Why all the Commotion about Invariant Theory?

D. Gregory Arnold
Air Force Research Lab
Target Recognition Branch, AFRL/SNAT
Bldg. 620, 2241 Avionics Circle, WPAFB, Ohio 45433-7321
garnold@mbvlab.wpafb.af.mil, (937)255-1115 x4388

Kirk Sturtz
Veridian Incorporated
5200 Springfield Pike, Suite 200, Dayton, Ohio 45431
ksturtz@mbvlab.wpafb.af.mil, (937)255-1115
http://www.atrinv.vdl.afrl.af.mil/

Abstract ***Recent research in the calculation and exploitation of object/image relations suggests new approaches to ATR. The resulting conceptual advances transform the ATR problem into a feature extraction and correspondence problem. This paper addresses several application areas and demonstrates how the corresponding research areas are related by a common theory.***

Keywords: Invariants, Object/Image Relations, ATR

1 Introduction

The use of invariant theoretic techniques in automatic target recognition (ATR) systems has long been promising but shortcoming in important practical applications. We strongly believe that with the new understanding of how invariant theory determines the object/image (O/I) relations that progress will advance rapidly in several areas of concern to the ATR community. The advantage of using O/I relations is that they allow one to circumvent estimating camera and object orientation parameters. This is significant since these parameters are often unknown. Fundamentally these O/I relations transform the object recognition problem to a feature extraction and correspondence problem. A short summary of what O/I relations are and how they are determined using invariant theory can be found in [1, 2, 3].

With this new understanding of how invariant theory can be used in the object recognition process, it has matured to the point that it is ripe for applications in many areas. This paper discusses several applications.

2 Application Areas

Here is an overview of ways that invariant theory, using O/I relations, can help in applications of interest to the object recognition community. The following examples are all possible by studying the transformation groups and imaging transformations of each system using Lie group analysis [4]. These examples are mostly unexplored but are readily accessible with the current techniques. An essential theme throughout these examples is that invariants naturally handle articulation and obscuration provided the features are sufficiently local.

Two types of invariants are interesting for ATR problems, geometric and photometric. Geometric invariants relate the object features to the sensed *locations*, whereas photometric invariants relate the object features to the sensed *intensity*. O/I relations of both types

will be necessary to maximize the discrimination.

2.1 Recognition

Before starting the specific examples, we give some general capabilities that invariants provide for the classification problem:

1. Register images of an object given 6 (full perspective is the worst case) corresponding points.

2. Optimize feature detector selection for general/specific recognition tasks.

2.1.1 EO (Visible)

Typically, EO invariants implies geometric invariants. These are functions of geometric features of the objects to be recognized. Stiller of Texas A&M and Weiss of the University of Maryland have developed an invariant function which can be used for recognizing general point configurations from a single image [5, 6]. We discuss the affine case because it is conceptually easier (they have developed the analogous theory for the full perspective camera). We concentrate on Weiss' algorithm since it has been applied to a small set of object recognition problems:

1. Extract invariant features from the *3-D* models. These features are a function of 5 points on the object and correspond to a single point in the invariant space. A point will exist in the invariant space for each combination of points of each object to be recognized.

2. Given a single image, extract all the features and group into sets of 5 points. From the image, each *2-D* -point set corresponds to a line in the *3-D* invariant space. Efficient methods can be used to determine which point sets intersect *3-D* model points in the invariant space.

3. Once an intersection has been found, a *3-D* model and viewpoint has been identified.

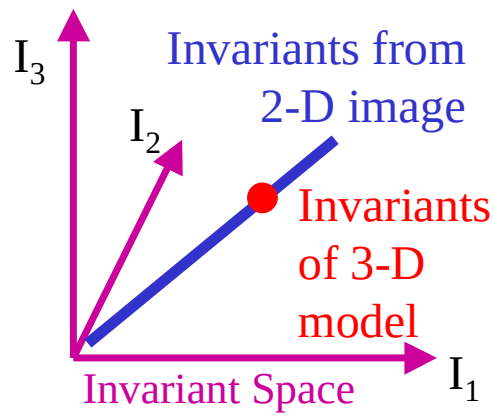

Figure 1: Object Recognition Example: A weak perspective 2-D image of 5 points yields 3 object invariants (big I's) and 4 image invariants (little i's). The *3-D* object represents a point in invariant space, and an image represents a line. If the point and the line intersect, then the image could have come from the object.

4. The *3-D* model and viewpoint can be used to project the entire model back into the original image. This image can then be compared in detail to verify a good match has been made.

The primary difficulty with this algorithm is the feature correspondence problem. It is currently handled by minimizing the number of features searched (by feature labeling) and exhausting the combinatorics. However, as it stands now, this algorithm is relatively fast and robust under very general circumstances.

2.1.2 IR/MS/HS (Infrared / Multispectral / Hyperspectral)

Two types of invariants are applicable to spectral data: geometric invariants and intensity invariants. Geometric invariants are the same as discussed above in the EO example. Intensity invariants are a generalization of the concept of color constancy also developed for EO imagery. The fundamental physical properties being measured are the reflectivity and emissivity of a material as a function of wavelength. The difficulty resides in the fact that these are

passive sensors (worse than a bistatic platform because the location, intensity, and waveform of the sources are unknown). Although we are just beginning to apply invariance theory toward hyperspectral imagery, Glenn Healey of U.C. Irvine has demonstrated some great results using an invariant space for resolving the reflectivity [7].

2.1.3 RADAR (monostatic Synthetic Aperture RADAR and High Range Resolution SAR mode)

In the SAR mode, we are primarily interested in the raw phase history data right off the sensor. A reasonable starting point is to assume object attached persistent scatterers are available. The invariant space will be simpler than the space for EO imagery (4 points minimum vs. 6 points minimum). Also, if an object is moving independently of the RADAR motion, then the correspondence problem is greatly reduced (see Tracking). Otherwise, feature labeling will be needed to help reduce the combinatorics problem discussed in EO. An invariant space can be found (as in EO) that indicates when a set of data points (which form a line in the invariant space) match a model (which is a point in invariant space). Also, (similar to EO), a *3-D* model can be reconstructed from multiple views of 4 or more points. This model can be used for recognition. Once the model has been built, it will greatly aid future tracking of the original object. The views of the object must be sufficiently different to resolve the *3-D* ambiguity (i.e. they must be numerically stable). See GMTI mode for a brief discussion on handling the degenerate cases. Stuff [8, 9] has been pursuing this idea and has had excellent success. Figure 2 shows an example of how this is done.

2.1.4 RADAR (monostatic High Range Resolution GMTI mode)

The invariant method in the GMTI mode is very similar to the method proposed in SAR mode. The primary differences are that in GMTI mode, one is typically tracking an object. By sending very rapid pulses the radar is effectively stationary and object motion can be separated from background clutter. Therefore, the feature correspondence problem is greatly simplified during each burst. However, it is unlikely that a single burst of pulses has long enough duration for significant object motion to occur. This implies that the object may not be completely modeled in *3-D* after one burst. Each burst may constrain the result to a line in the invariant space. If the line is sufficiently unique, recognition may be possible immediately. More likely, one would reason over several bursts. Features correspondence again becomes more difficult, but the increased time between views will likely produce enough variation in viewpoint to construct a full *3-D* model of the object being tracked. Once a *3-D* model has been built, recognition is easier and an image of the object may be produced from an arbitrary viewpoint.

2.2 Sensor Fusion

Invariant spaces are the silver bullet of sensor fusion. Each imaging modality/viewpoint can be mapped into the same invariant space if the geometry can be determined. This assumes the features are object attached. This assumption is restrictive but acceptable in many applications. However, generalizations of "object attached" can also be explored with invariant theory (for example, a specific transformation to relate a feature with the object to which it corresponds).

2.3 Tracking

This is probably one of the ripest areas for application, at least for the case where several images are taken of an object over a very short time period. Because the images will be very similar over a short time period, tracking a set of features through that time period should be relatively easy (given sufficient resolution and noise characteristics). The data can be transformed into the invariant space

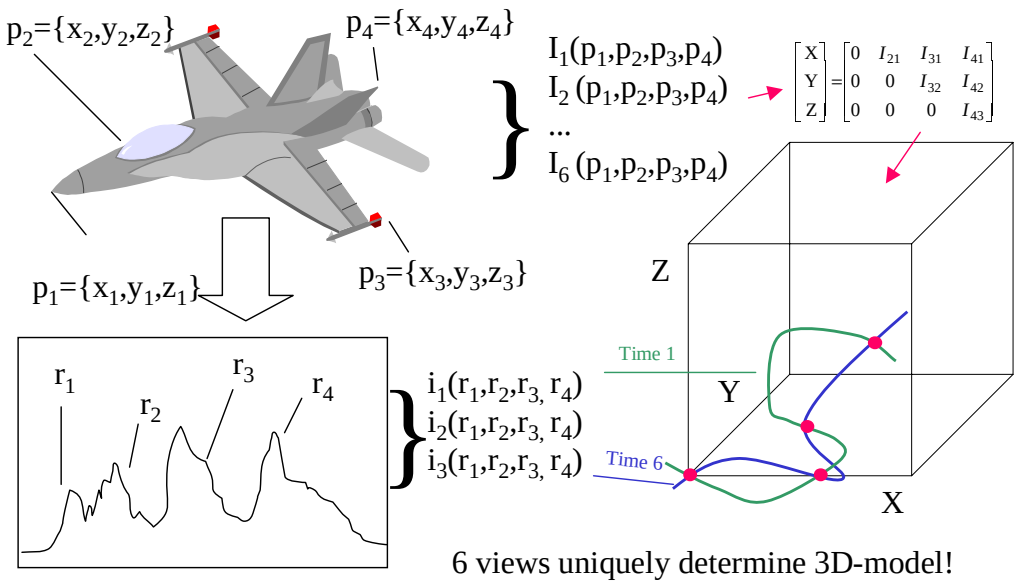

Figure 2: Example Application: Construction of a 3-D scattering center model from the raw RADAR data. The goal is to determine the big I's (6 in this case of 4 points) given a sequence of measurements (little i's). Each measurement will define a surface in 6-D (notionally shown here in *3-D*). Once sufficient measurements have been made, the intersection will define 4 unique points (the scattering centers in a standard position).

where a *3-D* model will be reconstructed from the data (multiple lines intersecting a common point in the affine case). This *3-D* model then serves as a direct means for continuing to track the object through an image sequence and also for recognizing the object. If the imagery has sufficient resolution, but feature extraction is difficult due to noise or obscuration, then the *3-D* invariant space can be used to determine if two sets of points could possibly correspond to the same object. Therefore, errant points can quickly be removed from consideration as a potential point on the object. Integrating invariants with existing methods for feature tracking will make it possible to perform geometric reasoning about features crossing and/or obscuring other features.

2.4 Database/Modeling

Invariants can be used to create model databases that contain all the essential information that will be used for a particular application. This model database can be used to develop theoretical bounds on class separability. The residual (after removing the invariant features) could be stored for noise analysis and for future developments. Furthermore, the invariant modeling process would naturally recognize similarities within the same model thus providing the capability to build models that automatically understand the geometric symmetry in the *3-D* model.

2.5 Simulation

One of the most useful proof techniques employed in analytic geometry is that by using invariance, a proof can be done in the "most advantageous position." For example, if one wanted to prove a theorem about the area of a triangle, one may use a triangle in a standard position. By invariance (with respect to the appropriate transformations) this position is equivalent to all other positions with respect to calculation of the area.

This concept applies to simulation. The general idea, which has been known for some time in the geometric invariance community, is called transference. Shashua [10] has used it

to transfer a picture from one orientation to another without any knowledge of the image and without solving for the underlying transformation. Obviously, there are limitations based on what is visible from what viewpoint. However, simulations would be nearly instantaneous by pre-computing an image within each unique region of the viewing sphere, and then transferring the standard position to the desired viewpoint. Furthermore, geometric reasoning may provide insight into a "smart transference" method such that occlusion could be handled using on the order of a single simulation. In other words, the simulation code would handle the physics associated with the sensor, and the invariance methods would handle the geometry associated with viewpoint.

2.6 Reasoning With Unknown Information

Just as trying to identify an object from its image is reasoning with unknown information (the z coordinate in the case of visible imagery), other reasoning tasks may be decomposed into knowns and unknowns. Invariant theory provides a technique to reduce the unknown information as much as is possible. The unknown information will represent surfaces in an invariant space. The known information will constrain the surface to a specific shape and region. In other words, the surface in this abstract space represents the range of possible outcomes due to the unknown information. This concept could be augmented with probabilities.

3 Summary

Object/image relations offer a fundamental approach to handling the geometry (location) and photometry (intensity) for ATR systems. Using invariant theory to determine O/I relations offers a unified viewpoint for understanding different sensors operating at different wavelengths.

For geometric problems, invariant theory transforms the object recognition problem into a feature extraction and correspondence problem. The geometric object/image relation expresses the geometric structure determined by the sensor. We have cited several application areas which can be addressed using O/I relations.

Our future research, in collaboration with the Ohio State University and the Air Force Institute of Technology, is directed at geometric problems where we are employing depth sensors. We suspect that photometric invariants will be necessary in conjunction with the geometric invariants in order to reduce the combinatorics and to do feature selection for this research.

References

[1] Greg Arnold and Kirk Sturtz. Complexity analysis of atr algorithms based on invariants. In *Proceedings Computer Vision Beyond the Visible Spectrum (CVBVS)*, Hilton Head, SC, June 2000. IEEE Computer Society. To Appear.

[2] Greg Arnold and Kirk Sturtz. Geometric object/image relations for radar. In E. Zelnio, editor, *Proceedings SPIE Int'l Conf.*, Algorithms for Synthetic Aperture Radar Imagery VII, Orlando, FL, Apr 2000.

[3] Greg Arnold, Kirk Sturtz, and Greg Power. Analyzing image sequences using invariants. In *CISST*, Las Vegas, Nevada, June 1999.

[4] Greg Arnold, Kirk Sturtz, and Vince Velten. Lie group analysis in object recognition. In *Proc. DARPA Image Understanding Workshop*, New Orleans, LA, May 1997. Morgan Kauffman.

[5] P.F. Stiller, C.A. Asmuth, and C.S. Wan. Single view recognition - the perspective case. In *Proceedings SPIE Int'l Conf., Vision Geometry V*, volume 2826, pages 226–235, Denver, CO, Aug 1996.

[6] Isaac Weiss. Model-based recognition of 3d objects from one view. In *Proc. DARPA Image Understanding Workshop*, pages 641–652, Monterey, CA, Nov 20-23 1998. Morgan Kauffman.

[7] David Slater and Glenn Healey. Exploiting an atmospheric model for automated invariant material classification in hyperspectral imagery. In *Proceedings SPIE Int'l Conf.*, Orlando, FL, Apr 1998.

[8] Mark A. Stuff. Three-dimensional analysis of moving target radar signals: Methods and implications for atr and feature aided tracking. In E. Zelnio, editor, *Proceedings SPIE Int'l Conf.*, Algorithms for Synthetic Aperture Radar Imagery VI, Orlando, FL, Apr 1999.

[9] Mark A. Stuff. Three dimensional invariants of moving targets. In E. Zelnio, editor, *Proceedings SPIE Int'l Conf.*, Algorithms for Synthetic Aperture Radar Imagery VII, Orlando, FL, Apr 2000.

[10] A. Shashua and M. Werman. Trilinearity of three perspective views and its associated tensor. In *Proceedings ICCV*, pages 920–925, Cambridge, MA, June 1995. IEEE, Computer Society. 20-23.