

# CHROMATIC STEREOPSIS

John R. Jordan III\*, Alan C. Bovik\*,  
and Wilson S. Geisler\*\*

\*Department of Electrical and Computer Engineering

\*\*Department of Psychology  
The University of Texas at Austin  
Austin, Texas 78712-1084

## Abstract

One approach to developing faster, more robust stereo algorithms is to seek a more complete and efficient use of information available in stereo images. The use of chromatic (color) information has been largely neglected in this regard. Motivations for using chromatic information are discussed, including *strong* evidence for the use of chromatic information in the human stereo correspondence process in the form of a novel psychophysical experiment which we have performed. Specifically, we measured the minimum disparity needed to produce a reversal in apparent depth in ambiguous chromatic "wallpaper" stereograms. Our results indicate that chromatic information plays an important role in the stereo correspondence process when luminance variations are present. To investigate the potential role of chromatic information in computational stereo algorithms, a novel chromatic matching constraint — the *chromatic gradient matching constraint* -- is presented. Then, a thorough analysis of the utility of this constraint in the PMF Algorithm is performed for a large range of sizes of the matching strength support neighborhood, and the performances of the algorithm with and without these constraints are directly compared in terms of disambiguation ability, matching accuracy and algorithm speed. The results demonstrate that chromatic information can greatly reduce the matching ambiguity, while significantly increasing both matching accuracy and algorithm speed.

## 1. Introduction

The development of automated vision systems capable of extracting three-dimensional scene information using a stereo camera geometry is an area of active research, with potential application in autonomous vehicle navigation, industrial assembly, etc. Although various techniques have been proposed, the absence of practical stereo systems indicates that aspects of the problem remain unsolved. The most critical task in stereo vision -- establishing correspondences — has received a considerable amount of attention, and the general principles regarding its solution are well understood. However, a solution which is

both accurate and computationally efficient over a wide variety of scenes remains elusive.

As with algorithms for other computer vision tasks, stereo algorithms have been influenced by observations and evidence regarding the human visual system. In particular, recent stereo algorithms have been strongly influenced by the edge-based approach developed by Marr and Poggio [Marr and Poggio, 1979, Grimson, 1981], which uses intensity zero crossings (ZCs) as fundamental matching primitives. In this approach, the solution is constrained by limiting the search space according to the bandwidth of the ZC operator used and enforcing smoothness of disparity. Other algorithms seek to increase accuracy by incorporating physical/psychophysical constraints ~ the most robust of which is perhaps the disparity gradient limit [Burt and Julesz, 1980, Pollard *et al.*, 1985].

Although promising results have been obtained, a faster, more robust solution to the correspondence problem is required if practical needs are to be met. Such an algorithm must deal with noise, occlusion, and transparency. One approach to developing such an algorithm is to seek new mathematical techniques; another approach - as is done here — is to seek a more complete and efficient use of available image information. One mode of information that has been largely neglected in stereo algorithms is chromatic information. Although the use of chromatic information as an aid in stereo has been mentioned and tested (in somewhat simplistic frameworks), a thorough study has never been performed.

There are several motivations for using chromatic information. First, chromatic intensity information is easily obtained. Second, while intensity variations in the visual environment are denser (subjectively) than the pattern of chromatic variations, a given luminance variation may occur many times in an image, causing a multitude of potential correspondences to be extracted. Since similar chromatic variations are typically less dense, they may be very useful in eliminating false correspondences. Thus, chromatic information can potentially improve algorithm speed and accuracy and assist in scenes where shadows, occlusion, transparency, or repetitive patterns are present. Third, new psychophysical evidence indicates that chromatic information is used in human stereopsis.

This material is based in part upon work supported by the Texas Advanced Research Program under Grant No. 3546.

## 2. Psychophysical Evidence for the Use of Color in the Human Stereo Correspondence Process<sup>1</sup>

Many experiments to determine the role of chromatic information in the process of human stereopsis have been performed; the general consensus has been that chromatic information is utilized to a much lesser degree than luminance information.

Initial studies showed that depth perception was possible for some subjects in opposite-contrast random-dot stereograms if, simultaneously, the hues of the corresponding dots were positively correlated in the two images [Treisman, 1962, Julesz, 1971]. Since subjects are unable to fuse achromatic opposite-contrast random-dot stereograms [Julesz, 1960], one must conclude that color plays at least some role in human stereopsis.

Most subsequent psychophysical research concerning the role of color in human stereopsis has been directed towards determining whether chromatic information alone is adequate for stereopsis [Lu and Fender, 1972, Gregory, 1977, De Weert and Sadza, 1983, Grinberg and Williams, 1985]. Initial results using isoluminant random-dot stimuli indicated that chromatic information alone is not adequate [Lu and Fender, 1972, Gregory, 1977], while studies using slightly different stimulus conditions found that depth perception was possible but very weak at isoluminance [De Weert and Sadza, 1983, Grinberg and Williams, 1985].

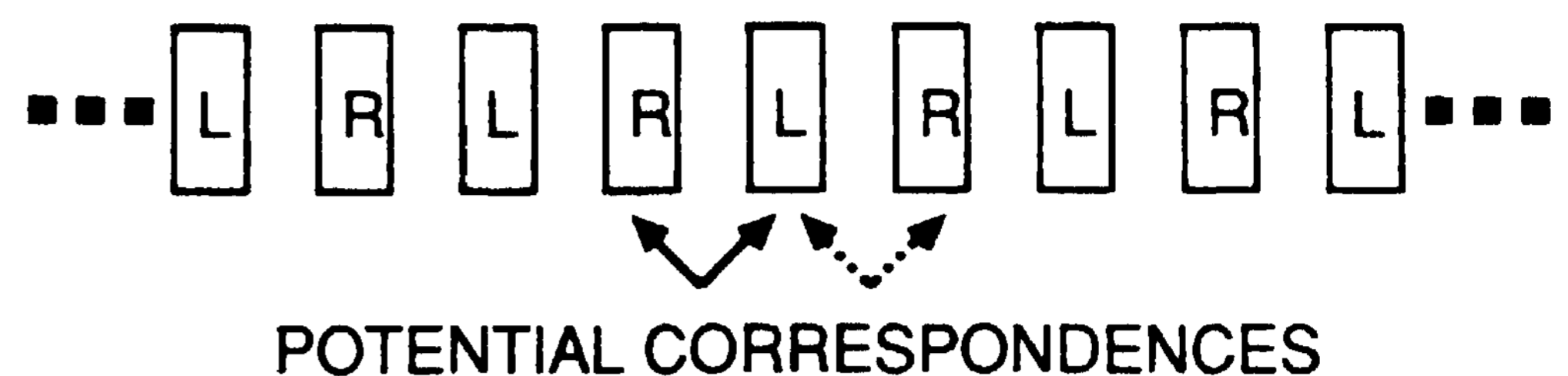
Studies using isoluminant figural stimuli have yielded more consistent results — all have found that depth perception for isoluminant figural stimuli is possible [Gregory, 1977, De Weert and Sadza, 1983]. Furthermore, studies using figural stimuli which isolate the short wavelength (blue) cone system have demonstrated reliable depth perception [Grinberg and Williams, 1985]. However, most of these investigators have noted that the quality of perceived depth for isoluminant stimuli is quite poor [De Weert and Sadza, 1983, Grinberg and Williams, 1985].

The isoluminant stimuli results have prompted the view that chromatic information is relatively unimportant in stereopsis [Lu and Fender, 1972, Grinberg and Williams, 1985]. However, the fact that chromaticity by itself yields poor stereopsis does not preclude the possibility that chromaticity could have a more significant effect in the presence of luminance changes, as was observed in the initial studies [Treisman, 1962, Julesz, 1971]. This is an important possibility to investigate because chromatic variations in natural scenes are almost always accompanied by luminance variations.

### 2.1 Methods

In order to test whether chromatic information is used in the presence of luminance variations, one must conceive of situations where using chromatic features would be necessary in order to solve the correspondence problem. One possibility is to use stereograms for which luminance information yields ambiguous results — more than one possible depth organization - but for which there are fewer possible matches if chromatic information is used. Here, we

chose to use ambiguous "wallpaper" stereograms comprised of equally spaced vertical bars which extend beyond the field of view. The superimposed "cyclopean" image of an example bar sequence with equiluminant, isochromatic bars and a disparity equal to exactly half the interbar distance is shown in Figure 1.



**Figure 1. Superimposed image of an ambiguous stereogram.**

Although many possible correspondences are possible, the ambiguity typically results in the perception of a single depth plane. In the absence of other cues, the observer's solution to the correspondence problem appears to be solved by proximity. By varying the luminances and chromaticities of the bars in a controlled manner, we may measure the relative contributions of luminance and chromaticity in the solution of the correspondence problem.

In our experiment, each stimulus consisted of a left and a right image of three vertical bar patterns on a dark background, as shown in Figure 2; the viewing distance was 45 cm, resulting in the pattern dimensions shown. With the exception of the relative displacements of the bar patterns, the two images were identical; hence, it suffices to discuss one image.

The middle bar pattern served as an unambiguous fixation pattern having zero disparity, while the upper and lower patterns constituted the test pattern. These patterns were identical and extended well beyond the field of view (imposed by blinders); furthermore, they had disparities of the same magnitude, but of opposite sign. For small disparity values, the three patterns were perceived at three different depth planes in either "forward-tilt ranked order" (top pattern appeared farthest from the subject) or "backward-tilt ranked order" (top pattern appeared closest to the subject). At larger disparity values, the relative perceived depths either assumed a different order or became rivalrous. By starting with zero disparity and adjusting the disparity values until the rank order of the perceived depth of the three patterns changed, the "rank-order range" could be measured for various stimuli.

During the entire experiment, each bar of the middle pattern was yellow and had a luminance of 14 cd/m<sup>2</sup>. However, the chromaticities and luminances of the bars of the upper and lower patterns were varied in an alternating pattern denoted as A- and B-type bars in Figure 2. Each bar type had one of two chromaticities (Red or Green) and one of three luminances (7, 14, or 21 cd/m<sup>2</sup>). It is important to note that stimuli in which the A- and B-type bars had the same luminance were *not* isoluminant since the background was dark. For notational convenience, we will refer to such patterns as "equiluminant"; analogously, we will refer to patterns with A- and B-type bars of the same chromaticity as "equichromatic".

<sup>1</sup>For a more detailed discussion see Jordan *et al.* [1989].

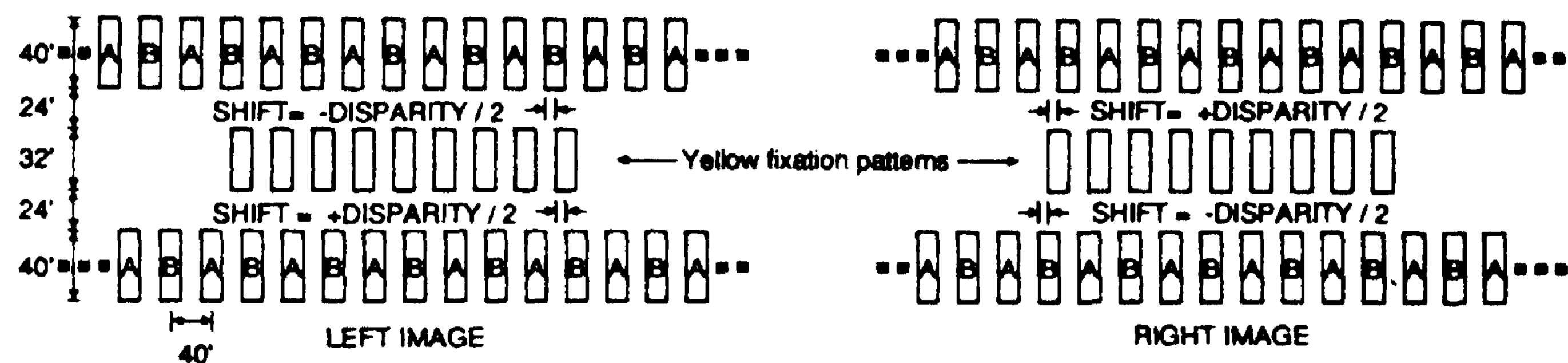


Figure 2. Stimulus pattern used for psychophysical experiment. See text for details.

The yellow fixation pattern remained present throughout the experiment. The stimulus patterns (like those in Figure 2) were presented for 2 seconds once every 7 seconds. Between stimulus presentations, the subject adjusted the disparity values of the upper and lower bar patterns, starting with zero disparity and adjusting the disparity values until the rank order of the perceived depth of the three patterns changed. Generally, this occurred when either (a) one or both of the upper and lower bar patterns flipped in depth relative to the center pattern or (b) the patterns became rivalrous. After satisfying the adjustment criterion, the subject indicated completion of the measurement by pressing an appropriate button.

## 2.2 Results

Mean results for 4 subjects (two naive) are shown in Figure 3; equichromatic and chromatic data are represented by squares and circles respectively. The data points for the equichromatic stimuli represent the average for red and green bar patterns. The equiluminant measurements provide the most direct test of whether chromaticity influences the stereo correspondence process. In the equiluminant case, the mean disparity at which the apparent depth ordering of the equichromatic bar patterns broke down or reversed was 17.1' (as expected when proximity is the only cue), while that for the chromatic bar patterns was 27.5'. This indicates that chromaticity is playing a substantial role in the solution of ambiguous wallpaper stereograms. In fact, this 10.4' difference is nearly one-third the non-diplopic fusion range reported for bars of similar width. Note that (a) a similar result was obtained regardless of the luminance ratio; and (b) the luminance ratio of the bars also had an effect on the rank-order range; however, even a 3:1 luminance ratio appears less influential than the present chromaticity difference.

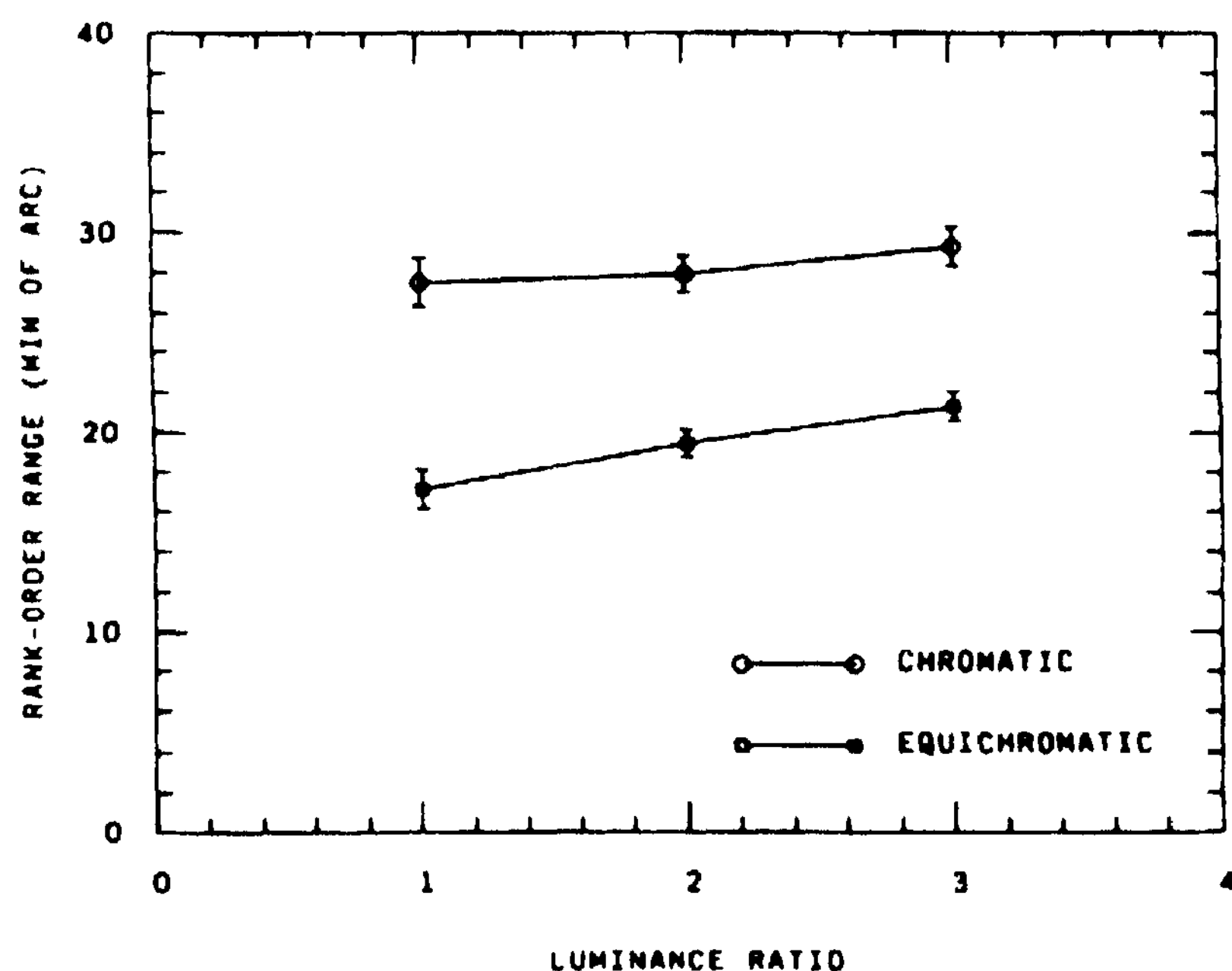


Figure 3. Rank-order range as a function of luminance ratio.

## 3. Using Chromatic Information in the PMF Algorithm<sup>2</sup>

After a stereo image pair has been acquired, intensity features corresponding to significant physical changes in the environment must be determined. In the PMF Algorithm, ZCs computed from the Laplacian of Gaussian (LoG) filtered intensity images are used -- the LoG is given by  $\nabla^2 g_{\sigma}(x,y)$ , where  $g_{\sigma}(x,y)$  is the 2-D rotationally symmetric Gaussian function with space constant  $\sigma$ .

The PMF Algorithm [Pollard *et al.*, 1985] is comprised of two distinct stages: (a) *match extraction*, where potential matches are extracted, followed by (b) *disparity selection*, where the correspondence problem is solved by selecting from the potential matches. The disparity selection stage consists of two sequential steps: (1) calculate matching strengths for each potential match by summing matching strength contributions from all neighboring potential matches within a circular region which we will denote the 'support neighborhood'; and (2) select matches by performing discrete relaxation on the basis of matching strengths.

The PMF Algorithm generally performs quite well, due primarily to the robust nature of the disparity gradient limit constraint; however it often performs poorly for scenes which include repetitive patterns, transparent surfaces, and complex occlusions. Though not always the case, this degradation in performance is often related to a failure to effectively disambiguate between potential matches and hence extract potential matches in the extraction stage. Thus, chromatic information is perhaps best used in the extraction stage.

In this study, color information has been used *only* to characterize spatially coincident zero crossings in the LoG-filtered intensity image. The reason for this is twofold. First, the use of chromatic information as a ZC attribute isolates the effects of including chromatic information; thus, the performances of the PMF Algorithm both with and without these attributes may be directly compared in terms of disambiguation ability and matching accuracy. Second, assuming appropriate chromatic matching constraints, the use of chromatic information as an attribute allows the stereo algorithm to perform as an intensity-only algorithm in scenes in which there is little or no chromatic variation.

### 3.1 The Incorporation of Chromatic Information into the Match Extraction Stage

Various linear and nonlinear transformations of the R, G, and B values are possible; however, most are not appropriate for extracting chromatic *matching* attributes be-

<sup>2</sup>For a more detailed discussion see Jordan and Bovik [1989].

cause either (a) noise in the chromatic images leads to instabilities in the transformed values; or (b) the estimation of noise (and hence matching constraint noise sensitivity) in the transformed values is intractable. Also, the non-transformed RGB values may be treated simply as intensity values taken by cameras which have various spectral sensitivities. Hence, the principles which apply to grey-level image analysis and noise estimation apply to the individual chromatic images as well, particularly in light of their strong positive correlation with the intensity spectrum.

In order to use chromatic information to characterize ZCs, it seems most appropriate to choose a feature which accurately represents local variations in the spectral values. There are two features which seem both conceptually and computationally feasible: chromatic edges and chromatic gradients. We have observed that the ZCs from the LoG-filtered chromatic images are rarely very different from those obtained from the intensity image; this is not surprising, as the chromatic images each have a strong intensity factor and are highly correlated. However, we have also observed that the individual ZCs are often not exactly coincident between the images; hence, a problem with accurately representing local variations arises. Furthermore, because of noise considerations, ZC attributes (orientation, gradient magnitude) are often derived from first order derivatives anyway — the chromatic gradient. Thus, we have used chromatic gradients.

For each chromatic image, the horizontal and vertical components of the chromatic gradient at each pixel are computed by Gaussian filtering (for noise suppression) and taking the first derivatives. Momentarily considering the Gaussian-smoothed Red image and letting  $R_x(x,y)$  and  $R_y(x,y)$  denote its horizontal and vertical derivatives, the chromatic gradient orientation is simply  $\tan^{-1}(R_y(x,y)/R_x(x,y))$  and the gradient magnitude is given by  $\sqrt{(R_x(x,y))^2 + (R_y(x,y))^2}$ . Green and Blue attributes are defined analogously.

The strong positive correlation of the intensity and RGB values suggests that chromatic gradient orientations may often be similar to the intensity chromatic gradient orientation, particularly at the intensity ZCs. Thus, a matching constraint on the chromatic gradient orientation (or on the signs of the individual horizontal and vertical components) may not provide much disambiguation ability. Nonetheless, variations in color are readily apparent; this suggests that a constraint on the chromatic gradient magnitude may provide greater disambiguation ability. Moreover, matching constraints on the values of the individual horizontal and vertical components (incorporating both sign *and* magnitude) may provide the greatest disambiguation ability. Deriving such constraints is difficult, however, since the effects of geometric distortion cannot be ignored.

Fortunately, the effect of geometric distortion has recently received considerable attention in the development of dense-feature stereo algorithms. Basically, the effect which geometric distortion has upon intensity gradient matching constraints may be examined by equating the intensity functions of a stereo image pair, assuming that corresponding points have the same image intensity and that

the epipolar constraint is imposed. The resulting equations are then differentiated in the horizontal and vertical directions, resulting in an expression for the horizontal and vertical distortions between two corresponding points. Finally, using the disparity gradient limit, upper bounds on the approximate horizontal and vertical distortions are found. Momentarily considering gradients of the Gaussian-smoothed Red image ( $\alpha = 1$ ), the resulting chromatic matching constraints are

$$|R_{l,i,j} - R_{r,i',j}| \leq \frac{\alpha}{2} |R_{l,i,j} + R_{r,i',j}| \quad (1)$$

and

$$|R_{l,i,j} - R_{r,i',j}| \leq \frac{\alpha}{2} |R_{l,i,j} + R_{r,i',j}| \quad (2)$$

where the first subscript indicates the image (left, right), the second subscript denotes the component direction,  $(i,j)$  and  $(i',j)$  are the coordinates of the ZCs in the left and right images respectively, and  $\alpha=1.0$  is the disparity gradient limit. Matching constraints for the horizontal and vertical components of the chromatic gradients of the Green and Blue spectra are defined analogously.

In images of real scenes, image noise is always present and will adversely affect the stereo algorithm performance unless incorporated into the matching constraints. Thus, assuming additive homogeneous Gaussian noise, a noise estimate term,  $k\sigma_x$  (or  $k\sigma_y$ ), is added to the matching constraints, where  $k$  is a positive number (chosen probabilistically) and the variances ( $\sigma_x$ ,  $\sigma_y$ ) be estimated by examining the histograms of the chromatic gradient components. The resulting matching constraints for the Red chromatic gradients are:

$$|R_{l,i,j} - R_{r,i',j}| \leq \frac{\alpha}{2} |R_{l,i,j} + R_{r,i',j}| + k\sigma_x \quad (3)$$

and

$$|R_{l,i,j} - R_{r,i',j}| \leq \frac{\alpha}{2} |R_{l,i,j} + R_{r,i',j}| + k\sigma_y \quad (4)$$

Chromatic matching constraints for the Green and Blue are defined analogously; in all cases, we have used  $k=2$ .

### 3.2 The Incorporation of Chromatic Gradient Matching Constraints

After ZC detection, the ZC's contrast sign and gradient orientation are recorded. To minimize noise effects only those ZCs which remain after thresholding are used. Furthermore, for simplicity, the problem is reduced to a 1-D matching problem by assuming a nonconvergent imaging geometry and the epipolar constraint.

Chromatic information is incorporated into the PMF Algorithm by imposing the chromatic gradient matching constraints at the match extraction stage. With these constraints, the match extraction stage may be stated as follows: for each non-horizontal ZC in the left image determine the set of corresponding candidate matches in the right image which: (1) lie within the search space; (2) have the same contrast sign; (3) have roughly the same orientation ( $\pm 30^\circ$ ); and, (4) satisfy the chromatic gradient matching constraints for *each* color. Note that the fourth constraint can be imposed upon the grey-level intensity gradients as well.

### 3.3 Experimental Results

A stereo image pair of a tilted planar surface covered by variously colored pins was acquired using a Panasonic CCD camera and Tiffen color filters; Figure 4 shows the intensity images for "Pins".

After image noise estimation, the PMF Algorithm was applied to the image pair using four types of application. For the first type of application (Type 1), only the first three matching constraints were used; for each of the remaining types, the fourth matching constraint was used as well. The matching constraint parameter and type of information upon which it was imposed in each case were: Type 2 application:  $a = 1.0$ , intensity gradient only; Type 3 application:  $a = 1.0$ , each chromatic gradient; and Type 4 application:  $a = 0.50$ , each chromatic gradient. The second application was included for fair comparison of the algorithm with and without chromatic information, while the fourth application indicates how the chromatic matching constraints work when there is little distortion (as in the case presented here).

For each type of application,  $a$  was approximately equal to 1.0 (pixel) and the search space size was  $\pm 60$  pixels. The performance of the PMF Algorithm depends upon the support neighborhood size; thus, the algorithm was applied using support neighborhood radii ranging from 3 to 10 pixels. After each application, the results of the algorithm were interactively verified for correctness.

Figure 5 shows intensity encoded disparity maps for "Pins"; intensity values represent *relative* disparities where higher intensity values correspond closer objects/surfaces. The disparity maps shown include the ground truth disparity map as well as the results of the Type 1 and Type 3 applications for a support neighborhood radius of 5 pixels. The intensity encoded display of disparities in these cases is somewhat misleading; verification and analysis of the results reveals a larger difference in the performance of the algorithms.

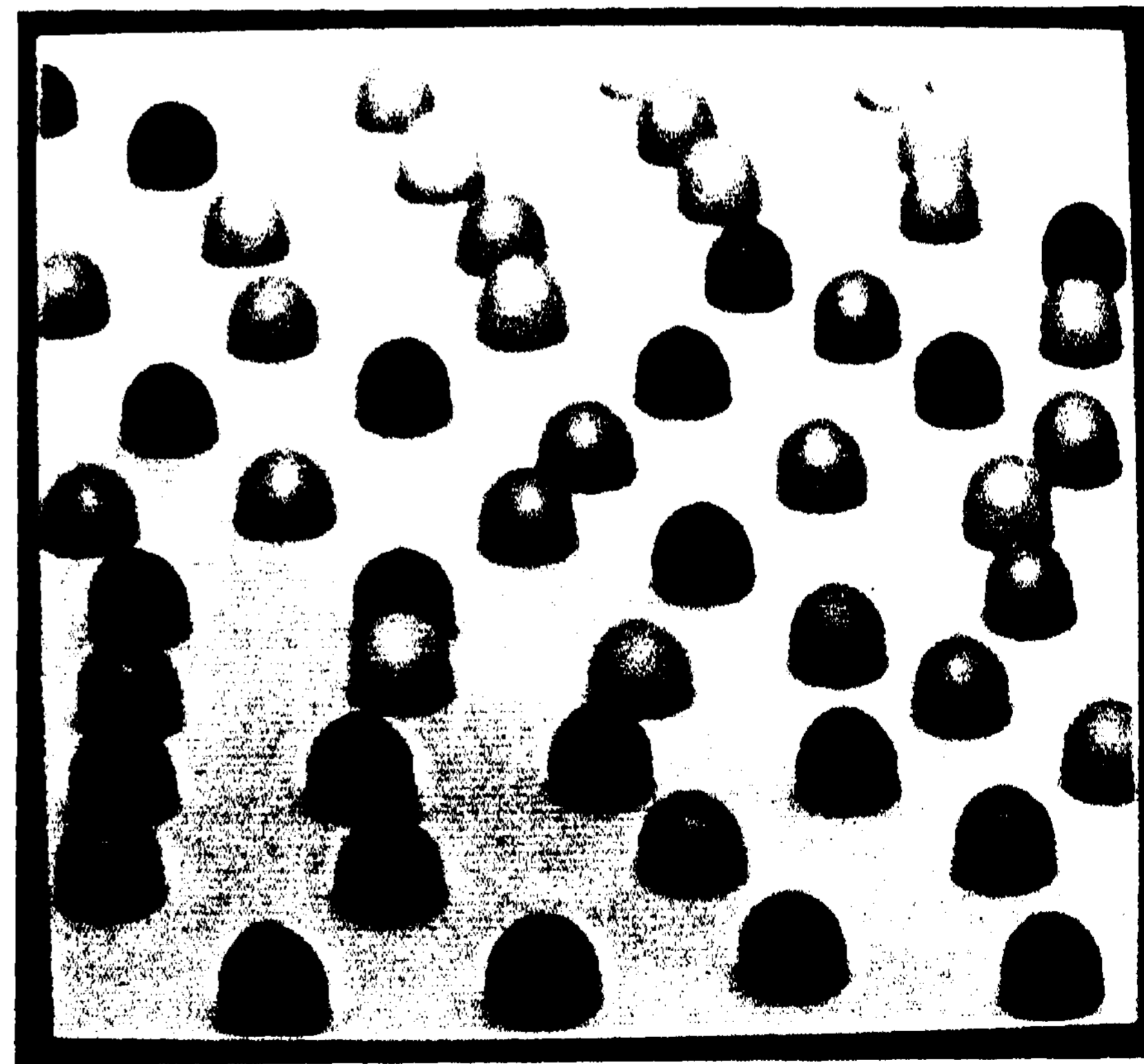
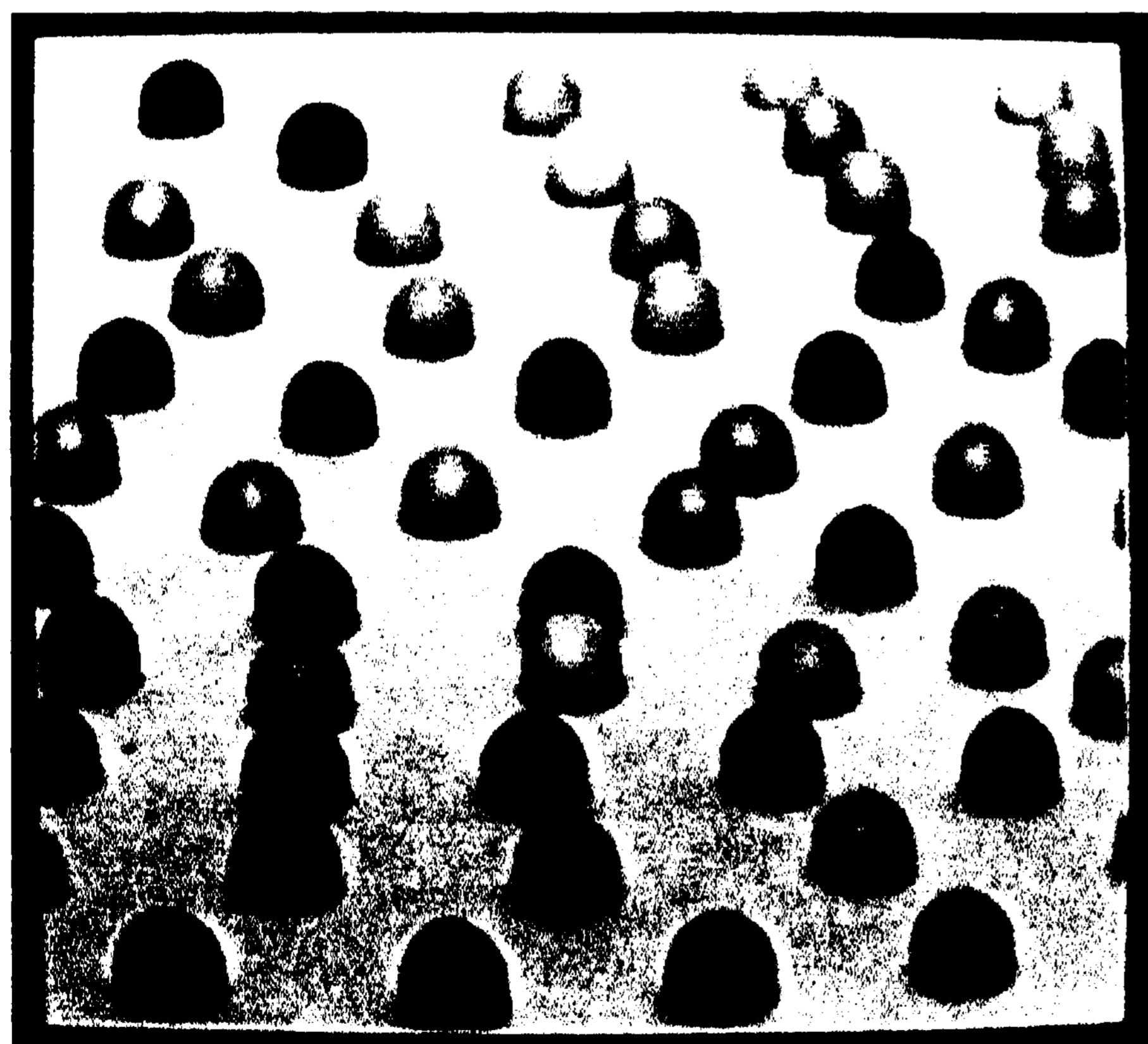


Figure 4. Left and right intensity image for the stereo image pair "Pins".

Analysis of the match extraction stage results gives a rough indication of the disambiguation ability of the chromatic matching constraints. At the extraction stage, the total number of candidate matches found for all ZCs decreased between 41.4% and 50.6% when the chromatic matching constraints were applied, indicating that the inclusion of chromatic information can greatly assist in the disambiguation of potential candidate matches. Presumably, this reduction in potential matches leads to increased speed and accuracy in the selection stage.

#### 3.3.1 Algorithm Accuracy Analysis

There are two different criteria by which accuracy may be analyzed: (1) by determining the number of matches which are correct; or (2) by determining the percentage of matches which are correct. The most advantageous algorithm is one which performs well for both of these criteria; however, strong performances for these two criteria are not necessarily correlated. Thus, it is appropriate to analyze the results for both criteria.

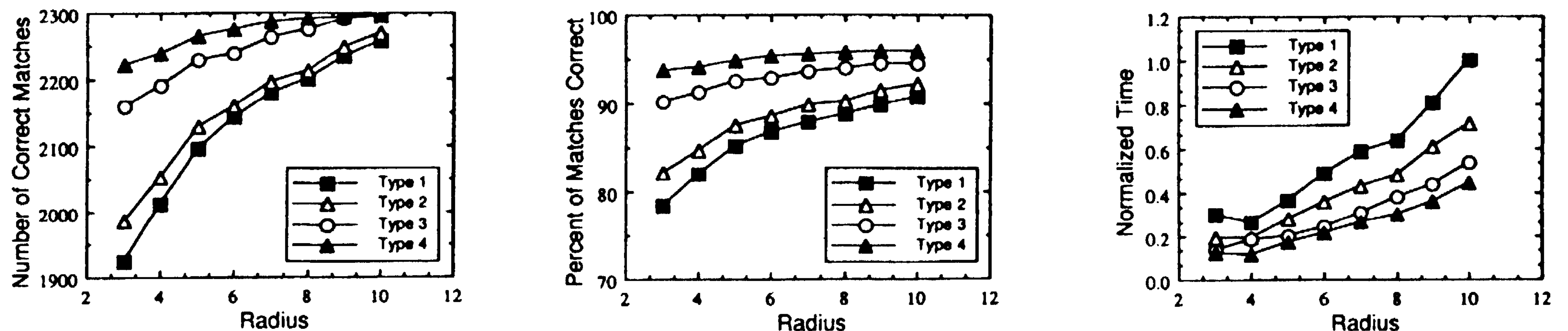
For each application of the PMF algorithm, the number of correct matches is plotted versus support neighborhood radius in Figure 6. Clearly, the inclusion of chromatic matching constraints results in a greater number of correct matches. This greater accuracy is possible even with a much smaller support neighborhood size than required for the Type 1 application. Figure 7 shows a plot of the percentage of matches which are correct versus support neighborhood radius for each application of the algorithm. Again, the inclusion of chromatic matching constraints clearly results in significantly increased accuracy relative to the Type 1 application. In both cases, the application of the gradient matching constraint to the intensity gradient (Type 2) results in little improvement

#### 3.3.2 Algorithm Speed Analysis

Although the PMF Algorithm performs reasonably well for many scenes, the selection stage computation can be quite time-consuming; hence, the effect which the use of



**Figure 5. Intensity-encoded depth maps for "Pins": (a) Ground truth disparity map; (b) Results of Type 1 application of PMF Algorithm; (c) Results of Type 3 Application of PMF Algorithm. See text for discussion of accuracy.**



**Figures 6, 7, and 8. Number of correct matches, Percent of matches correct, and Normalized selection stage times for "Pins".**

chromatic information has upon the speed of this stage is of interest.

A plot of normalized disparity selection stage time versus support neighborhood radius for each application of the PMF Algorithm to "Pins" is shown in Figure 8. Note that these times have been normalized by the maximum time which was required. For larger support neighborhood radius sizes, the applications using chromatic matching constraints show marked decrease in selection stage time relative to the Type 1 application.

A more direct indication of increased algorithm speed is the 'speedup factor' — the increase in algorithm speed for a Type 2, 3, or 4 application relative to that of the Type 1 application for the *same* radius value. Average speedup factors for each application of the algorithm were: Type 2 - 1.37, Type 3 -- 1.84, Type 4 -- 2.25. Thus, the inclusion of chromatic information clearly results in increased algorithm speed.

## References

- [Burt and Julesz, 1980] P. Burt and B. Julesz. A disparity gradient limit for binocular fusion. *Science*, 208:615-617, 1980.
- [De Weert and Sadza, 1983] C.M.M. De Weert and K. J. Sadza. New data concerning the contribution of colour differences to stereopsis. In *Colour Vision* (Edited by J. D. Mollon and L. T. Sharpe), pp. 553-562. Academic Press, New York, 1983.
- [Gregory, 1977] R. L. Gregory. Vision with isoluminant colour contrast: 1. A projection technique and observations. *Perception*, 6:113-119, 1977.
- [Grimson, 1981] W.E.L. Grimson. A computer

implementation of a theory of human stereo vision. *Phil. Trans. R. Soc. of Lond.*, B292:217-253, 1981.

- [Grinberg and Williams, 1985] D. L. Grinberg and D. R. Williams. Stereopsis with chromatic signals from the blue-sensitive mechanism. *Vision Res.*, 25:531-537, 1985.
- [Jordan and Bovik, 1989] J. R. Jordan, A. C. Bovik. On Using Color in Edge-Based Stereo Algorithms. Technical Report, Computer and Vision Research Center, The University of Texas at Austin, June, 1989.
- [Jordan *et al.*, 1989] J. R. Jordan, W. S. Geisler, and A. C. Bovik. Color as a Source of Information in the Stereo Correspondence Process. Submitted to *Vision Research*.
- [Julesz, 1960] B. Julesz. Binocular depth perception of computer-generated patterns. *Bell Sys. Tech. Journal*, 39:1125-1162, 1960.
- [Julesz, 1971] B. Julesz. *Foundations of Cyclopean Perception*. University of Chicago Press, Chicago, 1971.
- [Lu and Fender, 1972] C. Lu and D. H. Fender. The interaction of color and luminance in stereoscopic vision. *Invest. Ophthalm. Visual Sci*, 11:482-490, 1972.
- [Marr and Poggio, 1979] D. Marr and T. Poggio. A computational theory of human stereo vision. *Proc. R. Soc. Lond.*, B204:301-328, 1979.
- [Pollard *et al.*, 1985] S. Pollard, J. Mayhew, and J. Frisby. PMF: A stereo correspondence algorithm using a disparity gradient limit, *Perception*, 14:449-470, 1985.
- [Treisman, 1962] A. Treisman. Binocular rivalry and stereoscopic depth perception. *Quarterly Journal of Exp. Psych.*, 14:23-37, 1962.