Invited Paper

An Extension to the Rank Constraint for Improving the Reliability of Stereo Matching Using the Rank Transform

Jasmine Banks*, Mohammed Bennamoun*, Kurt Kubik* & Peter Corke**

*Space Centre for Satellite Navigation,
Queensland University of Technology,
Brisbane 4000, Queensland, Australia.
{j.banks, m.bennamoun, k.kubik}@qut.edu.au

**Centre for Mining Technology and Equipment,
Pinjarra Hills 4069,
Queensland, Australia.
pic@cat.csiro.au

ABSTRACT

The rank transform is one non-parametric transform which has been applied to the stereo matching problem. The advantages of this transform include its invariance to radiometric distortion, and its amenability to hardware implementation. This paper describes the derivation of the *rank constraint* for matching using the rank transform. Previous work has shown that this constraint was capable of resolving ambiguous matches, thereby improving match reliability. A new matching algorithm incorporating this constraint was also proposed. This paper extends on this previous work, by proposing a matching algorithm which uses a 2-dimensional match surface, in which the match score is computed for every possible template and match window combination. The principal advantage of this algorithm is that the use of the match surface enforces the left–right consistency and uniqueness constraints, thus improving the algorithm's ability to remove invalid matches. Experimental results for a number of test stereo pairs show that the new algorithm is capable of identifying and removing a large number of incorrect matches, particularly in the case of occlusions.

1. INTRODUCTION

A fundamental problem faced by stereo vision algorithms is that of determining correspondences between two images which comprise a stereo pair. The rank transform is one non-parametric transform which has been applied to the stereo matching problem[1]. Constraints such as the left–right consistency criterion and removal of anomalous disparities have been widely used by matching algorithms in order to identify and remove invalid matches. Although powerful, most of these constraints have little theoretical basis. This paper first of all describes previous work[2] in which one constraint which must be satisfied for a correct match was derived. This was termed the *rank constraint*. Experimental results have shown that this constraint is capable of resolving ambiguous matches, thereby improving matching reliability.

This previous work was based on computing the SAD match score for a template window and a series of candidate windows, resulting in a 1-dimensional match function. This paper proposes an extension to this previous work, in which the SAD match score is computed for every possible template and candidate window combination. This results in a 2-dimensional match surface. The main advantage of this method is that the structure of the match surface enforces the left–right and uniqueness constraints, resulting in improved ability to reject invalid matches. Some experimental results obtained from this algorithm are shown, and compared with those obtained from the previous algorithm. Finally, conclusions arising from this work are summarised.

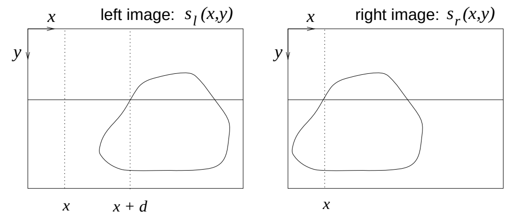


Fig. 1 Stereo pair to be matched. The left and right images are denoted $s_l(x, y)$ and $s_r(x, y)$ respectively.

2. MATCHING USING THE RANK TRANSFORM

Image Notation

As shown in Figure 1, the left and right images which comprise the stereo pair are denoted $s_l(x, y)$ and $s_r(x, y)$ respectively. Assuming the images are epipolar[3], a simple model for the relationship between corresponding image points is given by [4]:

$$s_r(x, y) = As_l(ax + d, y) + B + N(x, y) \qquad (1)$$

where A and B are the contrast and brightness factors for radiometric distortion, and N represents noise. The terms a and d represent geometric distortion, and in particular, d is the disparity difference we wish to find.

Matching Using the Rank Transform

The matching process accepts an epipolar aligned stereo pair as input, and produces a *disparity map* as output. Figure 2 illustrates the process. This section describes each step of this process in detail.

The Rank Transform: The rank transformation process involves passing the rank window over the image, and at each point, counting the number of pixels in the rank window whose value is less than the centre pixel. The rank

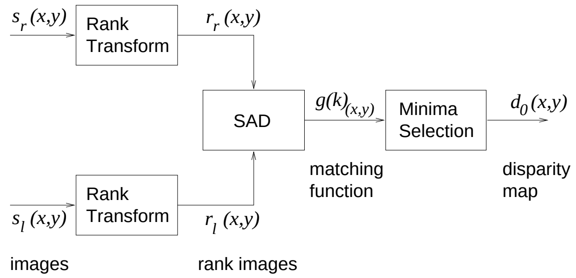


Fig. 2 The matching process using the rank transform.

transform may be expressed as:

$$r(x,y) = R - \sum_{(i,j)\in W} U\big[s(x+i,y+j) - s(x,y)\big] \qquad (2)$$

where $U[t]$ is the unit step function, R is the number of pixels in the rank window and $(i,j) \in W$ indicates the neighbourhood of the rank window.

Computation of SAD Matching Metric: The Sum of Absolute Differences (SAD) matching metric[5] provides a measure of the *similarity* between pixel regions. Given a *template* window, centred on $r_r(x,y)$, the SAD metric is computed for a series of *candidate* windows, centred on $r_l(x+k,y)$, where the test disparity k is varied in integer increments from 0 to d_{max}. This series of SAD scores is referred to as a *match function*, and is computed as follows

$$\begin{aligned} g(k)_{(x,y)} &= \mathrm{SAD}\big(r_r(x,y), r_l(x+k,y)\big) \\ &= \sum_{(m,n)\in M} \big|r_r(x+m,y+n) - r_l(x+k+m,y+n)\big| \end{aligned} \qquad (3)$$

where $(m,n) \in M$ indicates summation of over the match window. Example of match functions are shown in Figure 4.

Selection of Minima: This involves selection of the disparity at which the match function is a minimum, which is the disparity at which the template and candidate windows are most similar. The selection of the minima is expressed as

$$\begin{aligned} d_0(x,y) = k \mid \big[&\big(\forall p\,(p<k)\, g(k)_{(x,y)} < g(p)_{(x,y)}\big) \\ &\text{AND } \big(\forall p\,(p>k)\, g(k)_{(x,y)} \le g(p)_{(x,y)}\big)\big] \end{aligned} \qquad (4)$$

The disparity values $d_0(x,y)$ together comprise a *disparity map*. It is desirable that these disparity values correspond to the d term in Equation (1), the true difference in location of the pixel patterns in $s_l(x,y)$ and $s_r(x,y)$.

3. THE RANK CONSTRAINT

Derivation

An analytic expression for the match function is formulated, and is used to derive one possible constraint for a correct match. The rank transformed images are computed from Equation (2) as follows

$$r_l(x,y) = R - \sum_{(i,j)\in W} U\big[s_l(x+i,y+j) - s_l(x,y)\big] \qquad (5)$$

$$r_r(x,y) = R - \sum_{(i,j)\in W} U\big[s_r(x+i,y+j) - s_r(x,y)\big] \qquad (6)$$

Substituting Equations (5) and (6) into the match function of Equation (3) results in

$$g(k)_{(x,y)} = \sum_{(m,n)\in M} \Big| \sum_{(i,j)\in W} U(D_l) - U(D_r) \Big| \qquad (7)$$

where

$$\begin{aligned} D_l = &\ s_l(x+k+m+i, y+n+j) \\ &- s_l(x+k+m, y+n) \end{aligned} \qquad (8)$$

$$D_r = s_r(x+m+i, y+n+j) - s_r(x+m,y+n) \qquad (9)$$

The optimum disparity is selected as the disparity k at which Equation (3) is a minimum. Differentiating Equation (3) with respect to the test disparity, k results in

$$\begin{aligned} g'(k)_{(x,y)} = \sum_{(m,n)\in M} \Big\{ &sgn\Big(\sum_{(i,j)\in W} U(D_l) - U(D_r) \Big) \\ &\Big(\sum_{(i,j)\in W} \delta(D_l)(D'_l) \Big) \Big\} \end{aligned} \qquad (10)$$

where D'_l is given by

$$\begin{aligned} D'_l = &\ s'_l(x+k+m+i, y+n+j) \\ &- s'_l(x+k+m, y+n) \end{aligned} \qquad (11)$$

The first derivative will be zero if the *sgn* term is zero for all (m,n,i,j). Since the only possible values for the function $U[x]$ are 0 or 1, the conditions for the *sgn* term to be zero are

$$\begin{matrix} U[D_l] = 0 & \text{or} & U[D_l] = 1 \\ U[D_r] = 0 & & U[D_r] = 1 \end{matrix} \qquad (12)$$

for all (m,n,i,j). Substituting Equations (8) and (9) into (12) yields

$$\begin{aligned} &s_l(x+k+m, y+n) > s_l(x+k+m+i, y+n+j), \\ &s_r(x+m, y+n) > s_r(x+m+i, y+n+j) \end{aligned} \qquad (13)$$

$$\begin{aligned} &s_l(x+k+m+i, y+n+j) \ge s_l(x+k+m, y+n), \\ &s_r(x+m+i, y+n+j) \ge s_r(x+m, y+n) \end{aligned} \qquad (14)$$

Equations (13) and (14) together form one constraint for a correct match. This has been termed the *rank constraint*. The constraint depends on the relative ordering of pixels in rank windows centred on every pixel of the template and candidate windows. It could be re-interpreted as follows

if (left image window centre pixel > neighbourhood pixel)
then (right image window centre pixel must be
> neighbourhood pixel)

or

if (left image window neighbourhood pixel ≥ centre pixel)
then (right image window neighbourhood pixel must be
≥ centre pixel)

Essentially, this constraint enforces consistency in ordering with respect to the centre pixel, between template and candidate windows.

Testing and Results

A *constraint evaluation score* may be computed by examining the pixel values in the template and candidate windows, and incrementing a score each time the constraint is *not satisfied*. This score may be computed for each test disparity, resulting in a *constraint evaluation function*. This is derived from the conditions of Equation (12) as follows:

$$\begin{aligned} &P(k)_{(x,y)} \\ &= \sum_{(m,n)\in M} \sum_{(i,j)\in W} \overline{((\overline{U[D_l]} \cdot \overline{U[D_r]}) \oplus (U[D_l] \cdot U[D_r]))} \\ &= \sum_{(m,n)\in M} \sum_{(i,j)\in W} U[D_l] \otimes U[D_r] \qquad (15) \end{aligned}$$

where (x, y) is the centre of the template window, and k is the test disparity. The symbols $\cdot$, $\oplus$ and $\otimes$ denote the boolean AND, OR and XOR operations, while a bar above an item indicates the NOT operation. Equation (15) is designed to compute a measure of the extent to which the rank constraint is satisfied. When the constraint is perfectly satisfied, Equation (15) will return zero.

Initial testing of the constraint was carried out using a contrived stereo pair consisting of two images displaced by a known amount. Since the image is matched with itself, the true disparity, d, is precisely known, and furthermore, there is no noise or radiometric distortion, ie, $N(x, y) = 0$, $a = 1$, $A = 1$ and $B = 0$. For all test cases, the constraint was always completely satisfied at the correct disparity, thus confirming the validity of the constraint[7].

Experimental results were also obtained using the stereo pair of Figure 3. Since this is an actual stereo pair, the values of the noise and distortion terms, a, A, B, and N, are undetermined. Figure 4 shows the match functions and constraint functions obtained for a template windows located at $(x, y) = (280, 139)$ and $(168, 151)$ respectively. In (b), the constraint was most satisfied at the test disparity 20, which is the same disparity for which the corresponding match function of (a) is a minimum. This in fact corresponds to the correct disparity in this case. Another point to note is that the constraint function does not actually reach zero at this disparity, due to the presence of unknown values of geometric distortion, radiometric distortion and noise.

Figure 4(c) and (d) show an ambiguous situation where the match function has two main minima, at disparities of 3 and 23, where 23 is in fact the correct disparity. Unfortunately, the minima at a disparity of 3 is slightly less than the one at 23, which would result in an incorrect match being returned. However, the constraint evaluation function has one dominant minima at a disparity of 23, and is thus able to be used to resolve the ambiguous match. Further testing has shown that the rank constraint is able to resolve a large number of cases of ambiguous matches, not only for the stereo pair of Figure 3, but for other test pairs used. More detail of the derivation and testing of the rank constraint is given in [7].

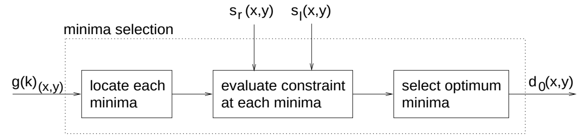
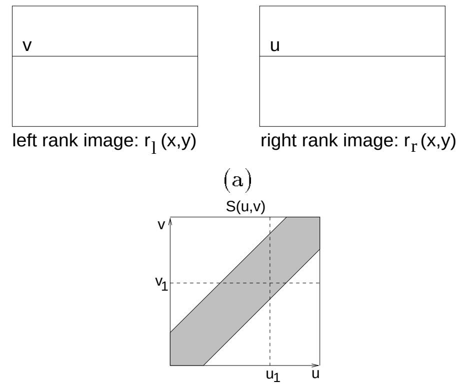


Fig. 5 Modified minima selection stage, incorporating the rank constraint.



(b)

Fig. 6 Computation of the match surface. (a) Corresponding epipolar lines in the rank transformed stereo pair. (b) Match surface consisting of an array of SAD scores. SAD scores need only be computed for $d_{min} \leq d \leq d_{max}$.

Matching Algorithm Incorporating the Constraint

The rank order constraint has been incorporated into the minima selection stage of the matching algorithm of Figure 2. This is shown in Figure 5. Instead of selecting the smallest minima of the match function, the modified algorithm selects the minima for which the constraint is most satisfied. The modified algorithm for minima selection is as follows:

compute match function

for each minima of the match function

    compute the constraint score

end

select the minima with the optimum constraint score

4. A NEW MATCHING ALGORITHM

The Match Surface

In the matching algorithm of Figure 2, the SAD score is computed for a template window and a series of candidate windows, resulting in a match function. However, given a pair of epipolar lines, it is possible to compute the SAD score for every possible template and candidate window combination. This results in a 2-dimensional match surface, as shown in Figure 6.

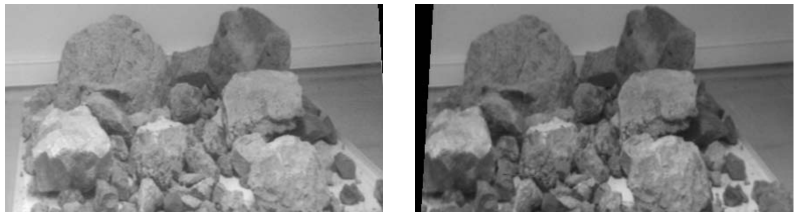


Fig. 3: Stereo pair of rocks[6].

The match surface shown in Figure 6 is defined as $S(u,v)$, where the u and v axes represent the x coordinate of the window in the right and left images respectively. A location (u,v) in the match surface is the SAD score obtained from the windows centred on $r_r(u,y)$ and $r_l(v,y)$. The disparity is computed from the difference in x location of the two windows, ie, $d = v - u$.

A horizontal line in the match surface is equivalent to a match function. For example, a vertical line situated at $u = u_1$ in Figure 6 is equivalent to a match function for a template window of $r_r(u_1,y)$ and candidate windows ranging from $r_r(u_1+d_{min},y)$ to $r_r(u_1+d_{max},y)$, where d_{min} and d_{max} are the minimum and maximum disparity values. The match surface in fact need only be computed for a diagonal band of values, ie, $d_{min} \leq d \leq d_{max}$. This means that a large proportion of values in the match surface array will be unused. A more efficient method of storing the array is described in [8]. However, for simplicity the discussion of the matching algorithm in this paper will be in terms of the match score array shown in Figure 6.

The Extended Rank Constraint

It is possible to re-define the rank constraint derived in Section 3 in terms of the match surface. The match surface may be expressed as

$$\begin{aligned} S(u,v) &= \mathrm{SAD}(r_r(v,y), r_l(u,y)) \\ &= \sum_{(m,n)\in M} \left| \sum_{(i,j)\in W} U(D_l) - U(D_r) \right| \end{aligned} \tag{16}$$

where

$$D_l = s_l(u+m+i, y+n+j) - s_l(u+m, y+n) \tag{17}$$

$$D_r = s_r(v+m+i, y+n+j) - s_r(v+m, y+n) \tag{18}$$

Also, u and v are related by

$$u = v + d \tag{19}$$

where d is the test disparity. For the rank constraint to be satisfied, the conditions of Equation (12) must hold. Consider the case where v is constant, ie, $v = v_1$. In this case the template window is located in the left image at (v_1,y) and the candidate window moves in integer increments in the right image from (v_1+d_{min}) to (v_1+d_{max}). This is actually equivalent to a match function, which is a horizontal line in the match surface of Figure 6. Substituting Equation (19) into Equations (17) and (18) yields

$$\begin{aligned} D_l = {} & s_l(v_1+d+m+i, y+n+j) \\ & - s_l(v_1+d+m, y+n) \end{aligned} \tag{20}$$

$$D_r = s_r(v_1+m+i, y+n+j) - s_l(v_1+m, y+n) \tag{21}$$

The conditions for the rank constraint to be satisfied therefore become

$$\begin{aligned} s_l(v_1+d+m, y+n) &> s_l(v_1+d+m+i, y+n+j) \\ s_r(v_1+m, y+n) &> s_r(v_1+m+i, y+n+j) \end{aligned} \tag{22}$$

$$\begin{aligned} s_l(v_1+d+m+i, y+n+j) &\geq s_l(v_1+d+m, y+n) \\ s_r(v_1+m+i, y+n+j) &\geq s_r(v_1+m, y+n) \end{aligned} \tag{23}$$

These conditions for a correct match can also be expressed in terms of a vertical match function in the match surface. In this case u is constant, ie, $u = u_1$. The template window will be located in the right image at (u_1,y), and the candidate window moves in integer increments in the left image from (u_1-d_{min}) to (u_1-d_{max}). This is equivalent to a match function, which is a vertical line in the match surface of Figure 6. Substituting Equation (19) into Equations (17) and (18) yields

$$D_l = s_l(u_1+m+i, y+n+j) - s_l(u_1+m, y+n) \tag{24}$$

$$\begin{aligned} D_r = {} & s_r(u_1-d+m+i, y+n+j) \\ & - s_l(u_1-d+m, y+n) \end{aligned} \tag{25}$$

The conditions for the rank constraint to be satisfied therefore become

$$\begin{aligned} s_l(u_1+m, y+n) &> s_l(u_1+m+i, y+n+j) \\ s_r(u_1-d+m, y+n) &> s_r(u_1-d+m+i, y+n+j) \end{aligned} \tag{26}$$

$$\begin{aligned} s_l(u_1+m+i, y+n+j) &\geq s_l(u_1+m, y+n) \\ s_r(u_1-d+m+i, y+n+j) &\geq s_r(u_1-d+m, y+n) \end{aligned} \tag{27}$$

The conditions of Equations (22) and (23) apply to a horizontal match function in the match surface of Figure 6, while those of Equations (26) and (27) are for a vertical match function. The conditions for the horizontal and vertical match functions are essentially equivalent, differing only in whether they are expressed in terms of the left or right image x coordinate, v or u respectively. In each case, the disparity d is the variable.

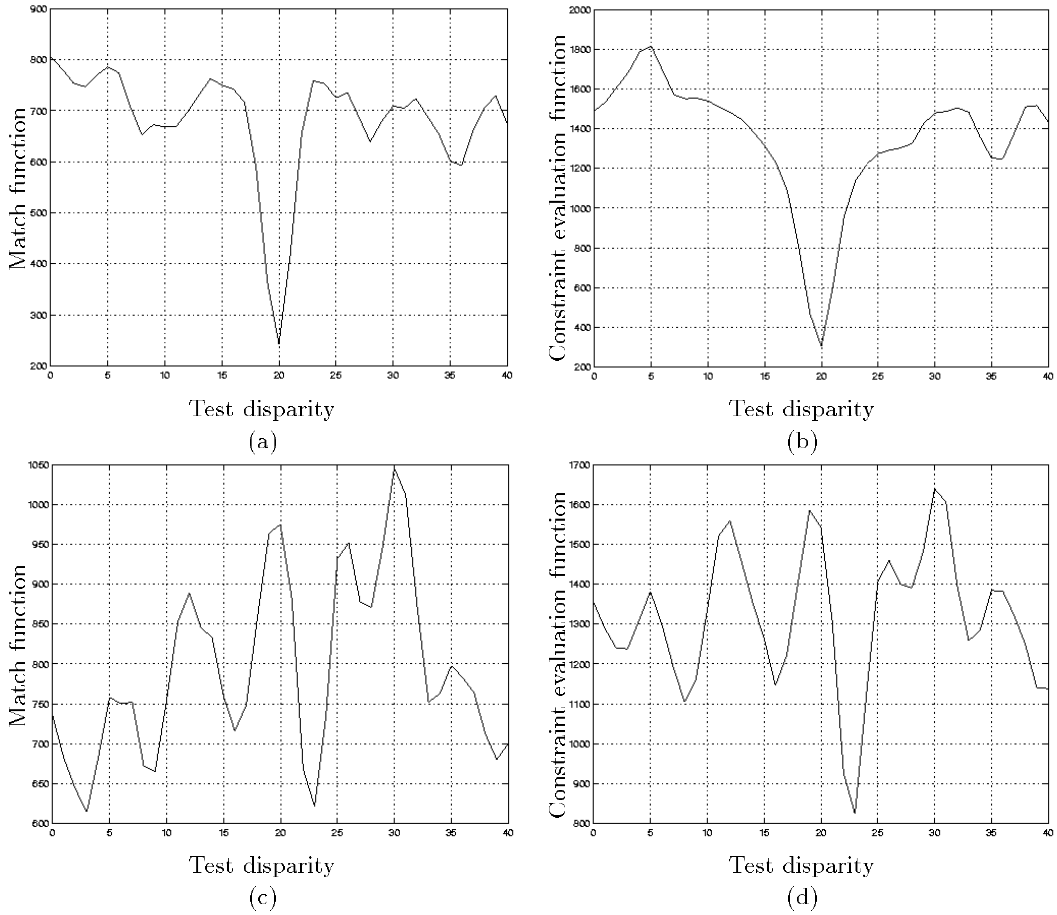


Fig. 4: Examples of match functions and constraint evaluation functions obtained from the stereo pair of Figure 3. (a) match function (b) constraint evaluation function obtained using a template match window centred $(280, 139)$. (c) match function (d) constraint evaluation function obtained using a template window centred on $(168, 151)$.

Matching Algorithm Using Match Surface

A new matching algorithm was devised which uses the concept of the match surface in addition to the rank constraint. The process is illustrated in Figure 7. Given a pair of corresponding epipolar lines in the rank transformed images, this algorithm first of all computes the match surface. This involves computing the SAD score for all possible template and candidate window combinations. Next, minima are identified in the match surface, as locations which are less than their 4-neighbours. Such points are minima of the horizontal and vertical intersecting match functions. Each minima at $S(u, v)$ corresponds to a possible match between $s_r(u, y)$ and $s_l(v, y)$. Once minima have been identified, the constraint score is then computed at each minima.

Next, a set of valid matches are selected. A match is selected as valid if it has the minimum constraint score in both horizontal and vertical directions. Selection of valid matches in this manner additionally enforces the following constraints:

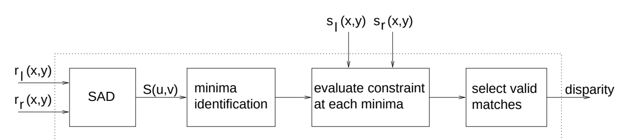
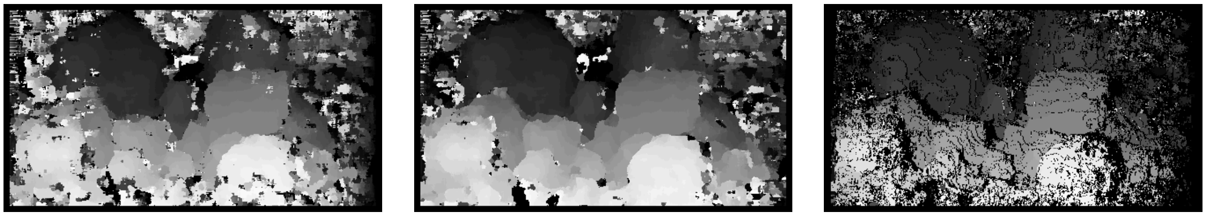


Fig. 7 Extended matching algorithm, which uses the match surface instead of a match function.

1. **left–right consistency [8]** If a location $s_r(x, y)$ matches to a location $s_l(x + d, y)$, then the location $s_l(x + d, y)$ must match back to $s_r(x, y)$. This constraint is implicitly enforced by only selecting matches as valid if the constraint score is minimum in both the horizontal and vertical directions.
2. **uniqueness** A particular location can only have one match. This is enforced by ensuring that there is only one match along each horizontal and vertical line.

Once a set of valid matches is selected, the disparity of each location along the epipolar line is found using the relationship of Equation (19).



(a) (b) (c)

Fig. 8: Disparity results obtained for the stereo pair of Figure 3, using (a) original rank matching algorithm of Figure 2, (b) algorithm incorporating rank constraint of Figure 5 and (c) algorithm using the match surface of Figure 7.

5. RESULTS

The matching algorithms described in this paper were tested using a number of stereo pairs, including the stereo pair of Figure 3. In all cases, a match window of size 11×11 and a rank window of size 5×5 were used. Figure 8 shows the disparity results obtained using (a) the rank transform based matching algorithm of Figure 2, and (b) the algorithm of Figure 5 incorporating the rank constraint. The disparity map results obtained for the algorithm using the 2-dimensional match surface instead of the 1-dimensional match function are shown in Figure 8(c). In all these images, the grey scale value indicates the value of disparity at that point. In Figure 8(c), black areas indicate locations where no valid match was returned, due to the failure of the left–right consistency criterion. In such cases, there was no minima of the match surface having the best constraint score in both horizontal and vertical directions.

6. DISCUSSION AND CONCLUSION

This paper has first of all described the derivation of the *rank constraint*, which must be satisfied for a correct match. Experimental work has proved that this constraint is capable of resolving ambiguous matches, thus improving match reliability. From a visual inspection of the disparity results of Figure 8(a) and (b), it can be seen that the algorithm using the rank constraint has correctly matched some areas which were not matched using the original algorithm. Most of these areas corresponded to ambiguous matches.

This paper has also presented an extension to this algorithm, which uses a 2-dimensional match surface instead of a one-dimensional match function. The principal advantage of this method is that selection of a set of valid matches from the match surface implicitly enforces the left–right consistency and uniqueness constraints, thus improving the ability of the matching algorithm to reject invalid matches. The disparity results of Figure 8(c) show that a significant number of invalid matches were removed. Most of these correspond to occlusions. However, a number of invalid matches remain, most of which are in the vicinity of the bland wall in the background of the scene. Therefore, one possible extension to this algorithm could involve identification of bland regions in the images, which could then be excluded from the matching process.

It also can be seen from Figure 8(c) that a significant number of valid matches are also removed by this matching algorithm. In such cases, instead of just flagging the match as invalid, it may be possible to select an alternative match, by more fully utilising the information computed for the match surface. For example, if there is no minima which has the best constraint score in both the horizontal and vertical directions, then another minima of the match surface may be selected as the correct match. Investigation of an appropriate method of selecting an alternative match is another area for future work associated with this algorithm.

5. REFERENCES

[1] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. *3rd European Conference on Computer Vision*, Stockholm, 1994.

[2] J. Banks, M. Bennamoun, P. Corke, and K. Kubik. A constraint to improve the reliability of stereo matching using the rank transform. *ICASSP'99*, Phoenix, Arizona, Mar. 1999, pp. 3321–3324.

[3] S. Barnard and M. Fischler. Computational stereo. *Computing Surveys*, Vol. 14, No. 4, Dec. 1982, pp. 553–572.

[4] C. Tomasi and R. Manduchi. Stereo matching as a nearest neighbour problem. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 3, Mar. 1998, pp. 333–340.

[5] P. Aschwanden and W. Guggenbühl. Experimental results from a comparative study on correlation-type registration algorithms. *Robust Computer Vision* : Wickmann, 1993, pp. 268–289.

[6] R. Bolles, H. Baker, and M. Hannah. The JISCT stereo evaluation. *Image Understanding Workshop*, DARPA, 1993, pp. 263–274.

[7] J. Banks, M. Bennamoun, P. Corke, and K. Kubik. Reliability analysis of the rank transform for stereo matching. *submitted to IEEE Transactions in Image Processing*, 1999.

[8] S. Cochran and G. Medioni. 3-D surface description from binocular stereo. *IEEE Trans. on pattern analysis and machine intelligence*, Vol. 14, No. 10, Oct. 1992, pp. 981–994.