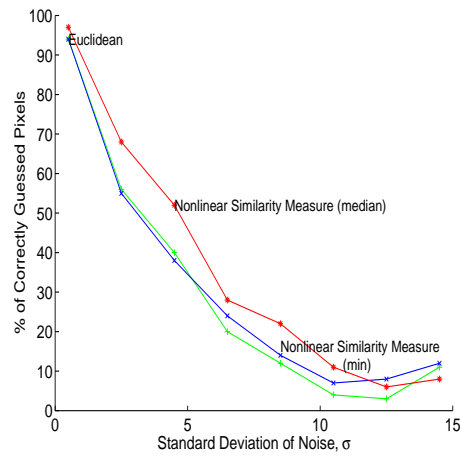
Some more experimental results for the images in Figure 6 are shown in Figures 7 and 8. While there is some variation across the images, and all methods



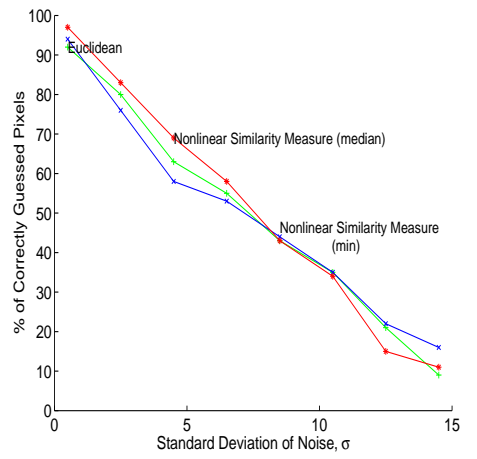
(a) Comparison for Barbara Image.



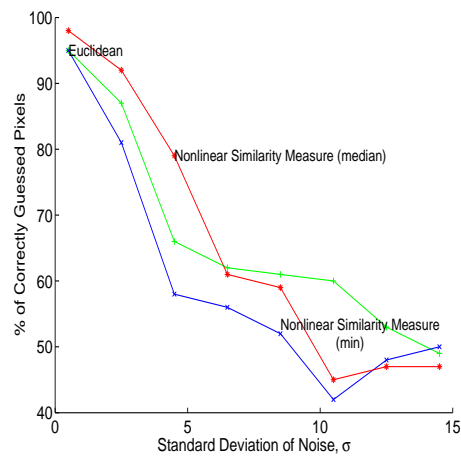
(b) Comparison for Building Image.



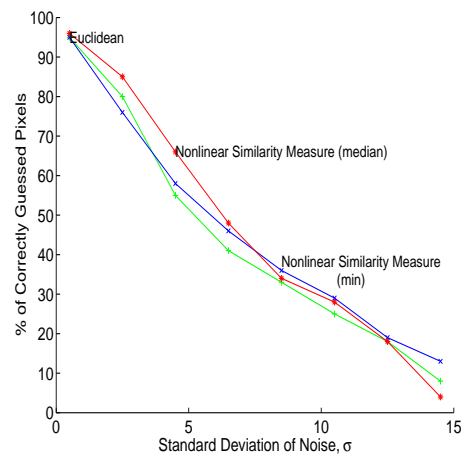
(c) Comparison for Girl Image.



(d) Comparison for Lighthouse Image.

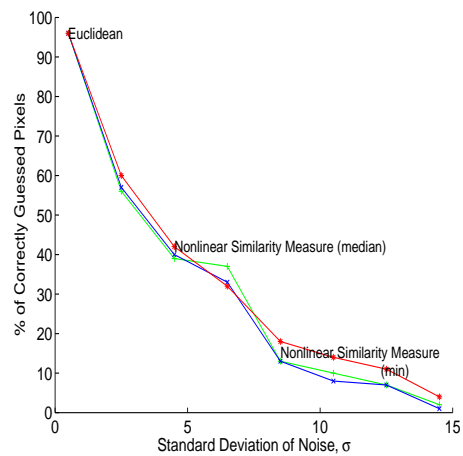


(e) Comparison for Gentleman Image.

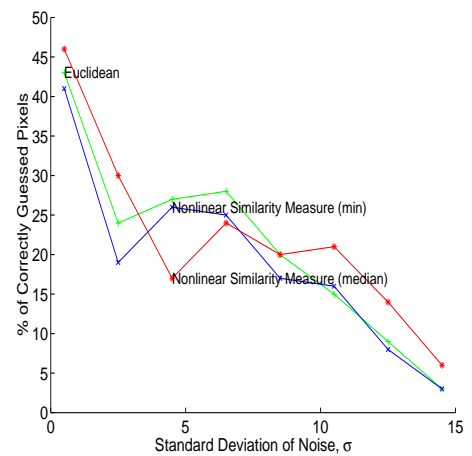


(f) Comparison for Palm Tree Image.

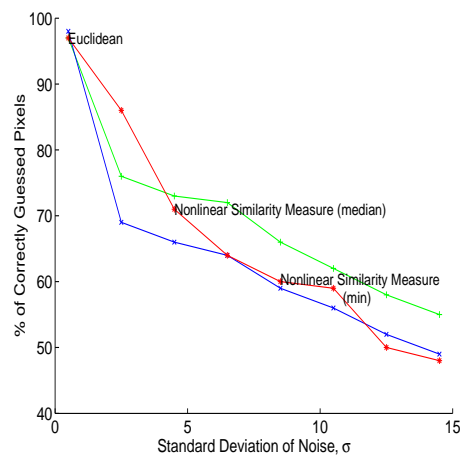
Figure 7. Comparison for images used in experimentation.



(a) Comparison for Shuttle Image.



(b) Comparison for Xray Image.



(c) Comparison for Lenna Image.

Figure 8. Comparison for images used in experimentation(Continued).

perform worse as noise increases, in most cases the proposed technique with median performs noticeably better, in terms of percentage of correct matches, than either of the other matching techniques (the proposed technique with minimum and conventional Euclidean distance). In only a few cases does our matching approach with median perform equally with, or slightly worse than one of the other techniques.

5. Conclusions

The proposed method is novel. Although it is a little more computationally expensive than matching using a Euclidean distance measure, its matching performance is in most cases appreciably greater. The experimental results show that our approach is less sensitive to noise, which is a common phenomenon in images. For future work, it would be very interesting to investigate in more detail why this method does perform better than euclidean distance measure, and whether the matching performance can be increased using steps of a fuzzy system model. It may be possible to make it more tolerant to noise by following fuzzy steps as done in the noise reduction in images by fuzzy filtering [8].

References

- [1] Steven M. Vajdic, Jeremy Graham, and Shaun Voigt. An application of image matching/fusion in medicine. *1996 IEEE*, 1:42–47, 1996.
- [2] Clark F. Olson. Maximum-likelihood image matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:853–857, 2002.
- [3] Tian Jinwen, Huan Jianzhong, Liu Jian, and Li Dchua. Image matching based on fuzzy information. *3rd International Conference on Signal Processing*, 2:946–949, 1996.
- [4] Moravec H.P. Towards automatic visual obstacle avoidance. *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, page 584, 1977.
- [5] Zhang Zhijia, Huang Shabai, and Shi Zelin. A fast strategy for image matching using Hausdorff distance. *2003 IEEE International Conference on Robotics, Intelligent Systems and Signal Processing*, 2:915–919, 2003.
- [6] Canny J. A computational approach to edge detection. *PAMI*, 8:679–698, 1986.
- [7] Abdul Ghafoor, Rao Naveed Iqbal, and Shoab Khan. Robust image matching algorithm. *4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications*, 1(7):155–160, January 2003.
- [8] Dimitri Van De Ville, Mike Nachtegael, Dietrich Van der Weken, Wilfried Philips, and Ignace Lemahieu. Noise reduction by fuzzy image filtering. *IEEE Transactions on Fuzzy Systems*, 11:429–436, 2003.
- [9] Muhammad Sirajul Islam and Les Kitchen. Image mathching using combo fuzzy measure. <http://www.cs.mu.oz.au/~msislam/cfm.pdf>.
- [10] Volker Gerdes. Stereo images with ground truth disparity and occlusion. <http://www-dbv.cs.uni-bonn.de/stereo.data>.
- [11] Christian Heipke. Overview of image matching techniques. http://phot.epfl.ch/workshop/wks96/art_3.1.html, 1996.