Computational Stereo Vision Using Color

John R. Jordan III and Alan C. Bovik

ABSTRACT: An investigation of the efficacy of color or chromatic information as an aid in solving the stereo correspondence problem is described. Two very simple stereo algorithms differing only in the use of chromatic information are applied to images containing various chromatic characteristics. Briefly, the Intensity Matching Algorithm matches the zero crossings in a Laplacian-of-Gaussian filtered image, while the Color Matching Algorithm includes chromatic gradients that characterize the intensity zero crossings. The stereo algorithms were defined with a very simple matching structure in an effort to isolate the effects of including chromatic information in early stereo processing. The results obtained indicate that the use of chromatic information can significantly reduce the ambiguity between potential matches while increasing the accuracy of the resulting matches.

Introduction

The examination of visual cues in an image—such as shading and occlusion—often yields information about the relative distances of objects in a scene; however, it cannot provide a quantitative measurement of the distances to objects. When the scene is imaged simultaneously from two locations, stereo correspondence between the resulting images can be used to determine the distance of objects. For example, Fig. 1 illustrates the imaging geometry for a binocular stereo imaging system comprised of cameras that have parallel optical axes, a common focal length f, and a baseline distance d. Let \( (x_L, y_L) \) and \( (x_R, y_R) \) denote the image coordinates of a three-dimensional scene coordinate \( (x, y, z) \) projected onto the image planes of the left and right cameras, respectively. Then, from perspective geometry, the image coordinates are related to the scene coordinates via the following relations:

\[
\frac{x_L}{f} = \frac{x + d/2}{z} \quad \frac{x_R}{f} = \frac{x - d/2}{z} \quad \frac{y_L}{f} = \frac{y}{z}
\]

If physically meaningful features can be identified at \( (x_L, y_L) \) and \( (x_R, y_R) \), and they are found to project from the same surface feature, then the spatial coordinates of the surface feature are as shown, where \( \Delta x = x_L - x_R \) is the disparity between the respective image coordinates [1]:

\[
x = \frac{d(x_L + x_R)}{2\Delta x} \quad y = \frac{d(y_L + y_R)}{2\Delta x} \quad z = \frac{df}{\Delta x}
\]

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 applications in autonomous vehicle navigation, industrial assembly and inspection, and three-dimensional microscopy. Although various techniques have been proposed for the extraction of such information, the absence of practical stereo vision systems indicates that certain aspects of the problem remain unsolved. The most critical task in stereo vision—finding correspondences between features detected in the left and right images—has received a considerable amount of attention, and the general principles regarding its solution are well understood. However, a solution that is both accurate and computationally efficient remains elusive. Current stereo vision algorithms have been strongly influenced by the feature-based approach developed by Marr and Poggio [2] and Grimson [3, 4], which uses intensity zero crossings—transitions through zero in the second derivative of smoothed intensity—as fundamental matching primitives. In their approach, the solution is constrained by limiting the search space according to the bandwidth of the zero-crossing operator used. More recent approaches have sought to increase accuracy by incorporating various physical/psychophysical constraints such as smoothness of disparity [5, 6], limits on the disparity gradient [6-8], and figural continuity [2, 9]. Other algorithms use more abstract concepts, such as surface interpolation [10], disparity functions [11], three views [12], and dynamic programming [13].

Although many promising results have been obtained through such investigations, a simpler, faster, and more robust solution to the correspondence problem is necessary if the needs of practical systems are to be met. Such a solution implies the need for faster computer architectures and, more importantly, a higher degree of accuracy, which can be obtained only through a more complete and efficient use of the information available in the stereo images. One mode of information that has been largely neglected in computational stereo algorithms is chromatic information. Although the use of chromatic information as an aid in computational stereo has been mentioned [14], a thorough investigation of its utility has never been performed. In this paper, we demonstrate that chromatic information can be used to significantly reduce the ambiguity between potential matches, while increasing the accuracy of the resulting matches. The motivations for using chromatic information are simple. First, chromatic intensity information is an easily obtained source of information; second, current psychophysical evidence indicates that chromatic information is used in

Fig. 1. Parallel axis imaging geometry for a binocular stereo imaging system.
human stereo fusion, although its specific role is unclear [15]-[21].

In this paper, the types of intensity and chromatic information available in visible images are discussed briefly, followed by descriptions of the simple algorithms used in these experiments. The paper concludes with a presentation of the results of applying each algorithm to three stereo image pairs and a discussion of the interpretation of these results. In keeping with largely accepted terminology, “intensity” will denote the camera output at each pixel. For this investigation, an intensity image is acquired for each of the red, green, and blue (RGB) spectra. The pixel-by-pixel sum of these spectral images yields the “gray-level” intensity image, which is typically called the “intensity image.” Hereafter, “Intensity Matching Algorithm” will refer to the stereo matching algorithm that uses only intensity information, while “Color Matching Algorithm” will refer to the stereo matching algorithm that uses chromatic information as well.

Description of the Experimental Stereo Algorithm

The approach that has been used to determine the potential usefulness of chromatic information in stereo algorithms is based on two very similar stereo matching algorithms differing only in the use (or lack of use) of chromatic information. In order to isolate the effects of including chromatic information, it has been used only to characterize spatially coincident zero crossings in the Laplacian-of-Gaussian (LoG) filtered gray-level (intensity) image. In this way, the performance of the two algorithms can be compared directly in terms of their disambiguation ability and matching accuracy. Very simple algorithms have been used intentionally so that the effects of including chromatic information in the early disambiguation stages can be illustrated most clearly. Despite their simple nature, these algorithms are important because they are comprised of operations that are used by a large class of stereo algorithms, namely those that follow the feature-based approach [2]-[8].

Intensity Information

After the left and right intensity images have been acquired, intensity features corresponding to significant physical changes in the environment must be determined. For the algorithms presented herein, zero crossings computed from the LoG-filtered intensity images are used. The LoG, which optimizes the trade-off between spatial localization and bandwidth, was chosen since it offers simplicity and uniformity while reducing noise and yielding closed zero-crossing contours [22]. The two-dimensional rotationally symmetric Gaussian used in this investigation is

\[ g_\sigma(s, y) = -K_\sigma \exp \left( -\frac{s^2 + y^2}{2\sigma^2} \right) \]  

where \( \sigma \) is the space constant of the Gaussian and \( K \) a scale factor. Thus, the LoG is

\[ \nabla^2 g_\sigma(x, y) = K \left( 2 - \frac{x^2 + y^2}{\sigma^2} \right) \exp \left( -\frac{s^2 + y^2}{2\sigma^2} \right) \]  

Zero crossings in the LoG-filtered images are detected by searching in both the horizontal and vertical directions for either a change of sign in adjacent pixel values or a transition through zero. In addition to its position, the zero crossing’s contrast sign and gradient orientation are recorded. To minimize noise effects, only those zero crossings that remain after the application of an adaptive double-thresholding procedure are used by the matching algorithm [23].

Intensity Matching Algorithm

After zero-crossing thresholding, potential matches between zero crossings from the left and right images must be found. For simplicity, the problem is reduced to a one-dimensional matching problem by assuming the images to be vertically registered (the epipolar constraint). Furthermore, a nonconvergent (parallel axis) imaging geometry is assumed so that the search space in the right image is constrained to lie to the left of the transformed left-image coordinates. Hence, the search space is both one-dimensional and unidirectional.

There is no apparent consensus concerning the appropriate search-space size to use when looking for candidate matches in the right image. Some algorithms have used statistical techniques to predict the location of zero crossings [2], while others have tried to emulate the human visual system by calculating an appropriate Panum’s fusional area [7]. In this study of the utility of chromatic information in solving the correspondence problem, it is not necessary that the search space be defined rigorously. Therefore, the minimum allowable disparity has arbitrarily been set to 20 percent of the image width. If anything, this search space is too large, making disambiguation more difficult; however, the algorithms still perform reasonably well.

Using the defined search space, the Intensity Matching Algorithm may be stated as follows: For each nonhorizontal zero crossing in the left image (horizontal zero crossings are eliminated because of their inherent ambiguity), determine the set of corresponding candidate matches in the right image that:

1. Lie within the search space;
2. Have the same contrast sign;
3. Have roughly the same orientation (±30 deg).

After determining all possible candidate matches for each zero crossing in the left image, the algorithm determines all zero crossings that have unique candidate matches—those that have only one candidate match and whose candidate match is not claimed by any other zero crossings. Zero crossings having unique candidate matches are considered matched and the algorithm terminates. In passing, it is important to note that a unique match is not necessarily a correct match.

Incorporation of Chromatic Information into the Stereo Algorithm

Chromatic Information

There are two methods of incorporating chromatic information into a stereo algorithm that seem both conceptually and computationally feasible: (1) find zero crossings in each of the color spectra images as well as in the intensity image, then (separately) match each set of zero crossings; or (2) simply use chromatic information to characterize the spatially coincident zero crossings from the intensity image. The first approach seems computationally inefficient since the matching is done effectively several times. Moreover, we have found that the zero crossings obtained from the chromatic spectra are rarely very different from those obtained from the intensity image; this is not surprising, as the chromatic spectra each have a strong intensity factor and are highly correlated [24]. Thus, the second approach has been chosen since it will require less computation and will work well over a more diverse range of images.

In order to use chromatic information to characterize zero crossings, an appropriate chromatic feature must be selected. The chromaticity at each zero crossing could be used; however, edges do not have chromaticities—rather they indicate a possible change in chromaticity. Therefore, it seems most appropriate to use chromatic gradients. The next question that must be answered is: What kind of chromatic information should be used to determine the chromatic gradients? Using Gaussian filtered intensity val-
ues from the red, green, and blue spectra is initially appealing; however, this approach has some disadvantages. First, the spectral intensity values may be sensitive to variations between the two cameras. Second, the red, green, and blue intensity values incorporate gray-level intensity information that already used elsewhere (in the zero crossings). An attractive alternative is to use normalized RGB values; since the chromaticity is approximately independent of the intensity, normalize the RGB values, factoring out the intensity. The normalized RGB values can be expressed in terms of the raw spectral intensity values as follows:

$$r(x, y) = \frac{R(x, y)}{R(x, y) + G(x, y) + B(x, y)}$$

(5a)

$$g(x, y) = \frac{G(x, y)}{R(x, y) + G(x, y) + B(x, y)}$$

(5b)

$$b(x, y) = \frac{B(x, y)}{R(x, y) + G(x, y) + B(x, y)}$$

(5c)

In order to fully characterize the manner in which the chromaticity is varying at each intensity zero crossing, this investigation uses three normalized difference spectra: red minus green ($D_{rg}$), green minus blue ($D_{gb}$), and blue minus red ($D_{br}$). Momentarily considering the red minus green spectrum

$$D_{rg}(x, y) = r(x, y) - g(x, y)$$

$$= \frac{R(x, y) - G(x, y)}{R(x, y) + G(x, y) + B(x, y)}$$

(6)

it is clear that the gradient of this spectrum at each zero crossing provides the manner and direction in which the spectrum is varying most rapidly. Thus, the sign and orientation of the gradient may be used as attributes of the chromatic gradient. The sign of the chromatic gradient is defined as the sign of the directional variation

$$\frac{\partial}{\partial x}(g_x * D_{rg}) = \frac{\partial}{\partial x} g_x * D_{rg}$$

(7)

computed along horizontal lines. At a given point in the image, a positive chromatic gradient sign for this spectrum indicates that the relative intensity of red to green is increasing (left to right) across the image, while a negative chromatic gradient sign indicates that it is decreasing. The orientation of the chromatic gradient is simply

$$\tan^{-1}\left(\frac{\partial}{\partial y}(g_x * D_{rg})/\partial x (g_y * D_{rg})\right)$$

(8)

The green-minus-blue and blue-minus-red difference spectra are defined analogously, as are their chromatic gradient signs and orientations, providing several descriptive attributes by which coincident zero crossings may be characterized.

**Color Matching Algorithm**

The only difference between the Intensity Matching Algorithm and the Color Matching Algorithm is the inclusion of a simple chromatic attribute constraint—the intensity zero crossings are still used as the matching features. The simplest constraint is to require that the candidate matches in the right image have the same chromatic gradient signs as the zero crossing in the left image. Thus, the Color Matching Algorithm may be stated as follows: For each nonhorizontal zero crossing in the left image, determine the set of corresponding candidate matches in the right image that:

1. Lie within the search space;
2. Have the same intensity contrast sign;
3. Have roughly the same orientation (±30 deg);
4. Have the same chromatic gradient sign for each of the normalized difference spectra.

After determining all possible candidate matches for all zero crossings in the left image, the algorithm determines those zero crossings that have unique candidate matches. Zero crossings that have unique candidate matches are considered matched and the algorithm terminates.

While it is possible to require that the left-image zero crossing and its candidate matches have roughly the same chromatic gradient orientation (e.g., within 30 deg), this information has not been used in the present investigation.

**Experimental Results**

Following feature extraction and zero-crossing thresholding, the Intensity Matching Algorithm and the Color Matching Algorithm were applied to three stereo image pairs—images of a model city scene taken in the Calibrated Imaging Laboratory at Carnegie-Mellon University, a Rubik’s cube, and a set of hand tools.

Figures 2-6 show intensity images, normalized difference spectra, thresholded zero crossings, nonhorizontal thresholded zero crossings, and intensity-encoded disparity maps for the stereo image pair “City” ($\alpha = 1.414$, search-space size = 48 pixels). The disparity maps are shown for each of the matching algorithms; intensity values represent relative disparities where higher intensity values correspond to larger disparities. The results of applying the stereo matching algorithms to this image pair are tabulated in Table 1 (there were 5937 nonhorizontal zero crossings in the left image seeking candidate matches). In this table, “Number of Zero Crossings Matched” denotes the number of zero crossings for which one or more candidate matches were found; “Total Number of Candidate Matches” denotes the total number of candidate matches found for all zero crossings; “Number of Unique Matches” refers to the number of zero crossings that were uniquely matched; and “Percent Unique Matches Correct” refers to the percentage of unique matches that were correct. Some incorrect matches are expected to occur due to the extremely simple nature of the matching algorithms.

A comparison of the results of the Intensity Matching Algorithm and the Color

![Fig. 2. Left and right intensity images for stereo image pair “City.”](image-url)
Fig. 3 Normalized difference spectra for the left image of the stereo image pair "City": (a) red minus green, (b) green minus blue, and (c) blue minus red.

Matching Algorithm for the image pair "City" reveals the obvious utility of chromatic information in solving the stereo correspondence problem. Four pronounced effects are observed. First, the total number of zero crossings having at least one candidate match decreased 39.53 percent when the Color Matching Algorithm was applied. Such a reduction probably indicates that the inclusion of chromatic information helps eliminate incorrect matches. Second, the total number of candidate matches decreased 60.78 percent when the Color Matching Al-
The Color Matching Algorithm was applied, indicating that the inclusion of chromatic information can assist greatly in the disambiguation of potential candidate matches. This reduction in candidate matches might provide significant computational savings in second-pass disambiguation algorithms. Third, the number of unique matches increased 70.85 percent when the Color Matching Algorithm was applied. Fourth, the percent of correct unique matches increased 0.86 percent when the Color Matching Algorithm was applied. While this is a rather small percentage, the fact that a large increase in the number of unique matches is accompanied by an equal or greater percentage of correct unique matches is itself an important indication of increased accuracy.

The corresponding percent changes in measured quantities for stereo image pairs “Cube” and “Tools” are summarized in Table 1. The rather large variations in the results are not surprising since the degree to which chromatic information assists stereo correspondence depends upon several factors, including zero-crossing density, zero-crossing similarity, and the images’ chromatic content. The greatest contribution of chromatic information is likely to occur in situations in which several candidate matches within the search space have the appropriate intensity attributes (contrast sign and gradient orientation) but only one has the appropriate chromatic gradient signs. Of course, such a situation is atypical—a point to which the three image pairs attest. In the “City” image pair, for example, there are many similar zero crossings within the search space; however, the chromatic information varies little and, thus, fails to disambiguate between candidate matches. In the “Cube” image pair, the zero crossings in the filtered image have similar attributes and the chromatic information is different at each zero crossing; however, the zero-crossing density is low enough that only one zero crossing with the appropriate intensity attributes lies within the search space. Thus, the large chromatic variations present in the image are less useful. Finally, in the “Tools” image pair, there are similar zero crossings within the search space; however, each has different chromatic signs, resulting in significant increases in the number of unique matches as well as the number of correct unique matches. Thus, the potential benefit of chromatic information in stereo matching algorithms depends on several factors; the presence of large variations in chromatic information does not, by itself, guarantee improved performance.

Finally, it is important to note that the increase in computational cost for the Color Matching Algorithm is negligible—assuming that the intensity and chromatic information is processed in parallel—since each zero crossing within the search space must be checked for potential correspondence regardless of the information used. The potential computational benefit of using chromatic information lies in subsequent stages of stereo processing, where the effect of a smaller number of candidate matches and a larger number of unique matches will become apparent.

**Table 1**

Summary of Results for Stereo Image Pair “City”

<table>
<thead>
<tr>
<th>Measured Quantity</th>
<th>Intensity Matching Algorithm</th>
<th>Color Matching Algorithm</th>
<th>Percent Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Zero Crossings Matched</td>
<td>4490</td>
<td>2715</td>
<td>-39.53</td>
</tr>
<tr>
<td>Total Number of Candidate Matches</td>
<td>9246</td>
<td>3626</td>
<td>-60.78</td>
</tr>
<tr>
<td>Number of Unique Matches</td>
<td>820</td>
<td>1401</td>
<td>+70.85</td>
</tr>
<tr>
<td>Percent Unique Matches Correct</td>
<td>77.56</td>
<td>78.23</td>
<td>+0.86</td>
</tr>
</tbody>
</table>

**Table 2**

Summary of Percent Changes in Measured Quantities for Stereo Image Pairs “Cube” and “Tools”

<table>
<thead>
<tr>
<th>Stereo Image Pair</th>
<th>Measured Quantity</th>
<th>“Cube”</th>
<th>“Tools”</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-14.87</td>
<td>-40.76</td>
</tr>
<tr>
<td></td>
<td>Total Number of Candidate Matches</td>
<td>-25.20</td>
<td>-60.44</td>
</tr>
<tr>
<td></td>
<td>Number of Unique Matches</td>
<td>+1.12</td>
<td>+163.25</td>
</tr>
<tr>
<td></td>
<td>Percent Unique Matches Correct</td>
<td>+1.00</td>
<td>+15.36</td>
</tr>
</tbody>
</table>

**Concluding Remarks**

This paper has described a preliminary investigation into the utility of chromatic information in solving the stereo correspondence problem. Two stereo matching algorithms differing only by the inclusion of simple chromatic information were implemented and applied to several stereo image pairs. The results of the investigation strongly indicate that using chromatic information can: (1) significantly reduce the ambiguity between candidate matches, (2) increase the number of unique matches, and (3) reduce the number of incorrect matches. As such, the use of chromatic information may greatly reduce the computational burden associated with current multistage stereo algorithms while increasing the accuracy of the resulting matches. We are currently experimenting with one well-accepted algorithm—the PMF Algorithm [7]—in order to investigate these issues.

**Acknowledgment**

The stereo image pair “City” was provided by the Calibrated Imaging Laboratory at Carnegie-Mellon University, sponsored by the NSF, DARPA, and NASA.

**References**


John R. Jordan III was born in Boston, Massachusetts, in 1962. He received the B.S. degree in electrical engineering (with highest honors) from the University of Notre Dame in 1984 and the M.S. degree in engineering from the University of Texas at Austin in 1985. He is currently a Research Assistant in the Laboratory for Vision Systems in the Department of Electrical and Computer Engineering at the University of Texas at Austin. His current research interests include computer vision, digital image processing, and psychophysical aspects of human visual perception. Mr. Jordan is a member of Tau Beta Pi, Eta Kappa Nu, and Phi Kappa Phi.

Alan C. Bovik was born in Kirkwood, Missouri, in 1958. He received the B.S. degree in computer engineering in 1980, and the M.S. and Ph.D. degrees in electrical and computer engineering in 1982 and 1984, respectively, all from the University of Illinois, Urbana-Champaign. He is currently an Assistant Professor in the Departments of Electrical and Computer Engineering and Biomedical Engineering at the University of Texas at Austin, where he is the Director of the Laboratory for Vision Systems in the Computer and Vision Research Center. His current research interests include the application of nonlinear statistical methods to problems in digital signal and image processing, biomedical image processing, computer vision, and computational aspects of biological visual perception. Dr. Bovik currently holds the Robert & Francis Stark Centennial Fellowship in Engineering.

1989 ACC—Call for Papers


The technical program will cover a wide range of topics relevant to the theory and practical implementation of control and industrial automation, and to university education in controls. Topics of interest will include linear and nonlinear systems, identification and estimation, signal processing, multivariable systems, large-scale systems, robotics and manufacturing systems, guidance and control, sensors, simulation, adaptive control, optimal control, expert systems, and control applications. Both contributed and invited papers will be included in the program of this conference.

**Preliminary Call for Contributed Papers:**

Prospective authors of regular papers should submit six (6) copies of the complete manuscript. Regular papers should be marked "1989 ACC" and should be submitted by September 15, 1988, to the Program Vice Chairman. If the authors request, the regular papers will be considered for journal publication provided the journal rules of submission for prospective manuscripts are followed. Short papers should consist of a 700-word (2-3 page) summary. The short papers, which will not be considered for journal publication, should be submitted by September 15, 1988, to the Program Vice Chairman for Contributed Papers. All regular and short papers should clearly describe the problem, the analytical or experimental techniques employed, and results obtained and relevant references.

**Preliminary Call for Invited Sessions:**

The Program Committee is soliciting proposals for invited sessions for the conference. Prospective organizers should contact the Program Vice Chairman for Invited Sessions before September 15, 1988.

General Chairman: H. Vincent Poor Coordinated Science Laboratory University of Illinois at Urbana-Champaign 1101 West Springfield Ave., Urbana, IL 61801; Phone: (217) 333-6449.

Program Chairman: Marija Illic, LEECS, Bldg. 10-029, Massachusetts Institute of Technology, Cambridge, MA 02139; Phone: (617) 253-4682.

Continued on p. 53.