Progress in Computer Vision at the University of Massachusetts

Edward M. Riseman, Allen R. Hanson
Robert Collins, Bruce Draper, Richard Weiss

Computer Vision Research Laboratory
Dept. of Computer and Information Science
University of Massachusetts
Amherst, MA 0l003

Abstract[1]

This report summarizes progress in image understanding research at the University of Massachusetts over the past year. Many of the individual efforts discussed in this paper are further developed in other papers in this proceedings. The summary is organized into several areas:

1. Image Exploitation under RADIUS
2. Learning in Vision
3. Mobile Vehicle Navigation
4. Recovering 3D Structure
5. The Image Understanding Architecture

The research program in computer vision at UMass has as one of its goals the integration of a diverse set of research efforts into a system that is ultimately intended to achieve real-time image interpretation in a variety of vision applications.

1. Image Exploitation under RADIUS

1.1. Project Goal

The goal of the Radius project (Research and Development for Image Understanding Systems) is to develop image understanding algorithms that support model-based aerial image exploitation. A flexible set of IU modules is being developed to acquire, extend and refine 3D geometric site models using known extrinsic and intrinsic camera parameters. In support of this effort, ISR3 (see Section 2.5) is being ported to the RCDE. Although individual modules are automated as much as possible, the image analyst maintains control of the site modeling session by deciding when and where each module is applied. A recent experiment on site modeling using the RADIUS Model Board 1 images and these algorithms is described in [Collins et. al. 1994].

1.2. Finding Roof Structures in the RADIUS Model Board Imagery

A new method for extracting flat, polygonal rooftops in monocular aerial imagery has been developed [Jaynes et. al. 1994]. Through a combination of bottom-up and top-down construction of perceptual groups, polygons in a single aerial image are robustly extracted.

Building roof detection begins by using vanishing point information in the 2D image [Collins and Weiss 1990] to generate a set of correlation masks for detecting features in the image corresponding to 3D orthogonal roof corners in the scene. Compatible pairs of roof corners initiate a search for supporting edge data, with all extracted corners and supporting edges entered into a feature relation graph. Perceptual grouping processes connect nodes in the graph representing corners and lines according to their geometric compatibility. These grouping processes invoke top-down feature verification routines so that features, and the relational links between features, are all verified with local information in the image and weighted in the graph according to their underlying image support.

Potential building roof hypotheses are detected as cycles in the feature relation graph. Virtual features may be hypothesized for the perceptual completion of partially occluded rooftops appearing as open chains in the graph. Extraction of the 'best' grouping of features into a set of building rooftop hypotheses is posed as a graph search problem. The maximally-weighted, independent set of cycles in the graph is extracted as the final set of roof boundaries. This method essentially arbitrates between many (possibly

[1]This research has been supported in part by the Advanced Research Projects Agency under TACOM contract number DAAE07-91-C-R035, HDL contract number DAAL02-91-K-0047, TEC contract number DACA76-92-C-0041, and Colorado State University sub-contract number DAAG04-93-G-0422, by the National Science Foundation under grant CDA-8922572, IRI-9113690, and IRI-9208920, and by Rome Labs under contract number F30602-94-C-0042.

conflicting) groupings by using information contained in the image to measure the certainty of each grouping choice. Note that all the grouping processes at this stage operate in the image domain, and further refinement will take place next at the 3D representation during the information fusion stage across multiple images.

1.3. Determining the 3D Structure of Buildings

Following polygonal roof detection, simultaneous multi-image, multi-line epipolar matching and triangulation algorithms are run to locate geometric support in other images (often with a great variety of viewpoints), and to determine the precise size, shape, and position of the building roof in the local site 3D coordinate system. At the moment, only flat, horizontal rooftops are detected.

The constraint of horizontal roofs allows the use of a very efficient and robust histogram-based algorithm to gather supporting line evidence from other images in the site database. This is achieved using the known poses to backproject into the scene each polygonal edge of the roof hypothesis in the first image to determine a planar wedge in which the corresponding 3d horizontal roof edge must lie. This wedge is truncated at a predetermined maximum and minimum height values for the site, and projected into a second image to form an epipolar search region. This quadrilateral search area is scanned for possible matching line segments at the expected orientation, each potential match implying a different roof height in the scene. All heights computed from potential polygon edge matches in the second image are stored as votes in a height histogram. Each vote is weighted by compatibility of the match in terms of expected line segment orientation and length. A single global histogram accumulates height votes from multiple images, and for multiple edges in a rooftop polygon. After all votes have been tallied, the histogram bucket containing the most votes yields a set of corresponding polygonal rooftop line segments in multiple images and a coarse estimate of the roof height in the scene.

A multi-image line triangulation routine again uses known camera poses to fuse the line segment correspondences determined by epipolar matching into a wireframe rooftop structure in the scene. During triangulation, each wireframe edge is characterized by the parameters of the algebraic infinite 3D line that coincides with it. In turn, the location and orientation of this 3D line is characterized by an objective function measuring how well its projection aligns with the 2D image segments that correspond to it. A number of different objective measures are being considered; the current one is a function of the sum of squared distances from each projected infinite line to the endpoints of corresponding 2D line segments in the image. The triangulation algorithm iteratively refines the position of each wireframe edge so that the sum-of-squares of objective measures between each algebraic line and corresponding image line segment is minimized. All lines in a rooftop polygon are thus triangulated simultaneously. Object-level geometric constraints such as assumed perpendicularity, parallelism, and coplanarity of polygonal rooftop edges are also imposed during triangulation for more reliable results.

The resulting 3D roof polygon is extruded down to the terrain to form a volumetric building model. The ground height at any location is specified by a digital terrain map that is either provided with the site database, or created from a pair of stereo images using the UMass Terrain Reconstruction System [Schultz 1994]. The final reconstructed building is then added to the current site model to form a more complete scene representation that can be used for future image registration and interpretation tasks.

1.4. Projective Intensity Mapping

To provide added realism for visual displays and as a means of symbolic storage, we have developed mechanisms for projectively warping image intensities onto polygonal building faces. Since multiple images are used, intensity information from all faces of the building polygon can be recovered, even though they are not all seen in any single image. By storing surface intensity information with each building object, the rendered intensity mappings provide a convenient storage method for later symbolic extraction of detailed surface structures like windows, doors and roof vents without needing to know ahead of time which information will be desired. Furthermore, this subsequent processing becomes greatly simplified. For example, rectangular lattices of windows or roof vents can be searched for in the unwarped intensity maps without complication from the effects of perspective distortion. Secondly, specific surface structure extraction techniques can be applied only where relevant, i.e. window and door extraction can be focused on building wall intensity maps, while roof vent computations are performed only on roofs. Future work will be directed towards combining intensity information from multiple views of each polygonal building facet to remove visual artifacts caused by shadows and occlusion and to potentially increase the clarity of surface intensity maps using super-resolution fusion techniques.

1.5. Camera Resection and Feature Correspondences

We continue to investigate new methods for performing camera resection and correspondence matching [Cheng et. al. 1994]. Given N images, and a set of corresponding 3D-to-2D point features for each image, a robust algorithm for simultaneously resecting intrinsic and extrinsic camera parameters for all views has been developed. A-priori constraints, such as the knowledge that all views were taken by the same camera, can be incorporated. Unknown camera parameters are determined from the standard colinearity equations using the Levenberg-Marquardt algorithm. This iterative least-squares (LS) algorithm minimizes an objective function that measures how closely projected 3D point features fall on their corresponding 2D image features in terms of squared residual error. It is well known that LS optimization techniques can fail catastrophically when outliers (called "blunders" in the photogrammetry literature) are present in the data (for example, see [Kumar 1994]). For this reason, the basic LS resection algorithm is embedded inside a least median squares (LMS) procedure that repeatedly samples subsets of correspondences to find one devoid of outliers by minimizing the median-squared residual distance error. LMS is robust over data sets containing up to 50% outliers. The combination of Levenberg-Marquardt with least median squares techniques provides a powerful and robust approach to camera resection.

Alternatively, given a pair of images for which camera pose is known, and a set of points extracted from each image, an algorithm has been developed to automatically establish image point correspondences and simultaneously triangulate their corresponding 3D point positions [Cheng et. al. 1994]. Each point may or may not have a correspondence in the other image; potential correspondences are detected using an "affinity" measure devised by [Scott and Longuet-Higgens 1991]. Ambiguities are resolved by transforming the correspondence problem into a network flow problem over a bipartite graph, and solving for the maximal flow. This approach is currently being extended to handle N-images, and to work for line features as well as points.

2. Learning in Vision

2.1. Project Goal

The goal of this project is the construction of a system for learning object-specific recognition strategies from a library of image understanding algorithms and training images [Draper et. al. 1993a]. Such a system will serve as an automatic integration system, combining IU algorithms (and representations) that were developed separately into executable object recognition systems, with only limited human interaction for locating the target objects in the training images. It will also serve as a generator of task-specific vision systems for applications ranging from the Unmanned Ground Vehicle to Radius to flexible manufacturing environments.

The Schema Learning System (SLS) involves a paradigm of learning with limited user interaction - just the interactive specification of goals via an example training set - and a library of vision routines (possibly a large collection). The vision routines can be as simple as a low-level operator, more complex algorithms, or an entire complex subsystem as a module. The key is that each has a specification of its input and its output, and therefore the learning algorithm (or compiler) knows how they can be linked across the different levels of representation to achieve the desired goal given the specified input. We believe SLS is break-through technology in allowing the many mini-theories, and the multitude of good vision algorithms that have been developed across the field, to be put together into real operational specialized vision systems - without the system builder (i.e. user) having to laboriously construct it by hand. While there is a computationally expensive off-line learning and compilation phase, the result is an efficient run-time system.

Work on the early stages of the schema learning project has focused on creating a representative library of image understanding algorithms, and on developing an object-oriented version of SLS that will allow us to compare and contrast different machine learning technologies in the context of learning recognition strategies.

2.2. Developing an IU Library

The IU algorithm library currently consists of procedures from three sources: KBVision tasks

[Williams 1990], Khoros modules [Rasure and Kubica 1994], and algorithms developed locally at UMass. Integrating algorithms from these three sources into a single library required creating a data exchange mechanism that would allow data created by one algorithm to be read by any other. This was accomplished by extending ISR3 [Draper et. al. 1994a,b] to read and write KBVision's *im* and *tks* formats and Khoros' *viff* format (in addition to its own native data formats) and then using ISR3 to convert data from one format to another as needed. Currently, our library consists of approximately 40 KBV tasks, 30 Khoros tasks, and we are in the process of adding another 20-30 UMass modules. Major experimental tests will be run using this set of approximately 100 algorithms.

We hope that eventually the library will include algorithms from across the IU community. In the future, other possible sources of algorithms and modules are the IUE, RCDE, etc.

2.3. Redesigning SLS

The original Schema Learning System (SLS) learned object-specific recognition strategies by a combination of logic-based symbolic inference (in the propositional calculus) and decision trees [Draper 1993b]. However, other machine learning technologies are available that might lead to improved object recognition schemas.

The object-oriented redesign of SLS is aimed at allowing an investigator to "mix and match" inference and classification algorithms, and to run controlled comparisons between them. The aim is not only to produce a system that learns useful recognition strategies, but to understand which learning algorithms are most effective for learning visual control strategies and why. In particular, we are examining alternatives to the symbolic inference algorithm in SLS in terms of an explanation-based learning or reinforcement learning approach. Similarly, the decision tree classifiers are being extended with either backpropagation neural networks or instance-based classifiers. Recent work by Brodley [Brodley 1994] suggests that combinations of these classifiers may lead to still better results for certain subproblems. Thus alternative learning methodologies and strategies should improve the power of SLS.

2.4. Object Classification from Multiple Features

Of particular interest has been the use of non-parametric pixel-level classifiers as focus-of-attention (FOA) mechanisms in outdoor scenes. FOA mechanisms allow a system to focus its resources on critical portions of the data, and are an important part of UMass's contribution to the CSU/Alliant/UMass RSTA effort [Beveridge 1994]. For the same reason, they are an important tool for the Schema Learning System to have at its disposal if it is to automatically learn efficient object recognition strategies.

To be effective in an outdoor context, a pixel-level classifier must be able to compensate for natural changes in outdoor illumination, something many traditional classifiers are not good at. We have been working with linear machine decision trees (LMDTs) for classifying pixels from their RGB values in outdoor images; the piecewise-linear functions in color space learned by the LMDTs approximate the change in apparent color of the target under natural lighting conditions. In our most challenging test of this technology, we are using LMDTs to identify the color of camouflage (on vehicles) in the publicly available Ft. Carson data set [Buluswar and Draper 1994]; preliminary results are extremely promising. In addition, we have developed algorithms for training LMDTs to optimize utility functions that prefer false positive responses (alerting the system to the presence of an object that isn't there) to false negative ones (failing to notice an object), thus making them more suitable for use as FOA mechanisms [Draper et. al. 1994c]

2.5. ISR3

During earlier research on knowledge-based approaches to image understanding, we designed and constructed an image token database, called the Intermediate Symbolic Representation (ISR), which supports integration of vision algorithms by providing user-definable data structures and common access routines. The ISR was optimized for common processing tasks in mid- and high-level vision, such as defining symbolic token types, adding and removing tokens, manipulating sets of tokens, and retrieving token sets associatively by feature value and spatial location. Recently the ISR was redesigned to take into account the constraints imposed by real-time systems such as those found in autonomous navigation and flexible manufacturing. This new version, ISR3, differs from the previous ISR in that it stores native C data structures, is memory-based rather than file-based, and provides primitives for memory management and multiprocess synchronization [Draper et. al.

1994b]. ISR3 is implemented in C++ and is also being ported to the RCDE.

3. Mobile Vehicle Navigation for UGV

3.1. Project Goal

In the past, the focus of the UMass ARPA-sponsored UGV effort has been on vehicle navigation using passive imaging techniques. In support of this effort, UMass built the Mobile Perception Laboratory using an Army HMMWV as the basic chassis. The research effort produced an integrated software system for vehicle navigation, including subsystems for driving control, road following (CMU's ALVINN system), obstacle detection, reflexive obstacle avoidance, vehicle positioning using visually prominent landmarks, vehicle servoing for cross country navigation (compass and landmark), real-time local path planning, etc. These subsystems were integrated in a behavior-based manner through a finite machine controller executing a descriptive behavior language [Rochwerger et. al. 1994]. Near real time performance was achieved; the vehicle was capable of moving approximately 6-10 mph. This effort terminated with the delivery of the MPL and associated software to TACOM in April 1994. Since then, we have been developing the concept and initial implementation of a system for stealth navigation in support of UGV RSTA missions as described below.

3.2. Stereo for Navigation

Obstacle detection is an important safety issue in mobile robotics, and must be done quickly and efficiently. Three different algorithms for obstacle detection based on stereo imagery have been developed, each based on a different set of assumptions concerning the amount information available. The first two algorithms were designed for yes/no responses without indicating which points are obstacles. They have the advantage of fast determination of the existence of obstacles based on the solvability of a linear system of equations. The first algorithm uses information about the ground plane, while the second algorithm only assumes that the ground is planar. The third algorithm continuously estimates the ground plane, and based on that determines the height of each matched point in the scene. The algorithms and experimental results are presented in [Zhang et. al. 1994] for real and simulated data, and the performance of the three algorithms under different noise levels are compared in simulation.

3.3. Reflexive Obstacle Avoidance

A practical real-time system (~2Hz for 256 x 240 images) for passive obstacle detection and avoidance has been developed and tested on the UMass MPL. Obstacle points are obtained from stereo imagery under the assumption that the ground plane is locally planar and that the cameras are calibrated. The local disparity map is thresholded against the expected disparity of the ground plane and points more than k feet above the ground are assumed to represent obstacles. These points are projected onto the ground plane, creating an instantaneous obstacle map (IOM). The algorithm modulates the steering and speed of the vehicle, which is assumed to be under the control of another process, such as the ALVINN road follower [Pomerleau 1992]. The IOM is transformed into a steering command to modify the direction and speed of the vehicle in a reflexive manner.

3.4. Constructing Terrain Models for Stealth Planning

A new correlation-based terrain reconstruction system has been developed that is capable of producing terrain maps that have much higher resolution than 30 meter DTED [Schultz 1994]. It uses multiple images taken from different perspective viewpoints and produces highly accurate reconstructions under a wide variety of viewing conditions, including camera base-to-height viewing ratios greater than two-to-one, and camera vergence angles in excess of 90 degrees. The system utilizes two or more images along with known camera acquisition parameters to create a rectified epipolar image. Multi-image stereo correlation matching is then performed using a novel set of weighted correlation masks that produce direct, subpixel disparity estimates, and are robust to image intensity distortions caused by large differences in perspective viewpoint. The output of the system is a dense array of elevations in rectilinear ground coordinates, and a registered orthonormal intensity image.

The unique features of this system are the implementation of weighted correlation masks that produce match scores that are less sensitive to perspective distortion and random pixel noise, and subpixel image shift operators that enable subpixel disparity estimates that are far more accurate than previous techniques. While the system produces very accurate terrain reconstructions, it may require up to an order of magnitude more convolution operations and storage elements than a traditional stereo correlation algorithm.

A simulated fly-through of the texture-mapped (i.e. rendered) terrain recovered from only two high-resolution aerial images of the Martin-Marietta UGV Demo B site is available on videotape. The algorithm for terrain reconstruction appears to be very accurate, discriminating about 1 to 2 foot height differences in the Martin-Marietta Demo B site images; ditches 3-5 deep feet are recovered, and cars in a parking lot stand out clearly. Current work involves quantifying the performance of the algorithm and determining its error characteristics.

3.5. Stealth Planning Using Reconstructed Terrain

Under a stealth navigation scenario, the goal is for a vehicle to navigate from a starting position to a final position while remaining hidden from a target. The criterion is to minimize the visibility of the vehicle from the target position while still allowing the vehicle to perform its RSTA functions. One situation that we are analyzing is where a hill separates the vehicle from the target, and a horizon line is visible behind the hill [Ravela et. al. 1994]. It is assumed that the initial position of the target is known approximately from a previous sighting or from aerial surveillance.

The path planning system is applied to the reconstructed digital elevation map (DEM) to find a path from the starting position to a position from which the target will be first visible. Visibility analysis based on an expected or determined enemy location will allow regions of visibility to be marked for avoidance. A path is then planned from the start to end position within the stealth region using a very fast non-holonomic path planner based on harmonic functions [Connolly 1992]. The end of the path is located at a terrain position that allows clear target acquisition.

Such a path planning system was demonstrated for vehicle concealment and a variety of other RSTA/Stealth navigation scenarios and behaviors at UGV Demo B on a Sparc workstation. It allows a user to specify a vehicle starting and ending position and the location of the target vehicle on the DEM recovered using the terrain reconstruction system.

3.6. Stealth Behaviors for Navigation

The last component of the stealth navigation scenario is the development of low-level navigational control algorithms to provide stealth capabilities during critical periods of a RSTA mission. In order to perform this scenario, the scout vehicle must be able to: 1) identify distinctive terrain features, including but not limited to horizon lines (e.g. ridge lines) near the target area, 2) re-identify these same features from new viewpoints, and 3) control the vehicle's speed, orientation, and precise positioning as the vehicle approaches the observation point where exposure must be minimized.

Several scenarios for stealth navigation are being developed, but those of most interest involve prior views of the terrain around the target area, either from the terrain model reconstructed from the aerial views, or from previous RSTA activities. Somewhat prior to the vehicle becoming visible, low-level servoing procedures would control vehicle mobility to minimize vehicle exposure from the target location. For example, one of the best terrain features for stealth navigation are distinctive horizon lines. These have the advantage of being higher than the enemy's position, and so they come into view before the vehicle or its sensor platform are exposed. When the vehicle is climbing a hill, it would slow to a crawl as terrain features known to be above the target area come into view, servoing on the relative angles of the target and horizon line. It would come to a stop exactly at the point at which its sensor platform can view the enemy, but the rest of the vehicle remains hidden behind the hill.

These behaviors utilize curve matching and tracking. Matching curves is based on the algorithm of [Kalvin et. al. 86] and is invariant to translation and rotation in the plane. Curve tracking is based on normalized cross-correlation and steerable filters.

3.7. Integrated Color CCD, FLIR, and LADAR Based Object Modeling and Recognition

This is a joint effort between laboratories at Colorado State University, the University of Massachusetts, and Alliant Techsystems, Inc. [Beveridge et. al. 1994] The goal of this program is the development of advanced automatic target recognition systems which fuses data from multiple visual modalities (color CCD, FLIR, and LADAR sensors). The approach is based on target hypothesis generation using features drawn from the three sensors, followed by a target verification stage by matching the hypothesized models to the sensor data. Target hypothesis generation is designed to solve the problem of indexing into large databases by using a quick indexing scheme to select potential target models. Indexing is based on both spatial, chromatic, and relational features extracted from the image. The verification stage employs a state of the art geometric model matching system which matches model features to image features in order to determine the optimal correspondence between the model and data.

4. Recovery of 3D Structure

4.1. 3D Reconstruction from Profiles

Giblin and Weiss [1994] have some new theoretical results on modeling curved surfaces from their occluding contours (profiles) in a sequence of images with known camera motion. The reconstruction algorithm was introduced by [Giblin and Weiss 87] for a simple class of motions and was generalized in (e.g. see [Cipolla and Blake 92]) to arbitrary motion. For the general motion case, epipolar correspondence played an important role in matching points from one profile to the next. This approach leads naturally to a representation of the surface by parametric patches based on epipolar curves, which makes explicit the geometry of the surface.

In general, the reconstruction of surfaces from profiles leaves gaps. These gaps are bounded by two types of curves. The first type, which we have analyzed, consists of 'frontier points' where the epipolar plane is tangent to the surface. This is one of the cases where the epipolar parameterization breaks down, and the epipolar curves on the surface are singular. The structure of the epipolar curves and contour generators at these points on the surface can be completely determined. The primary tool is the spatio-temporal surface which can also be used to facilitate the reconstruction. In addition, there may be gaps that are due to occlusion and the boundaries of these regions occur at T-junctions.

4.2. An Approach to 3D Reconstruction from Motion

J. Oliensis [1994] has developed a new approach to the problem of multiframe structure from motion, which essentially generalizes the earlier linear algorithm of Tomasi [1992] from orthographic to full perspective. It operates on tracked feature points and yields fast, accurate reconstruction algorithms which can be applied either in batch or recursive mode. These algorithms also can cope with the problem of failed tracking---tracked points missing from some images. They are well suited to image sequences such as are typically obtained in autonomous vehicle navigation. In addition, this work generalizes from the work of Heeger and Jepson [1992] on recovering translational motion from optical flow; it gives an effective method for determining the translation from sparse as well as dense optical flow. The approach has been tested on two real image sequences: the Rocket Field sequence [Dutta 1989], which involves straight-ahead translation and significant perspective effects, and the difficult PUMA sequence [Sawhney 1990]. Fast and accurate reconstructions were obtained for both these sequences on a DecStation 5000 using MATLAB. The approach is based on finding suitable approximations of the problem, leading to linear methods for recovering structure and motion that closely approximate maximum likelihood estimates.

4.3 Measuring Affine Transformation Using Gaussian Filters

Image deformations due to relative motion between an observer and an object may be used to infer 3-D structure. Up to first order these deformations can be written in terms of an affine transform. A new technique for measuring affine transforms from image patches, which correctly handles the problem of corresponding deformed patches, has been developed by Manmatha [1994]. No correspondence is required. Image patches are filtered using gaussians and derivatives of gaussians and the filters deformed according to the affine transform. The problem of finding the affine transform is therefore reduced to that of finding the appropriate deformed filter to use. The method is local, can handle larger affine deformations than previous techniques and requires fewer iterations to obtain the correct answer. The technique can be extended to points and lines and provides a uniform technique for measuring affine transforms using points, lines or image patches. In related work, Manmatha and Sawhney are examining the use of gaussian and derivative of gaussian filters to detect local symmetries in images.

5. The Image Understanding Architecture

5.1 TRP Award for Commercial Development of the IUA

A Technology Reinvestment Project award to the IUA development team of Amerinex Artificial Intelligence Inc., Hughes Research Laboratories, and UMass (through ACSIOM, the UMass technology transfer arm) will result in the development of two new generations of the IUA hardware and software over the next three years.

These systems will form a commercial off-the-shelf family of products for use in both civilian and military applications as embedded and stand-alone systems.

5.2 Status of IUA

The development of the prototype of the second generation IUA and the third generation proposed in the TRP is reported in [Weems et. al. 1994a].

5.3 UMass Research Effort

Research in parallel architectures for vision has continued in three areas, the IU Benchmark, a parallel vision architecture design and evaluation environment, and software support for heterogeneous parallel systems.

ARPA Parallel IU Benchmark

As recommended by the DARPA IU Benchmark Workshop participants [Weems et. al. 1988, 1991] we have developed a preliminary second level benchmark, which incorporates tracking of moving objects over a sequence of images.

In order to reuse as much of the static benchmark as possible, the new benchmark operates on the same type of scenes -- a 2.5 dimensional mobile of rectangles with chaff, but in the new benchmark, the mobile and chaff are blown by an idealized wind to produce predictable motion. The goal of the new benchmark is to test system performance over a longer period of time so that, for example, caches and page tables will be filled. The benchmark also explores I/O and real-time capabilities of the systems under test, and involves more high-level processing.

Parallel Architecture Evaluation

UMass has developed a system for capturing traces of programs written in the C++ class library as they execute on an abstract parallel machine. The traces are then fed to a simulation system that models hardware architectures with different features and parameters. The system, called ENPASSANT (Environment for Parallel System Simulation and Analysis Tools), allows us to gather real performance data for different architectural configurations, and to analyze the data statistically. It uses a new methodology, called trace compilation, to reduce the time and space required by typical trace-based analysis systems by a factor of 20 to 30. This enables us to generate performance data for a much wider range of architectural configurations. The performance data can then be contrasted with cost estimates for the different configurations to produce a specification for parallel vision architectures for different applications [Herbordt and Weems 1993].

Compilation Framework for Heterogeneous Parallel Software

We have developed a concept and initial design for a flexible compilation framework that will be able to manage multiple source programming languages, software module implementations, and target machine architectures in order to construct a working and optimized executable on a heterogeneous parallel system. The key element of the framework is a combination of a library of translation components (transformations, analyzers, optimizers, etc.) and a compilation director that plans the application of the components for each compilation. Development has begun on a multi-lingual front end that allows the user to annotate code with hints to be used by the compilation director [Weems et. al. 1994b]

References

Beveridge, J., Hanson, A., and Panda, D., "RSTA Research of the Colorado State, University of Massachusetts and Alliant Techsystems Team," in these proceedings.

Brodley, C., "Recursive Automatic Bias Selection for Classifier Construction,", submitted to Machine Learning.

Buluswar, S., Draper, B., Hanson, A., and Riseman, E. "Non-parametric Classification of Pixels under Varying Outdoor Illumination," in these proceedings.

Cheng, Y. Q., Collins, R., Hanson, A., and Riseman, E., "Triangulation Without Correspondences (I)", in these proceedings.

Cipolla, R. and A. Blake, "Surface shape from the deformation of apparent contours," IJCV 9 1992, 83-112.

Collins, R., Hanson, A., and Riseman, E., "Site Model Acquisition Under the UMass RADIUS Project", in these proceedings.

Collins, R. and Weiss, R. "Vanishing Point Calculation as a Statistical Inference on the Unit Sphere," Proc. ICCV, Osaka, Japan, December 1990, 400-403.

Connolly, C., "Applications of Harmonic Functions to Robotics," International Symposium on Intelligent Control, Glasgow, Scotland, August 1992.

Draper, B., Hanson, A., and Riseman, E,. "Integrating Visual Procedures for Mobile Perception", in Experimental Environments for Computer Vision and Image Processing (H. I.

Christensen and J. L. Crowley, eds.), World Scientific, 1994a, 183-193.

Draper, B., Kutlu, G., Hanson, A., and Riseman, E., "ISR3: Communication and Data Storage for an Unmanned Ground Vehicle", Proc. ICPR, Jerusalem, Israel, October 1994b.

Draper, B., Brodley, C., and Utgoff, P., "Goal-Directed Classification Using Linear Machine Decision Trees," T-PAMI, Vol. 16(9) to appear, 1994c.

Draper, B., Hanson, A., and Riseman, E., "Learning Blackboard-Based Scheduling Algorithms for Computer Vision," IJPRAI, Vol. 7, No. 2, 309-328, 1993a.

Draper, B. "Learning Object Recognition Strategies," Ph.D. Thesis, University of Massachusetts, TR93-50, 1993b.

Dutta, R., Manmatha, R., Williams, L., and Riseman, E., "A data set for quantitative motion analysis," Proc. CVPR, San Diego, CA, June 4-8, 1989, 159-164.

Giblin, P., and Weiss, R., "Epipolar Curves on Surfaces", in these proceedings.

Giblin, P. J. and R. S. Weiss, "Reconstruction of surface from profiles", Proc. ICCV, London, 1987, 136--144.

Heeger, D. J. and A. D. Jepson, "Subspace methods for recovering rigid motion I: Algorithm and implementation, Proc. IJCV 7, 1992, 95-117.

Herbordt, M. and C. Weems, "An Environment for Evaluating Architectures for Spatially Mapped Computation: System Architecture and Initial Results," Proc. Computer Architectures for Machine Perception (CAMP '93), New Orleans, LA, December 1993, 191-201.

Jaynes, C., Stolle, F., and R. Collins, "Task Driven Perceptual Organization for Extraction of Rooftop Polygons," in these proceedings.

Kalvin, A., Schonberg, E., Schwartz, J., and Sharir, M. "Two-Dimensional, Model-Based, Boundary Matching Using Footprints," International Journal of Robotics Research, 5(4):38--55.

Kumar, R., and A. Hanson, "Robust Methods For Estimating Pose and a Sensitivity Analysis", Computer Vision, Graphics, and Image Processing--Image Understanding, to appear, 1994.

Manmatha, R., "A Framework for Recovering Affine Transforms Using Points, Lines or Image Brightness", Proc. CVPR, Seattle, WA, June 1994, 141-146.

Oliensis, J., "A Linear Solution for Multiframe Structure From Motion", in these proceedings.

Pomerleau, D. A., Gowdy, J. and C. E. Thorpe, "Combining Artificial Neural Networks and Symbolic Processing for Autonomous Robot Guidance," Proc. ARPA IUW, San Diego, CA, Jan. 1992, 961-967.

Rasure, J. and S. Kubica, "The KHOROS Application Development Environment," in Experimental Environments for Computer Vision and Image Processing (H. I. Christensen and J. L. Crowley, eds.), World Scientific, 1994, 1-32.

Ravela, C., Weiss, R., Draper, B., Hanson, A., and Riseman, E., "Stealth Navigation: Planning and Behaviors", in these proceedings.

Rochwerger, B., Fennema, C., Draper, B., Hanson, A., and Riseman, E., "An Architecture for Reactive Behavior", Proc. ICPR, Jerusalem, Israel, October 1994, to appear.

Sawhney, H. S., Oliensis, J. and Hanson, A., "Description and Reconstruction from Image Trajectories of Rotational Motion", Proc. ICCV, Osaka, Japan, December, 1990, 494-498.

Schultz, H. "Terrain Reconstruction from Oblique Views", in these proceedings.

Scott, G. and Longuet-Higgins, H. "An Algorithm for Associating the Features of Two Patterns," Proc. Roy. Soc. London, Vol. B244, 1991, 21-26.

Tomasi, C. and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," Proc. IJCV 9, 1992, 137-154.

Weems, C., Williams, T., and D. Schwartz, "The Next Generation Image Understanding Architecture", in these proceedings, 1994a.

Weems, C., Weaver, G., and Dropsho, S., "Linguistic Support for Heterogeneous Parallel Processing: A Survey and an Approach," Proc. IEEE Heterogeneous Computing Workshop, Cancun, Mexico, April 26, 1994b.

Weems, C., Riseman, E., Hanson, A., and A. Rosenfeld, "The ARPA Image Understanding Benchmark for Parallel Computers," Journal of Parallel and Distributed Computing, 11, 1991, 1-24.

Weems, C., Riseman, E., Hanson, A., and A. Rosenfeld, "IU Parallel Processing Benchmark," Proc. CVPR, Ann Arbor, MI, June, 1988, 673-689.

Williams, T. "Image Understanding Tools,", Proc. ICPR 10, Atlantic City, N.J., June 1990, 106-610.

Zhang, Z., Weiss, R., and Hanson, A., "Obstacle Detection Based on Partial 3D Reconstruction," in these proceedings.