

①

VISION IN DYNAMIC ENVIRONMENTS

Contract DAAB07-86-K-F073

Final Technical Report

AD-A213 434

Sponsoring Agency: Defense Advanced Research Projects Agency
Information Sciences Technology Office
1400 Wilson Blvd.
Arlington, VA 22209-2308

Monitoring Agency: U.S. Army Center for Night Vision and Electro-Optics
Attn: DELNV-AC/G (G. Jones)
Fort Belvoir, VA 22060

Contractor: Computer Vision Laboratory
Center for Automation Research
University of Maryland
College Park, MD 20742-3411

Principal investigators: John (Yiannis) Aloimonos
Larry S. Davis
Azriel Rosenfeld

Period covered: June 30, 1986 – June 30, 1989

August 15, 1989

VISION IN DYNAMIC ENVIRONMENTS

Contract DAAB07-86-K-F073

Final Technical Report

Sponsoring Agency: Defense Advanced Research Projects Agency
Information Sciences Technology Office
1400 Wilson Blvd.
Arlington, VA 22209-2308

Monitoring Agency: U.S. Army Center for Night Vision and Electro-Optics
Attn: DELNV-AC/G (G. Jones)
Fort Belvoir, VA 22060

Contractor: Computer Vision Laboratory
Center for Automation Research
University of Maryland
College Park, MD 20742-3411

Principal investigators: John (Yiannis) Aloimonos
Larry S. Davis
Azriel Rosenfeld

Period covered: June 30, 1986 – June 30, 1989

August 15, 1989

A-1

A-1

23 1963

SECRET

Research conducted on the contract was primarily concerned with real-time, three-dimensional computer vision and image understanding. The results of the research were documented in 62 Technical Reports. This Final Technical Report consists of the abstracts of the earlier reports.

- [1] Ken-ichi Kanatani and Tsai-Chia Chou, "Tracing Finite Motions Without Correspondence." CAR-TR-211, CS-TR-1689, August 1986.

ABSTRACT: The 3D motion of an object having planar faces is traced, starting from a known position, from a sequence of 2D perspective projection images *without using any knowledge of point-to-point correspondence*. Computation is based on the shape of the region on the image plane corresponding to a planar face of the object. Given two images, a heuristic guess about the motion is first computed, and one image is transformed according to this estimated motion so that it is positioned close to the other image. Then, the motion that accounts for the remaining small discrepancy is estimated by measuring numerical *features* of the planar regions. The scheme is based on the optical flow due to infinitesimal motion, and estimation is done by solving a set of simultaneous *linear* equations. This process is iterated; after each estimate of the motion, one image is transformed according to the estimated motion so that it is positioned closer and closer to the target image. Various practical issues such as choice of features, constrained motions, face identification, and computation of features without actually transforming the images are discussed. Some numerical examples are also given.

- [2] Ken-ichi Kanatani, "Group Theoretical Methods in Image Understanding." CAR-TR-214, CS-TR-1692, August 1986.

ABSTRACT: This work is a brief summary of mathematics, in particular group representation theory, relevant to the study of image understanding and computer vision. We introduce fundamentals of group representation theory, theory of invariants and theory of Lie groups and Lie algebra, especially representations and invariants of the 2D and 3D groups of rotations $SO(2)$, $SO(3)$, in relation to image understanding and computer vision.

- [3] John (Yiannis) Aloimonos and Azriel Rosenfeld, "Monocular Stereopsis: Theory and Applications." CAR-TR-218, CS-TR-1698, August 1986.

ABSTRACT: A theory of monocular depth perception is presented. A moving cyclopean observer uses motion information to recover the depth of an object. The problem is studied for both orthographic and perspective projection and closed form solutions for the absolute depth functions are developed. Finally, an application of the theory to a vibrating camera is presented. In particular, our results are:

- (1) A moving cyclopean observer that does not know his motion can recover the absolute depth of an object from its shape and the induced optic flow field. Under orthography, a closed form solution for the absolute depth

function is given.

- (2) A moving cyclopean observer that knows his motion can recover the absolute depths of objects using only the spatiotemporal derivatives of the image intensity function. This result gives rise to useful applications, for example, the recovery of depth from a vibrating camera.

- [4] Muralidhara Subbarao, "Interpretation of Visual Motion: A Computational Study." CAR-TR-221, CS-TR-1706, September 1986.

ABSTRACT: A changing scene produces a changing image or *visual motion* on the eye's retina. The human visual system is able to recover useful three-dimensional information about the scene from this two-dimensional visual motion. This report is a study of this phenomenon from an information processing point of view. A *computational theory* is formulated for recovering the scene from visual motion. This formulation deals with determining the local geometry and the rigid body motion of surfaces from spatio-temporal parameters of visual motion. In particular, we provide solutions to the problem of determining the shape and rigid motion of planar and curved surfaces and characterize the conditions under which these solutions are unique. The formulation is generalized to the case of non-uniform (i.e. accelerated) and non-rigid motion of surfaces. This serves to address the two fundamental questions: *What scene information is contained in the visual motion field? How can it be recovered from visual motion?* The theory exposes the well known fact that the general problem of visual motion interpretation is inherently ill-posed. Furthermore, it indicates the minimum number of additional constraints (in the form of assumptions about the scene) necessary to interpret visual motion. It is found that, in general, the assumption that objects in the scene are rigid is sufficient to recover the scene uniquely.

A *computational approach* is given for the interpretation of visual motion. An important characteristic of this approach is a uniform representation scheme and a unified algorithm which is both flexible and extensible. This approach is implemented on a computer system and demonstrated on a variety of cases. It provides a basis for further investigations into both understanding human vision, and building machine vision systems.

- [5] Isaac Weiss, "3-D Shape Representation by Contours." CAR-TR-222, CS-TR-1707, September 1986.

ABSTRACT: The question of 3-D shape representation is studied on the fundamental and general level. The two aspects of the problem, (i) the reconstruction of a 3-D shape from a given set of contours, and (ii) finding "natural" coordinates on a given surface, are treated by the same theory. We first state a few basic principles that should guide any shape reconstruction mechanism, regardless of its physical implementation. Second, we propose a new mathematical procedure that complies with these principles and offers several advantages over the existing *ad hoc* treatments. Some general results are derived from this procedure, which conform very well with human visual perception.

- [6] David Harwood, Susan Chang and Larry S. Davis, "Interpreting Aerial Photographs by Segmentation and Search." CAR-TR-223, CS-TR-1709, September 1986.

ABSTRACT: A knowledge-based system for interpreting aerial photographs. Picture Query (PQ), first segments an image into primitive, homogeneous regions, then searches among combinations of these to find instances which satisfy definitions of object types. If primary evidence is insufficient, there may be an hypothesis-based search for the supporting evidence of related objects. This secondary search is restricted to windows by expected spatial relations. First instances are improved by searching for overlapping variants having better goodness-of-figure. The process may be repeated using re-estimated parameters of object definitions based on instances found previously. Results are reported for images of suburban neighborhoods, including roads, houses, and their shadows.

- [7] Isaac Weiss, "Curve Fitting with Optimized Mesh Point Placement." CAR-TR-224, CS-TR-1710, September 1986.

ABSTRACT: A recent theory of 3-D surface interpolation has been implemented numerically for the special case of curve fitting, in the plane and in 3-D space. The implementation demonstrates some of the major advantages of the theory, such as the ability of the curve to turn sharp corners, by concentrating mesh points (knots) in high curvature parts of the curve. This knot placement, which is a universal difficulty in previous interpolation or smoothing techniques, is done here automatically by the same principle of energy minimization that is used to fit the curve to the data.

- [8] Isaac Weiss, "Straight Line Fitting in a Noisy Image." CAR-TR-234, CS-TR-1727, November 1986.

ABSTRACT: The conventional least squared distance method of fitting a line to a set of data points is notoriously unreliable when the amount of random noise in the input (such as an image) is significant compared with the amount of data correlated to the line itself. Points which are far away from the line are usually just noise, but they contribute the most to the distance averaging, skewing the line from its correct position. We present a statistical method of separating the data of interest from random noise, based on a maximum likelihood principle.

- [9] Behrooz Kamgar-Parsi, "An Efficient Line Search Algorithm for Optimization of Multivariate Functions." CAR-TR-239, CS-TR-1738, November 1986.

ABSTRACT: An efficient line search algorithm for the minimization of multivariate functions is presented. The main element of this algorithm is a new interpolation technique. This interpolation technique is a cubic interpolation which, in addition to the three pieces of information used for a quadratic interpolation, employs the function value in the middle of the interval of interest. In a program based on a quasi-Newton method with the BFGS update formula, we tested the new line search routine against a line search routine that utilizes standard quadratic and cubic interpolation techniques. We used ten standard small-to-medium-size test functions ($2 \leq N \leq 30$). The results indicates that, on the

average, when using the new line search routine, the minimization of a function achieved during every two iterations *that require a line search* is equivalent to the minimization of the function which is achieved during three iterations when using a line search routine that is based on standard interpolation techniques. (This suggests that the new line search routine is 33% more efficient.) As regards the general impact of the new interpolation technique on the program, we found the following; reduction in the number of iterations by almost 10%; reduction in the number of gradient evaluations by over 8%; and although the new interpolation technique requires an extra function evaluation, the number of function evaluations, too, showed a decline (albeit small: 2-3%).

- [10] Behrooz Kamgar-Parsi and Roger Eastman, "Calibration of a Stereo System With Small Relative Angles," CAR-TR-240, CS-TR-1739, November 1986.

ABSTRACT: Practical difficulties in the calibration of a two camera stereo system in an uncontrolled environment are studied for the case where the relative orientation angles are small and the distance between the two cameras is known. This is done by deriving explicit analytical solutions for the relative pan, tilt and roll angles in terms of the world pan angle (often referred to as gaze angle) and the coordinates of the image points used in their computation. These solutions allow us a better understanding of the intricacies of the problem of calibration in general. The purpose of this work has been twofold, both practical and theoretical. Its practical purpose is to provide us with a reliable method for the computation of camera orientations when the relative rotation angles are small. Its theoretical purpose is to provide us with insight as to how errors due to quantization and uncertainty in the image center location can affect the computation of rotation angles, so that we can look for ways to minimize their impact. (These findings are likely to be of use even when the relative rotation angles are not small.) In particular it is shown that the sensitivity of the computation of the relative pan and roll angles to the above sources of error greatly depends on the choice of image points used for the computation of these angles, whereas the sensitivity of the computation of the relative tilt angle to the error due to image center position is only marginally affected by our choice of the image points. All of the analytical findings have been supported by extensive simulation.

- [11] John (Yiannis) Aloimonos and Behrooz Kamgar-Parsi, "Correspondence from Correspondence," CAR-TR-260, CS-TR-1769, January 1987.

ABSTRACT: The problem of image matching is investigated from a theoretical point of view. We study the problem of computation of visual correspondence, given that we already know some values of the correspondence function. We study the mathematical constraints that will enable us to grow a solution for the correspondence function from a point where its value is known, using the image intensity function. The results are applicable to many image matching problems, such as stereo image interpretation, object analysis, motion analysis, change detection and the like.

- [12] John (Yiannis) Aloimonos and Anargyros Papageorgiou, "On the Kinetic Depth Effect: Lower Bounds, Regularization and Learning." CAR-TR-261, CS-TR-1770, January 1987.

ABSTRACT: The problem of the kinetic depth effect is revisited. We investigate how many points in how many views are necessary and sufficient to recover structure. The constraints in the cases where the velocities of the image points are known, and the positions of the image points are known with the correspondence between them established, are different and have to be studied separately. In the case of two projections of any number of points there are infinitely many solutions but if we regularize the problem we get a unique solution under some assumptions. Finally, an algorithm is discussed for learning this particular kind of regularization.

- [13] John (Yiannis) Aloimonos, "Combining Shading and Motion to Compute Shape and Light Source Direction." CAR-TR-262, CS-TR-1771, October 1986.

ABSTRACT: Most of the basic problems in computer vision, as formulated, admit infinitely many solutions. An example of this is the shape from shading problem. But vision is full of redundancy and there are several sources of information that if combined can provide unique solutions for a problem. In this paper, we combine shading and motion to uniquely recover the light source direction and the shape of the object in view.

- (1) We develop a constraint among retinal motion displacements, local shape, and the direction of the light source. It is worth noting that this constraint does not involve the albedo of the imaged surface. This constraint is of importance in its own right, and can be used in related research on computer or human vision.
- (2) We develop a constraint between retinal displacements and local shape. Again, this constraint is important on its own, and it lies at the heart of the algorithms presented in this paper.
- (3) We present algorithms for the unique computation of the lighting direction and the shape of the object in view.
- (4) We present experimental results, using synthetic images, that test the theory.

- [14] Minas E. Spetsakis and John (Yiannis) Aloimonos, "Closed Form Solution to the Structure from Motion Problem from Line Correspondences." CAR-TR-274, CS-TR-1798, March 1987, Revised February 1988.

ABSTRACT: A theory is presented for the computation of three dimensional motion and structure from dynamic imagery, using only line correspondences. The traditional approach of corresponding microfeatures (interesting points—highlights, corners, high curvature points, etc.) is reviewed and its shortcomings are discussed. Then, a theory is presented that describes a closed form solution to the motion and structure determination problem from line correspondences in three views. The theory is compared with previous ones that are based on

nonlinear equations and iterative methods.

- [15] Tsai-Chia Chou, John (Yiannis) Aloimonos and Azriel Rosenfeld, "Correspondenceless Model Based and Active Perception of Shape from Contour." CAR-TR-275, CS-TR-1800, March 1987.

ABSTRACT: The problem of shape from contour is examined. In traditional passive perception approaches, this problem has infinitely many solutions; and special assumptions or ad hoc heuristics must be employed in order to reduce the space of solutions to a unique value. There is excellent research on this topic [6]. An alternative approach is to consider an active observer, i.e. an observer that moves in a known way or employs some a priori knowledge which will enable a unique computation of shape. The theory described here shows how to recover shape from contour by utilizing invariant properties of contours in different perspective projections. Correspondence of features among images is not used. In particular, the results are: 1) A monocular observer can uniquely determine from one view the shape of a planar surface which contains multiple contours (at least two), provided the 3D area and length ratio of contours are given. 2) A monocular observer can determine the shape of a planar contour from two views. 3) If the 3D area and length are given (model based applications), the shape of a planar contour can be determined from one view by a monocular observer. Finally, some experimental results are presented.

- [16] Minoru Asada, "Cylindrical Shape from Contour and Shading without Knowledge of Lighting Conditions or Surface Albedo." CAR-TR-276, CS-TR-1801, March 1987.

ABSTRACT: This paper presents an algorithm for reconstructing the shape of a cylindrical object from contour and shading without knowing the surface albedo of the object or the lighting conditions of the scene. The input image is segmented into spherical, cylindrical, or planar surfaces by analyzing local shading. The cylindrical surface is characterized by the direction of the generating lines, determined from spatial derivatives in the image. The brightest generating line has strong constraints on the shading analysis on the cylindrical surface and leads to a simplification of the equation which represents the relation between the contour shape and the shading. Although there remains one degree of freedom between the surface normal of the base plane and the slant angle of the generating line, we can uniquely recover the cylindrical shape from this solution (up to reflection). Experimental results for a synthetic image are shown.

- [17] Anup Basu and John (Yiannis) Aloimonos, "A Robust Algorithm for Determining the Translation of a Rigidly Moving Surface without Correspondence, for Robotics Applications." CAR-TR-279, CS-TR-1818, March 1987.

ABSTRACT: A method is presented for the recovery of the three-dimensional translation of a rigidly translating object. The novelty of the method consists of the fact that four cameras are used in order to avoid the solution of the correspondence problem. The method is immune to low levels of noise and has good behavior when the noise increases. The noise immunity is so high that even

though the algorithm is intended only for translating objects, its accuracy is very high even if the object is rotating (with a small rotation) as well.

- [18] John (Yiannis) Aloimonos and Michael Swain, "Shape from Patterns: Regularization." CAR-TR-283, CS-TR-1826, April 1987.

ABSTRACT: We present a theory for the recovery of the shape of a surface covered with small elements (texels). The theory is based on the apparent surface-pattern distortion in the image and fits the regularization paradigm, recently introduced in computer vision by Poggio et al. (1985). A mapping is defined based on the measurement of the local distortions of a repeated unknown texture pattern due to the image projection. This mapping maps an apparent shape on the image to a locus of possible surface orientations in gradient space. The analysis is done under an approximation of the perspective projection called paraperspective. The resulting algorithm is applied to several synthetic and real images to demonstrate its performance.

- [19] Randal C. Nelson and John (Yiannis) Aloimonos, "Finding Motion Parameters from Spherical Flow Fields (Or the Advantages of Having Eyes in the Back of your Head)." CAR-TR-287, CS-TR-1840, April 1987.

ABSTRACT: A theory is developed for determining the motion of an observer given the flow field over a full 360 degree image sphere. The method is based on the fact that the foci of expansion and contraction for an observer moving without rotation are 180 degrees opposed; and on the observation that if the flow field on the sphere is considered around three equators defining the three principal axes of rotation, then the effects of the three rotational motions decouple. The three rotational parameters can thus be determined independently by searching, in each case, for a rotational value for which the derotated equatorial flow field can be partitioned into disjoint 180 degree arcs of clockwise and counter-clockwise flow. The direction of translation is obtained as a by-product of this analysis. Since this search is two dimensional in the motion parameters, it can be performed relatively efficiently. Because information is correlated over large distances, the method can be considered a pattern recognition rather than a numerical algorithm. The algorithm is shown to be robust and relatively insensitive to noise and to missing data. Both theoretical and empirical studies of the error sensitivity are presented. The theoretical analysis shows that for white noise of bounded magnitude M , the expected error is at worst linearly proportional to M . Empirical tests demonstrate negligible error for perturbations of up to 20% in the input, and errors of less than 20% for perturbations of up to 200%.

- [20] Ramesh Sitaraman and Azriel Rosenfeld, "Probabilistic Analysis of Two Stage Matching." CAR-TR-294, CS-TR-1858, June 1987.

ABSTRACT: In this paper, we study two stage matching procedures as applied to labelled graphs and other domains relevant to computer vision. We do not require that the match be exact but only that it satisfy a specified error criterion. We show that it is computationally more efficient to initially match a subgraph and check the rest of the graph only when this match succeeds. A probabilistic

analysis of the expected cost of this procedure is given with the aim of determining the optimum subgraph size which minimizes this cost. The results are extended to graph matching with geometric constraints as well as to templates.

- [21] David Shulman and John (Yiannis) Aloimonos, "(Non)Rigid Motion Interpretation: A Regularized Approach." CAR-TR-295, CS-TR-1860, June 1987.

ABSTRACT: Determining 3-D motion from a time-varying 2-D image is an ill-posed problem: Unless we impose additional constraints, an infinite number of solutions is possible. The usual constraint is rigidity, but many naturally occurring motions are not rigid and not even piecewise rigid. A more general assumption is that the parameters (or some of the parameters) characterizing the motion are approximately (but not exactly) constant in any sufficiently small region of the image. If we know the shape of a surface we can uniquely recover the smoothest motion consistent with image data and the known structure of the object, through regularization [17],[18],[19]. This paper develops a general paradigm for the analysis of nonrigid motion. The variational condition we obtain includes maximizing isometry [6], rigidity [9], and planarity [2, 4] as special cases. If the variational condition is applied at multiple scales of resolution, it can be applied to turbulent motion [33]. Finally, it is worth noting that our theory does not require the computation of correspondence (optic flow or discrete displacements), and it is effective in the presence of motion discontinuities.

- [22] Tsai-Chia Chou and Ken-ichi Kanatani, "Recovering 3D Rigid Motion Without Correspondence." CAR-TR-297, CS-TR-1864, June 1987.

ABSTRACT: Given the perspective images of an object before and after a 3D rigid motion of finite magnitude, together with the depth information of the object before motion, a method is presented to recover the motion parameters without having to solve the *correspondence problem*. Let image *features* be defined as *functionals* over images containing outstanding points, line segments, or surface regions on the object. The infinitesimal variation of an image feature can be expressed as a *linear* constraint on the motion parameters. Thus, an infinitesimal motion can be estimated by solving a set of simultaneous linear equations. Hence, if an appropriate initial value is given, a finite motion can be recovered by *iteratively* applying above infinitesimal estimation; after each iteration, the image is transformed, according to the estimated motion, to get closer and closer to the target image. The appropriate initial values can be chosen *finitely* from a bounded parameter space which is obtained from the given images. Both synthetic data and real images are demonstrated in experiments.

- [23] Behzad Kamgar-Parsi and Behrooz Kamgar-Parsi, "A Nonparametric Method for Fitting a Straight Line to a Noisy Image." CAR-TR-315, CS-TR-1903, September 1987.

ABSTRACT: In fitting a straight line to a noisy image, the least squares method becomes highly unreliable either when the noise distribution is non-normal or when it is contaminated by outliers. We propose a nonparametric method, the Direct Linear Plot, to overcome these difficulties. This method is free of

assumptions about the noise distribution, and is insensitive to outliers. It is efficient and its implementation does not involve practical difficulties, such as local minima or poor convergence of iterative procedures.

- [24] Behrooz Kamgar-Parsi and Behzad Kamgar-Parsi, "Evaluation of Quantization Error in Computer Vision." CAR-TR-316, CS-TR-1904, September 1987.

ABSTRACT: Due to the important role that digitization error plays in the field of computer vision, a careful analysis of its impact on the computational approaches used in the field is necessary. In this paper we develop the mathematical tools for the computation of the average error due to quantization. They can be used in estimating the actual error occurring in the implementation of a method. Also derived is the analytic expression for the probability density of error distribution of a function of an arbitrarily large number of independently quantized variables. The probability of the error of the function to be within a given range can thus be obtained accurately. In analyzing the applicability of an approach one must determine whether the approach is capable of withstanding the quantization error. If not, then regardless of the accuracy with which the experiments are carried out the approach will yield unacceptable results. The tools developed here can be used in the analysis of the applicability of a given algorithm, hence revealing the intrinsic limitations of the approach.

- [25] John (Yiannis) Aloimonos, Isaac Weiss and Amit Bandyopadhyay, "Active Vision." CAR-TR-317, CS-TR-1905, August 1987.

ABSTRACT: We investigate several basic problems in vision under the assumption that the observer is active. An observer is called active when engaged in some kind of activity whose purpose is to control the geometric parameters of the sensory apparatus. The purpose of the activity is to manipulate the constraints underlying the observed phenomena in order to improve the quality of the perceptual results. For example a monocular observer that moves with a known or unknown motion or a binocular observer that can rotate his eyes and track environmental objects are just two examples of an observer that we call active. We prove that an active observer can solve basic vision problems in a much more efficient way than a passive one. Problems that are ill-posed and nonlinear for a passive observer become well-posed and linear for an active observer. In particular, the problems of shape from shading and depth computation, shape from contour, shape from texture and structure from motion are shown to be much easier for an active observer than for a passive one. It has to be emphasized that correspondence is not used in our approach, i.e., active vision is not correspondence of features from multiple viewpoints. Finally, active vision here does not mean active sensing. This paper introduces a general methodology, a general framework in which we believe low-level vision problems should be addressed.

- [26] Eiki Ito and John (Yiannis) Aloimonos, "Determining Three Dimensional Transformation Parameters from Images: Theory." CAR-TR-318, CS-TR-1906, August 1987.

ABSTRACT: We present a theory for the determination of the three dimensional transformation parameters of an object from its images. The input to this process is the image intensity function and its temporal derivative. In particular, our results are:

- 1) If the structure of the transforming object in view is known, then the transformation parameters are determined from the solution of a linear system. Rigid motion is a special case of our theory.
- 2) If the structure of the object in view is not known, then both the structure and transformation parameters may be computed through a hill climbing or simulated annealing algorithm.

- [27] John (Yiannis) Aloimonos, "Shape from Texture." CAR-TR-319, CS-TR-1907, August 1987.

ABSTRACT: A central goal for visual perception is the recovery of the three-dimensional structure of the surfaces depicted in an image. Crucial information about three-dimensional structure is provided by the spatial distribution of surface markings, particularly for static monocular views: projection distorts texture geometry in a manner that depends systematically on surface shape and orientation. To isolate and measure this projective distortion in an image is to recover the three dimensional structure of the textured surface. For natural textures, we show that the uniform density assumption (texels are uniformly distributed) is enough to recover the orientation of a single textured plane in view, under perspective projection. Furthermore, when the texels cannot be found, the edges of the image are enough to determine shape, under a more general assumption, that the sum of the lengths of the contours on the world plane is about the same everywhere. Finally, several experimental results for synthetic and natural images are presented.

- [28] John (Yiannis) Aloimonos and Michael Swain, "Paraperspective Projection: Between Orthography and Perspective." CAR-TR-320, CS-TR-1908, August 1987.

ABSTRACT: We study an approximation of perspective projection, called paraperspective. It turns out that it is a very good approximation of perspectivity under a variety of situations, and it can be used very successfully in texture, contour and motion analysis, as well as object recognition. In this paper we analyze paraperspective projection, compare it with orthography and perspective, apply it to problems that have been addressed in a different way in the literature and use it to discover invariant geometric relations that were unknown up to now. A version of paraperspective projection first appeared in [1]. The main contribution here lies in the conclusion that very good results are obtained by applying to perspective images algorithms developed from a computational theory based on paraperspective projection. This, along with the simplicity of paraperspective and the fact that this projection leads to the discovery of perspective invariants,

motivates the study of paraperspective projection in the context of image understanding.

- [29] Randal C. Nelson and John (Yiannis) Aloimonos, "Using Flow Field Divergence for Obstacle Avoidance in Visual Navigation." CAR-TR-322, CS-TR-1914, September 1987.

ABSTRACT: The practical recovery of quantitative structural information about the world from visual data has proven to be a very difficult task. In particular, the recovery of motion information which is sufficiently accurate to allow practical application of theoretical shape from motion results has so far been infeasible. Yet a large body of evidence suggests that use of motion is an extremely important process in biological vision systems. It has been suggested by Thompson that qualitative visual measurements can provide powerful perceptual cues, and that practical operations can be performed on the basis of such clues without the need for a quantitative reconstruction of the world. The use of such information is termed "inexact vision". This paper describes the investigation of one such approach to the analysis of visual motion. Specifically, the use of certain measures of flow field divergence were investigated as a qualitative cue for obstacle avoidance during visual navigation. It is shown that a quantity termed the *directional divergence* of the 2-D motion field can be used as a reliable indicator of the presence of obstacles in the visual field of an observer undergoing generalized rotational and translational motion. Moreover, the necessary measurements can be robustly obtained from real image sequences. A simple differential procedure for robustly extracting divergence information from image sequences which can be performed using a highly parallel, connectionist architecture is described. The procedure is based on the twin principles of directional separation of optical flow components and temporal accumulation of information. Experimental results are presented showing that the system responds as expected to divergence in real world image sequences, and the use of the system to navigate between obstacles is demonstrated.

- [30] Behzad Kamgar-Parsi and Behrooz Kamgar-Parsi, "An Efficient Model of Neural Networks for Optimization." CAR-TR-326, CS-TR-1922, September 1987.

ABSTRACT: Hopfield and Tank have shown that neural networks can be used in solving very complicated computational problems if they are formulated as optimization problems. Furthermore, they have shown that to obtain good solutions it is necessary to use analog networks rather than digital networks. Simulations of analog networks involve the solution of many coupled differential equations and therefore can be time consuming. For software implementation we propose a model of an analog network that does not involve differential equations, and thus is much more efficient. This is accomplished without compromising the quality of the solutions. Like Hopfield and Tank we use the Traveling-Salesman Problem as an example.

- [31] Rand Waltzman, "Finding Symmetries of Polyhedra." CAR-TR-333, CS-TR-1937, October 1987.

ABSTRACT: This paper presents a representation for polyhedra that does not depend on any external coordinate system. The representation contains the complete metrical as well as topological information from which a polyhedron can be reconstructed. Moreover, the representation is unique. This paper also presents algorithms for finding all of the rotational and reflectional symmetries of a polyhedron using this representation. These algorithms do not perform any numerical computation, they are practical and have been implemented in Franz Lisp.

- [32] John (Yiannis) Aloimonos and Anup Basu, "Combining Information in Low-Level Vision." CAR-TR-336, CS-TR-1947, November 1987.

ABSTRACT: Low level modern computer vision is not domain dependent, but concentrates on problems that correspond to identifiable modules in the human visual system. Several theories have been proposed in the literature for the computation of shape from shading, shape from texture, retinal motion from spatiotemporal derivatives of the image intensity function and the like.

The problems with some of the existing approaches are basically the following:

- (1) The employed assumptions are usually very strong (they are not present in a large subset of real images), and so some of the algorithms fail when applied to real images.
- (2) Usually the constraints from the geometry and the physics of the problem are not enough to guarantee uniqueness of the computed parameters. In this case, strong additional assumptions about the world are used, in order to restrict the space of all solutions to a unique value.
- (3) Even if no assumptions at all are used and the physical constraints are enough to guarantee uniqueness of the computed parameters, then in most cases the resulting algorithms are not robust, in the sense that if there is a slight error in the input (i.e. a small amount of noise in the image), this results in a catastrophic error in the output (computed parameters), and this is observed from experiments.

It turns out that if several available cues are combined, then the above mentioned problems disappear in most cases; the resulting algorithms compute robustly and uniquely the intrinsic parameters (shape, depth, motion, etc.).

In this paper the problem of machine vision is explored from its basics. A low level mathematical theory is presented for the unique and robust computation of intrinsic parameters. The computational aspect of the theory envisages a cooperative highly parallel implementation, bringing in information from five different sources (shading, texture, motion, contour and stereo), to resolve ambiguities and ensure uniqueness of the intrinsic parameters.

- [33] Isaac Weiss. "Projective Invariants of Shapes." CAR-TR-339, CS-TR-1965, January 1988.

ABSTRACT: A major goal of computer vision is object recognition, which involves matching of images of an object, obtained from different, unknown points of view. Since there are infinitely many points of view, one is faced with the problem of a search in a multidimensional parameter space. A related problem is the stereo reconstruction of 3-D surfaces from multiple 2-D images. We propose to solve these fundamental problems by using geometrical properties of the visible shape that are invariant to a change in the point of view. To obtain such invariants, we start from classical theories for differential and algebraic invariants not previously used in image understanding. As they stand, these theories are not directly applicable to vision. We suggest extensions and adaptation of these methods to the needs of machine vision. We study general projective transformations, which include both perspective and orthographic projections as special cases.

- [34] Eiki Ito and John (Yiannis) Aloimonos, "Is Correspondence Necessary for the Perception of Structure from Motion?." CAR-TR-340, CS-TR-1966, January 1988.

ABSTRACT: The fundamental assumption of almost all existing computational theories for the perception of structure from motion is that moving elements on the retina correspond projectively to identifiable moving points in three-dimensional space. Furthermore, these computational theories are based on the fundamental idea of retinal motion, i.e. they use as their input the velocity with which image points are moving (optic flow or discrete displacements). In this research, we investigate the possibility for the development of computational theories for the perception of structure from motion that are not based on the concept of the velocity of individual image elements, i.e. they do not use optic flow or displacements as input.

- [35] Behrooz Kamgar-Parsi and Behzad Kamgar-Parsi, "Simultaneous Fitting of Several Planes to Point Sets Usings Neural Networks." CAR-TR-346, CS-TR-1975, January 1988.

ABSTRACT: It is a simple problem to fit one line to a collection of points in the plane. But when the problem is generalized to two or more lines then the problem complexity becomes exponential in the number of points because we must decide on a partitioning of the points among the lines they are to fit. The same is true for fitting lines to points in three-dimensional space or hyperplanes to data points of high dimensions. Although the problem is NP-complete we show that it can be formulated as an optimization problem for which very good, but not necessarily optimal, solutions can be found by using a dedicated neural network. Furthermore, we show that given a tolerance one can determine the *number* of lines (or planes) that should be fitted to a given point configuration. This problem is prototypical of a class of problems in computer vision, pattern recognition and data fitting. For example, the method we propose can be used in reconstructing a planar world from range data or in recognizing point patterns in an

image.

- [36] Jacob Beck and Richard Ivry, "On the Role of Figural Organization in Perceptual Transparency." CAR-TR-347, CS-TR-1976, January 1988.

ABSTRACT: Metelli (1974) made an important contribution by identifying order and magnitude restrictions for a pattern of intensities and showing that when they are satisfied the perception of transparency readily occurs. These restrictions were derived from a physical model of transparency. We argue that the visual system does not use intensity information to compute indices of transmittance and reflectance analogous to what an optical engineer might do in describing a physical instance of transparency. Rather, a lightness pattern affects perceptual transparency, just as geometric properties do, through processes that impose an organization on sensory information rather than through processes that recover quantitative descriptions. In the absence of depth cues, such as stereopsis and motion parallax, the perception of transparency occurs when the lightness relations in a pattern favor the perception of a continuous boundary across x-junctions. We present evidence for two kinds of violations of the order and magnitude restrictions, simple and strong. Transparency judgments, though reduced in number, still occur for simple violations of the order and the magnitude restrictions. Transparency judgments occur relatively infrequently for strong violations. A physical model of transparency fails to capture the difference between simple and strong violations of the order and magnitude restrictions. We discuss (a) the basis for differentiating between simple and strong violations of the order and magnitude restrictions, (b) how simple and strong violations affect the perception of transparency, and (c) the occurrence of transparency with and without color constancy, i.e., the color seen through the transparent surface looks or fails to look the same as the color seen directly.

- [37] Avraham Margalit and Azriel Rosenfeld, "Using Probabilistic Domain Knowledge to Reduce the Expected Computational Cost of Template Matching." CAR-TR-355, CS-TR-2008, March 1988.

ABSTRACT: Matching of two digital images is computationally expensive, because it requires a pixel-by-pixel comparison of the pixels in the image and in the template. If we have probabilistic models for the classes of images being matched, we can reduce the expected computational cost of matching by comparing the pixels in an appropriate order. In this paper we show that the expected cumulative error when matching an image and a template is maximized by using an ordering technique. We also present experimental results for digital images, when we know the probability densities of their gray levels, or more generally, the probability densities of arrays of local property values derived from the images.

- [38] David Shulman and John (Yiannis) Aloimonos, "Boundary Preserving Regularization: Theory Part I." CAR-TR-356, CS-TR-2011, April 1988.

ABSTRACT: Many problems in low-level vision and in several other scientific or engineering disciplines are ill-posed in the sense that their solutions do not exist, are not unique, or do not depend continuously on the data. We approach these problems with Tikhonov regularization. That means we seek a solution that is a compromise between the requirements of consistency with constraints imposed by the data and of consistency with a priori smoothness assumptions. Unfortunately, the solution obtained blurs boundaries and makes it hard to recognize where the real world variables change sharply. We approach this difficulty by assuming the errors (the inconsistency between data and solution) at nearby points are correlated and we first deblur the errors before regularizing. Similarly we have to deblur the smoothness term of our variational condition before we can apply regularization theory. In general decorrelation is a hard problem, but making special assumptions about the blurring kernel (e.g. the kernel is Gaussian or more generally Levy stable), we can recover the magnitude of the deblurred error (or smoothness) as a linear expression in terms of the original error (or smoothness) and its derivatives. We are, in effect, imposing a requirement that not only the error but also its derivatives should tend to be small (because noise is often far from being white). The resulting variational condition is not the optimal condition but the Euler-Lagrange equations will be linear if the constraints are.

We also suggest a convex approximation technique for solving the piece-wise smooth interpolation problem which results in a convex condition if the original constraints were linear. The paper is written for the non-mathematically oriented reader.

- [39] John (Yiannis) Aloimonos and Jean-Yves Hervé, "Correspondenceless Detection of Depth and Motion for a Planar Surface." CAR-TR-357, CS-TR-2021, April 1988.

ABSTRACT: We show that a binocular observer can recover the depth and three-dimensional motion of a rigid planar patch, without using any correspondences between the left and right image frames (static) or between the successive dynamic frames (dynamic). We study uniqueness and robustness issues with respect to this problem and we provide experimental results from the application of our theory to synthetic and real images. We introduce and work in an enriched Marr paradigm consisting of four levels: computational theory, algorithms (representation), stability (robustness), and implementation.

- [40] John Sullins, "Boolean Learning in Neural Networks." CAR-TR-359, CS-TR-2023, May 1988.

ABSTRACT: Most methods of determining the weights of a connectionist network are based on gradient descent algorithms that attempt to minimize the difference between the expected and actual input-output behaviors. The successes of these methods have been limited due to the fact that global optimization for an arbitrary function is not possible today. An alternative system is presented, one that relates the input-output behavior of a connectionist network

to a Boolean expression in disjunctive normal form. Each hidden unit of the network learns to detect one of the conjunctive parts of the expression by starting with a single input configuration that correctly activates an output and generalizing to a conjunctive set of inputs—an *and-set*—by deleting inputs that do not affect the correctness of the input-output behavior at that unit. Unlike gradient descent methods, which may become trapped in local minima, or simulated annealing methods, which may need an infinite amount of time to reach a good state, this system determines a correct solution to many problems very quickly.

- [41] David Harwood, Raju Prasannappa and Larry Davis, "Preliminary Design of a Programmed Picture Logic." CAR-TR-364, CS-TR-2048, June 1988.

ABSTRACT: The objective of the PPL project is to design and implement a general and modular logic-programmed system for two-dimensional interpretation of image theories in image structures obtained by image analysis. Important subsystems include heuristic search for object instances with optimization of goodness-of-figure, and procedures for computing basic image components, locales for searches, and predicates. We illustrate some of these in an application to aerial images of suburban neighborhoods.

- [42] Avraham Margalit, "A Parallel Algorithm to Generate a Markov Random Field Image on a SIMD Hypercube Machine." CAR-TR-365, CS-TR-2050, June 1988.

ABSTRACT: Generating a Markov random field image is a computationally very expensive process on a sequential processor. We present here a parallel algorithm to perform this task on a SIMD hypercube machine. The problem of implementing such a parallel algorithm is discussed and the implementation of the algorithm on the *Connection Machine* along with some of our results are presented. We show from theoretical and experimental results that a 40% degree of parallelism is optimal for this algorithm. In our implementation we demonstrate a 40% degree of parallelism and an effective speedup of more than 70 times over the sequential implementation on a Vax 11/785 running Unix.

- [43] Subbarao Kambhampati, "An Approach to Flexible Reuse of Plans." CAR-TR-367, CS-TR-2054, June 1988.

ABSTRACT: The value of enabling a planning system to remember the plans it generates for later use was acknowledged early in planning research. The systems developed, however, were very inflexible as the reuse was primarily based on simple strategies of generalization via variablization and later unification. We propose an approach for flexible reuse of old plans in the presence of a generative planner. In our approach the planner leaves information relevant to the reuse process in the form of annotations on every generated plan. To reuse an old plan in solving a new problem, the old plan along with its annotations is mapped into the new problem. A process of *annotation verification* is used to locate applicability failures and suggest refitting strategies. The planner is then called upon to carry out the suggested modifications—to produce an executable plan for the new problem. This integrated approach obviates the need for any extra domain

knowledge (other than that already known to the planner) during reuse and thus affords a relatively domain-independent framework for plan reuse. We will describe the realization of this approach in two disparate domains (blocks world and process planning for automated manufacturing) and propose extensions to the reuse framework to overcome observed limitations. We believe that our approach to plan reuse can be profitably employed by generative planners in many applied domains.

- [44] Lee Spector, James A. Hendler, John Canning and Azriel Rosenfeld, "Symbolic Model/Image Matching in Expert Vision Systems." CAR-TR-370, CS-TR-2060, July 1988.

ABSTRACT: Existing expert vision systems generally match models to images using only numeric "goodness-of-fit" measures. The computation of such measures usually involves the combining of incommensurate quantities and the loss of low level knowledge that could be useful at higher levels. The methods employed, and hence the software developed, often cannot be generalized for use within other domains or at other levels of abstraction. We feel that there is a need for a more general symbolic image/model matching paradigm, and for the development of software tools that implement it. In this report we outline motivations for the development of a general purpose symbolic matcher, present an overview of a current implementation, and discuss several important requirements that any such system ought to meet. We also present a detailed example showing our matcher in action on real-world image data. A User's Guide for our system is included as an Appendix.

- [45] Randal C. Nelson, "Visual Navigation." CAR-TR-380, CS-TR-2087, August 1988.

ABSTRACT: Visual navigation is a major goal in machine vision research, and one of both practical and basic scientific significance. The practical interest reflects a desire to produce systems which move about the world with some degree of autonomy. The scientific interest arises from the fact that navigation seems to be one of the primary functions of vision in biological systems. Navigation has typically been approached through reconstructive techniques since a quantitative description of the environment allows well understood geometric principles to be used to determine a course. However, reconstructive vision has had limited success in extracting accurate information from real-world images. This report argues that a number of basic navigational operations can be realized using qualitative methods based on inexact measurement and pattern recognition techniques.

Navigational capabilities form a natural hierarchy beginning with simple abilities such as orientation and obstacle avoidance, and extending to more complex ones such as target pursuit and homing. Within a system, the levels can operate more or less independently, with only occasional interaction necessary. This report considers three basic navigational abilities: *passive navigation*, *obstacle avoidance*, and *visual homing*, which together represent a solid set of elementary, navigational tools for practical applications. It is demonstrated that all

three can be approached by qualitative, pattern-recognition techniques. For passive navigation, global patterns in the spherical motion field are used to robustly determine the motion parameters. For obstacle avoidance, divergence-like measurements on the motion field are used to warn of potential collisions. For visual homing an associative memory is used to construct a system which can be trained to home visually in a wide variety of natural environments. Theoretical analyses of the techniques are presented, and implementation and testing of working systems described.

- [46] John Canning, "A Note on Mask-Based Least Squares Line Fitting." CAR-TR-384, CS-TR-2095, August 1988.

ABSTRACT: A method to improve the estimate of least squares line fits to thin stripes in images is proposed. By using the geometry of local gray level patterns and their contrasts, the accuracy of the least squares line fits can be improved markedly. The improved method's performance is compared to that of the Canny line detector.

- [47] Ramesh Kumar Sitaraman, "The Ordered Matching Problem." CAR-TR-387, CS-TR-2098, August 1988.

ABSTRACT: We consider the problem of optimally ordering a set of operations, the outcomes of which are random. In Sections 1 and 2, we introduce the problem and illustrate it with the example of template matching. In Sections 3 and 4, we give procedures for finding the optimal dynamic strategy and the optimal static strategy respectively. In Section 5, we consider a constrained form of the problem and show that it has a simple optimal strategy. In Section 6, we investigate the complexity issues involved in finding optimal strategies. In Section 7, we discuss directions for future research.

- [48] Minas E. Spetsakis and John (Yiannis) Aloimonos, "Optimal Computing of Structure from Motion Using Point Correspondences in Two Frames." CAR-TR-389, CS-TR-2101, September 1988.

ABSTRACT: One of the problems associated with any approach to the structure from motion problem using point correspondences, i.e. recovering the structure of a moving object from its successive images, is the use of least squares on dependent variables. We formulate the problem as a quadratic minimization problem with a non-linear constraint. Then we derive the condition for the solution to be optimal under the assumption of Gaussian noise in the input, in the Maximum Likelihood Principle sense. This constraint minimization reduces to the solution of a non-linear system which in the presence of modest noise is easy to approximate. We present two efficient ways to approximate it and we discuss some inherent limitations of the structure from motion problem when two frames are used that should be taken into account in robotics applications that involve dynamic imagery. In addition, our formulation introduces a framework in which previous results on the subject become special cases.

- [49] John (Yiannis) Aloimonos and Dimitris P. Tsakiris, "On the Mathematics of Visual Tracking." CAR-TR-390, CS-TR-2102, September 1988.

ABSTRACT: A mathematical theory for visual tracking of a three-dimensional target of known shape moving rigidly in 3-D is presented here and it is shown how a monocular observer can track an initially foveated object and keep it stationary in the center of the visual field. Our attempt is to develop correspondence-free tracking schemes and get rid of the limitations inherent in the optical flow formalism. Moreover, a general tracking criterion, the *Tracking Constraint*, is derived, which reduces tracking to an appropriate optimization problem. The connection of our tracking strategies with the Active Vision Paradigm is shown to provide a solution to the Egomotion problem under the assumption of knowledge of shape.

In this work, tracking strategies based on the recovery of the 3-D motion of the target are devised under the above assumption. A correspondence-free scheme is derived, which depends on global information about the scene (provided by linear features of the image) in order to bypass the ill-posed problem of computing the spatial derivatives of the image intensity function, and amounts to the solution of a linear system of equations in order to estimate the 3-D motion of the target. An important feature of these tracking strategies is that they do not require continuous segmentation of the image in order to locate the target. Supposing that the target is sufficiently textured, dynamic segmentation using temporal derivatives of the linear features provides sufficient information for the tracking phase. Therefore, this approach is expected to perform best when previous ones fail, namely in a complex visual environment.

Experimental results for the algorithms presented here demonstrate their robustness in the presence of noise.

- [50] Radu S. Jasinschi, "Towards a Theory of Apparent Visual Motion." CAR-TR-394, CS-TR-2117, October 1988.

ABSTRACT: The existence of two separate mechanisms for the processing of apparent motion, the short- and long-range processes, as proposed by Braddick in 1974, has been analyzed through many different psychophysical experiments. In particular the fact that for the short-range process there exists an upper bound for the spatial displacement and temporal interstimulus interval between successive stimulus presentations was confirmed by several of these experiments.

In order to gain a more formal understanding of these issues, we analyze the phenomenon of apparent motion from the point of view of a reconstruction problem. This allows us to use the sampling theorem to analyze the problem of temporal (spatial) reconstruction of uniformly translating patterns. In the case where the velocity field can only be extracted with uncertainty, it can be shown that there exists a maximum temporal (spatial) sampling interval, such that aliasing does not occur. We argue that, in the case of the short-range process, due to its temporal (spatial) reconstruction ability, a similar effect could intervene in the limitation of its activity to a small spatio-temporal scale.

- [51] Menashe Brosh, Behrooz Kamgar-Parsi and Behzad Kamgar-Parsi, "The Reliability of the Closed-Form Solution to the Image Flow Equations for 3D Structure and Motion (Quadric Patch)." CAR-TR-397, CS-TR-2123, October 1988.

ABSTRACT: Relative motion between objects and the viewer generates a time-varying image which, in principle, can be used as a source of 3D information about the structure of the objects and the relative motion. One approach to obtaining 3D information from time-varying imagery is to utilize the image flow field and its derivatives. The characteristics of the image flow field depend both on the relative motion and the surface of the object. Thus, given the image flow field, in theory, one can invert the problem and recover the relative motion and the structure of the object. In this paper we analyze the intrinsic reliability of such an approach, i.e. assuming that the image flow field is known accurately, except for quantization error, we derive closed-form expressions for the error due to quantization in the recovered 3D motion and structure parameters. These expressions are essential for revealing the intrinsic limitations of the approaches used for the recovery of the 3D parameters from a given image flow field and are thus of great practical importance. Also presented are several illustrative examples.

- [52] Minas Spetsakis and John (Yiannis) Aloimonos, "A Multi-Frame Approach to Visual Motion Perception." CAR-TR-407, CS-TR-2147, November 1988.

ABSTRACT: The main issue in the area of motion estimation given the correspondences of some features in a sequence of images is sensitivity to error in the input. The main way to attack the problem is redundancy in the data. Up to now all the algorithms developed either used two frames or depended on restrictive assumptions and ad hoc techniques. We present in this paper an algorithm based on multiple frames that employs only the rigidity assumption, is simple and mathematically elegant, extremely flexible and, most importantly, is a major improvement over the two-frame algorithms. The algorithm does minimization of the mean square error which we prove equivalent to an eigenvalue minimization problem. One of the side effects of this mean square method is that the algorithm can have a very descriptive physical interpretation in terms of the "loaded spring model".

- [53] John Sullins, "Distributed Learning: Motion in Constraint Space." CAR-TR-412, CS-TR-2166, December 1988.

ABSTRACT: Most methods of learning in distributed environments are based on gradient descent algorithms that involve changing the weights of the network in order to minimize the difference between the expected and actual input-output behaviors. The successes of such "motion in weight space" methods have been limited due to their inability to capture the implicit constraints of the behavior and properly distribute them among the units of the network. An alternative system is presented, one based on *motion in constraint space*. It relates the input-output behavior of a connectionist network to a Boolean expression in disjunctive normal form, where each hidden unit of the network learns to detect one

of the conjunctive parts of the expression. The potential constraints at a processor are the states of an input configuration that correctly activates the outputs. These constraints are added and removed from the processors in such a way that the correctness of the behavior of the network is maximized. Unlike gradient descent methods, which may become trapped in local minima, or simulated annealing methods, which may need an infinite amount of time to reach a good state, this system determines a correct solution to many problems very quickly. Unlike most traditional "machine learning" algorithms, this system can learn concepts in parallel, is capable of continuously adapting to new information, and is highly resistant to feedback error. Applications to problems such as recognizing (learning) 2-D shapes (such as fish tails, for example) show the potential of the applicability of the method to practical problems.

- [54] Behzad Kamgar-Parsi, Behrooz Kamgar-Parsi and William A. Sander. "Quantization Error in Spatial Sampling: Comparison Between Square and Hexagonal Pixels." CAR-TR-415, CS-TR-2171, January 1989.

ABSTRACT: Square and hexagonal spatial samplings, because of their processing ease, are used most widely in image and signal processing. However, no rigorous treatment of the quantization error due to hexagonal sampling has appeared in the literature. In this paper we develop mathematical tools for estimating quantization error in hexagonal sensory configurations. These include analytic expressions for the average error and the error distribution of a function of an arbitrarily large number of hexagonally quantized variables. The two quantities, the average error and the error distribution, are essential in assessing the reliability of a given algorithm. For comparison we also present the corresponding expressions for square spatial sampling, so that they can be used in comparing the magnitude of the error incurred in hexagonal versus square quantization for a given algorithm. They can thus be used to determine which sampling technique would result in less quantization error for a particular algorithm. Such a comparison is important due to the paramount role that quantization error plays in computational approaches to computer vision. Some general observations in regard to the relative accuracy of hexagonal versus square quantization are also presented. It is hoped that the expressions derived in this paper will have an impact on both sensor design and the assessment of the reliability of a given algorithm under hexagonal as well as square quantization.

- [55] Ken-ichi Kanatani, "Hypothesizing and Testing Geometric Properties of Image Data." CAR-TR-416, CS-TR-2172, January 1989.

ABSTRACT: A general formulation is given for testing particular geometrical configurations of image data. The procedure consists of hypothesizing and testing: We first estimate an ideal geometrical configuration which supposedly exists, and then check to what extent the original edge data must be displaced to support the hypothesis. Thus, all types of tests are reduced to computing a single measure of edge displacement without involving ad-hoc measures and threshold values depending on the problem. Also, no explicit forms of probability distribution need be introduced. All the procedures are described by explicit algebraic

expressions in unit vectors which represent points and lines on the image plane, so that no computational overflow occurs and no searches or iterations are required.

- [56] Behzad Kamgar-Parsi, J. Anthony Gualtieri, Judith E. Devaney and Behrooz Kamgar-Parsi, "Clustering in Parallel with Neural Networks." CAR-TR-417, CS-TR-2173, January 1989.

ABSTRACT: Partitioning a set of N patterns in a d -dimensional metric space into K clusters—in a way that those in a given cluster are more similar to each other than the rest—is a problem of interest in image analysis, astrophysics and other fields. As there are approximately $\frac{K^N}{K!}$ possible ways of partitioning the patterns among K clusters, finding the best solution is beyond exhaustive search when N is large. We show that this problem in spite of its exponential complexity can be formulated as an optimization problem for which very good, but not necessarily optimal, solutions can be found by using a neural network. To do this the network must start from many randomly selected initial states. The network is simulated on the NASA MPP (a 128×128 SIMD array machine), where we use the massive parallelism not only in solving the differential equations that govern the evolution of the network, but also in starting the network from many initial states at once thus obtaining many solutions in one run. We obtain speed-ups of two to three orders of magnitude over serial implementations.

- [57] Radu S. Jasinschi, "Intrinsic Constraints in Space-Time Filtering: A New Approach to Representing Uncertainty in Low-Level Vision." CAR-TR-425, CS-TR-2201, February 1989.

ABSTRACT: This paper describes how, in the process of extracting the optical flow through space-time filtering, we have to take into account constraints associated with the motion uncertainty, as well as with the spatial and temporal sampling rates of the temporal sequence of images. The motion uncertainty is shown to satisfy an inequality, as a consequence of the use of the Cramér-Rao inequality, which is a function of the filter parameters. On the other hand, the spatial and temporal sampling rates have lower bounds, which depend on the motion uncertainty, the maximum support in the frequency domain and the estimated optical flow. These lower bounds on the sampling rates and on the motion uncertainty are constraints which constitute an intrinsic part of the computational structure of space-time filtering. They are of a different nature than the ones used in regularization theory, because they do not dictate any arbitrary constraints on the parameters being computed, but instead arise as a natural consequence of the estimation process. By conjugating these constraints, we are able to devise an algorithm which describes an adaptive procedure of estimating the various parameters involved in space-time filtering. This corresponds to an instance of an adaptive system, through which the variables involved in the process of space-time filtering are allowed to vary inside a range which is consistent with the various intrinsic constraints governing the process.

- [58] Dong Yoon Kim, J. John Kim and Azriel Rosenfeld, "A Robust Method for Fitting a Straight Line to a Noisy Image." CAR-TR-428, CS-TR-2212, March 1989.

ABSTRACT: In fitting a straight line to a noisy image, the least square method becomes unreliable if non-Gaussian outliers are present. We introduce the Least Median Square (LMS) method, which provides:

- protection against distortion by up to 50% of contaminated data;
- good efficiency in the presence of various type of noise;
- an amount of computation comparable with the least square method.

- [59] Behzad Kamgar-Parsi, Behrooz Kamgar-Parsi and Menashe Brosh, "Exact Results for the Sum of Uniform Random Variables." CAR-TR-429, CS-TR-2226, April 1989.

ABSTRACT: We derive exact analytic expressions for the distribution function, the probability density function, and the mean deviation of the sum $Y = \sum_{i=1}^N a_i X_i$, where X_i are independent random variables with uniform distributions, for an arbitrary number of variables N and arbitrary parameter values a_i . We also investigate the approach of the sum to the Central Limit.

- [60] Menashe Brosh, Behrooz Kamgar-Parsi and Behzad Kamgar-Parsi, "Reliability Analysis of the Closed-Form Solution to the Image Flow Equations for 3D Structure and Motion (Planar Patch)." CAR-TR-431, CS-TR-2228, April 1989.

ABSTRACT: The idea of obtaining 3D information about the structure of the object and its relative motion with respect to the viewer, from the time-varying optic field at the image plane, has attracted the attention of a large number of researchers for many years. As a result a number of papers have appeared in the literature deriving formulas for computation of shape and motion parameters. However, no rigorous assessment of the reliability of such approaches has appeared in the literature. In a recent paper, we analyzed the reliability of the approach for a curved surface in motion and did not find it encouraging. In this paper, we analyze the intrinsic reliability of such an approach for a planar patch in motion. More precisely, as was the case in the error analysis of curved surfaces, the assumption is that except for the quantization error, the image flow field is known accurately. That is, we derive closed-form expressions for the error (due to quantization) in the recovered 3D motion and structure parameters. These expressions are essential for revealing the intrinsic limitations of the approaches used for the recovery of the 3D parameters from a given image flow field and are thus of great practical importance.

[61] Anup Basu and John (Yiannis) Aloimonos, "Approximate Constrained Motion Planning." CAR-TR-435, CS-TR-2234, April 1989.

ABSTRACT: The problem of finding a collision-free path connecting two points (start and goal) in the presence of obstacles, with constraints on the curvature of the path, is examined. This problem of curvature-constrained motion planning arises when (for example) a vehicle with constraints on its steering mechanism needs to be maneuvered through obstacles. Though no lower bound on the difficulty of the problem in 2-D is known, exact algorithms given so far for the reachability question are exponential. We obtain a simple polynomial time algorithm for obtaining an approximation scheme for the above problem. The approximation scheme can be used for obtaining the minimum curvature path or minimum length path satisfying a given curvature constraint. A probabilistic analysis of the scheme is given to analyze its usefulness. The method is easily generalizable to 3-D.

[62] John R. Sullins, "Distributed Learning of Texture Classification." CAR-TR-444, CS-TR-2254, May 1989.

ABSTRACT: A large number of statistical measures have been postulated for the description and discrimination of textures. While most are useful in some situations, none are totally effective in all of them. An alternative approach is to *learn* which measures are best for particular circumstances. In this paper the distributed learning system of *constraint motion* is used to learn relevant texture descriptors from a set of well-known first and second order grey-level statistics. Using this system, a network of distributed units partitions itself into sets of units that detect one and only one of the given classes of textures. Each of these sets is further partitioned into individual units that detect natural subtypes of these texture classes, ones which do not necessarily produce the same types of statistics at the local level. Together, these units form a network capable of determining the texture classification of an image.

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b RESTRICTIVE MARKINGS N/A		
2a SECURITY CLASSIFICATION AUTHORITY N/A			3 DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited		
2b DECLASSIFICATION/DOWNGRADING SCHEDULE N/A			5 MONITORING ORGANIZATION REPORT NUMBER(S) N/A		
4 PERFORMING ORGANIZATION REPORT NUMBER(S) N/A			7a NAME OF MONITORING ORGANIZATION U.S. Army Center for Night Vision and Electro-Optics		
6a NAME OF PERFORMING ORGANIZATION University of Maryland		6b OFFICE SYMBOL (If applicable) N/A	7b ADDRESS (City, State, and ZIP Code) Fort Belvoir, VA 22060		
6c ADDRESS (City, State, and ZIP Code) Center for Automation Research College Park, DM 20742-3411			9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAAB07-86-K-F073		
8a NAME OF FUNDING/SPONSORING ORGANIZATION Defense Advanced Research Projects Agency		8b OFFICE SYMBOL (If applicable) IPSO	10 SOURCE OF FUNDING NUMBERS		
8c ADDRESS (City, State, and ZIP Code) 1400 Wilson Blvd. Arlington, VA 22209-2308			PROGRAM ELEMENT NO	PROJECT NO	TASK NO
			WORK UNIT ACCESSION NO		
11 TITLE (Include Security Classification) VISION IN DYNAMIC ENVIRONMENTS -- Contract DAAB07-86-K-F073 -- Final Technical Report					
12 PERSONAL AUTHOR(S) Azriel Rosenfeld					
13a TYPE OF REPORT Final		13b TIME COVERED FROM 6/30/86 TO 6/30/89		14 DATE OF REPORT (Year, Month, Day) August 15, 1989	
15 PAGE COUNT 26					
16 SUPPLEMENTARY NOTATION					
17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP			
19 ABSTRACT (Continue on reverse if necessary and identify by block number) Research conducted on the contract was primarily concerned with real-time three-dimensional computer vision and image understanding. The results of the research were documented in 62 Technical Reports. This Final Technical Report consists of the abstracts of the earlier reports.					
20 DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS			21 ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a NAME OF RESPONSIBLE INDIVIDUAL			22b TELEPHONE (Include Area Code)		22c OFFICE SYMBOL