EVALUATION OF DIFFERENT METHODS FOR USING COLOUR INFORMATION IN GLOBAL STEREO MATCHING APPROACHES

Michael Bleyer[1], Sylvie Chambon[2], Uta Poppe[1], Margrit Gelautz[1]

[1]Institute for Software Technology and Interactive ystems
Vienna University of Technology
Favoritenstrasse 9-11/188/2, A-1040 Vienna, Austria
[bleyer, poppe, gelautz]@ims.tuwien.ac.at

[2]Laboratoire Central des Ponts et Chaussées (LCPC)
Route de Pornic, BP 4129
44341 Bouguenais Cedex, Nantes, France
chambon@lcpc.fr

Commission III/2

KEY WORDS: Stereoscopic Matching, Stereoscopic Reconstruction, Colour Analysis, Experimental Analysis;

ABSTRACT:

Global algorithms currently represent the state-of-the-art in dense stereo matching. These methods first set up an energy function. The energy function is then subject to optimization, which is typically achieved via graph-cuts or belief propagation. In this paper, we concentrate on the energy modelling aspect. An experimental study that focuses on the role of colour in stereo energy functions is presented. We evaluate the performance of various forms for using colour and compare it against grey-scale matching. Colour is thereby represented in nine different colour systems. The L_1 and L_2 distances are evaluated for computing the colour differences in the selected systems. We embed the resulting energy functions into two stereo algorithms and test them on 30 ground truth test image pairs. The results of our benchmark show that colour information, in general, leads to a significant performance gain over using intensity only. According to our evaluation results, the selection of the applied colour space is of specific importance in global stereo matching.

1 INTRODUCTION

During the last couple of years, global stereo approaches have gained increased attention in the stereo vision community for their excellent performance in stereo algorithm evaluation studies such as the Middlebury Benchmark (Scharstein and Szeliski, 2002). These methods formulate the stereo problem in terms of an energy function, which is typically in the form of

$$E = E_{data} + E_{smooth}. \quad (1)$$

Here, the data term E_{data} assesses the agreement of the disparity solution with the input images by computing a match measurement, while the smoothness term E_{smooth} imposes a penalty on spatially neighbouring pixels carrying different disparity labels.

There has been a significant amount of work on minimizing the energy of (1), which is nowadays typically accomplished using graph-cuts (Boykov et al., 2001) or belief propagation (Sun et al., 2003). However, it has often been overlooked that the energy functions under consideration might represent suboptimal models for the stereo problem. For example, (Meltzer et al., 2005) have shown that, despite the NP-hardness of their optimization problem, an exact optimum can be obtained for some standard benchmark stereo pairs using reweighted message passing. Nevertheless, even the global energy minimum has led to disparity maps that show relatively large errors in comparison to the ground truth image. This clearly indicates that real progress in global stereo matching can rather be achieved by improving the energy functions than by concentrating on the optimization component.

The contribution of this paper lies in a systematic evaluation study on stereo energy functions. We thereby focus on the role of colour in the energy functions' data terms. The intuition why colour information should lead to an improved energy model over using intensity information only is relatively clear. Colour is expected to reduce one of the major problems in stereo matching, namely matching ambiguity. For example, suppose that we are solely using the intensity information in the matching process. Then a red pixel of the left image matches a green point and the correct red pixel of the second view equally well, if the red and green colours both project to the same intensity value. This ambiguity is obviously resolved by using the colour information.

Our work is motivated by the observation that a lot of stereo researchers still simply convert the stereo pairs to grey-scale images, although colour is typically available (Mayer, 2003). Since it is unclear if colour shows positive effects when using global methods, the colour information is thereby often discarded deliberately. Therefore, this work concentrates on two questions. First, does colour help to improve the performance of global stereo matching approaches? Second, in which form should colour be used to maximize the algorithms' quantitative performance?

In the context of prior work, colour evaluation studies have already been conducted in (Okutomi and Tomita, 1992, Koschan, 1993, Chambon and Crouzil, 2005). All of these studies are restricted to local methods (window-based correlation), and a corresponding study for global algorithms is still missing. Some experiments focusing on different match measurements have been reported in (Scharstein and Szeliski, 2002). The role of colour has, however, not been investigated. Two stereo evaluation studies have also appeared very recently. As opposed to our work, they address radiometric invariant dissimilarity measurements (Hirschmüller and Scharstein, 2007) and different aggregation methods in local stereo (Wang et al., 2006).

The remainder of this paper is organized as follows. We start by formulating the energy functions evaluated in this paper in section 2. Section 3 then presents two stereo algorithms that incorporate these energy functions and are used in this benchmark. Section 4 presents the test data and provides details on the disparity computation. Finally, we report our results in section 5.

2 ENERGY FUNCTIONS

Let $\mathcal{I}$ denote the set of all pixels in the left image and $\mathcal{D}$ be the set of allowed disparity labels. A disparity solution D maps each pixel $p \in \mathcal{I}$ to a disparity $d_p \in \mathcal{D}$. The goodness of a disparity map D is evaluated by an energy function, which we define by

$$E(D) = \underbrace{\sum_{p \in \mathcal{I}} m(p, p - d_p)}_{E_{data}} + \underbrace{\sum_{(p,p') \in \mathcal{N}} s(d_p, d_{p'})}_{E_{smooth}}. \quad (2)$$

Here, $m(p, p - d_p)$ is a function that computes the colour dissimilarity between a pixel p and its matching point $p - d_p$ in the second view. The smoothness function $s(d_p, d_{p'})$ penalizes neighbouring pixels that are assigned to different disparities. The neighbourhood structure $\mathcal{N}$ thereby contains all pixel pairs (p, p') of the left view with p and p' being spatial neighbours in four-connectivity. The functions $m(\cdot)$ and $s(\cdot)$ of data and smoothness terms are defined in the following.

2.1 Data Term

The data term is the subject of our study. This term computes the colour dissimilarities between corresponding pixels in a predefined colour space. The choice of the colour system thereby has direct influence on the estimation of the data term. This is why, apart from using intensity information only, we evaluate the performance of nine different colour systems. The investigated colour spaces are categorized into

- Primary systems: RGB and XYZ;
- Luminance-chrominance systems: LUV, LAB, AC_1C_2 and YC_1C_2;
- Perceptual systems: HSI;
- Statistical independent component systems: $I_1I_2I_3$ and $H_1H_2H_3$.

Table 1 plots corresponding conversion formulas. For the references associated to each colour system, the reader is referred to (Chambon and Crouzil, 2005).

We evaluate two difference measurements in order to compute the pixel dissimilarity in a given colour space. The first one is the L_1 distance, which represents the summed-up absolute differences of colour channels. The corresponding dissimilarity function is defined by

$$f^{L_1}(p, q) = \sum_{1 \le i \le 3} |p_i - q_i| \tag{3}$$

with p and q being pixels of the left and right images, respectively. The subscript i denotes the ith colour channel in the selected colour system. As a second difference measurement, we compute the L_2 distance. This represents the Euclidean distance between two points in the colour space. In this case, the dissimilarity function is given by

$$f^{L_2}(p, q) = \sqrt{\sum_{1 \le i \le 3} (p_i - q_i)^2}. \tag{4}$$

There are two exceptions where the difference measurements of equations (3) and (4) are not suitable. First, for grey-scale matching, we define the dissimilarity function as

$$f^{Grey}(p, q) = |p_I - q_I| \tag{5}$$

with the subscript I denoting the intensity channel. Second, for the HSI space, we adopt the measurement proposed by (Koschan, 1993). This measurement is computed by

$$\begin{aligned} f^{HSI}(p, q) &= \sqrt{(p_I - q_I)^2 + p_S^2 + q_S^2 + 2p_S q_S \cos\theta} \\ \theta &= \begin{cases} |p_H - q_H| & \text{if } |p_H - q_H| < \pi \\ 2\pi - |p_H - q_H| & \text{otherwise.} \end{cases} \end{aligned} \tag{6}$$

Here, the subscripts H, S and I denote the corresponding colour channels of the HSI system.

As a final step in the calculation of the data term, we apply the measurement of (Birchfield and Tomasi, 1998) to reduce the negative effects of image sampling on the disparity reconstruction performance. We modify this measurement to make it applicable on colour pixels. We therefore compute the colour values of a pixel p^- that lies in between p and its spatially left neighbour p^l. This is accomplished by linear interpolation so that

$$p_i^- = \frac{p_i + p_i^l}{2} \qquad 1 \le i \le 3 \tag{7}$$

with the subscript i being the ith channel in the chosen colour system. Analogously, p^+ is determined as the colour point that is located between p and its right neighbour p^r. For symmetry, we also compute q^- and q^+ for the matching point q in the second view. The final function $m(p, q)$ of equation (2) is then given by

$$\begin{aligned} m(p, q) = \min(&f(p, q), f(p^-, q), f(p^+, q), \\ &f(p, q^-), f(p, q^+)) \end{aligned} \tag{8}$$

with $f(\cdot)$ being one of the difference functions in equations (3-6) that is suitable for the selected colour space. Our energy functions differ in the way $f(\cdot)$ is computed. Using two difference measurements for eight different colour systems and a single distance measurement for $Grey$ and HSI, this sums up to 18 energy functions investigated in this study.

Name	Definition
XYZ	$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$
LUV	$L = \begin{cases} 116\,(Y/Y_w)^{\frac{1}{3}} - 16 & \text{if } Y/Y_w > 0.01 \\ 903.3\; Y/Y_w & \text{otherwise} \end{cases}$ $U = 13L(u' - u'_w)$ with $u' = \frac{4X}{X+15Y+3Z}$ $V = 13L(v' - v'_w)$ with $v' = \frac{9Y}{X+15Y+3Z}$ X_w, Y_w, Z_w: white reference components
LAB	$A = 500(f(X/X_w) - f(Y/Y_w))$ $B = 200(f(Y/Y_w) - f(Z/Z_w))$ $f(x) = \begin{cases} x^{1/3} & \text{if } x > 0.008856 \\ 7.787x + \frac{16}{116} & \text{otherwise} \end{cases}$
AC_1C_2	$\begin{pmatrix} A \\ C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{\sqrt{3}}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{-1}{2} & \frac{-1}{2} & 1 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$
YC_1C_2	$\begin{pmatrix} Y \\ C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & \frac{-1}{2} & \frac{-1}{2} \\ 0 & \frac{-\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$
HSI	$I = \frac{R+G+B}{3}$, $S = 1 - 3\frac{\min(R,G,B)}{R+G+B}$ $H = \begin{cases} \arccos H_1 & \text{if } B \le G \\ 2\pi - \arccos H_1 & \text{otherwise} \end{cases}$ $H_1 = \frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}$
$I_1I_2I_3$	$\begin{pmatrix} I_1 \\ I_2 \\ I_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{-1}{2} \\ \frac{-1}{4} & \frac{-1}{4} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$
$H_1H_2H_3$	$\begin{pmatrix} H_1 \\ H_2 \\ H_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ \frac{-1}{2} & 0 & \frac{-1}{2} \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$

Table 1: Conversions from RGB to the investigated colour spaces.

2.2 Smoothness Term

The smoothness term of our energy function is implemented as a modified version of the Potts model. It is better suited for handling slanted surfaces than the standard Potts model. We define the smoothness function $s(d_p, d_{p'})$ of equation (2) by

$$s(d_p, d_{p'}) = \begin{cases} 0 & \text{if } d_p = d_{p'} \\ P_1 & \text{if } |d_p - d_{p'}| = 1 \\ P_2 & \text{otherwise} \end{cases} \quad (9)$$

with P_1 and P_2 being user-defined penalties. A small penalty P_1 serves to penalize small jumps in disparity that do not exceed a value of one pixel. P_1 is motivated by the fact that the standard Potts model tends to overpenalize such small jumps and consequently performs poorly in the reconstruction of slanted surfaces. The second penalty P_2 with $P_2 > P_1$ accounts for penalizing large jumps in disparity. Such jumps occur at the disparity borders. In order to align disparity discontinuities with the edges of the reference view, we vary the value of P_2 depending on the colour gradient. This is accomplished by

$$P_2 = \begin{cases} P_3 \cdot P_2' & \text{if } \sum_{1 \le i \le 3} |p_i - p_i'| < T \\ P_2' & \text{otherwise.} \end{cases} \quad (10)$$

The colour gradient is thereby computed in RGB space. Throughout our test runs, we use a fixed value of 2 for P_3 and T is set to 30. Parameter settings for P_1 and P_2' are discussed in section 4.

3 STEREO ALGORITHMS

We embed the energy functions of section 2 into two different stereo algorithms. Optimization of the energy in equation (2) is known to be NP-complete. Both algorithms therefore only provide an approximation of the energy minimum. We have chosen to apply two different stereo methods in order to also investigate the influence of different optimization schemes on the results of our evaluation.

The first stereo algorithm is referred to as graph-cut (GC) method. The algorithm uses the graph-cut-based α-expansion algorithm (Boykov et al., 2001) in order to optimize energy (2). We include this method into our benchmark, since graph-cuts represent a widely adopted standard method for minimizing energy functions of this type. To derive an implementation of the α-expansion algorithm, we incorporate our energy functions into the MRF framework provided by (Szeliski et al., 2006). Note that our energy functions do not account for the occlusion problem. We will therefore only evaluate the matching error in non-occluded image regions in order not to corrupt our benchmark results by large error percentages in occluded areas.

The second stereo method (Bleyer and Gelautz, 2008) is referred to as dynamic programming (DP) method. It relies on a special form of tree-based dynamic programming to approximate the energy minimum of equation (2). The method builds two separate tree graphs for each individual pixel p of the reference view as shown in Figure 1. The global energy minima for both trees rooted on p are then efficiently computed using dynamic programming. Finally, the results of these trees are combined to derive p's disparity. Although we do not evaluate the error in occluded regions, it is pointed out that, as opposed to the GC method, the DP algorithm incorporates a method for occlusion handling. We have chosen this method, since it is less affected by the scanline streaking problem than other dynamic programming approaches. In comparison to the GC method, the DP algorithm has the advantage of being significantly faster.

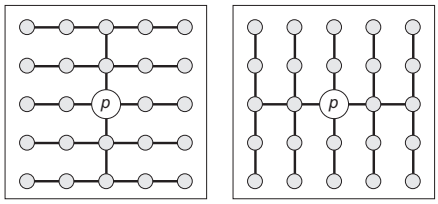


Figure 1: Trees applied by the DP method for a single pixel p. Nodes represent image pixels, while edges indicate that the smoothness function $s(\cdot)$ operates between adjacent pixels.

4 TEST DATA AND DISPARITY COMPUTATION

Figure 2 shows the test image pairs along with the corresponding ground truth images that are used in our study. All images are taken from the Middlebury Ground Truth Database (Scharstein and Szeliski, 2002, Hirschmüller and Scharstein, 2007). Our test data includes the four stereo pairs that are currently used in the Middlebury Benchmark (Tsukuba, Venus, Teddy and Cones). We refer to these image pairs as the 2003 sets. In addition, we include six stereo pairs of the 2005 data sets and 20 pairs of the 2006 sets. This sums up to 30 test pairs, which provides a reasonable amount of test data. The images of the 2005 and 2006 sets have been generated to be more challenging than those of the 2003 sets that can virtually be regarded as solved. This is why algorithms produce high error percentages on the new test sets. These data sets should therefore be well suited to discriminate the performance of different colour methods.

A crucial point when using global stereo methods is parameter tuning. For our energy functions, we have to find appropriate values for the parameters P_1 and P_2'. These parameters balance the data term against the smoothness term. This balance is obviously changed by applying different data terms. For this reason, we estimate an individual setting of P_1 and P_2' for each of the 18 investigated energy functions. Moreover, we use two different parameter settings depending on which stereo method is applied. Regarding our test data, we have chosen to use a single setting for the 2003 sets. A second setting is estimated for the 2005 data and a third one for the 2006 data. This results into $18 \cdot 2 \cdot 3$ settings of P_1 and P_2'. To determine the parameters, we compute the disparity maps for the corresponding test sets using varying values for P_1 and P_2'. (We have tested approximately 100 different parameter combinations.) We then select this parameter setting that shows the smallest average error percentage, which is computed by comparison against the ground truth images. Although this parameter tuning step is a tedious task, it is required to keep our comparison of energy functions fair.

5 EXPERIMENTAL RESULTS

Figure 3 plots a comparison of grey-scale matching with colour matching based on the RGB and LUV colour spaces. We show results for both stereo methods used in combination with the L_1 difference measurement. The applied error metric is the percentage of wrong pixels in unoccluded regions having an absolute disparity error larger than one pixel. It is surprising that, according to the plots of Figure 3, it is a good option not to use colour at all for the four stereo pairs of the 2003 data sets that are currently used in the Middlebury Benchmark. At least for the Venus and Teddy sets, colour even seems to worsen the results regardless of the applied stereo algorithm. In fact, similar experiences on these images, which are extremely popular in the stereo community, might have been a reason that has led researchers away from using colour in global stereo. However, results of both stereo algorithms on the 26 new test images of the 2005 and 2006 test sets

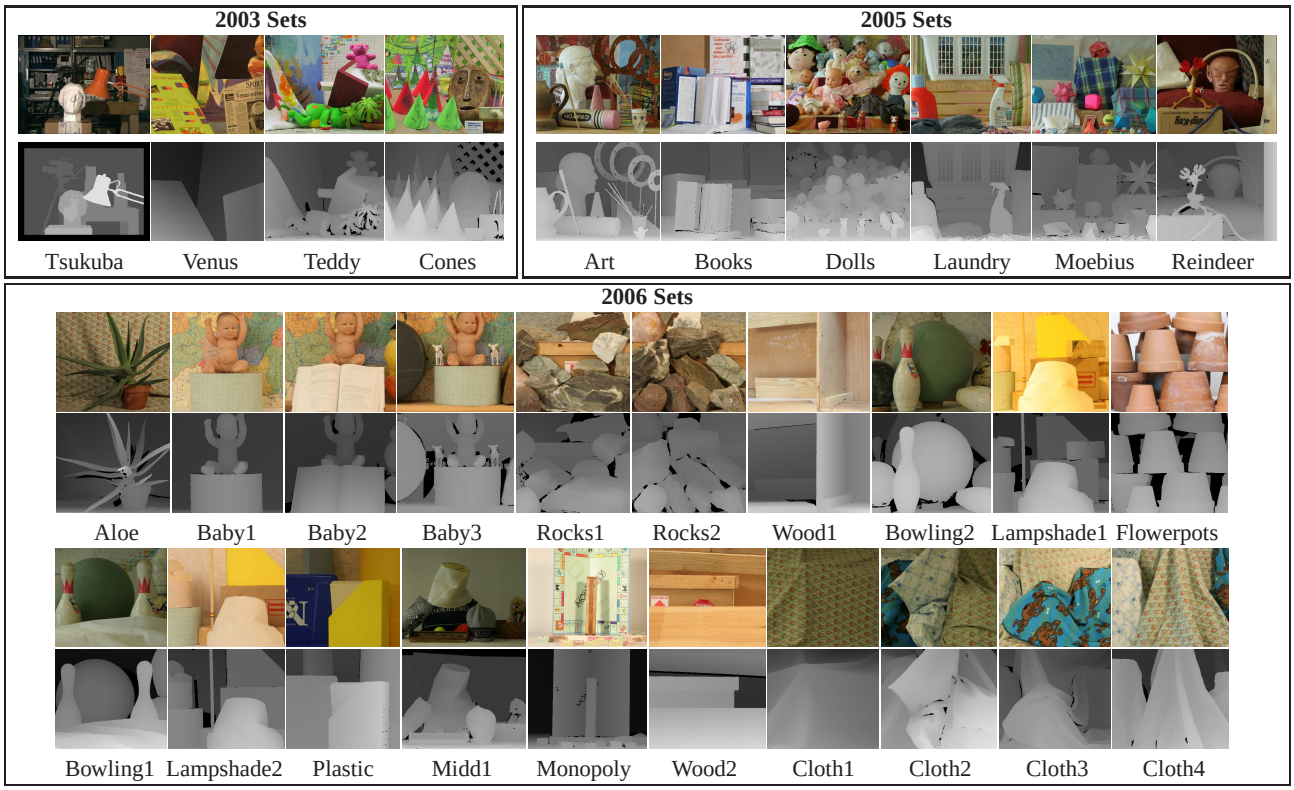


Figure 2: Test sets used in this study. Left views of the stereo pairs and corresponding ground truth images are shown.

clearly speak a different language. The results of intensity-based matching is clearly inferior in comparison to RGB and LUV. Especially, the LUV space performs well on these images.

We show disparity and error maps for three selected stereo pairs computed by the DP (Figure 4) and the GC methods (Figure 5) that are used in conjunction with the L_1 difference measurement. The error maps are derived by plotting pixels whose absolute disparity error is larger than one pixel. Black pixels in the error maps represent errors in non-occluded regions, while grey pixels correspond to errors in occluded areas. It is seen from these error maps that grey-scale matching generates less wrong pixels on the Teddy image pair than RGB and LUV. For the Dolls and Reindeer test images, the opposite observation is made.

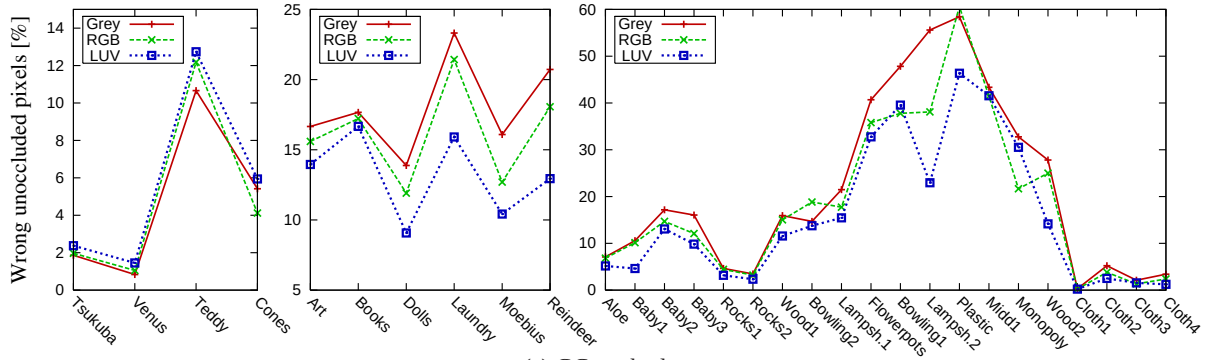
Figure 6 summarizes the overall results of our colour evaluation study. The tables plot four different combinations of stereo algorithms and difference measurements. Two performance measurements are computed. First, the error percentages of pixels exceeding an error threshold of one in unoccluded image areas are determined. Second, for each test pair, we rank the colour spaces according to their error percentages, so that the colour space with the lowest error percentage receives rank 1, while the worst-performing one is given rank 10. We compute the average errors given in % (*Avg. Error*) and average ranks (*Avg. Rank*) over all 30 test sets (*All Sets*) and considering the three test sets separately (*2003*, *2005* and *2006 Sets*). Similar to the Middlebury Online Table, the tables are sorted to derive a performance ranking of the investigated colour spaces. We sort the tables according to the values of *Avg. Rank* using the data sets *All Sets*.

It is seen that the order in which the colour spaces appear in the four tables of Figure 6 remains almost constant. The luminance-chrominance colour spaces LUV, AC_1C_1 and YC_1C_1 show the best performance according to our benchmark. This result is also interesting from a psychological point of view, since luminance-chrominance systems are very close to human perception. The $I_1I_2I_3$ system performs slightly better than RGB. The fifth rank for the RGB system is, however, surprisingly poor when considering that RGB is typically chosen for integrating colour into global approaches. Ranks 6 to 8 are taken by XYZ, $H_1H_2H_3$ and LAB. The worst performance is obtained for grey-scale matching, although HSI performs inferior in some cases.

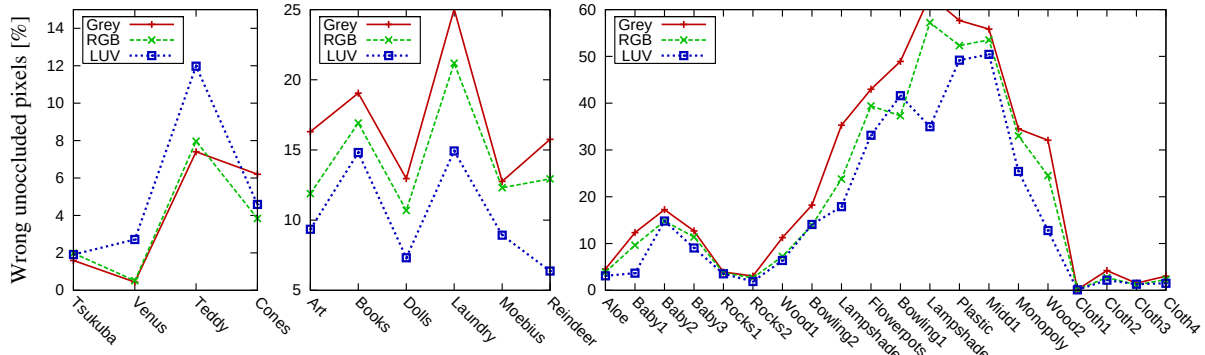
The tables of Figure 6 also allow measuring the amount of improvement that is achieved by using colour information. Let us therefore focus on the L_1 difference that is most commonly applied in stereo algorithms. When looking at the average error percentages (*Avg. Error*) for the GC method in Figure 6a, we obtain a value of 18.5% on *All Sets* for grey-scale matching in comparison to 13.8% for LUV. In fact, this results into 25.4% lower error rates, which represents a significant improvement. For the DP method (Figure 6a), we even derive an improvement of 29.0%. When determining the performance improvement of LUV in comparison to RGB, we determine 14.8% lower error rates for the GC method and 17.0% for the DP algorithm. The tables also allow for a comparison of the L_1 and L_2 difference measurements. This is accomplished by comparing the average error percentages (*Avg. Error*) of the tables in Figures 6a and 6b against those of Figures 6c and 6d. As opposed to the use of different colour systems, there are only relatively small differences in the performance of L_1 and L_2. It is, however, recognized that L_1 seems to perform slightly better. Surprisingly, as can be depicted from the tables, also the application of two different stereo algorithms seems to have much less influence on the error percentages than the choice of the colour system. The important decision in global stereo matching therefore seems to be the selection of a well-suited colour space.

6 CONCLUSIONS AND FUTURE WORK

This paper has investigated the role of colour information in global stereo matching approaches. We have provided a systematic experimental evaluation of various colour energy functions that are

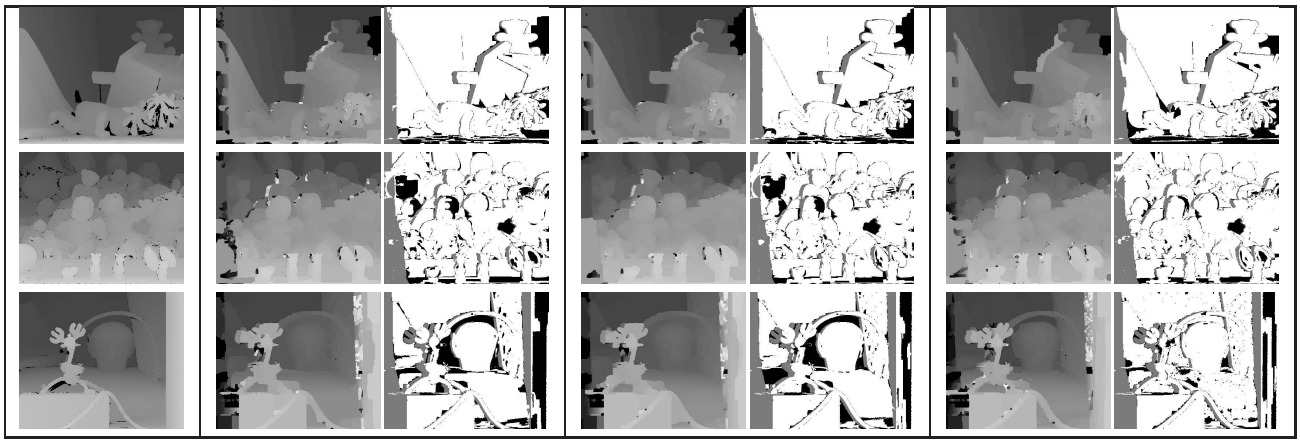


(a) GC method.



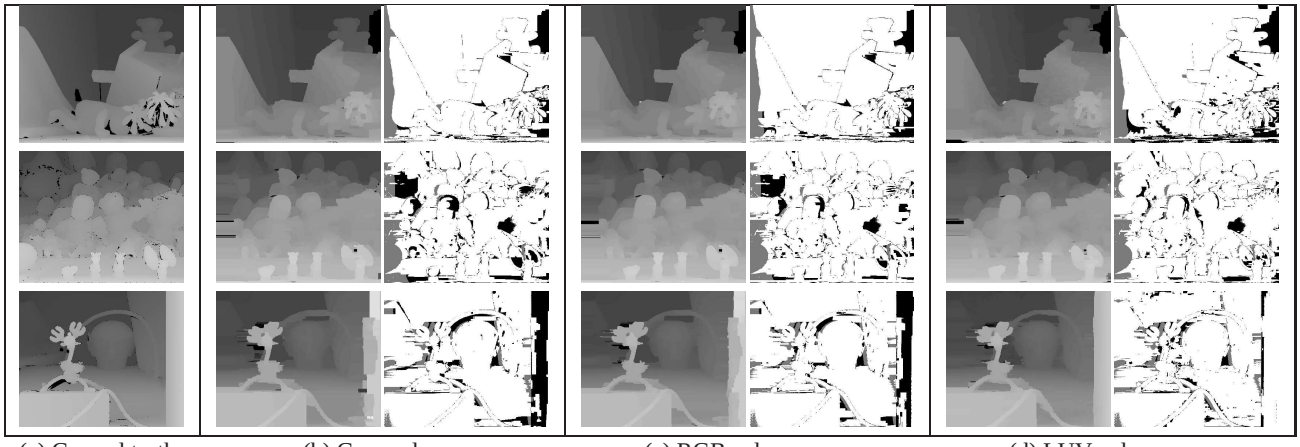
(b) DP method.

Figure 3: Error plots for the GC and DP methods using the L_1 difference for the 2003 data set (left column), the 2005 data set (middle column) and the 2006 data set (right column).



(a) Ground truth (b) Grey values (c) RGB colour space (d) LUV colour space

Figure 4: Disparity and error maps of the GC method using the L_1 difference applied on the Teddy, Dolls and Reindeer test sets.



(a) Ground truth (b) Grey values (c) RGB colour space (d) LUV colour space

Figure 5: Disparity and error maps of the DP method used in combination with the L_1 difference for Teddy, Dolls and Reindeer.

Colour Space	All Sets		2003 Sets		2005 Sets		2006 Sets	
	Avg. Rank↓	Avg. Error	Avg. Rank	Avg. Error	Avg. Rank	Avg. Error	Avg. Rank	Avg. Error
LUV	3.3_1	13.8_1	9.5_{10}	5.6_{10}	1.5_1	13.2_1	2.7_1	15.6_1
AC_1C_2	3.3_1	14.6_2	7.0_9	5.2_8	2.8_3	14.4_3	2.7_1	16.6_2
YC_1C_2	3.5_3	15.0_3	6.8_7	5.2_9	2.2_2	13.9_2	3.3_3	17.4_4
$I_1I_2I_3$	4.1_4	15.1_4	6.3_6	5.0_7	3.7_4	14.9_4	3.8_4	17.2_3
RGB	5.5_5	16.2_5	5.0_5	4.8_5	6.2_6	16.2_5	5.4_5	18.6_5
$H_1H_2H_3$	5.7_6	16.8_7	2.8_1	4.6_1	6.0_5	16.2_6	6.3_6	19.5_7
XYZ	6.1_7	16.5_6	3.8_3	4.7_3	6.7_7	16.2_7	6.4_7	18.9_6
LAB	7.2_8	18.4_8	3.5_2	4.6_2	7.7_8	16.7_8	7.9_8	21.7_{10}
Grey	8.0_9	18.5_{10}	3.8_3	4.7_4	9.5_{10}	18.1_9	8.5_{10}	21.4_9
HSI	8.2_{10}	18.4_9	6.8_7	4.9_6	8.8_9	18.6_{10}	8.3_9	21.1_8

(a) GC method with L_1 distance.

Colour Space	All Sets		2003 Sets		2005 Sets		2006 Sets	
	Avg. Rank↓	Avg. Error	Avg. Rank	Avg. Error	Avg. Rank	Avg. Error	Avg. Rank	Avg. Error
LUV	2.8_1	13.7_1	7.8_{10}	5.3_{10}	1.0_1	10.3_1	2.3_1	16.3_1
AC_1C_2	3.4_2	14.8_3	5.8_5	4.2_8	2.5_2	11.9_3	3.3_3	17.8_3
YC_1C_2	3.4_2	14.7_2	6.8_8	4.3_9	2.5_2	11.8_2	3.1_2	17.6_2
$I_1I_2I_3$	4.7_4	15.3_4	6.5_6	4.0_6	4.0_4	12.6_4	4.6_4	18.4_4
RGB	4.8_5	16.5_5	4.0_3	3.6_2	5.7_5	14.3_5	4.7_5	19.7_6
XYZ	5.9_6	16.5_6	6.8_8	3.8_4	6.7_7	14.9_6	5.5_6	19.6_5
$H_1H_2H_3$	6.1_7	17.7_7	3.8_2	3.7_3	6.3_6	15.1_7	6.5_7	21.3_7
LAB	7.4_8	18.8_9	3.3_1	3.4_1	8.0_8	16.0_8	8.0_8	22.7_9
HSI	8.0_9	18.2_8	6.5_6	4.0_7	8.7_9	16.5_9	8.1_9	21.6_8
Grey	8.6_{10}	19.3_{10}	4.0_3	3.9_5	9.7_{10}	17.0_{10}	9.2_{10}	23.1_{10}

(b) DP method with L_1 distance.

Colour Space	All Sets		2003 Sets		2005 Sets		2006 Sets	
	Avg. Rank↓	Avg. Error	Avg. Rank	Avg. Error	Avg. Rank	Avg. Error	Avg. Rank	Avg. Error
LUV	3.2_1	14.5_1	9.3_{10}	5.8_{10}	1.8_1	13.4_1	2.5_1	16.6_1
YC_1C_2	3.8_2	15.6_3	8.3_9	5.2_9	2.3_2	14.5_2	3.4_2	18.0_3
AC_1C_2	4.0_3	15.5_2	6.8_8	5.1_8	3.7_3	14.8_3	3.6_3	17.8_2
$I_1I_2I_3$	4.8_4	16.2_5	6.3_7	4.9_6	4.0_4	15.3_4	4.7_4	18.8_5
RGB	5.1_5	16.7_6	3.3_1	4.8_4	5.0_5	16.1_5	5.5_6	19.2_6
XYZ	5.4_6	16.0_4	5.5_6	5.0_7	6.3_7	16.4_7	5.1_5	18.1_4
$H_1H_2H_3$	6.6_7	17.9_7	3.5_2	4.8_3	6.2_6	16.2_6	7.4_7	21.0_8
LAB	6.8_8	18.3_9	3.5_2	4.6_1	6.7_8	16.5_8	7.5_8	21.6_{10}
Grey	7.6_9	18.5_{10}	3.5_2	4.7_2	9.2_9	18.1_9	8.0_{10}	21.4_9
HSI	7.7_{10}	17.9_8	5.3_5	4.9_5	9.8_{10}	18.3_{10}	7.5_9	20.4_7

(c) GC method with L_2 distance.

Colour Space	All Sets		2003 Sets		2005 Sets		2006 Sets	
	Avg. Rank↓	Avg. Error	Avg. Rank	Avg. Error	Avg. Rank	Avg. Error	Avg. Rank	Avg. Error
LUV	3.1_1	13.7_1	8.8_{10}	6.2_{10}	1.0_1	10.5_1	2.6_1	16.2_1
YC_1C_2	3.5_2	14.9_2	6.5_6	4.4_8	2.3_2	12.3_2	3.3_3	17.7_3
AC_1C_2	3.5_3	14.9_3	7.0_8	4.4_9	2.7_3	12.3_3	3.1_2	17.7_2
RGB	4.4_4	16.1_5	4.0_4	3.6_3	5.0_4	13.5_5	4.3_4	19.3_6
$I_1I_2I_3$	5.1_5	16.0_4	7.3_9	4.1_7	5.0_4	13.3_4	4.7_5	19.2_4
XYZ	5.4_6	16.2_6	5.5_5	3.8_4	5.7_6	14.0_6	5.3_6	19.3_5
$H_1H_2H_3$	6.4_7	17.7_7	2.8_1	3.5_2	6.5_7	14.9_7	7.2_7	21.4_7
LAB	7.2_8	18.6_9	2.8_1	3.4_1	7.8_8	15.4_8	8.0_9	22.5_9
HSI	7.8_9	18.2_8	6.8_7	4.0_6	9.2_9	16.5_9	7.7_8	21.6_8
Grey	8.5_{10}	19.3_{10}	3.8_3	3.9_5	9.8_{10}	17.0_{10}	9.1_{10}	23.1_{10}

(d) DP method with L_2 distance.

Figure 6: Quantitative performance of the investigated colour spaces. Average ranks and average error percentages are plotted. The subscripts represent the rank of a value in the table. More explanation is given in the text.

generated by combining different distance measurements with several colour systems in the energy functions' data terms. These energy functions are embedded into two stereo algorithms and tested on 30 stereo image pairs for which ground truth data is available.

Our results show that colour, in general, improves the results of global stereo methods. In fact, the performance gain in our benchmark is relatively high. We report approximately 25% less disparity errors when using the best-performing colour system instead of grey-scale matching. However, it has also been recognized that colour does not necessarily improve the performance on four frequently used image pairs, namely the current Middlebury Evaluation sets. The best-performing method for incorporating colour rather depends on the selected colour system than on the applied difference measurement. The best-performing colour spaces of our study are three luminance-chrominance systems LUV, AC_1C_2 and YC_1C_2. This is specifically interesting, since they are very close to human perception. An interesting result is also that RGB, which is the most popular colour representation in stereo matching, only gives results of average quality.

Our current study concentrates on a single prior in the energy formulation. We have investigated this prior, since it is frequently used in the literature. Although we believe that this will not substantially change the outcome of our study, other priors need to be investigated in future work.

ACKNOWLEDGEMENT

Michael Bleyer would like to acknowledge the Austrian Science Fund (FWF) for financial support under project P19797.

REFERENCES

Birchfield, S. and Tomasi, C., 1998. A pixel dissimilarity measure that is insensitive to image sampling. TPAMI 20(4), pp. 401–406.

Bleyer, M. and Gelautz, M., 2008. Simple but effective tree structures for dynamic programming-based stereo matching. In: VISAPP, Vol. 2, pp. 415–422.

Boykov, Y., Veksler, O. and Zabih, R., 2001. Fast approximate energy minimization via graph cuts. TPAMI 23(11), pp. 1222–1239.

Chambon, S. and Crouzil, A., 2005. Colour correlation-based matching. Int J Robot Autom 20(2), pp. 78–85.

Hirschmüller, H. and Scharstein, D., 2007. Evaluation of cost functions for stereo matching. In: CVPR, pp. 1–8.

Koschan, A., 1993. Dense stereo correspondence using polychromatic block matching. In: CAIP, Vol. 719 of LNCS, pp. 538–542.

Mayer, H., 2003. Analysis of means to improve cooperative disparity estimation. In: IAPRSIS, Vol. XXXIV, (Part 3/W8), pp. 25–31.

Meltzer, T., Yanover, C. and Weiss, Y., 2005. Globally optimal solutions for energy minimization in stereo vision using reweighted belief propagation. In: ICCV, Vol. 1, pp. 428–435.

Okutomi, M. and Tomita, G., 1992. Color stereo matching and its application to 3-d measurement of optic nerve head. In: ICPR, Vol. 1, pp. 509–513.

Scharstein, D. and Szeliski, R., 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. IJCV 47(1/2/3), pp. 7–42.

Sun, J., Zheng, N. and Shum, H., 2003. Stereo matching using belief propagation. TPAMI 25(7), pp. 787–800.

Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tappen, M. and Rother, C., 2006. A comparative study of energy minimization methods for markov random fields. In: ECCV, Vol. 2, pp. 19–26.

Wang, L., Gong, M., Gong, M. and Yang, R., 2006. How far can we go with local optimization in real-time stereo matching. In: 3DPVT, pp. 129–136.