STEREO MATCHING BY HIERARCHICAL, MICROCANONICAL ANNEALING

Stephen T. Barnard

Artificial Intelligence Center
SRI International, Menlo Park, California 94025

April 16, 1987

ABSTRACT

An improved stochastic stereo-matching algorithm is presented. It incorporates two substantial modifications to an earlier version: a new variation of simulated annealing that is faster, simpler, and more controllable than the conventional "heat-bath" version, and a hierarchical, coarse-to-fine-resolution control structure. The Hamiltonian used in the original model is minimized, but far more efficiently. The basis of microcanonical annealing is the Creutz algorithm . Unlike its counterpart, the familiar Metropolis algorithm, the Creutz algorithm simulates a thermally isolated system at equilibrium. The hierarchical control structure, together with a Brownian state-transition function, tracks ground states across scale, beginning with small, coarsely coded levels. Results are shown for a 512 x 512 pair with 50 pixels of disparity.

I INTRODUCTION

The paper describes a stochastic optimization approach to stereo matching. An early version is described in [1]. Two major extensions have been made that permit substantially improved performance: First, a new variety of simulated annealing is employed which is more efficient, more easily implemented, and offers more control over the annealing process than the standard form. Secondly, the system is extended to operate over several levels of resolution, exploiting relatively quickly computed minima at lower levels of resolution to initialize its state at higher levels. This method leads both to more efficiency and to the ability to deal with much larger ranges of disparities.

The basic representation remains unchanged. The state of the system encodes a dense map of discrete horizontal disparities, defined over the left image, which specify corresponding points in the right image. In the improved design, this state is relative to a particular level of resolution.

The energy of a lattice site is composed of two terms, each of which expresses a constraint important in stereo matching:

$$E_{ij} = |I_L(i,j) - I_R(i,j + D(i,j))| + \lambda|\nabla D(i,j)| \ .$$

I_L and I_R are the left and right image intensities, or some simple function of them (such as a Laplacian). Subscripts i and j range over all sites in the left image lattice. D is the disparity map. The first term is the absolute difference in intensity between corresponding points (where the correspondences are defined by the current state of the system), and the second is proportional to the local spatial variation in the current state (the magnitude of the gradient of the disparity map). The constant λ is used to balance the terms.[1] This equation expresses two competing constraints: first, the image intensities of corresponding points should be more-or-less equal, and second, the disparity map should be more-or-less continuous.

The Hamiltonian,

$$E = \sum E_{ij} \ ,$$

is unchanged; therefore, the ground states that we seek to determine or to approximate are the same. The new design is strictly concerned with improved performance. The techniques used to obtain it are not restricted to stereo matching, and should be considered for any simulated-annealing approach to low-level vision problems.

II MICROCANONICAL ANNEALING

The conventional simulated annealing algorithm uses an adaptation of the Metropolis algorithm to bring a system to equilibrium at decreasing temperatures. The Metropolis algorithm [2] defines a Markov process that generates a sequence of states, such that the probability of occurrence of any particular state is proportional to its Boltzman weight,

$$P(S) \propto \exp(-\beta E(S)) \ ,$$

where $E(S)$ is the energy of state S and β is the inverse temperature of the system. This process generates samples from the *canonical ensemble*; that is, the system is considered to be immersed in a heat bath with a controllable temperature. Annealing is accomplished by imposing a *schedule* for reducing the temperature, $1/\beta$, so as to keep the system close to equilibrium.

Creutz has described an alternative technique that simulates the *microcanonical ensemble* [3]. In this method, the total energy remains constant (for some fixed point in the schedule). Instead

Support for this work was provided by the Defense Advanced Research Projects Agency under contracts DCA 76-85-C-0004 and MDA903-86-C-0084.

[1] The performance of the system is not very sensitive to the value chosen for λ. In the example of Section IV, as well as all three examples in [1], $\lambda = 5$ was used.

of simulating a system immersed in a heat bath, the Creutz algorithm simulates a thermally isolated system in which energy is conserved. It performs a random walk through state space, constraining states to a surface of constant energy. The simplest way to accomplish this is to augment the representation with an additional degree of freedom, called a demon, that carries a variable amount of energy, E_D. The total energy of the system is now:

$$E = E(S) + E_D \ .$$

Normally, the demon is constrained to have nonnegative energy, although the possibility of giving it negative energy is useful in annealing, as will be discussed below.

In the Metropolis algorithm a potential new state S' is chosen randomly, and is accepted or rejected based on the change in energy:

$$\Delta E = E(S') - E(S) \ .$$

If ΔE is negative, the new state is accepted; otherwise, it is accepted with probability $\exp(-\beta \Delta E)$.

The Creutz algorithm is quite similar. If ΔE is negative, the new state is accepted, and the demon energy is increased ($E_D \leftarrow E_D - \Delta E$). If ΔE is nonnegative, however, acceptance of the new state is contingent upon E_D: if $\Delta E < E_D$ the change is accepted, and the demon energy is decreased ($E_D \leftarrow E_D - \Delta E$); otherwise, the new state is rejected. Clearly, the total energy of the system remains constant.

The Creutz algorithm has several advantages. Unlike the Metropolis algorithm, it does not require the evaluation of transcendental functions. Of course, in practice these functions can be stored as tables, but we would like our algorithm to be adaptable to fine-grained parallel processors such as the Connection Machine. The small amount of local memory in such machines makes lookup tables unattractive. The Creutz algorithm can easily be implemented with only integer arithmetic — again, a significant advantage for fine-grained parallel processors and for VLSI implementation. Experiments indicate that the Creutz method can be programmed to run an order of magnitude faster than the conventional Metropolis method for discrete systems [4]. A further important advantage is that microcanonical simulation does not require high-quality random numbers.

In conventional simulated annealing, we control the process by specifying the temperature. In the microcanonical version, however, temperature is not a control parameter; it is a statistical feature of the system. In fact, standard arguments can be used to show that at equilibrium the demon energies have a Boltzman distribution:

$$P(E_D) \propto \exp(-\beta E_D) \ .$$

If the demon energies are non-negative, the inverse temperature can be determined from the mean value:

$$\beta = 1/ < E_D > \ .$$

Control of microcanonical annealing is accomplished by periodically removing energy from the system. The method used to generate the results in Section IV is as follows:

1. Assume that the process begins in a random state S_0 of high energy with respect to the ground state. Call this $E_0 = E(S_0)$. The initial demon energy is zero.

2. Remove a fixed proportion of this energy, δE, by reducing the demon energy ($E_D \leftarrow E_D - \delta E$). The results of Section IV were obtained with $\delta E = E_0/300$.

3. Run the Creutz algorithm until the system reaches equilibrium. A reasonable test for equilibrium is to consider the rate of accepted moves to states of higher energy. If the system is large enough, this rate will increase steadily until it approximates the rate of moves to states of lower energy. We terminate the process for a particular energy level when the rate of accepted moves to higher energy states decreases (measured over N site visits).

4. Repeat steps (2) and (3) until no further improvement is observed.

This procedure has only one free parameter: the ratio $\delta E/E_D$. While $E_D < 0$, the Creutz method operates as a "greedy" algorithm, accepting only moves to lower energy states and increasing E_D. When E_D becomes positive (which happens quickly if δE is small), the demon begins to exchange energy between lattice positions. As the ground state is approached, E_D remains negative because it cannot absorb more energy from the lattice.

The algorithm described above is sequential, and therefore is not suitable for parallel processing. Each local state transition can change E_D, which in turn can affect the next transition. Fortunately, the technique can be modified to a parallel one by using a separate demon for each lattice site. (As we add demons, the technique moves toward a canonical-ensemble simulation. In fact, if the number of demons is very large compared to the number of sites, the technique specializes to the Metropolis algorithm [3].) Preliminary experiments with one demon per site indicate good performance, although the results of Section IV were generated with the single-demon algorithm.

III HIERARCHICAL ANNEALING

In the original model [1], annealing was performed only at the level of resolution of the stereo images. In some cases, the images were first bandpass-filtered to remove low-frequency components — in effect, a simple photometric correction for inconsistent sensor gains or film development. Lower and upper bounds on disparity were specified in advance, and all state transitions were considered with equal probability. As the range of disparity became large, this scheme required much larger amounts of computation. If we have m permissible disparities and N pixels, the size of the state space is m^N. If we double the range of disparity, the size of the state space increases by a factor of 2^N. Since N is a rather large number (2^{18} in the example shown in Section IV), we see that the state space grows explosively with increasing disparity range.

A natural extension of the method is to adopt a hierarchical, coarse-to-fine control structure. At a coarse level of resolution, the number of lattice sites (i.e., the number of pixels) and the range of disparity are small; therefore the size of the state space is relatively small.[2] We should be able to compute an approximate ground state quickly, and then use it to initialize the annealing process at the next, finer level of resolution.

[2] Although the state space may be large in absolute terms.

The Laplacian pyramid, originally developed as a compact image-coding technique [5], offers an efficient representation for hierarchical annealing. In a Laplacian pyramid, an image is transformed into a sequence of bandpass-filtered copies, $I^0, I^1, I^2, ..., I^n$, each of which is smaller than its predecessor by a factor of 1/2 in linear dimension (a factor of 1/4 in area), with the center frequency of the passband reduced by one octave. This transform can be computed efficiently by recursively applying a small generating kernel to create a Gaussian (low-passed) pyramid, and then differencing successive low-passed images to construct the Laplacian pyramid.

After constructing Laplacian pyramids from the original stereo images, disparity is reduced by a factor of 1/2 in successive levels. Therefore, at some level, disparity is small everywhere. For typical stereo images, we can take this to be level $n-4$. [For example, if the original images were a power of 2 in linear dimension, the Laplacian images at level $n-4$ would be 16 x 16 pixels. Disparities in the range of 0 to 63 pixels in a pair of 512 x 512 images would be reduced to the range of 0 to 1 pixel, with truncation, at the $(n-4)$th level.] We shall start annealing at this level, find an approximate ground state, and then expand the solution to the next level. To make this coarse-to-fine strategy work, however, we need two further modifications. We must use a different process for generating state transitions, and we must specify how a low-resolution result is used to start the annealing process at the next-higher level.

In the original model, the probability of choosing a new disparity for consideration as a new state was uniformly distributed over the prior range of disparities:

$$P(d_j \rightarrow d_k) = \frac{1}{m} .$$

This is not compatible with our intention of guiding the process with lower-resolution results, however. Because we are now assuming the disparity of a lattice site to be close to its correct value, a more effective generating process is to restrict the disparities to increase or decrease by one pixel:

$$P(d_j \rightarrow d_k) = \begin{cases} 0.5 & \text{if } |d_j - d_k| = 1 \\ 0 & \text{otherwise} \end{cases} ,$$

with the further restriction that the disparities are not allowed to specify corresponding points outside the boundary of the right image. In this scheme, the system undergoes Brownian motion through state space. An additional feature of this state-transition function is that it is no longer necessary to specify bounds on disparity in advance.

Expanding a low-resolution result to the next level is slightly more complicated. Obviously, one should begin by simply doubling the size of the low-resolution lattice and doubling the disparity values. Having done this, however, we find that the new state has an artificially low energy because every odd disparity value is "unoccupied," and the new map is therefore more uniform than it should be. A spurious symmetry is imposed on the new state that is solely due to the quantization of the previous result, which is likely to place the system near a metastable state (a local minimum) from which it cannot recover. Fortunately, there is an easy solution to this problem: Destroy this symmetry by adding heat. One effective way is as follows:

1. Compute the energy E_0^k of the initial state at level k. (This state has been determined by doubling the result at level $k+1$.)
2. Add a fixed proportion of this energy, ΔE^k, by increasing the demon temperature ($E_D \leftarrow E_D + \Delta E^k$). The results of Section IV were obtained with $\Delta E^k = E_0^k/10$.
3. Run the Creutz algorithm until the system reaches equilibrium.
4. Repeat steps (2) and (3) until the number of accepted higher-energy states exceeds the number of rejected higher-energy states.

IV RESULTS

The images in Figure 1 are an aerial stereo pair (512 x 512 pixels each) covering part of Martin Marietta's test site for the Autonomous Land Vehicle Project near Denver. The terrain is dominated by a long, steep "hogback" (running diagonally across the middle of the images), that separates two broad valleys. The highest terrain is in the upper left. Several roads and a few buildings may be seen. Disparity ranges from zero to 50 pixels.

The results of applying the microcanonical, hierarchical annealing algorithm to these data are illustrated in Figure 2, which shows the approximate ground state at the highest level of resolution. Disparity values are displayed in 15 pseudocolors with wraparound.

The largest problem solved by the previous version of the system was on the order of a 256 x 256 pair with 15 levels of disparity. Solving this problem required about 12 hours on a Symbolics 3600. The result shown here is a 512 x 512 pair with 50 levels of disparity, and was computed in about the same amount of time. Consider the size of the state space of the two problems. The state space of the current problem[3] exceeds the size of the old problem by a factor of more than 10^{36000}.

V CONCLUSIONS

Two major improvements to a stochastic stereo-matching system have been described: first, a simulation of the microcanonical ensemble, as opposed to the canonical ensemble of conventional simulated annealing; and second, the extension to a hierarchical control structure based on Laplacian pyramids. Both techniques are rather general in nature and can probably be used in other simulated-annealing applications. Together, they permit solutions of stereo-matching problems that are far beyond the competence of the original system.

We are currently evaluating the quantitative increase in performance attributable to each new technique. Qualitatively, the use of microcanonical annealing yields perhaps an order of magnitude increase in efficiency. A potentially more important benefit is that its computation is simpler than the computation of standard annealing, and is therefore more readily implemented in fine-grained parallel systems. The hierarchical control structure, combined with the Brownian state-transition function, contributes most of the increased performance.

[3] Considering the range of disparity to be fixed at 50 pixels.

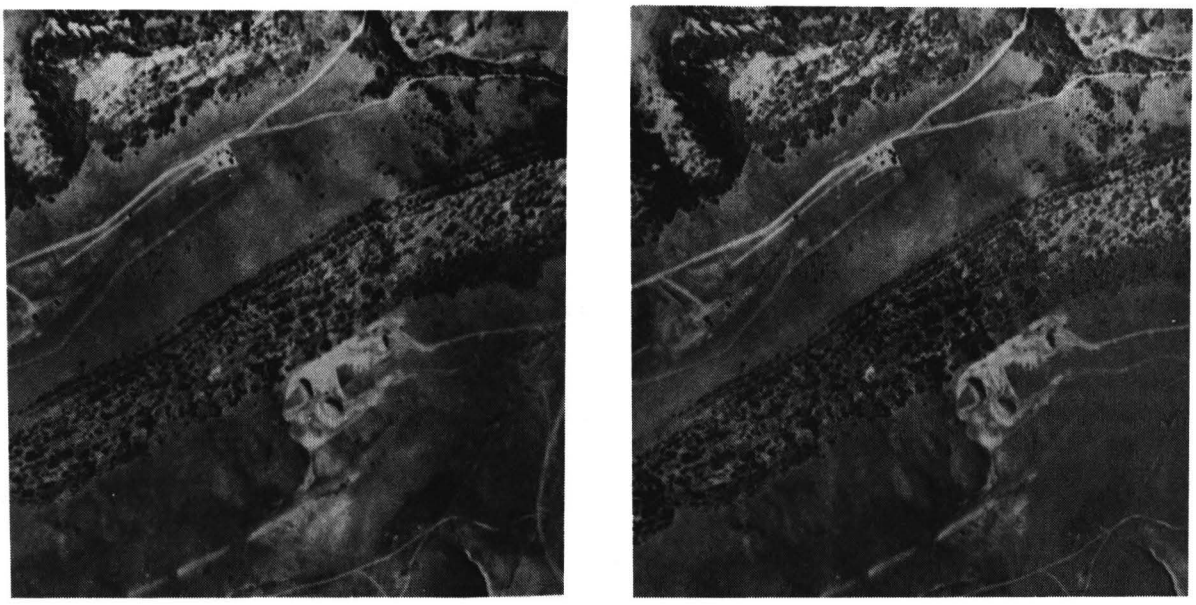

Figure 1: Aerial stereogram of the ALV test site.

References

[1] S. Barnard, A stochastic approach to stereo vision, *Proceeding of the 5th National Conference on Artificial Intelligence* (Philadelphia, August 1986), vol. 1, American Association of Artificial Intelligence, Menlo Park, California, pp. 676-680.

[2] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, Equations of state calculations by fast computing machines, *Journal of Chemical Physics*, vol. 21, no. 6, June 1953, pp. 1087-1092.

[3] M. Creutz, Microcanonical Monte Carlo simulation, *Physical Review Letters*, vol. 50, no. 19, May 9, 1983, pp. 1411-1414.

[4] G. Bhanot, M. Creutz, and H. Neuberger, Microcanonical simulation of Ising systems, *Nuclear Physics*, B235[FS11], 1984, pp. 417-434.

[5] P. Burt, The Laplacian pyramid as a compact image code, *IEEE Transactions on Communications*, vol. COM-31, no. 4, April 1983, pp. 532-540.

Figure 2: Approximate ground state.