Histogram-Based Scene Matching Measures

O. Sjahputera, J.M. Keller, P. Matsakis, P. Gader, and J. Marjamaa

Department of Computer Engineering and Computer Science
University of Missouri-Columbia
Columbia, MO 65211

Abstract -*Fuzzy set theory has been used to handle uncertainty in various aspects of image processing, pattern recognition and computer vision. High-level computer vision applications hold a great potential for fuzzy set theory because of its natural language capabilities. Scene description, a language-based interpretation of regions and their relationships, is one such application that has used fuzzy sets with some success. This paper extends our earlier and on-going work in scene description in the following sense. If we have a linguistic description (from the system or from a human), and we revisit the scene, perhaps from a different orientation, can we match the scene objects and their relationships to be confident that we are indeed in the same place. We develop a scene matching methodology to accomplish this using histograms of forces between objects.*

1. Introduction

In earlier work, Keller and Wang [1] used a fuzzy rule-base to generate linguistic description of relative position between two image objects, and ultimately, to produce a complete description of the scene. The fuzzy rule-base received confidence values of the four main directional relations (LEFT, ABOVE, RIGHT, BELOW) and SURROUND based on the histogram of angles. Another approach for spatial relationship estimation is due to Matsakis and Wendling [2] who postulated an axiomatic framework for functions from which they generated "histograms of forces" to represent spatial relations between a pair of 2D image objects. By selecting particular functions, they generated various histograms, ranging from the histogram of angles to a histogram of gravitational forces. Keller and Matsakis [3] utilized the histogram of forces to generate numeric features from multiple force histograms that were used to generate a linguistic description of a scene using a rule-base. This approach encompasses the earlier paradigm and leads to a richer language for scene description. Due to the complementary nature of the histograms of forces, it is even possible to construct a self assessment measure for each linguistic description between a pair of image objects.

In this paper, we describe an approach to compare pairs of 2D image objects in different scenes and assign a confidence value indicating how similar they are. This is a part of an ongoing research in the area of linguistic scene description for applications in recognition technology in general, and automatic target recognition in particular. Each histogram set serves as evidence that must be satisfied if we attempt to find the same object in other digital images. Hence, we need to measure the degree of similarity between two sets of histograms of forces to determine how well a match is satisfied. We must also take into account that the two descriptions to be compared may have different orientations. We use the notion of "main direction" of the histogram of forces and normalize for orientation by shifting the histograms. This approach is tested on a controlled set of synthetic images and on real objects extracted from LADAR images of a powerplant scene.

2. Spatial Relation Methods

The fuzzy relative position between 2D objects is often represented by a histogram of angles [4, 5]. The histogram of angles associated with any pair (A,B) of crisp and digitized objects is a function Ang^{AB} from $\mathbb{R}$ into $\mathbb{N}$. For any direction θ, the value $Ang^{AB}(\theta)$ is the number of pixel pairs (p,q) belonging to $A \times B$ such that p is in direction θ of q. In [2], Matsakis and Wendling introduced the notion of the histogram of forces. It generalizes and supersedes that of the histogram of angles. It ensures rapid processing of raster data as well as of vector data, and of crisp objects as well as of fuzzy objects. It also offers solid theoretical guarantees, and allows explicit accounting of metric information. The histogram of forces associated with (A,B) via F, or the F–histogram associated with (A,B), is a function F^{AB} from $\mathbb{R}$ into $\mathbb{R}_+$. Like Ang^{AB}, this function represents the relative position of A with regard to B. For any direction θ, the value $F^{AB}(\theta)$ is the total weight of the arguments that can be found in order to support the proposition "A is in direction θ of B". More precisely, it is the scalar resultant of elementary forces. These forces are exerted by the A points on those of B, and each tends to move B in direction θ. Actually, the letter F

denotes a numerical function. Let r be a real. If the elementary forces are in inverse ratio to d^r, where d represents the distance between the points considered, then F is denoted F_r. For instance, the F function associated with the universal law of gravitation is F_2. The F_2–histogram and F_0–histogram (histogram of constant forces) have very different and very interesting characteristics. The latter, very similar to the histogram of angles, gives a global view of the situation. It considers the closest parts and the farthest parts of the objects equally, whereas F_2–histogram focuses on the closest parts.

3. Histogram Compatibility Measure

Each pair of objects from the reference scene and the target scene are represented by relative position histograms. Hence, we need a method to determine how well the target objects match those of the reference. Histogram compatibility measure between histogram H1 and H2 is defined as:

$$C_{H_1H_2} = \max_{\theta}\left(\frac{\left|H_2^*(\theta)\cap H_1^*\right|}{\left|H_2^*(\theta)\cup H_1^*\right|}\right) \quad \theta\in[-180^o,180^o]$$

Both histograms are normalized, noted as H^*. Hence, both histograms can be seen as fuzzy sets, where at least one histogram is a normal fuzzy set. H_2 is shifted by some angle θ to best match H_1. The angle θ that produces maximum match is called the shift angle. This shift angle can be associated with the amount of rotation in 2D space (or rotation around the z-axis for 3D space) as shown below. We choose to use the standard min and max operator as our intersection and union operator. Here $C_{H1H2}\in[0,1]$ where the histogram compatibility measure C_{H1H2} is maximized to 1 when $H_1 = H_2$.

4. Synthetic Images

In this paper we use 8 images (Image1-Image8) with Image1 as the template. Each image contains a reference (dark gray) and an argument object (the light gray object). The rotational variants are obtained by rotating the original image from 0^o to 90^o with 5^o increment, and the declination variants are obtained by varying the declination angle from -45^o to 45^o with 5^o increment. The original form (zero rotation and zero declination) of these images are given in Figure 1. For these synthetic images, the goal is to determine from the spatial relationship histograms whether the assignment of light and dark gray objects from test images (2-8) to template image (1) correspond to a scene match. The test images simulate conditions such as variation in object shape and size (Image2-Image7), and variation due to individual object rotation (Image3, Image8). These could be due to problems in segmentation or occlusion in the actual 3-D scene.

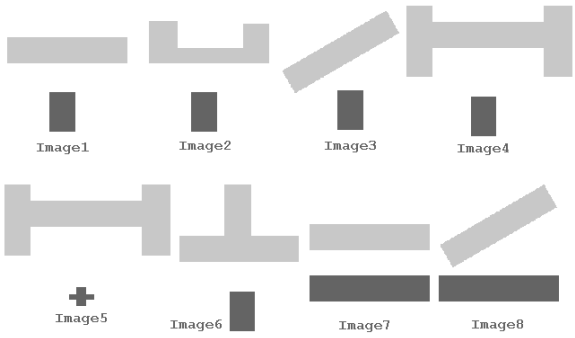

Figure 1: Image1 to Image8 in original orientation.

5. Experiments with Synthetic Images

To calculate the spatial relation compatibility measure between object pair A and B, we first generate the histograms: F0 (constant forces), F2 (gravitational forces), and A0 (angles) that represent the relative positions of the two objects. Then we calculate the histogram compatibility measure for each corresponding pair of histograms between the test image and the template. Hence, we will have C^{F2}, C^{F0}, and C^{A0}. The aggregation of these values is defined as the spatial relation compatibility measure between object pair A and B. In this paper we selected the Choquet integral [6] with a λ fuzzy measure as our aggregation method. The degree of optimism of these histogram compatibility measures can be given in the order of F2, F0, and A0, with F2 being the most optimistic. Therefore, we heuristically chose the density values of 0.6, 0.9, and 0.99 for F2, F0 and A0 respectively. This will reduce the influence of F2, while utilizing more of the pessimistic one (A0). The results of each histogram measures and their fusion using Choquet integral from Image1 are shown in Figure 2.

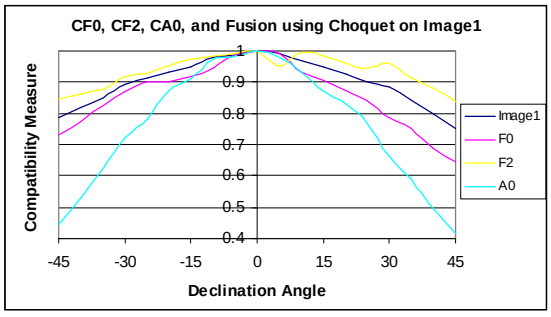

Figure 2: Results of C^{F0}, C^{F2}, C^{A0} and their Choquet fusion on declination variants of Image1 at 0 rotation.

5.1 Effect of Image Rotation

We found that image rotation does not affect the spatial relation compatibility measure. All rotational variants of Image1 receive compatibility measure above 0.98 at any rotation angle, where variants of other test images obtain 0.8 or less compatibility value, except for Image3, which will be explained in the following section. The average shift angle approximates the actual rotation angle applied to the images. The only exceptions are found on Images 5 and 8, where their shift angles lead the actual rotation by about 5^o. This is due to non-symmetric object positioning in Image5 and Image8 resulting in rather skewed, non-symmetric histograms, while other test images have relatively symmetrical histograms.

5.2 Effect of Shape and Size Variation

The test images used here (Image2, Image4-Image7) were obtained by altering the shape and size of one or both objects in an attempt to modify various aspects of the histograms (F0, F2, A0). These histogram aspects include shape, magnitude and support. We define our targets as declination variants (constructed by computer graphics) of each test image with Image1 as the template. The results are shown in figure 3.

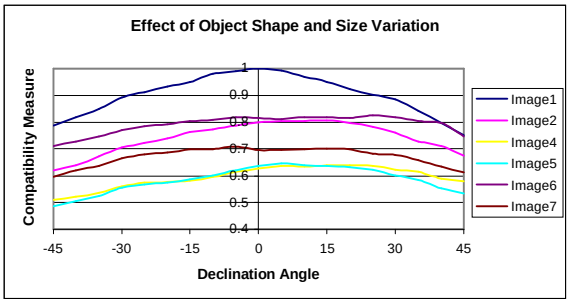

Figure 3: Effect of shape and size variations in objects with Image1 as template.

The declination variants of Image1 always receive the highest compatibility measures at any declination angle. Also, when the shape distortion does not affect the support of histograms (as in Image2, Image6 and Image7) the compatibility measure is higher than those which do (i.e., Image4 and Image5). The highest compatibility measure obtained by a target image is achieved by Image6 around 25^o-30^o declination. Note that the argument object in Image6 contains an additional structure perpendicular to the center of the original argument object. As the declination angle increases from 0^o, perspective will have the effect of making this structure look smaller, hence reducing its influence over the histogram. This makes the image look more like the template. However, perspective also makes the two horizontally parallel objects look closer at higher declination angles. Eventually, the latter perspective effect becomes significant enough to cause the compatibility measure of Image6 to drop back down.

5.3 Effect of Individual Object Rotation

We have stated that image rotation has little effect on spatial relation compatibility measure. However, we may run into cases where only one of the participating objects is "rotated", perhaps due to perspective distortion, occlusion, or segmentation problems. Consider Image3 where the argument object is rotated from the template (Image1) by 30^o around its centroid. The results from comparing declination variants against the template are shown in figure 4. Declination variants of Image3 receive very high compatibility measures. In fact, at declination angle of 35^o or higher, these variants achieved higher compatibility measures than their counterparts from Image1. This condition can be attributed to the reference object's symmetry (close to a square) and size (smaller than argument). This condition causes the reference object to be evaluated like a point with respect to the argument. Thus, rotating only the argument by some angle creates similar effects to the histograms as rotating the whole image by the same angle and direction. This is shown by the average shift angle for declination between -45^o to 45^o, which is found around 24^o, approximating the actual 30^o rotation we applied. For large positive declination angle, the size of the argument object in declination variants of Image1 is reduced in size uniformly due to its horizontal orientation. However, for Image3, the size reduction effect due to perspective is experienced the most by the object tip pointing away from us. On the other hand, perspective also makes the other tip (pointing toward us) look larger than the original. So, at high positive declination angle, the variants of Image3 match the features found in the histograms of the original template better than declination variants of the template itself.

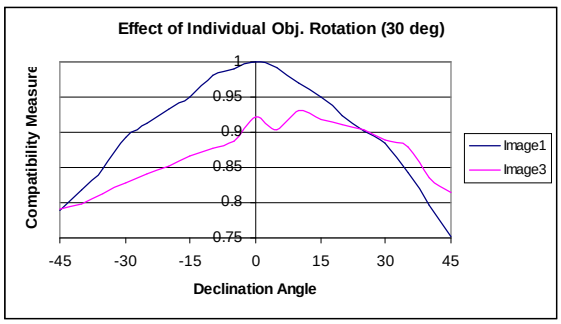

Figure 4: Effect of argument object rotation (Image3) with template=Image1.

To see whether our previous result also holds for situation where the reference object is not compact and is comparable in size to the argument, we created Image7 where the reference object is identical to the argument. For the target, we rotated the argument object 30° around its centroid and designated it as Image8. From Figure 5, we see that declination variants of Image8 also receive high compatibility measures, exceeding those of the declination variants of the template at extreme angles (-45°). However, the average shift angle of Image8 was measured around 8°, far from the actual angle of rotation applied to the argument object.

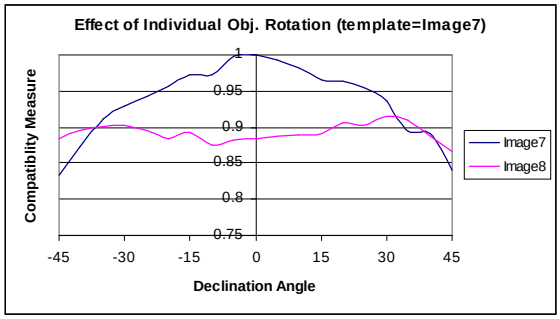

Figure 5: Effect of argument object rotation (Image8) with template=Image7.

6. Scene Matching with Real Images

We also used several hand segmented scenes taken from power-plant images generated using laser radar (LADAR). The LADAR range data is processed using the pseudo-intensity filter [1] defined as $[1+G_x^2+G_y^2]^{-0.5}$ where G_x and G_y are Sobel gradient magnitudes in a 3x3 window after median filtering. An example is shown in the top image of figure 6. The power-plant range data were provided by the Naval Air Warfare Center (NAWC), China Lake, Ca. Several flights were made from different directions over the power plant. The objects appear different between flights and at different times in a given flight. The bottom image in figure 6 shows 4 hand segmented objects in one of the images. We chose the objects from DS1 as our template image (figure 7). There were 10 images altogether (named DS1 - DS10), seven of which contained the same 4 objects, and three where the objects assignment was wrong. We ask two interrelated questions: 1. Are the objects in the given scene the same as those in the reference?; and 2. What is the best assignment of object labels to verify the answer to question 1?

Given two scenes with manually assigned references, the compatibility measure between the two scenes is defined as the average of the compatibility measures between each argument object (labeled 1 through 3) and the reference (labeled 0). Since there are only three argument objects in each scene, there are six possible label matching schemes between argument objects in the target scene and argument objects in the template scene. A brute force approach is currently used to determine the best matching scheme. Below, five test images are described. These images provide representative samples of all the images tested.

When comparing a scene to itself, the scene compatibility measure is expected to reach its maximum at 1. This hypotheses is verified when we tested DS1 against itself as the template.

The Scene, DS2, is taken from the same flight as the template, DS1, shown in figure 8. This explains the similarity in shapes and positioning. The major differences between the two images are scaling, due to the approach on the object, and slight variations in the shapes of objects due to the inaccuracy of segmentation. In spite of these differences, the compatibility measure between the two scenes remains rather high, 0.88.

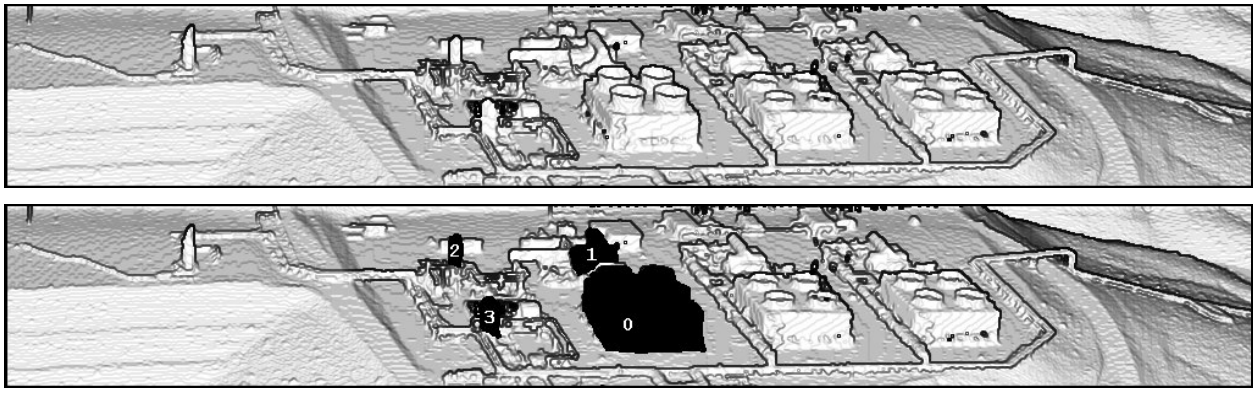

Figure 6. Pseudo-intensity image of power-plant scene as used in DS4 scene. Labeled objects are as shown in DS4.

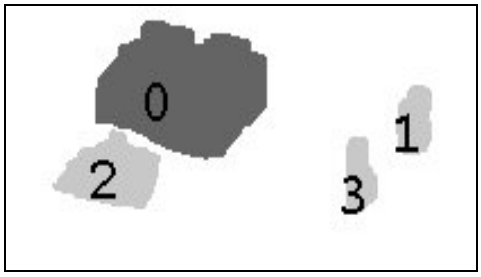

Figure 7. Objects from DS1 (template scene).

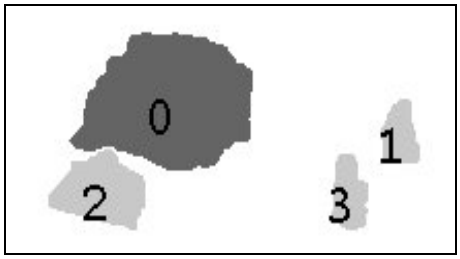

Figure 8. Objects from DS2 (same flight as DS1).

Now, a scene, DS5, that is obviously not the same as DS1 is tested (figure 9). Here, the maximum compatibility measure of any matching scheme is 0.37. This is quite low, and indicates that the assumption that the two scenes are the same is false.

Figure 9. Objects from DS5 (scaled to fit).

In figure 10 (DS6), the same four buildings as in DS1 are found, viewed from a different flight path and a different approach angle. The highest scene compatibility at 0.65 belongs to the correct matching scheme. The low compatibility measure of this test is partially due to the unusual changes in shape between DS6 and the template scene. Also, the angle between the original flight path and this one, in combination with the declination angle, causes a large distance alteration between the reference object and object 3.

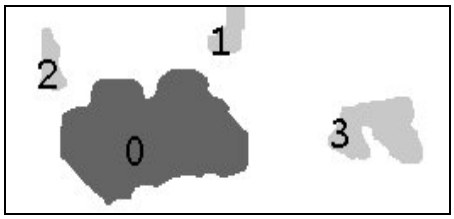

Figure 10. Objects from DS6.

The fifth test, DS7 (figure 11), shows the same scene from yet another path and approach angle. This flight is in the opposite direction from the original. The test yields a maximum scene compatibility of 0.79 with the correct matching scheme. This problem is easier than the previous because the difference in approach angle is about 180^{o}. At this angle, the vertical and horizontal distances between objects have about the same ratio as those in the template scene.

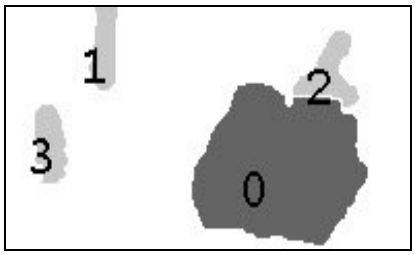

Figure 11. Objects from DS7.

7. Conclusions

Spatial relationship compatibility measures defined in this paper can be used to detect shape variation in objects and perform object/scene matching. It is also flexible enough to accommodate a limited range of declination variation. Several histogram compatibility measures are used to calculate spatial relation compatibility measure between two object pairs. In turn, several spatial relation compatibility measures are used to find the scene compatibility measure as shown using the power-plant scenes.

Acknowledgements

This work is supported by the Office of Naval Research (ONR) grant N00014-96-0439.

References

[1] J. Keller, X. Wang, "A Fuzzy Rule-based Approach for Scene description Involving Spatial Relationships", *Computer Vision & Image Understanding*, revised, 1999.

[2] P. Matsakis and L. Wendling, "A New Way to Represent the Relative Position between Areal Objects", *IEEE Trans. PAMI*, vol. 21, no. 7, pp. 634-643, 1999.

[3] J. M. Keller and P. Matsakis, "Aspects of High Level Computer Vision Using Fuzzy Sets", *Proceedings, FUZZ-IEEE'99*, Seoul, Korea, pp. 847-852, 1999.

[4] X. Wang and J. Keller, "Human-based Spatial Relationship Generalization Through Neural/Fuzzy Approaches", *Fuzzy Sets and Systems*, Vol. 101, No. 1, pp. 5-20, 1999.

[5] K. Miyajima and A. Ralescu, "Spatial Organization in 2D Segmented Images: Representation and Recognition of Primitive Spatial Relations", *Fuzzy Sets and Systems*, vol. 65, No. 2/3, pp. 225-236, 1994.

[6] M. Grabisch, T. Murofushi, M. Sugeno (Eds), *Fuzzy Measures and Fuzzy Integral: Theory and Application*, Physica-Verlag, Heidelberg, 2000.